



PhD Thesis

Algorithms for contactless fingerprint recognition

submitted in satisfaction of the requirements for the degree

Doctor of Science in Physics

of the TU Wien, Faculty of Physics

Dissertation

Algorithmen für kontaktlose Fingerabdruckerkennung

ausgeführt zum Zwecke der Erlangung des akademischen Grads

Doktor der technischen Wissenschaften

eingereicht an der TU Wien, Fakultät für Physik

Dipl.-Ing. **Laurenz Ruzicka**, BSc

Matr.Nr.: 01619916

Betreuung: Associate Prof. Dr.techn. Dipl.-Ing. **Clemens Heitzinger**
Institute of Information Systems Engineering
Machine Learning
Technische Universität Wien
Erzherzog-Johann-Platz 1, 1040 Wien, Österreich

Betreuung: Dr. rer. nat. **Bernhard Kohn**
Digital Safety & Security
Austrian Institute of Technology
Giefinggasse 4, 1210 Wien

Begutachtung: Prof. Dr. **Kiran Raja**
Department of Computer Science
Faculty of Information Technology and Electrical Engineering
Norwegian University of Science and Technology
Ametyst-bygget, 207, Gjøvik, Teknologivegen, Norway

Begutachtung: Prof. Dr. **Patrick Bours**
Department of Information Security and Communication Technology
Faculty of Information Technology and Electrical Engineering
Norwegian University of Science and Technology
T511 Topasbygget Kallerud, Gjøvik, Norway

Wien, im October 2025

Kurzfassung

Der Fingerabdruck ist eine der am weitesten verbreiteten biometrischen Modalitäten und eignet sich hervorragend zur Identitätsüberprüfung, vor allem aufgrund der Einzigartigkeit und Beständigkeit von Papillarleisten und deren Mustern. Fingerabdruckerkennung wird von Strafverfolgung und Grenzkontrolle bis hin zur Absicherung persönlicher Geräte und finanzieller Transaktionen benutzt. Während Fingerabdrücke traditionell durch direkten Kontakt erfasst werden, bieten kontaktlose Methoden Vorteile bezüglich Hygiene, Sicherheit und Benutzerfreundlichkeit. Die praktische Anwendung wird jedoch durch die freie Präsentation der Hand, wechselnden Aufnahmeumgebungen sowie der Notwendigkeit mit bestehenden, kontaktbasierten Fingerabdruckdatenbanken interoperabel zu bleiben, erschwert. Diese Dissertation präsentiert eine Reihe von Algorithmen und Methoden, um diese Probleme entlang der gesamten Erkennungspipeline zu adressieren. Die Beiträge umfassen Lösungen für die Bildnormalisierung, die Qualitätssicherung der Daten, die robuste Validierung und die Privatsphäre schützende Fingerabdruckvergleiche, wodurch Genauigkeit, Zuverlässigkeit und der Datenschutz von kontaktlosen Fingerabdrucksystemen verbessert werden.

Ein notwendiger Schritt in jedem kontaktlosen Sensorsystem ist die präzise Segmentierung der Fingerspitze vom Hintergrund, welcher komplex und dynamisch sein kann. Diese Arbeit stellt drei neue Deep-Learning-Architekturen für diese Aufgabe vor. Das erste Modell ist eine Variation der U-Net Architektur und wurde für bereits zugeschnittene Einzelfingerbilder entwickelt. Es übertrifft bestehende State-of-the-Art (SOTA) Segmentierungsmodelle in Genauigkeit [217]. Darauf aufbauend wurde *FingerUNeSt++* entwickelt, welches einen ResNeSt-Encoder mit einem UNet++-ähnlichen Decoder kombiniert und eine durchschnittliche Intersection-over-Union (mIoU) von 99% erreicht und damit bestehende SOTA Methoden erneut übertrifft. Das dritte Modell, *TipSegNet*, macht den vorgelagerten Fingerdetektionsschritt überflüssig, indem es alle vier Fingerspitzen direkt aus einem Bild der Hand segmentiert und klassifiziert. Durch die Nutzung eines ResNeXt-101-Backbones mit einem Feature Pyramid Network (FPN) erreicht *TipSegNet* eine SOTA-Leistung mit einer mIoU von 99% und Genauigkeit von 100%.

Um die Interoperabilität mit kontaktbasierten Systemen zu verbessern, müssen die geometrischen Verzerrungen bei kontaktlosen Aufnahmen korrigiert werden. Im Rahmen dieser Dissertation wurde eine Verarbeitungspipeline entwickelt, welche Rotationen sowohl in der Bildebene (Hand-Drehung) als auch aus der Bildebene heraus (Hand-Kippung) korrigiert und anschließend die Fingeroberflächentextur mittels parametrischer Entzerrungsmodelle ebnet. Dabei kann auf verschiedene parametrische Entzerrungsmodelle zurückgegriffen werden, welche das Abrollen verschiedener Fingerformen auf dem Sensor simulieren. Für die Drehkorrektur wird die im vorherigen Absatz beschriebene Segmentierungsmaske zur Bestimmung der zentralen Fingerachse verwendet. Für die Kippkorrektur kommt die Segmentierungsmaske erneut zum Einsatz um die Querschnittsbreite eines ellipsoiden Fingermodells zu bestimmen und um dann, mithilfe des detektierten Cores als Referenzpunkt, die Kippkorrektur durchzuführen. Bei der Evaluierung reduzierte eine fingerspezifische Anwendung dieser Prozessschritte die Equal-Error-Rate (EER) für den Abgleich von kontaktlosen mit kontaktbasierten Abdrücken um relative 36,9% (von 1,57% auf 0,99% EER). Eine zusätzliche empirische Analyse der Position des Cores von über 40.000 Fingerabdrücken zeigte, dass dessen Position nicht geometrisch zentriert ist und systematische, modalitätsbedingte Abweichungen sowie eine natürliche Variabilität von 6-12% der

Fingerbreite aufweist. Diese Studie quantifiziert eine grundlegende Grenze für die Genauigkeit von Präsentationskorrekturmethode, die ausschließlich auf dem Core als Referenzpunkt basieren, und identifiziert die nicht-zentrale Fischer-(NCF)-Verteilung als passendes statistisches Modell für dessen Position. Eine fingerspezifische Auswertung wurde ebenfalls durchgeführt.

Für die finale Identifikation eines Fingerabdrucks ist dessen Qualität, sowie die Qualität der Referenzaufnahme, entscheidend. Diese Arbeit begegnet dem Mangel an spezifischen Qualitätsmetriken für mobile kontaktlose Fingerabdrücke durch die Evaluierung von *MCLFIQ*, einer Adaption des etablierten NFIQ 2 Frameworks. Durch das erneute Trainieren des Random-Forest-Klassifikators von NFIQ 2 auf modalitätsspezifischen, synthetischen Daten zeigt *MCLFIQ* eine verbesserte Leistung bei der Vorhersage der Qualität von kontaktlos aufgenommenen Fingerabdrücke im Vergleich zum ursprünglichen NFIQ 2.2 und anderen Standardmetriken. Das angepasste Modell priorisiert Merkmale, die mit Bildschärfe und lokaler Klarheit der Papillarleisten zusammenhängen, welche die dominanten Qualitätsfaktoren bei mobilen Aufnahmen sind.

Darüber hinaus wird in dieser Dissertation ein neuartiges, self-supervised Framework zur Erkennung von Stücklungs-Artefakten vorgestellt. Diese Artefakte können während des Zusammensetzens von mehreren Finger-Einzelaufnahmen zu einem Master-Fingerabdruck entstehen. Dieses Problem wird von Standard-Qualitätsmetriken, wie zum Beispiel NFIQ und auch von adaptierten Metriken wie *MCLFIQ*, nicht erfasst. Durch die programmatische Erzeugung von Artefakten wurde ein Deep-Learning-Modell trainiert um diese Defekte zu erkennen, ohne manuelle Annotierungsarbeit zu benötigen. Der resultierende Detektor ist genau ($\text{IoU} > 90\%$), robust gegenüber anderen Qualitätsmängeln des Fingerabdrucks und generalisiert gut über verschiedene Fingerabdruckmodalitäten (kontaktlos, gerollt, flach).

Die Validierung dieser komplexen Algorithmen erfordert standardisierte und wiederholbare Testmethoden. Zu diesem Zweck wurde die Herstellung von physikalisch-synthetischen Fingerphantomen untersucht. Verschiedene Herstellungstechniken (direkte Lasergravur, CNC-Fräsen, 3D-Druck von Gussformen) und Materialien (Silikon, Gelatine, Elastomer) wurden getestet. Die Arbeit zeigt, dass hochpräzise, stabile und physikalisch realistische 3D-Phantome erfolgreich hergestellt werden können, insbesondere durch den Guss von Silikon in hochpräzise Gussformen. Diese Phantome replizieren erfolgreich einen Fingerabdruck und dienen als Werkzeuge für die objektive und wiederholbare Evaluierung der Leistungsfähigkeit von Sensoren und Verarbeitungsalgorithmen, ohne den Einsatz von echten Fingerabdrücken.

Um den Einsatz der Fingerabdruckererkennung in datenschutz-sensitiven Anwendungen wie dem Abgleich von Fahndungslisten an Grenzkontrollen zu ermöglichen, demonstriert diese Arbeit eine praktische, die Privatsphäre wahrende Lösung. Ein SOTA Minutien-basierter Vergleichsalgorithmus, SourceAFIS, wurde erfolgreich für die Ausführung in einem Multiparty Computation (MPC) Framework adaptiert. Durch algorithmische Umstrukturierung von Operationen welche im MPC Protokoll stark rechenintensiv sind und die Nutzung von Klartextberechnungen für nicht-sensitive Zwischendaten erreicht die Implementierung eine Abgleichzeit von ungefähr 17 Sekunden. Diese Geschwindigkeit wird unter Beibehaltung einer hohen, mit dem Klartext-Algorithmen vergleichbaren Erkennungsgenauigkeit erzielt, wodurch ein sicherer Fingerabdruckabgleich unter realistischen Bedingungen praktisch umsetzbar wird.

Abschließend zeigt diese Arbeit, dass die Herausforderungen, die mit einer vollständigen End-to-End-Pipeline für die sichere, kontaktlose Fingerabdruckererkennung einhergehen, gelöst werden können. Durch die Integration der entwickelten Lösungen wird der Weg von den erfassten Rohbildern zu standardisierten und interoperablen biometrischen Templates ermöglicht. Die Relevanz dieses integrierten Ansatzes liegt darin, Hochsicherheitsanwendungen auf kontaktlosen Systemen zu ermöglichen, wie beispielsweise den Abgleich mit Fahndungslisten an Grenzkontrollen durch einen Beamten mittels eines Standard-Smartphones.

Abstract

Fingerprint recognition is a widely used biometric modality because of its uniqueness and persistence of friction ridge patterns, which provide a reliable means of verifying identity. It is a key technology in applications ranging from law enforcement and border control to securing personal devices and financial transactions. While traditionally acquired through direct contact, contactless methods offer advantages in hygiene and user convenience. However, practical deployment is challenged by the unconstrained nature of the acquisition process, which introduces variations in finger pose, illumination, and scale, while still requiring to be interoperability with large, legacy contact-based fingerprint databases. This thesis presents a set of algorithms and methodologies to address these problems across the recognition pipeline, from image capture to secure template comparison. The contributions include solutions for image normalization, data assurance, robust validation, and secure deployment, which improve the accuracy, reliability, and privacy of contactless fingerprint systems.

A necessary step in any contactless pipeline is the accurate segmentation of the fingertip from its background, which is often complex and variable. This work proposes three novel deep learning architectures for this task. The first is a custom U-Net-based model for pre-cropped single-finger images that outperforms existing segmentation models [217]. This was improved with *FingerUNeSt++*, which combines a ResNeSt encoder with a UNet++-like decoder, achieving a mean Intersection-over-Union (mIoU) of 99% on the test set. The third model, *TipSegNet*, removes the need for a separate finger detection step by segmenting and labeling all four fingertips directly from a whole-hand image. Using a ResNeXt-101 backbone with a Feature Pyramid Network (FPN) to handle multi-scale objects, *TipSegNet* obtains an mIoU of 99% and an accuracy of 100%.

Interoperability with contact-based systems requires correcting the geometric distortions in contactless captures. This research developed a processing pipeline to correct for in-plane (yaw) and out-of-plane (roll) rotations, which then flattens the fingertip texture using parametric unwarping models. The pipeline uses the segmentation mask for yaw correction and an elliptical finger model with the detected core for roll correction. On an operational dataset, a finger-wise optimized application of the pipeline reduced the Equal Error Rate (EER) for contactless-to-contact-based comparison by a relative 36.9% (from 1.57% to 0.99%). A large-scale empirical analysis of the fingerprint core’s position across over 40,000 samples showed that its location is not geometrically centered, with systematic, modality-induced biases and a natural variability of 6-12% of the finger’s width. This study quantifies a limit on the accuracy of alignment methods that rely on the core and identifies the Non-Central Fischer (NCF) distribution as the best-fitting statistical model for its position for most fingers, including a finger-dependent analysis.

The quality of a captured sample affects recognition performance. This work addresses the absence of dedicated quality metrics for mobile contactless fingerprints by adapting the established NFIQ 2 framework, resulting in *MCLFIQ*. By retraining the NFIQ 2 random forest classifier on modality-specific synthetic data, *MCLFIQ* shows improved performance in predicting the utility of contactless samples compared to the original NFIQ 2.2 and other baselines. The new model prioritizes features related to image sharpness and local ridge clarity, which are key quality factors in mobile captures. Additionally, this thesis introduces a self-supervised framework to detect structural artifacts from fingerprint mosaicking, a problem not handled by standard quality

metrics. A deep learning model was trained on programmatically generated artifacts to detect these defects without manual annotation. The resulting detector is accurate (IoU > 0.9), robust to other quality defects, and generalizes across different fingerprint modalities (contactless, rolled, slap).

The validation of these algorithms requires standardized testing methods. This research investigated the manufacturing of physical synthetic fingerprint phantoms, evaluating various fabrication techniques (direct laser engraving, CNC machining, 3D printing of molds) and materials (silicone, gelatin, elastomer). The work shows that high-fidelity, stable, and realistic 3D phantoms can be produced, particularly using silicone cast in high-precision molds. These phantoms replicate a known ground-truth pattern and provide a repeatable method for evaluating sensor and algorithm performance.

For privacy-sensitive applications like watchlist checks, this thesis presents a practical privacy-preserving comparison solution. The minutiae-based comparison algorithm SourceAFIS was adapted for execution within a Multiparty Computation (MPC) framework. Through algorithmic refactoring and the use of plaintext computations for non-sensitive data, the implementation achieves a comparison time of approximately 17 seconds while maintaining high recognition accuracy comparable to its plaintext counterpart. This makes secure remote fingerprint comparison practical for real-world scenarios.

Finally, this thesis shows that the challenges accompanying a complete, end-to-end pipeline for secure contactless fingerprint recognition can be solved. By integrating the developed solutions, one can go from raw images captured to standardized and interoperable biometric templates. The relevance of this integrated approach is its ability to enable high-security applications on contactless systems, such as an agent performing a watchlist check at a border using a standard smartphone.

Statement of Author's Contribution

This thesis is based on eight publications. My specific contributions to each publication are detailed herein.

For the publication *Towards Real-Time Privacy-Preserving Minutiae Matching* [1], where I was a co-author, my primary contribution was the development and assessment of the experimental framework and optimization strategies for the MPC-based algorithm. This work involved implementing and evaluating multiple minutiae extractors to establish performance baselines, designing and analyzing optimization strategies such as minutiae quality thresholding, and conducting the comprehensive performance evaluation of the final hybrid pipeline. Additionally, I contributed significantly to the writing and revision of the manuscript.

In the work presented in *MCLFIQ: Mobile Contactless Fingerprint Image Quality* [2], my primary responsibility as a co-author was the extensive experimental evaluation of the proposed MCLFIQ model. My contribution included generating Idkit comparison scores used for the Error-vs-Discard Characteristic (EDC) curve analysis and establishing a critical performance baseline using the AIT sharpness score. I also assisted in the writing and review of the publication.

I served as the lead author on the remaining six publications. In this capacity, I was responsible for the entire research lifecycle for each of these works: from the conception of the core idea, through the algorithm design, software implementation, and experimental validation, to the final preparation and revision of the manuscript. These publications are: *FingerUNeSt++: Improving Fingertip Segmentation in Contactless Fingerprint Imaging using Deep Learning* [3], *TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging* [4], *Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints Using Pose Correction and Unwarping* [5], *Towards Fingerprint Mosaicking Artifact Detection: A Self-Supervised Deep Learning Approach* [6], *Toward Synthetic Physical Fingerprint Targets* [7], *Centrality of the Fingerprint Core Location* [8].

Own Publications

- [1] J. Mader, F. Wohner, L. Ruzicka, and T. Loruenser. “Towards Real-Time Privacy-Preserving Minutiae Matching”. In: *Proceedings of the 23rd Workshop on Privacy in the Electronic Society*. WPES '24. New York, NY, USA: Association for Computing Machinery, Nov. 21, 2024, pp. 98–108. ISBN: 979-8-4007-1239-5. DOI: 10.1145/3689943.3695049. URL: <https://dl.acm.org/doi/10.1145/3689943.3695049> (visited on 11/24/2024).
- [2] J. Priesnitz, A. Weißenfeld, L. Ruzicka, C. Rathgeb, B. Strobl, R. Lessmann, and C. Busch. “MCLFIQ: Mobile Contactless Fingerprint Image Quality”. In: *IEEE Transactions on Biometrics, Behavior, and Identity Science* 6.2 (Apr. 2024), pp. 272–287. ISSN: 2637-6407. DOI: 10.1109/TBIOM.2024.3377686. URL: <https://ieeexplore.ieee.org/document/10473152> (visited on 07/07/2025).
- [3] L. Ruzicka, B. Kohn, and C. Heitzinger. *FingerUNeSt++: Improving Fingertip Segmentation in Contactless Fingerprint Imaging using Deep Learning*. IET Biometrics, Nov. 2024.
- [4] L. Ruzicka, B. Kohn, and C. Heitzinger. “TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging”. In: *Sensors* 25.6 (Mar. 14, 2025), p. 1824. ISSN: 1424-8220. DOI: 10.3390/s25061824. URL: <https://www.mdpi.com/1424-8220/25/6/1824> (visited on 04/15/2025).
- [5] L. Ruzicka, D. Söllinger, B. Kohn, C. Heitzinger, A. Uhl, and B. Strobl. “Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints Using Pose Correction and Unwarping”. In: *IET Biometrics* 2023 (Dec. 18, 2023). Ed. by V. Conti. Publisher: Hindawi, p. 7519499. ISSN: 2047-4938. DOI: 10.1049/2023/7519499. URL: <https://doi.org/10.1049/2023/7519499>.
- [6] L. Ruzicka, A. Spenke, S. Bergmann, G. Nolden, B. Kohn, and C. Heitzinger. *Towards Fingerprint Mosaicking Artifact Detection: A Self-Supervised Deep Learning Approach*. Jan. 9, 2025. DOI: 10.48550/arXiv.2501.05034. arXiv: 2501.05034[cs]. URL: <http://arxiv.org/abs/2501.05034> (visited on 06/25/2025).
- [7] L. Ruzicka, B. Strobl, S. Bergmann, G. Nolden, T. Michalsky, C. Domscheit, J. Priesnitz, F. Blümel, B. Kohn, and C. Heitzinger. “Toward Synthetic Physical Fingerprint Targets”. In: *Sensors* 24.9 (Jan. 2024). Number: 9 Publisher: Multidisciplinary Digital Publishing Institute, p. 2847. ISSN: 1424-8220. DOI: 10.3390/s24092847. URL: <https://www.mdpi.com/1424-8220/24/9/2847> (visited on 04/15/2025).
- [8] L. Ruzicka, B. Strobl, B. Kohn, and C. Heitzinger. “Centrality of the Fingerprint Core Location”. In: *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies BIOSTEC*. Biostec. Vol. Volume 1: Biosignals, Biodevices, Bioinformatics, Bioimaging. Rome, Italy: Scitepress, 2024, pp. 713–720. ISBN: 978-989-758-688-0. DOI: 10.5220/0012309300003657. URL: <https://www.scitepress.org/PublicationsDetail.aspx?ID=k/DomVEyngU=&t=1>.

Acknowledgments

First and foremost, I would like to thank my supervisors, Bernhard Kohn at the Austrian Institute of Technology (AIT) and Clemens Heitzinger at the Technische Universität Wien, for their guidance, continuous support, and encouragement throughout my PhD research.

I would also like to thank all the colleagues and collaborators I had the pleasure of working with during this journey. The discussions, feedback, and shared efforts greatly helped both the scientific and personal aspects of my doctoral experience. A special thanks goes to Youssef for the many insightful and enjoyable office talks, which made the everyday work environment both productive and pleasant.

My sincere thanks go to the AIT for providing the opportunity and environment to conduct my PhD research. I am also grateful to the administrative and technical staff who provided support in countless ways, making research and everyday work much smoother.

I would also like to express my gratitude to my friends, whose presence and support brought joy and perspective throughout these years. In particular, I thank Paul for the countless discussions about mathematics, machine learning, and philosophy, which continually inspired new ideas and ways of thinking. I am equally grateful to Philipp for joining me in rope climbing and on the many climbing trips we shared.

Most importantly, I would like to thank my family, Anna, and my friends for their unwavering support and encouragement throughout the last years. Their patience, understanding, and belief in me have been a constant source of strength and motivation.

Contents

1	Introduction	12
1.1	Motivation and Problem Statement	12
1.2	Research Challenges and Contributions	15
1.3	Thesis Outline	15
2	Background	17
2.1	Fingerprint History	17
2.2	Anatomy and Fingerprint Features	18
2.3	Fingerprint Image Acquisition	20
2.4	Fingerprint Image Pre-processing	21
2.5	Fingerprint Feature Extraction	24
2.6	Fingerprint Comparison	25
2.7	Computer Vision	26
2.8	Deep Learning	28
2.8.1	Artificial Neural Networks	28
2.8.2	Training	29
2.8.3	Convolutional Neural Networks	32
2.8.4	Specialized Architectures	33
3	Development of Contactless Fingerprint Algorithms	35
3.1	Fingerprint Segmentation in Contactless Images	35
3.1.1	Challenges of Contactless Fingerprint Segmentation	35
3.1.2	Related Work	36
3.1.3	Proposed Segmentation Algorithm	41
3.1.4	Experimental Evaluation and Results	43
3.1.5	Discussion	47
3.2	Fingerprint Core Localization	51
3.2.1	Challenges of Accurate Core Detection	51
3.2.2	Related Works	51
3.2.3	A Deep Learning Approach for Fingerprint Core Detection	52
3.2.4	Empirical Analysis of Fingerprint Core Centrality and Distribution	55
3.2.5	Discussion	60
3.3	Pose Correction and Unwarping for Contactless Fingerprints	62
3.3.1	Challenges of Pose Variations in Contactless Fingerprints	63
3.3.2	Related Works	64
3.3.3	Proposed Pose Correction Algorithm	65
3.3.4	Proposed Unwarping Algorithm	73
3.3.5	Experimental Evaluation and Results	73
3.3.6	Discussion	78
3.4	Contactless Fingerprint Quality Assessment	80
3.4.1	Challenges of Quality Assessment for Contactless Fingerprints	80
3.4.2	Existing Approaches and Limitations	81

3.4.3	The NFIQ 2 Framework as a Basis	81
3.4.4	The MCLFIQ Algorithm: Adaptation and Training	82
3.4.5	Experimental Evaluation and Results	83
3.4.6	Discussion	89
3.5	Fingerprint Mosaicking Artifact Detection	90
3.5.1	Challenges of Fingerprint Mosaicking	90
3.5.2	Related Work	91
3.5.3	Proposed Mosaicking Artifact Detection Algorithm	92
3.5.4	Experimental Evaluation and Results	95
3.5.5	Discussion	98
3.6	Synthetic Finger Phantoms	100
3.6.1	Related Work	100
3.6.2	Challenges of Synthetic Fingerprints and Fingers	101
3.6.3	Manufacturing Methodologies	102
3.6.4	Experimental Evaluation and Results	106
3.6.5	Discussion	109
3.7	Privacy Preserving Fingerprint Comparison	112
3.7.1	Related Works	113
3.7.2	Fingerprint Comparison	114
3.7.3	Privacy Preserving Comparison with MPC	115
3.7.4	Discussion	119
4	Conclusion	121
4.1	Conclusion	121
4.1.1	Image Normalization and Interoperability	121
4.1.2	Data Assurance and Quality	122
4.1.3	Validation and Secure Deployment	123
4.1.4	The Complete Pipeline: A Holistic View	123
4.1.5	Limitations of the Proposed Methods	124

Chapter 1

Introduction

1.1 Motivation and Problem Statement

From unlocking smartphones to accessing restricted areas, verifying a person's identity or identifying an unknown individual is a common requirement in many systems. Traditional methods such as passwords, PINs, or physical tokens are widely used, but they can be forgotten, lost, stolen, or shared. These weaknesses create security risks and usability problems.

To address these issues, biometric authentication uses biological traits to either verify a claimed identity or determine who someone is. Biometrics are tied to the individual and cannot easily be transferred or replicated, which makes them a useful alternative. Common biometric traits include facial features, iris patterns, voice, gait, and DNA. Among these, fingerprint recognition is one of the oldest and most widely used methods. The process of identifying individuals by comparing fingerprints is known as dactyloscopy.

The use of dactyloscopy in criminal investigations has a rich history. In 1893 the Troup Committee, named after its chairman Charles Edward Troup was formed to investigate ways to identify habitual criminals in England. As a result, in 1894, all newly arrested criminals in England and Bengal had their fingerprints recorded. [11]

In 1892, in Buenos Aires, Argentina, a murder was solved using thumbprint evidence found at the crime scene. The Rojas murder case is considered to be the first homicide solved by fingerprint evidence, and Argentina became the first country to rely solely on fingerprints as a method of individualization. Another criminal case in Bengal in 1898 is considered to be the first case in which fingerprint evidence was used to secure a conviction. [11]

These early successes led to the widespread adoption of fingerprint databases by law enforcement agencies worldwide, enabling the identification of suspects, solving of crimes, and maintenance of criminal records. The uniqueness and permanence of fingerprint patterns allows authorities to match crime scene evidence to individuals with high accuracy.

Beyond criminal investigations, fingerprint recognition has become a key technology in border control and immigration. Many countries deploy fingerprint-based biometric systems at airports and border crossings to verify traveler identities [65], combat identity fraud [159, 198], and ensure national security. Moreover, fingerprint biometrics have been widely adopted in government-issued identity programs, such as India's biometric identity program called Aadhar. It voluntarily collects the facial image, both iris and the 10 finger prints of Indian citizens and acts as a digital identity document [104]. National identification systems, voter registration programs [36], and social welfare distribution schemes [178] increasingly rely on fingerprints to ensure that benefits reach the intended recipients while preventing fraud. In many developing nations, where traditional forms of identity documentation are less reliable, biometric identification provides a secure and scalable solution for establishing and verifying identity. Another application of fingerprint recognition is in financial transactions and banking. Biometric authentication has become a standard feature in mobile banking apps and payment verification systems. The integration of fingerprint recognition in access control is another major area of application [59].

Many organizations use biometric attendance systems and access control solutions to regulate entry into secure facilities [201]. Finally, consumer electronics and personal device security represent a rapidly growing domain for fingerprint biometrics. Smartphones, laptops, and smart home devices increasingly incorporate fingerprint recognition as a way to unlock the device.

The strength of fingerprint recognition lies in the unique characteristics of fingerprint patterns. These patterns of ridges and valleys are highly distinctive, even between identical twins [251], and remain stable throughout an individual's life [276]. Historically, the acquisition of fingerprints was a straightforward process, relying on ink, where a user would press or roll their inked finger on a piece of paper. The next step was the introduction of contact-based sensors, where a user would press or roll their finger against a platen. This simplicity of the contact-based sensor, coupled with the high accuracy achievable with automated biometric identification systems (ABIS), led to the widespread adoption of fingerprint recognition and the creation of massive, legacy databases by law enforcement and government agencies.

However, the traditional contact-based approach has limitations that have become increasingly apparent in recent years. With the onset of the COVID-19 pandemic, hygiene has become a major consideration. Contact-based sensors require users to share a common surface, creating a potential vector for the transmission of germs and infectious diseases [176]. This concern has created a strong demand for contactless alternatives across various sectors. Beyond hygiene, contact-based acquisition can also be inconvenient and uncomfortable. Users must carefully position and press their finger against the sensor, which can be challenging for individuals with dry, oily, or damaged skin, or in situations with limited dexterity. Moreover, the contact itself introduces non-linear distortions to the finger, and therefore to the fingerprint, which are hard to reverse [237]. Finally, contact-based methods are also prone to creating latent prints on the sensors, which create privacy issues and introduce artifacts into the recorded fingerprints.

These limitations together with advances in imaging technology and computer vision have led to a shift towards contactless fingerprint recognition. Contactless acquisition removes the risk of cross-contamination, improves user comfort, and reduces distortions. To realize these benefits for the use-case of identity identification, contactless systems must solve the interoperability problem with existing contact-based databases. This is important to allow for a smooth transitions without loss of historical information of already existing fingerprint databases [217]. This interoperability problem arises from differences in image acquisition. In contact-based sensing, the physical contact between the fingertip and platen in contact-based devices deforms the skin, introducing distortions. For contactless acquisition on the other hand, variations in finger pose (roll, pitch, and yaw are unconstrained) and hand distance can introduce geometric distortions [247] and scale inconsistencies. Scale inconsistencies make the same fingerprint appear different sizes across captures, and compensating for this variation is challenging [87]. Many fingerprint comparison algorithms, particularly those developed for legacy systems, are sensitive to scale differences, causing comparisons to fail, even for the same fingerprint [133]. Image quality consistency is also a challenge [133]. Uneven illumination, shadows, specular reflections, and focus issues degrade contactless images, obscure ridge-valley structures and add noise, further complicating comparison beyond the pose and scale issues.

With the introduction of deep learning approaches, which utilize large amounts of training data to solve a task, the importance of task-specific labeled fingerprint datasets grew. However, acquiring and annotating large-scale, high-quality fingerprint datasets is a labor intensive and expensive process. Real-world data collection is moreover also subject to strict privacy regulations. Furthermore, obtaining accurate ground truth labels for tasks like segmentation, core or minutiae point detection, and especially artifact detection can be extremely difficult, often requiring manual annotation by dactyloscopy experts. This data scarcity has inspired the development and use of synthetic fingerprint generators. These generators offer a way to create perfectly

labeled data, encompassing a wide range of variations of minutiae pattern, ridge orientations, fingerprint artifacts, and image quality. Building upon the concept of synthetic fingerprint images, the field is progressing towards the creation of physical, synthetic fingers. These 3D replicas, also called phantoms, can be generated using synthetic fingerprint data mapped to arbitrary finger geometries. They offer a standardized approach to biometric system evaluation and sensor development. Synthetic fingers eliminate the variability and ethical concerns associated with using real human subjects. They enable precise control over fingerprint characteristics, replicating common fingerprint types as well as edge cases and even skin properties like dryness or elasticity. Furthermore, if combined with a robotic operator, they allow control over more extensive properties such as presentation pose for contactless sensors and pressure, rolling speed or others for contact-based sensors.

However, the creation of realistic and effective synthetic fingers presents its own set of challenges. Different production methodologies, such as 3D printing, molding, and casting, result in varying levels of fidelity in replicating the details of fingerprints, including ridge structure, pores, and tactile properties.

For effective fingerprint enrollment, capturing a substantial portion of the fingertip surface is important. A larger initial enrollment image provides more biometric data, increasing the certainty when comparing fingerprints by future downstream applications. Furthermore, it allows to work more reliably with latent partial prints where only small sections of a finger are found, as can be the case for example at crime scenes. In traditional contact-based systems this is accomplished through rolled fingerprint acquisition, where the finger is rolled nail-to-nail, capturing the entire fingerprint surface. This extensive surface area, and the resulting increase in identifiable minutiae, make rolled fingerprints the preferred method for enrollment.

Contactless systems aim to replicate the information captured by rolled fingerprints. However, without physical contact to guide the finger, contactless acquisition either records a 3D representation of the fingertip [126], or uses a mosaicking process [49, 109, 163, 252] to stitch together a complete image of the fingerprint. This second approach involves stitching together multiple partial images, acquired as the finger is presented to simulate the rolling motion. This process is prone to errors, introducing mosaicking artifacts such as misaligned ridges, duplicated or missing minutiae, and blurred regions. These artifacts present a significant challenge in contactless rolled fingerprint acquisition, and to a lesser extent also in contact-based fingerprint acquisition, potentially compromising comparison accuracy and leading to misidentification or non-identification. A further complication is the lack of robust quality assessment metrics specifically for contactless fingerprints, and even more so for contactless rolled fingerprints. Existing metrics, like NIST Fingerprint Image Quality (NFIQ) 2 [246], are primarily designed and trained for contact-based images and do not address the unique challenges of contactless acquisition and mosaicking artifacts.

Despite the challenges described in the previous paragraphs, contactless fingerprint recognition holds many benefits for secure, hygienic and convenient authentication. However, as biometric systems become more integrated into applications, data privacy and data safety become more and more important. Unlike passwords, which can be reset when compromised, biometric data is immutable. This makes the transmission and sharing of biometric information particularly sensitive. It requires protection against both accidental leaks and malicious misuse. Traditional fingerprint authentication relies on storing and processing biometric data locally on secure hardware, relying on trusted partners when sharing the biometric data. However, this approach limits broader applications, such as remote identity verification or identity identification/verification with a non-trustworthy partner. Multi-party computation (MPC) offers an alternative by enabling secure fingerprint comparison without exposing raw biometric data [151]. However, MPC introduces significant computational overhead, especially for minutiae-based template

comparison algorithms. Despite these challenges, recent advancements in MPC protocols leading to reduced computational demands have made privacy-preserving biometric template comparison increasingly practical.

1.2 Research Challenges and Contributions

This thesis aims to address key technical challenges hindering the widespread adoption and interoperability of contactless fingerprint recognition systems, as described in section 1.1. The objective is to develop and evaluate novel algorithms and methods that improve the accuracy, robustness, and privacy of contactless fingerprint acquisition. To achieve this, the research is guided by the following questions:

- RQ1** How can segmentation of fingerprints from contactless hand images be made more accurate and robust across different environments?
- RQ2** What techniques can effectively correct pose variations in contactless fingerprints to enhance comparison accuracy?
- RQ3** How can core point detection in contactless images be improved to enable better alignment with contact-based databases?
- RQ4** What are the problems of established fingerprints quality assessment metrics and what can be changed to improve their performance on contactless fingerprint images?
- RQ5** How can we detect errors in the mosaicking process of contactless fingerprint sensor systems?
- RQ6** Can we produce realistic, synthetic 3D fingers and how do they perform compared to real fingers? What manufacturing techniques can be used?
- RQ7** How can we create a privacy preserving comparison approach that circumvents the risk of exposing biometric data to other parties?

These questions are explored throughout the thesis, with each chapter addressing a specific research challenge, which is linked to one to many research questions.

1.3 Thesis Outline

This thesis is structured as follows:

Chapter 2: Background provides a comprehensive overview of the fundamental principles of fingerprint recognition, including fingerprint history and anatomy, image acquisition, pre-processing, feature extraction, and comparison, as well as an introduction to the fundamental computer vision and deep learning techniques, such as artificial neural networks, training of deep learning models, convolutional neural networks and more specialized architectures, that build the basis on which the research was conducted.

Chapter 3: Development of Contactless Fingerprint Detection Algorithms is the core of the thesis, presenting the novel algorithms and methods developed to address the challenges outlined in the previous section 1.2. This chapter is divided into seven sections:

- **Section 3.1: Fingerprint Segmentation in Contactless Images** discusses the difficulties in accurately segmenting contactless fingerprint images due to variations in lighting, complex backgrounds, and skin tones. It then presents proposed novel segmentation algorithms (Custom U-Net based algorithm, FingerUNeSt++ and TipSegNet) and details their implementation, evaluation, and results.
- **Section 3.2: Fingerprint Core Location** focuses on the empiric position of the fingerprint core relative to the fingerprint pattern and an accurate way of detecting the core point in contactless fingerprints. The section includes a description of the empirical analysis and the outcomes, as well as the proposed deep learning algorithm to localize the core, the experimental setup and results of the core localization.
- **Section 3.3: Pose Correction and Unwarping for Contactless Fingerprints** addresses the challenge of pose variations (roll and yaw) and non-linear distortions in contactless images. It proposes a method combining geometric transformations and deep learning to correct these distortions. The section includes algorithm descriptions, experimental setup, results, and a critical discussion.
- **Section 3.4: Contactless Fingerprint Quality Assessment** discusses the need for a reliable quality assessment metric specifically designed for contactless fingerprints. It will describe the adaption of NFIQ 2 to contactless fingerprint images, leading to MCLFIQ, and validate the effectiveness of this approach.
- **Section 3.5: Fingerprint Mosaicking Artifact Detection** addresses the artifacts introduced during the mosaicking process. This process is used to capture a full nail-to-nail fingerprint in both contact-based and contactless scenarios. A novel deep learning-based detection method is proposed to identify and score these artifacts.
- **Section 3.6: Synthetic Finger Phantoms** addresses the production of synthetic 3D fingers based on synthetic fingerprints. Different manufacturing approaches are presented, compared and tested.
- **Section 3.7: Privacy Preserving Fingerprint Comparison** explores techniques for secure and privacy-preserving fingerprint comparison. It proposes an optimized minutiae-based comparison algorithm using Multi-Party Computation (MPC) for efficient encrypted fingerprint comparisons.

Chapter 4: Conclusion summarizes the key findings and contributions of the thesis. It additionally provides a fictive use-case showing the expected effect of combining the algorithms developed in this thesis on a real-world scenario.

Chapter 2

Background

This chapter provides an introduction to the field of fingerprint recognition as well as a short overview of the most important algorithmic discoveries in the field of computer vision and deep learning, which can be applied to the broader context of fingerprint recognition. It begins by reviewing the principles of fingerprint recognition, covering aspects from finger anatomy to image processing and fingerprint comparison. Subsequently, it continues by explaining the specific contactless fingerprint recognition, highlighting the evolution of the field, the state-of-the-art, and the persistent challenges that motivate this research. The goal is to provide a thorough and comprehensive overview of what is currently known, in order to show where improvement is possible.

Fingerprint recognition stands as one of the most established and reliable biometric modalities, playing an important role in forensic science, access control, and personal identification. It builds on the uniqueness and persistence of fingerprint patterns.

2.1 Fingerprint History

Friction ridge skin, which has unique patterns and stays mostly persistent throughout an individual's lifetime (except injury or disease) [276], forms the foundation for fingerprint-based biometric identification. This biological characteristic has a long and rich history, predating its widespread adoption in modern forensic science.

The use of fingerprints can be traced back to ancient civilizations. Archaeological evidence shows impressions, both accidental and intentional, on artifacts such as 6000-year-old earthenware discovered in northwest China. The earliest deliberate use is found in clay seals used for documents during China's Qin Dynasty (221-206 B.C.), demonstrating an understanding of fingerprints for proving authorship and preventing tampering. The Chinese recognized the value of friction ridge impressions for individualization perhaps as early as 300 B.C. and in Japan as early as A.D. 702. A Japanese domestic law enacted in A.D. 702 hints at a similar understanding. European scientists began detailed observation of friction ridge skin in the 17th and 18th centuries, with Dr. Nehemiah Grew first describing it in detail in 1684. [11]

The systematic classification of fingerprints began to take shape in the 19th century. Dr. Johannes E. Purkinje's 1823 thesis classified fingerprints into nine pattern types, a precursor to later, more comprehensive systems. Sir Francis Galton's groundbreaking work in the late 19th century, including his book *Finger Prints* (1892), established the uniqueness and permanence of fingerprints and introduced the now-standard terminology for minutiae (Galton details). Around the same time, Juan Vucetich developed a fingerprint classification system and first used fingerprints for criminal identification in Argentina. Sir Edward Richard Henry, building upon Galton's work, developed the Henry classification system in the late 19th century, which achieved widespread adoption. [11]

Concurrent with the development of classification systems, scientific advancements deepened our understanding of friction ridge skin. In 1880, Henry Faulds published on the value of

fingerprints for identification, including their potential use at crime scenes. Researchers like Hermann Welcker and Sir William James Herschel conducted studies on the permanence of friction ridges. Embryological studies by scientists such as Arthur Kollmann, Harris Hawthorne Wilder and Inez Whipple revealed the formation of friction ridges and volar pads during fetal development. [11]

In 1971, the first optical fingerprint capture machines began to appear [277] and in 1975 the FBI started funding the development of sensors and minutiae extraction technologies. Through the 1980s and 1990s, optical slap scanners (FTIR-based) became standard in law enforcement AFIS systems. Notably, in 1998 Siemens demonstrated the first smartphone with an embedded fingerprint sensor, spurring others (Motorola, Apple, etc.) to follow [277]. Apple's 2013 Touch ID (capacitive) and later developments brought fingerprint sensors to millions of consumers. In the late 2010s, under-display sensors emerged: first optical (e.g. Vivo in 2018) and soon ultrasonic in OLED screens (e.g. Samsung S10 in 2019) [277]. Thermal sensors (using heat patterns) and multispectral optical sensors (capturing sub-surface detail) also appeared in niche roles (car locks, government ID scanners). Recently (2023), agencies have even begun experimenting with contactless capture (e.g. smartphone-camera fingerprints in FBI pilot projects) [68].

2.2 Anatomy and Fingerprint Features

The physical structure of friction ridge skin is the basis for its use as a biometric trait. It consists of two primary layers: the epidermis (outer layer) and the dermis (inner layer). The epidermis itself has multiple layers, with the stratum corneum (surface layer) being composed of dead cells and the stratum basale (basal layer) being responsible for generating new skin cells. Friction ridges are formed by the intricate arrangement of raised ridges and furrows (valleys) on the surface of the skin. These patterns are not superficial, they are anchored by a double row of papillae pegging into the dermis. [149]

Fingerprint features are typically described at three levels of detail:

1. Level 1 Features (overall ridge flow): This level refers to the overall pattern formed by the flow of ridges. The primary pattern types are:
 - Arches: Ridges enter from one side of the finger and exit on the other, forming a wave-like pattern. Subtypes include plain arches (smooth, wave-like) and tented arches (with a sharp, upward thrust). Arches are the least common pattern type, appearing with low distribution rates.
 - Loops: Ridges enter and exit on the same side of the finger, curving around a central point. Loops are classified as radial (loop opens towards the thumb) or ulnar (loop opens towards the little finger), based on their orientation. Loops are more common patterns.
 - Whorls: Ridges form circular or spiral patterns around a central point. Subtypes include plain whorls, central pocket loop whorls, double loop whorls (two separate loop formations), and accidental whorls (combinations of patterns or irregular patterns). Whorls are less common than loops but more frequent than arches.

Within these patterns, two specific points of interest are used for classification and alignment. The approximate center of the fingerprint pattern, which is called core and second, the triangular areas where ridges from three different directions converge, also called deltas.

2. Level 2 Features (minutiae): These are the small, specific details within the ridge patterns, also known as Galton details. They are the primary features used for fingerprint comparison.

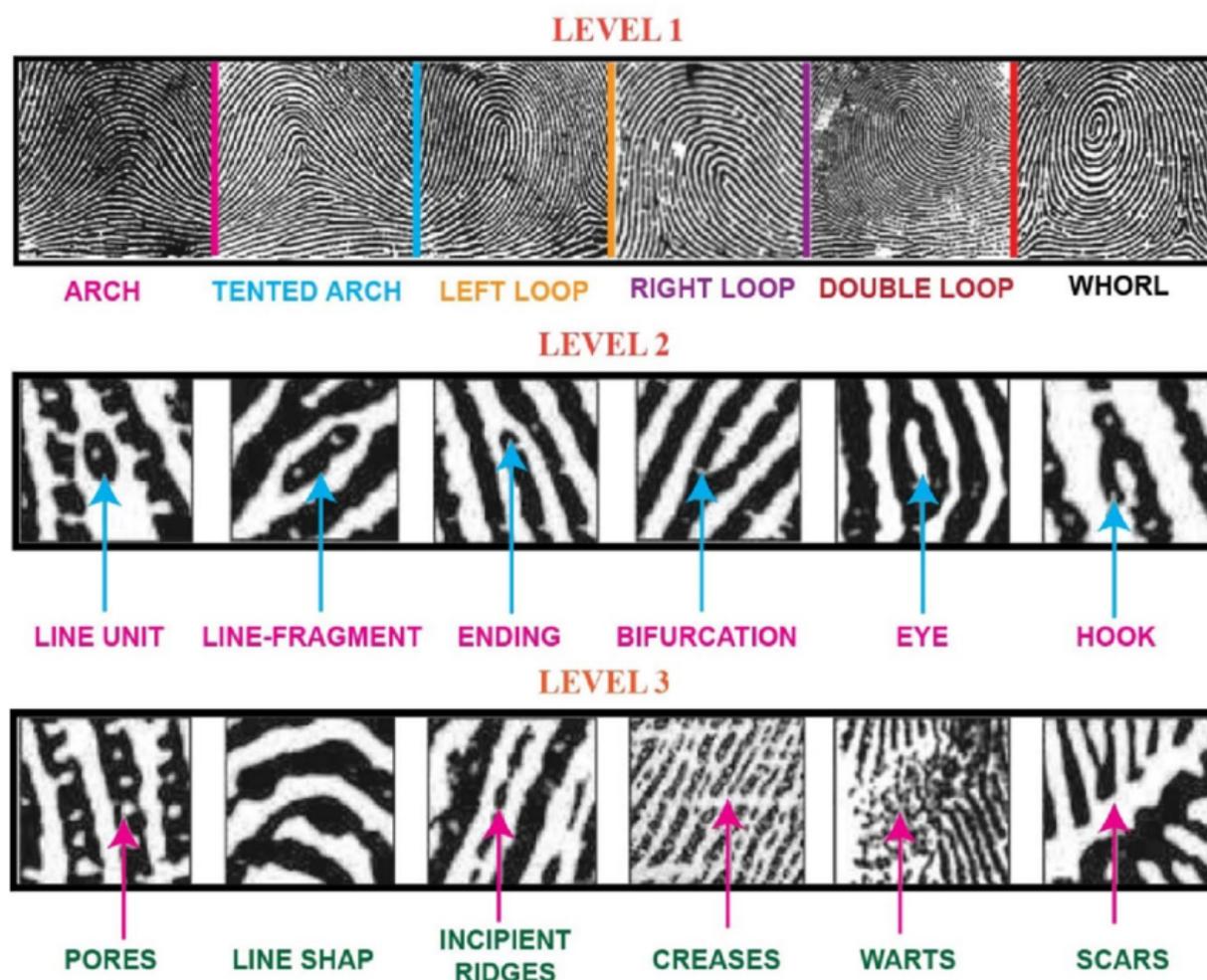


Fig. 2.1: Overview of the different feature levels with example images, taken from [13].

The two most common minutiae types are ridge endings: Points where a ridge terminates and bifurcations: Points where a ridge splits into two branches. Other, rarer, types include line fragments (very short ridges), line units (short, isolated ridges), enclosures/eyes (ridges that bifurcate and rejoin, forming an enclosed space), and hooks (short branches off a main ridge). Minutiae are defined by their location (x, y coordinates on the fingerprint image), their angle (the direction of the ridge at that point), and their type (ending, bifurcation, etc.). ABIS system often only work with ridge endings and bifurcations.

3. Level 3 Features (pores and ridge details): These are scars, warts ceases, as well as very fine details of the ridge itself. Those fine details require high-resolution imaging to be reliably captured and used. They consist of the positions, distributions, and shapes of sweat pores along the ridges or the shapes of the ridge edges (edgeoscopy). This includes features like ridge indentations, protrusions, and angles, as well as incipient ridges, which are thin, immature ridges that appear between the main ridges. Level 3 features are primarily used in high-resolution scenarios, such as comparing partial fingerprints, when image quality is exceptionally good. They can provide additional discriminating information to supplement Level 2 features.

The distinctiveness of fingerprints comes from the fact that no two individuals, not even identical twins, share the same fingerprint patterns at Level 2 and Level 3 detail. This individuality is a result of the complex interplay of genetic and environmental factors during fetal development. The genes provide the general blueprint for ridge patterns, but the precise arrangement of minutiae and pores is influenced by random events and the specific environment within the womb during weeks 10-17 of gestation, when the volar pads (transient swellings of tissue on the palms and soles) regress and friction ridges begin to form. [149]

The patterns formed during fetal development remain remarkably consistent throughout an individual's lifetime. Superficial injuries to the epidermis will result in the ridges regrowing in their original pattern. Only deep damage to the dermis, significant scarring, decomposition, or certain skin diseases can permanently alter the friction ridge pattern. This immutability is very important for fingerprint recognition's reliability. [267]

2.3 Fingerprint Image Acquisition

The accuracy and reliability of any fingerprint-based biometric system depends on the quality of the acquired fingerprint images. The method used to capture these images significantly influences their characteristics and subsequent processing steps required to use them in downstream applications such as ABISs [108, 276]. Fingerprint acquisition techniques can be broadly categorized into traditional inked methods, contact-based live scan technologies, and contactless methods.

Historically, fingerprint acquisition relied on the inked method, a manual process involving the application of ink to the fingertip and then pressing or rolling the finger onto a paper card [11]. While serving as a foundational technique, inked fingerprinting is inherently prone to distortions. These are a result of uneven ink distribution, smudging, and variations in applied pressure, making it unsuitable for modern, automated biometric systems.

Contact-based live scan technologies represent a significant advancement, utilizing sensors where the user places their finger directly on a platen surface. Several sensor types exist, each with distinct operating principles. Optical sensors, often employing Frustrated Total Internal Reflection (FTIR), use a light source and a prism. When a finger is placed on the platen, the ridges make contact with the surface, frustrating the total internal reflection of light, while the valleys remain separated by an air gap. This creates a contrast difference that forms the fingerprint image [16, 154, 196]. While FTIR-based optical sensors are well-established and widely used, they are susceptible to image degradation from dirt, latent fingerprint residues, and variations in skin condition, such as dryness or excessive moisture [270].

Capacitive sensors utilize arrays of tiny capacitors. The capacitance between the sensor and the finger varies depending on whether a ridge (touching the sensor) or a valley (separated by an air gap) is present [196, 253, 263, 273]. This variation in capacitance is then used to construct the fingerprint image. Capacitive sensors are known for their high precision and improved resistance to spoofing attacks compared to some optical methods [71]. However, they may be affected by extremely dry or damaged skin, and the sensor size can sometimes limit the captured fingerprint area.

Ultrasonic sensors operate by emitting ultrasonic pulses and measuring the resulting echoes. The echoes differ between ridges and valleys due to the varying acoustic impedance, allowing for the creation of a detailed, often 3D, map of the fingerprint [78, 153, 214]. Ultrasonic imaging can penetrate the outer layer of skin, resulting in high-quality images that are less susceptible to contamination from dirt and moisture. However, these sensors tend to be more expensive and potentially slower than other technologies.

Thermal sensors measure temperature differences between the ridges (which are in contact with the sensor and therefore warmer) and the valleys (which are cooler due to the air gap) [91]. Thermal sensors can operate effectively under a range of skin conditions, including both wet and dry fingers, and offer good security. However, they can be sensitive to changes in the ambient temperature and have a higher cost compared to some other sensor types.

Multispectral imaging captures fingerprint images at multiple wavelengths and polarization conditions [211, 285]. This approach allows for the acquisition of both surface and subsurface fingerprint features, making it robust to variations in skin condition, such as dirt, moisture, and even poor contact [211]. Multispectral imaging also has shown potential in detecting spoofing attempts. The main drawbacks are the increased complexity and higher cost of multispectral systems.

Contactless fingerprint acquisition can be seen as the next step, following after the progressive development of contact-based live-scan device. Contactless methods, typically optical, employ cameras and computer vision algorithms to capture fingerprint images without any physical contact between the finger and a sensor [87, 217]. Some contactless systems incorporate structured or specialized lighting to enhance image quality. The primary advantages of contactless acquisition are improved hygiene [176, 275] and a reduction in the non-linear skin distortions that occur when a finger is pressed against a platen [237]. Furthermore, 3D contactless systems take this a step further by creating a three-dimensional model of the fingertip, using techniques such as Time-of-Flight (TOF) sensors, structured light, stereo vision, or shape from focus/defocus [127, 169]. This 3D representation captures a more complete depiction of the fingerprint, which can potentially improve comparison accuracy, especially when dealing with rolled fingerprint equivalents [87]. Users also generally prefer contactless acquisition over contact based [176].

Latent fingerprint acquisition, while distinct from live scan methods, is another biometric use-case required in forensic science. Latent fingerprints are impressions left unintentionally, for example, at crime scenes [3, 37, 72]. Recovering these prints involves techniques such as dusting with fine powders or applying chemical enhancements like ninhydrin or cyanoacrylate fuming to make the invisible ridge patterns visible [128].

To ensure consistency and interoperability, several standards have been established for fingerprint image acquisition. The FBI's Appendix F specifies image quality requirements for large-scale fingerprint comparison and human expert comparison [61]. PIV-071006 provides a less strict standard, primarily intended for one-to-one verification scenarios [131]. Image resolution, typically measured in dots per inch (DPI), is one of the most important parameters specified in these standards [157].

2.4 Fingerprint Image Pre-processing

Raw fingerprint images, independently whether they were obtained through contact-based or contactless sensors, do seldom have the quality required for direct feature extraction. Factors such as the sensor's internal electronics [156, 270] or optical components, variations in skin conditions such as dryness, wetness, or the presence of contaminants [91, 156], and, specifically for contactless acquisition, uneven illumination and specular reflections [91] reduce the fidelity of the fingerprint recording. Pre-processing techniques are therefore used to reduce the impact of these factors. The goal of pre-processing is to enhance the clarity of the ridge and valley structures while avoiding introducing false information. This improves the accuracy and reliability of subsequent feature extraction stages and, consequently, the overall system performance.

Contact-based and contactless fingerprint acquisition methods present distinct challenges that require different pre-processing approaches. Contact-based fingerprint images are characterized

by non-linear distortions, which are however consistent over subsequent recordings and are therefore only relevant if comparing the fingerprint recordings with recordings from contactless sensor modalities. These distortions arise from the elastic deformation of the fingertip skin as it is pressed against the sensor platen [237, 275]. In contrast, contactless images, while free from pressure-induced distortions, are susceptible to issues like pose induced geometric distortion, uneven illumination, specular reflections (glare), variations in focus, and motion blur [91].

The typical fingerprint image pre-processing pipeline involves a sequence of operations, each addressing specific image degradations. Fingerprint images often show low contrast between ridges and valleys, hindering accurate feature detection. Image enhancement techniques therefore aim to improve this contrast using various approaches. Histogram equalization is a widely used technique that redistributes pixel intensities across the entire image, thereby enhancing overall contrast [35]. It operates by transforming the image's intensity histogram to approximate a uniform distribution. However, global histogram equalization may not be optimal for images with significant local variations in contrast. Adaptive histogram equalization (AHE) addresses this limitation by dividing the image into smaller regions and performing histogram equalization separately within each region [185]. This localized approach is particularly beneficial for contactless fingerprint images, which often exhibit uneven illumination due to variations in finger-to-sensor distance and ambient lighting conditions. Contrast Limited Adaptive Histogram Equalization (CLAHE) further improves upon AHE by limiting the amplification of noise in relatively homogeneous regions. Furthermore, variations in illumination or sensor sensitivity can result in images that are either too dark (under-exposed) or too bright (over-exposed). Brightness adjustment aims to normalize the overall image brightness to a standard level, making the ridge structures more discernible [253]. One common approach is to adjust the mean gray value of the image to a predefined target value [99]. This ensures consistency across different images and facilitates subsequent processing steps.

Noise, which will be present in any image acquisition process, can be seen as random variations in pixel intensities that can obscure the true fingerprint pattern. Noise reduction filters aim to suppress this noise while preserving image details. Gaussian filtering is a linear smoothing technique that employs a Gaussian kernel to perform a weighted average of pixel values within a neighborhood [35]. The Gaussian kernel's weights are determined by a Gaussian function, effectively blurring the image and reducing high-frequency noise components. The degree of smoothing is controlled by the standard deviation of the Gaussian function. Another common technique, median filtering, is a non-linear filtering technique that replaces each pixel's value with the median value of its neighboring pixels [108]. This approach is particularly effective at removing salt-and-pepper noise, characterized by isolated bright and dark pixels, while preserving edges better than linear smoothing filters. Beyond Gaussian and median filters, other noise reduction techniques are available, each with its own strengths and weaknesses. Anisotropic diffusion filters, for example, are designed to reduce noise while preserving edges and fine details by selectively smoothing along the direction of image gradients [99]. The selection of a suitable filter depends on the type and level of noise and the recording conditions.

Binarization is the process of converting a grayscale fingerprint image into a binary image, consisting of only black and white pixels. This simplifies the representation, typically assigning black pixels to ridges and white pixels to valleys. This process helps subsequent feature extraction by clearly delineating the ridge structure. Global thresholding techniques, like Otsu's method, are used to determine the separation. Otsu's method is a widely used automatic thresholding technique that selects a global threshold value based on the image's histogram [180]. It aims to find the threshold that minimizes the intra-class variance (variance within the foreground and background classes) or, equivalently, maximizes the inter-class variance (variance between the foreground and background classes). However, adaptive thresholding techniques, unlike

global thresholding, compute different threshold values for different regions of the image [185]. This adaptability makes them more robust to variations in illumination and contrast, which are commonly encountered in contactless fingerprint images. Adaptive thresholding methods typically calculate the threshold for each pixel based on the statistics of its local neighborhood. Binarization, in the context of fingerprints, is essentially a segmentation of the image, and it is therefore closely connected [286].

Segmentation is the process of isolating the region of interest (ROI), which is the fingerprint itself, from the background. Accurate segmentation is especially important in contactless recording scenarios where the background can be complex and variable, potentially containing non-fingerprint objects or textures [87, 116]. The segmentation performance has a direct impact on downstream tasks such as pose correction [217, 247] and feature-based comparison [87], both of which rely on a precise mask of the fingertip region to separate it from the background.

Traditionally, fingertip segmentation methods have focused on isolating a single fingertip from the background. These methods can be broadly categorized into color-, or brightness for grayscale images, -based, machine learning-based, and shape-based approaches. Color-based methods use the skin tone of fingertips to differentiate them from the background [18, 116]. Machine learning-based methods, on the other hand, use algorithms trained on labeled data to identify and delineate fingertip regions [164, 190]. Shape-based methods leverage the characteristic geometric properties of the fingertip, such as its elliptical shape and curvature, to perform segmentation [288]. More recent approaches in contactless fingerprint recognition have moved towards directly segmenting fingertips from hand images, bypassing the need for a separate fingertip detection step [164, 190]. This approach simplifies the processing pipeline and can improve efficiency. Deep learning, particularly the use of convolutional neural networks (CNNs), has significantly advanced the field of contactless fingerprint segmentation [80, 82].

Ridge thinning, also known as skeletonization, is a process that reduces the width of the ridges in a binary fingerprint image to a single pixel while preserving the connectivity and overall structure of the ridge pattern [170]. This simplified representation facilitates feature extraction, particularly the detection of minutiae. Thinning algorithms typically employ morphological operations, such as erosion and dilation, to iteratively remove pixels from the boundaries of ridges until only a one-pixel-wide skeleton remains [170]. The Zhang-Suen algorithm is a classic and widely used thinning algorithm that iteratively removes pixels based on specific connectivity criteria [41].

The orientation field represents the local direction of the ridges at each point in the fingerprint image. It is used in various fingerprint processing stages, including image enhancement, scale normalization, minutiae extraction, and fingerprint alignment during comparison [34, 268]. Gradient-based methods are often applied, where the orientation field is calculated by calculating the gradient of the image intensity at each pixel [262]. The gradient direction, perpendicular to the ridge direction, provides an estimate of the local ridge orientation. Slit-based methods are another option. These methods utilize specially designed filters, or slits, to analyze the local ridge pattern and determine the dominant orientation. Gabor filters, tuned to specific frequencies and orientations, are commonly used for this purpose.

Fingerprint images are typically acquired at different resolutions, depending on the sensor used. To ensure interoperability and compatibility with standard fingerprint databases and comparison algorithms, it is often necessary to normalize the image resolution to a standard value, as for example, specified in the FBI's Appendix F [61]. For systems which are not able to correct for scale variations, this implies scaling the fingerprint image such that the ridge distance is set to a fixed value.

The quality of pre-processing has a direct and significant impact on the accuracy and reliability of subsequent fingerprint analysis steps. Noise and artifacts in the image can be misinterpreted

as genuine minutiae, leading to false detections. Poor contrast or inadequate noise reduction can obscure genuine minutiae, causing them to be missed during feature extraction. Errors in the estimated orientation field can lead to incorrect fingerprint classification, which interfere with the comparison process.

2.5 Fingerprint Feature Extraction

Fingerprint feature extraction is the step that transforms the information of the preprocessed fingerprint image into a representation suitable for comparison and recognition. While various features can be extracted, this thesis primarily focuses on two key approaches: traditional minutiae extraction and the more recent development of learned fingerprint representations, often referred to as embeddings.

The process of minutiae extraction requires a preprocessed fingerprint image that has been binarized, thinned, and enhanced to improve the clarity of the ridge structure [99, 157]. Then, the so called crossing number is calculated. The crossing number is defined as the number of ridge transitions (changes from ridge to valley, or valley to ridge) along a small neighborhood around a given pixel [106, 154]. It is calculated by summing the differences between adjacent pixel pairs in an eight-connected neighborhood and dividing the result by two [157]. A pixel on a ridge ending will have a crossing number of 1, while a pixel on a bifurcation will have a crossing number of 3 [106, 157]. Pixels on continuous ridges have a crossing number of 2, and are therefore not minutiae.

Despite preprocessing, the initial minutiae extraction often includes spurious minutiae caused by noise, artifacts, or remaining imperfections in the image. Therefore, a post-processing stage is typically used to suppress these false minutiae. This stage commonly employs a combination of heuristics and rules based on geometric and topological relationships between the detected minutiae. For instance, distance-based rules dictate that minutiae located too close to each other are likely to be spurious, representing the same feature detected multiple times [157]. Angle-based rules can identify and eliminate pairs of ridge endings that point in opposite directions and are very close, as these likely represent a single broken ridge rather than two distinct endings. Minutiae located near the boundary of the fingerprint region are often unreliable due to incomplete ridge information and are therefore typically removed [71]. Furthermore, analyzing the surrounding ridge structure provides valuable context for validating minutiae, since genuine minutiae should be consistent with the local ridge flow [287].

Once a reliable set of minutiae has been extracted and refined, each minutia is represented by its location (x , y coordinates in the fingerprint image), its angle (the orientation or direction of the ridge at that point, typically measured in degrees or radians relative to a reference axis), and its type (classified as either a ridge ending or a bifurcation) [157]. This (x , y , angle, type) representation forms the basis for many traditional fingerprint comparison algorithms.

Regardless of the feature extraction method used, assessing the quality of the extracted features can improve the fingerprint comparison. Confidence measures are often assigned to individual minutiae, indicating the reliability of their detection. These confidence scores can be based on factors such as local ridge clarity, contrast, and the degree of agreement with the estimated orientation field [71, 117]. In addition to individual minutiae confidence, overall quality scores are often computed for the entire set of extracted features. These scores provide an indication of the overall reliability of the fingerprint representation [4].

While minutiae are the most widely used features, other fingerprint characteristics can also provide useful information. Singular points, namely cores and deltas, are important global features. Cores represent the approximate center of the fingerprint pattern, while deltas are

triangular regions where ridges from three different directions converge [17, 157]. Singular points are valuable for fingerprint classification, alignment, and indexing. The ridge count, defined as the number of ridges between two specific points (such as a core and delta, or two minutiae), is another useful feature [157]. The local ridge frequency, which is the reciprocal of the average inter-ridge distance, provides information about the density of the ridge pattern [99, 154]. Texture features, derived from the gray-level distribution in the fingerprint image, often using techniques like Gabor filter responses, can also capture discriminating information [107].

In recent years, a new, deep learning based fingerprint feature extraction approach was introduced: learned fingerprint representations, or embeddings. Unlike manually designed features like minutiae, these representations are automatically learned from data and converted into a fixed-sized vector space. The input to these learning-based methods can vary, it may be the raw or preprocessed fingerprint image, the frequency spectrum of the image [87], or even a set of extracted minutiae locations and orientations [32]. Deep learning architectures, such as Convolutional Neural Networks (CNNs) [32], Autoencoders [186, 282], or Transformers [88], are employed to learn hierarchical features from the input data. The output is typically a fixed-length vector, referred to as an embedding, which represents the fingerprint in a high-dimensional feature space [32, 284]. The goal is to create embeddings that are both compact, enabling efficient storage and comparison, and discriminative, ensuring that different fingerprints are mapped to distinct regions in the embedding space [32]. These fixed-size embeddings offer advantages such as fast encryption and efficient, privacy-preserving comparison operations [205].

2.6 Fingerprint Comparison

Fingerprint comparison, also frequently referred to as fingerprint recognition, is the process of determining the degree of similarity between two fingerprint impressions. The goal is to assess the likelihood that the two fingerprints originated from the same finger [157].

Fingerprint comparison systems are broadly categorized into two main types: verification and identification. Verification, also known as 1:1 comparison, involves comparing a query fingerprint against a single claimed identity's fingerprint [157]. This is commonly used in access control applications, where a user claims an identity and provides their fingerprint to verify that claim. Identification, or 1:N comparison, involves comparing a query fingerprint against a database of fingerprints to find the best match (or determine that no match exists) [157]. This is used in forensic investigations and large-scale identification systems, where the identity of the individual providing the query fingerprint is unknown.

The dominant approach to fingerprint comparison is minutiae-based comparison, which relies on comparing the sets of minutiae extracted from the two fingerprints [106, 157]. This process typically involves several key stages: alignment, pairing, scoring, and thresholding.

Alignment is the first step, aiming to find the optimal spatial transformation that brings the two minutiae sets into the best possible correspondence. This transformation typically involves translation (shifting the position) and rotation, and may sometimes include scaling [32, 106]. Two main alignment strategies exist: global and local. Global alignment seeks a single transformation that applies to the entire fingerprint [106]. This approach often leverages global features such as singular points (cores and deltas) [17] or the overall orientation field [209] to guide the alignment process. However, global alignment can be less effective when dealing with significant non-linear distortion, which is common in contact-based fingerprint images due to the elasticity of the skin. Local alignment, in contrast, attempts to mitigate the effects of non-linear distortion by applying multiple local transformations to different regions of the fingerprint [20, 209]. This approach

often utilizes local minutiae descriptors, which capture the characteristics of the ridge structure in the immediate vicinity of each minutia [70], or analyzes local ridge structures directly [97].

Once the minutiae sets have been aligned, the next step is minutiae pairing, which involves establishing correspondences between minutiae in the two sets [157]. This is not a trivial task, as one or both fingerprint impressions may contain missing minutiae (due to poor image quality or incomplete capture) or spurious minutiae (resulting from noise or artifacts in the preprocessing stages). Furthermore, the overlap between the two fingerprints may be limited, especially in partial fingerprint scenarios. Common methods for minutiae pairing include threshold-based pairing, which establishes correspondences based on distance and angle difference thresholds between minutiae [114], and relaxation methods, which iteratively refine the pairing based on the consistency of neighboring minutiae pairs [249].

After pairing, a similarity score is calculated to quantify the degree of similarity between the two minutiae sets [157]. This score typically takes into account several factors, including the number of mated minutiae, the distances between paired minutiae, the differences in their angles, the types of minutiae (ridge endings or bifurcations), and the consistency of ridge counts between corresponding minutiae pairs [106, 113]. Scoring approaches can range from simple weighted sums of these factors to more sophisticated statistical models.

Finally, a thresholding step is performed, where the calculated similarity score is compared to a predefined threshold to make a match/non-match decision [157]. The selection of this threshold is part system design, as it balances the false acceptance rate (FAR) – the probability of incorrectly declaring a match between fingerprints from different fingers – and the false rejection rate (FRR) – the probability of incorrectly declaring a non-match between fingerprints from the same finger.

While minutiae-based comparison is the most prevalent approach, other methods exist. Embedding-based comparison involves comparing learned fingerprint representations (embeddings) using a distance metric to determine the likelihood of a match [32, 87]. Correlation-based comparison directly compares fingerprint images using correlation measures, assessing the similarity of pixel intensity patterns [19]. However, this approach is generally sensitive to displacement and rotation between the fingerprint images. Image-based comparison utilizes image features, such as texture characteristics or ridge frequency, without explicitly extracting minutiae [173]. Pattern-based or ridge-based comparison leverages the entire ridge pattern, including Level 1 features like overall pattern type (arch, loop, whorl), for comparison [107].

The performance of any fingerprint comparison method is influenced by several factors. Poor image quality, characterized by noise, low contrast, or blurring, significantly reduces comparison accuracy [4]. Non-linear distortion, caused by the elastic deformation of the skin during contact-based acquisition, degrades performance, particularly for methods relying on global alignment [20, 208]. Partial fingerprints, where only a portion of the complete fingerprint is available, present a challenge due to the limited overlap and reduced number of comparable features [113]. Latent fingerprints, often encountered in forensic investigations, are particularly challenging due to their typically poor quality, partial nature, and presence of distortion and noise [110]. In the context of identification (1:N comparison), the size of the reference database also plays a role. Larger databases inherently increase the probability of false matches [54].

2.7 Computer Vision

Computer vision is a multidisciplinary field focused on enabling computers to extract and interpret information from digital images and videos [75, 195, 245]. It draws on methods and ideas from image processing, pattern recognition, machine learning, and artificial intelligence [238]. Applications of computer vision are wide-ranging and include object detection [290], image

classification [125], scene understanding [93], image restoration [254], medical imaging analysis [140], autonomous driving [26], robotics [231], and augmented reality [7].

Digital images are typically represented as matrices of numerical values. Each value in the matrix corresponds to a pixel location (x, y) and indicates either intensity (in grayscale images) or color (in color images) [85, 175]. Grayscale images assign a single value per pixel, usually in the range 0 (black) to 255 (white) [85]. Color images use multiple channels—commonly red (R), green (G), and blue (B)—to represent pixel color. This can be expressed as

$$I(x, y) = [R(x, y), G(x, y), B(x, y)], \quad (2.1)$$

where (x, y) denotes the pixel coordinates, and R , G , and B represent the respective color channel values.

Computer vision relies on basic image processing operations as foundational tools for more complex algorithms. These operations are generally divided into point operations and neighborhood operations. Point operations modify each pixel independently of its neighbors [85, 175]. Common examples include brightness adjustment (adding a constant to each pixel), contrast enhancement (expanding the range of intensity values), and thresholding (converting an image to binary by assigning pixel values based on a set threshold) [85]. A general form of a point operation is

$$g(x, y) = T[f(x, y)], \quad (2.2)$$

where $f(x, y)$ is the input pixel value, $g(x, y)$ is the output pixel value, and T is the transformation function.

Neighborhood operations, or filtering, update a pixel's value based on its surrounding pixels [85, 175]. A common example is linear filtering, where a kernel (a small matrix of coefficients) is convolved with the image [85]. The new pixel value is a weighted sum of neighboring values, determined by the kernel

$$g(x, y) = \sum_{i=-m}^m \sum_{j=-n}^n f(x-i, y-j) \cdot h(i, j), \quad (2.3)$$

where $h(i, j)$ are the kernel coefficients, and $m, n \in \mathbb{N}^+$ determine the half-width and half-height of the kernel, respectively. The kernel size is thus $(2m+1) \times (2n+1)$. Different kernels produce different effects. Smoothing filters (e.g., Gaussian and mean filters) reduce noise and blur the image. Sharpening filters (e.g., Laplacian and unsharp masking) enhance edges and details [278]. Edge detection filters, such as the Sobel [236], Prewitt [188], and Canny [31] operators, highlight boundaries between regions. In addition to filtering, geometric transformations alter the spatial layout of an image [85, 175]. These include scaling, rotation, translation, affine transformations (combinations of geometric modifications), and perspective transformations (to simulate changes in viewpoint) [92].

Building on basic image processing techniques, feature extraction focuses on identifying and representing distinctive elements within an image [148, 212]. The main goal is dimensionality reduction: to represent image content in a more compact form while retaining information needed for tasks like object recognition or image retrieval. As discussed in Section 2.5, fingerprint analysis often emphasizes minutiae extraction. In general computer vision, however, a broader range of features is considered. Common examples include edges, which indicate boundaries between regions with different intensities, corners, which mark points of rapid intensity change, and blobs, which refer to areas that differ from their surroundings in intensity or texture [139]. Advanced feature descriptors such as SIFT (Scale-Invariant Feature Transform) [148], SURF (Speeded-Up Robust Features) [15], and ORB (Oriented FAST and Rotated BRIEF) [212] are designed to be

robust to changes in scale, rotation, and lighting conditions. These methods usually involve two stages: detecting keypoints and computing descriptors, which are vector representations of the local image region around each keypoint. While similar in structure to fingerprint embedding techniques (Section 2.5), these general-purpose methods are not specifically designed for the ridge-valley patterns of fingerprints.

A key task in computer vision is object detection and recognition, which involves both locating objects in an image and identifying their categories [69, 260]. Early methods included template comparison, where a reference image is slid across the input image and a similarity score is computed at each location to find matches [29]. Feature-based methods extract local features (e.g., SIFT or HOG) and use classifiers like Support Vector Machines (SVMs) [48] or Adaboost [76] to identify objects based on those features. Recent advances rely on deep learning and can be divided into two main categories: two-stage and one-stage detectors. Two-stage detectors, such as Faster R-CNN [202], first generate candidate object regions and then classify and refine them. One-stage detectors, including YOLO (You Only Look Once) [200] and SSD (Single Shot MultiBox Detector) [143], perform detection in a single pass through the network, typically offering faster performance at the cost of some accuracy. Many of these methods use anchor boxes, which are predefined bounding boxes of different shapes and sizes, as initial estimates for object positions [202]. The network then predicts adjustments to these boxes and assigns confidence scores. More recently, anchor-free methods have been introduced that predict object locations directly without relying on predefined anchors [254, 290], offering simpler design and better adaptability. Non-maximum suppression (NMS) [171] is a common post-processing step used to remove duplicate detections by keeping only the highest-scoring box for each object and discarding overlapping boxes.

Image segmentation is another important task in computer vision. It involves dividing an image into regions that correspond to objects or parts of objects [161]. Traditional methods include thresholding, which assigns pixel classes based on intensity values compared to a threshold [180], and clustering techniques, such as k-means [150], which group pixels by similarity in features like intensity, color, or texture. While these methods are still useful, deep learning approaches have become dominant. Convolutional neural networks (CNNs) have achieved strong performance on segmentation tasks [161, 206], thanks to their ability to model complex features and contextual relationships within the image.

2.8 Deep Learning

Deep learning is a subfield of machine learning that uses artificial neural networks with multiple layers (hence deep) to automatically learn hierarchical representations from data [86, 129]. This representation learning capability distinguishes deep learning from traditional machine learning approaches, which often rely on manually engineered features. The motivation behind deep learning comes from the limitations of hand-engineered features, which can be time-consuming to design, suboptimal for complex tasks, and difficult to adapt to new data or domains. Deep learning automates the feature extraction process, allowing the network to learn relevant features directly from the input data [22]. The power of deep learning lies in its ability to model complex, non-linear relationships within large datasets, leading to significant advancements in various fields, including computer vision.

2.8.1 Artificial Neural Networks

Artificial Neural Networks (ANNs) form the foundation of deep learning. The basic building block of an ANN is the neuron (or perceptron) [207], which receives multiple inputs, performs a

weighted sum, adds a bias term, and then applies an activation function to produce an output. Mathematically, this can be represented as:

$$\text{output} = \text{activation} \left(\sum_i (w_i \cdot x_i) + b \right), \quad (2.4)$$

where x_i are the inputs, w_i are the corresponding weights, b is the bias, and activation is the activation function. Depending on the model architecture and also sometimes position of the neuron inside the network, the equation for the neuron can slightly change. For example, in modern transformer architectures and also often in convolutional layers, the bias term is typically left out. The reason for this is that the layers are mostly followed by a normalization layer, which effectively removes any bias term, and leaving the bias away reduces model overhead. Activation functions introduce non-linearity into the network, enabling it to learn complex patterns that cannot be modeled by linear functions alone. Common activation functions include the sigmoid function, the hyperbolic tangent (tanh) function, and the Rectified Linear Unit (ReLU) [167]. ReLU, defined as $f(x) = \max(0, x)$, is particularly popular due to its simplicity, computational efficiency, and ability to mitigate the vanishing gradient problem (discussed later) [84]. Neurons are organized into layers: an input layer that receives the initial data, one or more hidden layers that perform intermediate computations, and an output layer that produces the final result. In feedforward networks, the most common type of ANN, information flows in one direction, from the input layer through the hidden layers to the output layer, without any cycles or feedback connections [86].

2.8.2 Training

Training an ANN involves adjusting its weights and biases to minimize the difference between its predictions and the ground truth labels. This is achieved through a process that involves a loss function, backpropagation, and an optimization algorithm. The loss function (also called a cost function or objective function) quantifies the discrepancy between the network's output and the desired output (also called ground truth) for a given input [86]. The choice of loss function is task-specific. For example, mean squared error is commonly used for regression tasks, while cross-entropy is often used for classification. Regularization terms are often added to the loss function to prevent overfitting, where the model learns the training data too well and performs poorly on unseen data. L1 (also called Lasso regression) and L2 (also called Ridge regression) regularization, for example, add penalties based on the magnitude of the network's weights, thus encouraging smaller weights and simpler models [172].

Backpropagation is an algorithm for efficiently computing the gradient of the loss function with respect to the network's weights and biases [213]. It uses the chain rule to derive the gradient of the loss over each model weight or bias. This gradient indicates the direction and magnitude in which the weights and biases should be adjusted to reduce the loss. An optimization algorithm, such as Gradient Descent (GD), Stochastic Gradient Descent (SGD) [204], or Adam [120], is then used to update the weights and biases based on the computed gradient. These optimization algorithms differ primarily in how they compute and apply gradients.

Gradient Descent (Equation 2.5) calculates the gradient over the entire training set, resulting in a stable but potentially slow update. Stochastic Gradient Descent (Equation 2.6), on the other hand, estimates the gradient using only a small random subset of training samples (a mini-batch), which allows faster and more frequent updates at the cost of added noise. A common extension of SGD is the inclusion of momentum (Equation 2.7), which accumulates a velocity vector to smooth out updates and help the optimization process escape shallow local minima.

RMSProp (Equation 2.8) improves upon SGD by using a moving average of squared gradients to normalize the learning rate for each parameter, effectively damping oscillations in steep directions. Adam (Equation 2.9) combines momentum and RMSProp by computing both the first moment (mean) and the second moment (uncentered variance) of the gradients, and applies bias corrections to these estimates, resulting in an adaptive and efficient optimizer.

$$\text{Gradient Descent (GD): } \theta_{t+1} = \theta_t - \eta \cdot \nabla_{\theta} \mathcal{L}(\theta_t), \quad (2.5)$$

$$\text{Stochastic Gradient Descent (SGD): } \theta_{t+1} = \theta_t - \eta \cdot \nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t), \quad (2.6)$$

$$\begin{aligned} \text{SGD with Momentum: } v_t &= \gamma v_{t-1} + \eta \cdot \nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t) \\ \theta_{t+1} &= \theta_t - v_t, \end{aligned} \quad (2.7)$$

$$\begin{aligned} \text{RMSProp: } v_t &= \beta v_{t-1} + (1 - \beta) (\nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t))^2 \\ \theta_{t+1} &= \theta_t - \eta \cdot \frac{\nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t)}{\sqrt{v_t + \epsilon}}, \end{aligned} \quad (2.8)$$

$$\begin{aligned} \text{Adam: } m_t &= \beta_1 m_{t-1} + (1 - \beta_1) \nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t) \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) (\nabla_{\theta} \mathcal{L}_{\mathcal{B}}(\theta_t))^2 \\ \hat{m}_t &= \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \\ \theta_{t+1} &= \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \end{aligned} \quad (2.9)$$

In these equations, θ_t denotes the model parameters (weights and biases) at iteration t , and η represents the learning rate, a hyperparameter that controls the step size during optimization. The function \mathcal{L} refers to the loss, while $\nabla_{\theta} \mathcal{L}$ is the gradient of the loss with respect to θ , computed over the entire dataset or a mini-batch \mathcal{B} . For SGD with momentum, v_t is the accumulated velocity vector and γ is the momentum coefficient. In RMSProp, v_t tracks the exponentially weighted moving average of squared gradients, with β controlling the exponential decay rate. Adam extends this idea by introducing two separate moving averages: m_t , the first moment (mean) with decay rate β_1 , and v_t , the second moment (variance) with decay rate β_2 . The bias-corrected estimates \hat{m}_t and \hat{v}_t are used to ensure stable parameter updates, and ϵ is a small constant added to avoid division by zero.

Following the choice of optimization algorithm, normalization plays an important role in stabilizing and accelerating ANN training. Feature scaling ensures that input features have comparable value ranges, reducing the risk that features with larger magnitudes dominate the learning process. Common approaches include min-max scaling and standardization (z-score normalization). In addition to input normalization, Batch Normalization [105] is often applied to the activations within each mini-batch during training. This reduces internal covariate shift—changes in the distribution of layer inputs over time, helping to improve convergence speed and training stability. Layer Normalization [8] is a related method that normalizes across the features of individual training samples rather than across the batch dimension.

Training large and deep neural networks from scratch typically requires access to extensive datasets and substantial computational resources. Collecting and labeling such datasets can be expensive and time-consuming, and the computational demands of training may limit practical use, especially in resource-constrained settings. Two common strategies to mitigate these challenges are transfer learning and data augmentation.

Transfer learning [181] makes use of models pre-trained on large-scale datasets, such as ImageNet [55], and adapts them to new but related tasks. Instead of training a model from randomly initialized weights, training begins from a pre-trained model that has already learned a general set of features. This approach is particularly useful when the target dataset is small. The process of adapting the model to the new task is called fine-tuning. A common fine-tuning strategy is to freeze the early layers of the pre-trained model—those that typically extract low-level features such as edges or textures—and only update the later layers that are more task-specific. This allows the model to retain useful general representations while adapting to the new domain.

Even when using transfer learning, the quantity and variety of training data can significantly affect performance. Data augmentation [230] helps improve generalization by artificially increasing the size and diversity of the training dataset. This is done by applying label-preserving transformations to existing samples. Typical transformations include geometric changes such as rotation, flipping, cropping, scaling, and translation, adjustments to color properties like brightness, contrast, and saturation, and the addition of noise. These synthetic variations help the model learn to be invariant to minor input changes, which reduces overfitting and improves performance on unseen data.

Deep learning frameworks such as TensorFlow [1], PyTorch [184], and Keras [43] provide built-in support for both transfer learning and data augmentation, making these techniques accessible and widely adopted in practice.

An often overlooked yet important factor in training neural networks is the initialization of the model's weights. Poor initialization can lead to vanishing or exploding gradients, which may slow down or even prevent convergence during training. To address this, several initialization strategies have been proposed, each designed to preserve the variance of the activations and gradients across layers.

For layers using sigmoid or tanh activation functions, the Xavier (or Glorot) initialization [84] is commonly used. This method aims to maintain a consistent variance of activations between layers by drawing weights from a distribution with variance

$$\text{Var}(w) = \frac{2}{n_{\text{in}} + n_{\text{out}}}$$

where n_{in} and n_{out} are the number of input and output units of the layer, respectively. Weights are typically sampled either from a normal distribution

$$w \sim \mathcal{N}\left(0, \frac{2}{n_{\text{in}} + n_{\text{out}}}\right)$$

or from a uniform distribution

$$w \sim \mathcal{U}\left(-\sqrt{\frac{6}{n_{\text{in}} + n_{\text{out}}}}, \sqrt{\frac{6}{n_{\text{in}} + n_{\text{out}}}}\right)$$

For layers using ReLU or its variants (such as Leaky ReLU or PReLU), He initialization [96] (also known as Kaiming initialization) is preferred. ReLU-based activations zero out negative values, which affects the variance of the outputs. He initialization compensates for this by drawing weights from a distribution with variance

$$\text{Var}(w) = \frac{2}{n_{\text{in}}}$$

Common implementations sample from

$$w \sim \mathcal{N}\left(0, \frac{2}{n_{\text{in}}}\right) \quad \text{or} \quad w \sim \mathcal{U}\left(-\sqrt{\frac{6}{n_{\text{in}}}}, \sqrt{\frac{6}{n_{\text{in}}}}\right)$$

depending on whether a normal or uniform distribution is used.

For convolutional layers, the same initialization schemes are applied, with n_{in} and n_{out} corresponding to the number of input and output channels multiplied by the kernel size.

For recurrent neural networks (RNNs), weight initialization requires additional care due to the recursive nature of their computation. Orthogonal initialization is often used for recurrent weight matrices to help preserve long-term dependencies [224]. This method ensures that the gradient norms remain stable over time by initializing weights with orthogonal matrices drawn from a distribution that preserves the spectral radius.

Biases are typically initialized to zero, though in some cases (e.g., in LSTM networks), specific biases (such as the forget gate bias) are initialized to positive values to encourage gradient flow in early training.

Proper weight initialization improves training speed, stability, and the final performance of the network. Most modern deep learning frameworks support these initialization strategies out of the box, either as defaults or as configurable options for each layer.

2.8.3 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a widely used architecture in deep learning, particularly effective for processing data with a grid-like topology, such as images. The core component of CNNs is the convolutional layer [129], which applies a set of learnable filters (also referred to as kernels) to the input data. Each filter is a small matrix of weights that moves across the spatial dimensions of the input, performing an element-wise multiplication followed by a summation. This operation produces a single value in the output feature map at each location and is equivalent to the convolution operation defined in Equation 2.3, where $f(x, y)$ represents the input image (or feature map), $h(i, j)$ represents the filter weights, and $g(x, y)$ is the resulting output feature map.

The discrete 2D convolution with bias b can be written as

$$\text{Output}(x, y) = \sum_i \sum_j \text{Input}(x - i, y - j) \cdot \text{Filter}(i, j) + b \quad (2.10)$$

The stride determines how far the filter moves at each step. A stride of 1 results in a full coverage of the input, while larger strides reduce the spatial resolution of the output. Padding controls the spatial dimensions of the output feature map. Zero-padding, the most common form, adds rows and columns of zeros around the input border to preserve input dimensions when using stride 1.

Pooling layers typically follow convolutional layers and reduce the spatial dimensions of feature maps. This helps lower computational cost and introduces a degree of translational invariance, making the model less sensitive to small shifts in the input. Common pooling operations include max pooling, which selects the maximum value within a local region (e.g., a 2×2 window), and average pooling, which computes the mean value over that region. By aggregating information from nearby activations, pooling helps retain the presence of features even when their exact location varies slightly. For example, a 2×2 max pooling layer with a stride of 2 reduces the height and width of the feature map by half, while also increasing robustness to small translations in the input.

Fully connected layers are usually placed at the end of a CNN architecture. These layers operate like those in standard feedforward neural networks, where each unit is connected to every unit in the previous layer. They aggregate information extracted by preceding layers and are often used for classification or regression tasks.

The final output before classification is typically a set of logits, which are unnormalized scores for each class. For multi-class classification, these logits are passed through a softmax function to obtain class probabilities

$$P(y = c|\mathbf{x}) = \frac{e^{z_c}}{\sum_{k=1}^C e^{z_k}} \quad (2.11)$$

where z_c is the logit for class c , C is the total number of classes, and $P(y = c|\mathbf{x})$ is the probability of the input \mathbf{x} belonging to class c . For binary or multi-label classification, a sigmoid function is applied instead

$$P(y = 1|\mathbf{x}) = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.12)$$

where $\sigma(z)$ maps the logit z to a probability between 0 and 1.

CNN architecture has evolved through several notable developments. The Neocognitron [77] introduced early concepts such as local receptive fields and hierarchical feature extraction. LeNet-5 [129] applied convolutional and pooling layers to digit recognition and was one of the first CNNs trained with backpropagation. The success of AlexNet [55] on the ImageNet challenge highlighted the scalability of deep CNNs and introduced practices like ReLU activation, dropout, and data augmentation.

Subsequent models improved upon AlexNet in various ways. ZFNet [280] adjusted filter sizes and strides and offered visualizations to understand learned features. VGGNet [234] demonstrated the effectiveness of deeper architectures with smaller filters. GoogLeNet [244] introduced the Inception module, which combined filters of different sizes in parallel. ResNet [95] tackled the vanishing gradient problem by introducing residual connections, which allow gradients to propagate more effectively through very deep networks. DenseNet [101] extended this idea with dense connectivity between layers to encourage feature reuse.

Recent architectures have focused on improving computational efficiency. MobileNets [100, 222] use depthwise separable convolutions to reduce the number of parameters and operations. EfficientNet [248] scales network width, depth, and resolution in a balanced way using a compound scaling method. These architectures, while initially designed for image classification, are widely used as backbone networks for a variety of vision tasks.

2.8.4 Specialized Architectures

While CNNs have achieved strong performance in image classification, they face limitations in tasks that require precise spatial information or handling of objects at multiple scales. In particular, the repeated downsampling in standard CNNs reduces spatial resolution, making it difficult to localize object boundaries accurately in tasks such as semantic segmentation [146]. Additionally, relying on features from a single resolution, typically from the deepest layers, is often insufficient for detecting objects of varying sizes [137]. These limitations have led to the development of specialized architectures tailored to these challenges. Two widely used examples are U-Net and Feature Pyramid Networks (FPNs).

The U-Net architecture [206] was introduced for biomedical image segmentation, where precise pixel-level predictions are required and labeled data is often limited. Earlier encoder-decoder architectures for segmentation were able to extract high-level features but typically lost fine spatial details due to downsampling. This loss made it difficult to segment small or thin structures accurately.

U-Net addresses this problem using a symmetric encoder-decoder design with skip connections. The encoder, or contracting path, follows a standard CNN pattern of convolutional and pooling layers that reduce spatial dimensions while capturing higher-level context. The decoder, or expanding path, upsamples the feature maps using up-convolutions (also called transposed convolutions) and combines them with corresponding feature maps from the encoder through skip connections. These skip connections concatenate encoder feature maps with the decoder's, allowing the network to retain and reuse spatial detail that would otherwise be lost. This structure improves both localization accuracy and gradient flow during training. The architecture's layout forms a U-shape, which gives the model its name. Although originally developed for medical imaging, U-Net has been widely adopted for general semantic segmentation and related tasks.

Feature Pyramid Networks (FPNs) [137] were proposed to improve object detection, particularly for detecting objects at multiple scales. Standard CNNs extract increasingly abstract features at each layer, but deeper layers lack the spatial resolution needed for small objects, while earlier layers lack semantic richness.

FPNs build a multi-scale feature representation by combining features from different depths of a CNN. The architecture consists of three parts: a bottom-up pathway (the original backbone CNN), a top-down pathway, and lateral connections. The bottom-up pathway generates feature maps at progressively coarser spatial resolutions. The top-down pathway upsamples the deeper, semantically rich feature maps. Lateral connections merge these with higher-resolution feature maps from the bottom-up pathway using element-wise addition, typically after applying a 1×1 convolution to match dimensions. This produces a set of feature maps at multiple resolutions, each containing both semantic and spatial information. These maps are then used for detection at different object scales. FPNs are now a standard component in modern object detection and instance segmentation systems due to their improved performance on small and large objects alike.

Chapter 3

Development of Contactless Fingerprint Algorithms

3.1 Fingerprint Segmentation in Contactless Images

Capturing a fingerprint without physical contact offers advantages in convenience and hygiene, yet it comes with a problem: precisely identifying the fingertip within the captured image. This process, known as segmentation, involves digitally isolating the region of interest, the fingerprint area, from what is often a complex background. Figure 3.1 shows how an ideal segmentation algorithm would create an output mask 3.1b for a given, cropped single-finger input 3.1a. Note that this is the simplest case, where the cropped finger is provided as input.

The quality of this selection directly influences the success of subsequent stages, such as correcting the finger's three-dimensional pose and extracting features for comparison. Consequently, without reliable segmentation, the system's ability to accurately identify an individual is compromised.

This section first introduces the specific difficulties encountered in fingertip segmentation in contactless scenarios, followed by a review of existing approaches. Next, we will describe our proposed segmentation algorithms, explain the experimental setup and end with a discussion.

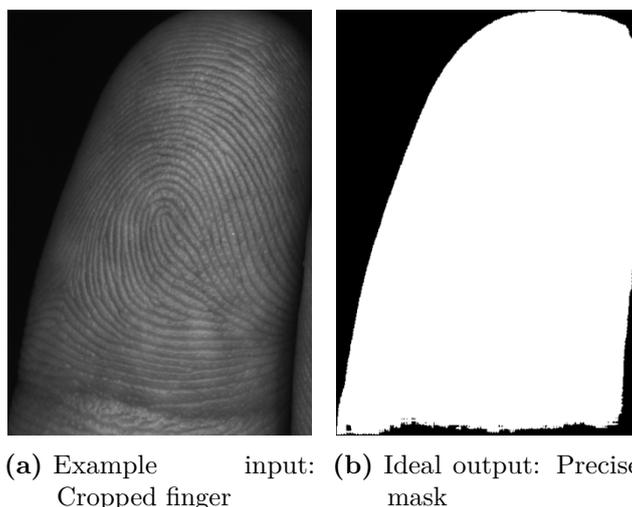


Fig. 3.1: Illustration of the input and desired precise output for single-finger segmentation.

3.1.1 Challenges of Contactless Fingerprint Segmentation

While contact-based approaches benefit from the finger being pressed against a platen, which typically ensures a relatively uniform background, contactless acquisition captures images in

unconstrained environments. This freedom results in highly variable and often complex backgrounds, potentially cluttered with arbitrary objects, diverse textures, or even other parts of the hand [116, 164, 165]. The effects of imprecise segmentation are substantial, as these inaccuracies directly propagate to subsequent processes. For instance, errors in the segmentation mask lead to a flawed 3D finger model, which is the basis for pose correction, thereby undermining the integrity of feature extraction and the ultimate comparison performance [217, 247].

Additionally, several factors inherent to contactless imaging further compound the segmentation challenge. The freedom in finger presentation allows for wide variations in roll, pitch, and yaw. Such diverse poses can lead to foreshortening, self-occlusion, and non-canonical views of the fingertip, making it difficult to define a consistent and recognizable shape for segmentation [215, 247]. Additionally, the visual appearance of the skin itself introduces variability. Differences in pigmentation or the presence of contaminants like dirt can alter pixel intensities within the fingertip region, complicating the delineation of a stable and clear boundary [164]. The visibility of the fingertip can also be compromised. It may be partially occluded by other fingers or not fully captured within the image frame, and such missing information makes it difficult to establish the true contour of the finger, again impacting segmentation accuracy [190]. Moreover, uncontrolled illumination conditions are a frequent issue. Uneven lighting, shadows cast by the hand or external objects, and specular reflections (glare) from the skin can obscure fingerprint details, create misleading false edges, or hide true boundaries, all of which can lead to significant segmentation errors [116]. Lastly, inconsistencies during the image acquisition process, such as subject movement or incorrect camera focus, can result in out-of-focus images or motion blur. These degradations reduce overall image quality, leading to blurred or indistinct fingertip boundaries that are inherently challenging to delineate with precision [190].

3.1.2 Related Work

In contact-based fingerprint acquisition, the finger is pressed against a sensor surface. While this provides a relatively controlled environment, segmentation is still necessary. One reason is to remove the noisy area at the borders of the image. Early, simple approaches often relied on pixel brightness thresholding techniques. For instance, Otsu's method [180] calculates a global threshold based on the image brightness histogram to separate foreground and background pixels [189]. This works well when there's a clear bimodal distribution in the histogram, as is sometimes the case with fingerprints, separating fingerprint and background [180]. However, variations in pressure, skin conditions, or sensor noise can make global thresholding less effective. Other traditional approaches involve analyzing local pixel features, like variance, mean, and coherence [19, 52]. These features can help distinguish the structured ridge-valley pattern of the fingerprint from the more random background noise. Rules based on combinations of those features can then be established to segment the image [53].

Contactless systems, which capture images using cameras or other optical devices, offer improved hygiene and usability [189], but suffer from issues that contact-based systems do not have, as described in the previous section 3.1.1.

To address these problems, researchers explored more sophisticated image processing techniques. One class of such methods are color-based approaches. Instead of operating directly on grayscale images, these methods leverage the color information present in contactless images [116]. The core idea is that skin exhibits a characteristic color signature, which can be used to differentiate it from the background. A common strategy is to transform the image from the standard RGB (Red, Green, Blue) color space to other color spaces like HSV (Hue, Saturation, Value), TSL (Tint, Saturation and Lightness) or others, where skin color tends to cluster more distinctly [116, 255]. For example, utilizing the TSL color space, a simple pre-filtering in the RGB image can be

done. Areas which are too blue/green and too dark, too yellow, or too green/blue compared to other colors, are filtered. Afterwards the image is transformed to the TSL color space. Then a normalization is done to get a normalized ST color space. Using upper and lower bounds on the S and T values, segmentation is performed [255]. Another method first transforms the image to the HSV color space, and then uses a normalized histogram, calculated from a reference region, to segment the image using back-projection [116]. Also an option is to use a Gaussian mixture model background subtraction, which models the background as a mixture of Gaussian distributions and then subtracts it from the image [116]. While these approaches can improve over simple thresholding, challenges related to consistent lighting and dynamic backgrounds remain. Edge detection using methods like the Canny edge detector [31] can also be considered, aiming to find the boundaries of the fingertip. However, noise sensitivity and incomplete edges in low-quality images are their limitations. Region growing algorithms, starting from seed points and expanding based on pixel similarity, offer another approach. But their reliance on good seed point selection and sensitivity to image variations makes them less stable.

More and more, the field has shifted towards machine learning, particularly deep learning, methods [164, 165, 190]. Convolutional Neural Networks (CNNs) have proven effective for image segmentation due to their capacity to learn hierarchical features from data. Ronneberger et al. [206] introduced U-Net, originally designed for biomedical image segmentation. Its encoder-decoder structure with skip connections allows for multi-scale feature fusion, leading to precise segmentation. Several U-Net variants have been proposed, each with specific improvements. U-Net++ [289] introduces nested and dense skip connections for enhanced feature fusion, while Squeeze-Unet [21] aims at reduced computational cost and memory footprint. EfficientUNet++ [232] utilizes the EfficientNet architecture to balance accuracy and efficiency. These U-Net variations have been applied to both single-finger segmentation [190, 215] and direct fingertip segmentation from hand images [216]. U-Net has also been adapted for 3D fingerprint enhancement [141].

Beyond U-Net, other powerful semantic segmentation models like DeepLab [40] and Mask R-CNN [94] have also been applied to contactless fingerprint segmentation [116]. A key building block of nearly all computer vision models has been convolution. A key variation of it is atrous convolution (dilated convolution). Unlike standard convolution, atrous convolution inserts spaces between kernel values. This expands the receptive field without increasing parameters or computation. The dilation rate controls the spacing, allowing the network to capture larger contextual information while preserving spatial resolution.

DeepLab, a semantic segmentation model, builds upon this. Early DeepLab versions used atrous convolutions to enlarge the receptive field. DeepLabv2 significantly advanced this with Atrous Spatial Pyramid Pooling (ASPP). ASPP uses parallel atrous convolutional layers, each with a different dilation rate, on the same input. Image pooling is also incorporated, for the global context. This captures multi-scale information. These features are combined, creating a representation for all object size variations. DeepLabv2 also utilized deeper backbone networks (e.g., ResNet-101) and a fully connected Conditional Random Field (CRF) [124] for boundary refinement, along changes of its learning rate policy.

While DeepLab and DeepLabv2 focuses on semantic segmentation, instance segmentation, which is identifying and delineating individual object instances, is addressed by the Mask R-CNN family of models. The progression starts with R-CNN (Regions with CNN features). R-CNN first uses a selective search algorithm to propose candidate object regions. Each region is then warped to a fixed size and fed into a convolutional neural network (CNN) to extract features. Finally, these features are used for classification and bounding box regression, leading to an object detection output.

Fast R-CNN improved upon R-CNN [83]. Instead of processing each region proposal separately, Fast R-CNN applies the CNN to the entire image to produce a feature map. Then, for each region proposal (still obtained using selective search), a Region of Interest (RoI) Pooling layer extracts a fixed-size feature vector from the feature map. This vector is used for classification and bounding box regression. This shared computation significantly speeds up the process.

Faster R-CNN further refined this by replacing the slow selective search with a Region Proposal Network (RPN) [202]. The RPN is a fully convolutional network that slides over the feature map. At each position, it predicts objectness scores and bounding box refinements for a set of anchor boxes (predefined bounding boxes). The second stage, identical to Fast R-CNN's RoI pooling, classification, and bounding box regression, then processes these RPN-generated proposals.

Mask R-CNN extends Faster R-CNN for instance segmentation [94]. It adds a mask branch, a fully convolutional network that predicts a binary mask for each proposal, indicating pixel-level object boundaries. Mask R-CNN also introduces RoIAlign. Unlike RoIPool (in Faster R-CNN), which can cause misalignments due to quantization, RoIAlign uses bilinear interpolation, preserving spatial accuracy and improving mask quality. Therefore Mask R-CNN detects objects and creates a pixel-accurate mask for each one.

The performance of these segmentation and detection models relies on the backbone network, which is the core architecture responsible for extracting features from the input image. One of the most influential ones is ResNet (Residual Network) [95]. ResNet introduced the concept of residual connections (or skip connections). These connections allow the gradient to flow more easily during training by adding the input of a block of layers to its output. This mitigates the vanishing gradient problem, where gradients become extremely small as they are backpropagated through very deep networks, hindering learning [95]. This allowed for the training of deeper networks.

As a next step, ResNeXt was developed [271]. ResNeXt introduces the concept of cardinality, which refers to the number of independent paths (transformations) within a building block. Instead of simply making the network deeper or wider, ResNeXt increases cardinality. Each path within a block performs the same operations (same topology), but on different subsets of the input channels. These parallel transformations are then aggregated. This split-transform-merge strategy allows the network to learn a richer set of features without a substantial increase in parameter count or computational complexity.

ResNeSt (Split-Attention Networks) further refines this concept [281]. It incorporates split-attention blocks. Within each block, the input is divided into groups (cardinality), and each group is further split into subgroups (radix). Attention mechanisms are then applied within each cardinal group, allowing the network to focus on different feature representations within each group. The outputs of these attentive groups are then aggregated. This allows for a form of channel-wise attention that is sensitive to different feature representations.

Another big architectural improvement came in the form of the Feature Pyramid Network (FPN) [137]. FPNs address the challenge of multi-scale object detection and segmentation a bit differently than the U-Net. A standard CNN produces a hierarchy of feature maps, with deeper layers representing higher-level semantic information but at lower spatial resolution. FPNs create a pyramid of feature maps, combining high-level, semantically strong features with low-level, high-resolution features. They do this using a top-down pathway and lateral connections. The top-down pathway upsamples the coarser, semantically richer feature maps, and the lateral connections merge them with the corresponding feature maps from the bottom-up pathway (the standard CNN feature extraction). This results in feature maps at multiple scales, all with rich semantic information, enabling the detection and segmentation of objects at various sizes. An example of the practical implementation is, for instance, TipSegNet, which leverages

a ResNeXt-101 backbone in combination with an FPN, demonstrating how these architectural components can be combined for specific tasks.

Processing pipelines have also evolved. Some approaches use a cascaded strategy, first employing object detection to locate fingers and then applying segmentation to those regions individually [116, 217]. While this simplifies the segmentation task, it depends on the initial detection accuracy. In contrast, direct fingertip segmentation aims to segment all fingertips directly from the hand image [216], eliminating the detection step and potentially increasing efficiency. In both cases, most approaches predict the segmentation mask. Priesnitz et al. in [190] however employed DeepLabv3+ to predict landmark points and construct circular fingertip segmentation masks based on the predicted hand features.

Evaluating segmentation performance, especially in scenarios like fingerprint analysis where the fingertip occupies a relatively small image portion, requires metrics beyond simple pixel-wise accuracy. Pixel-wise accuracy can be misleading due to significant class imbalance. The dominant background can lead to high accuracy even if the fingertip is entirely missed.

The Intersection over Union (IoU), also known as the Jaccard Index, addresses this issue. IoU is calculated by dividing the area of intersection between the predicted segmentation mask \hat{y} and the ground truth mask \bar{y} by the area of their union. This can be visualized as

$$\text{IoU} = \frac{|\hat{y} \cap \bar{y}|}{|\hat{y} \cup \bar{y}|} = \frac{\text{Area of Overlap}}{\text{Area of Union}}. \quad (3.1)$$

IoU values range from 0 to 1, with 1 indicating perfect overlap between prediction and ground truth. Because IoU considers both false positives and false negatives, it provides a balanced measure of segmentation quality, placing emphasis on accurately segmenting the smaller foreground region (the fingertip).

Several loss functions are commonly employed in segmentation, each with its own strengths and weaknesses.

One option is the Jaccard Loss, which is directly derived from the IoU metric. It is calculated as $1 - \text{IoU}$. A perfect overlap therefore corresponds to a Jaccard loss of 0. Because of its close relationship to IoU, Jaccard Loss naturally aims to optimize for the desired segmentation outcome. However, a potential drawback is that it can sometimes lead to unstable gradients during training, particularly in the initial stages. This instability stems from the fact that the Jaccard index (and thus the Jaccard Loss) is not smoothly differentiable when dealing with discrete pixel classifications (0 or 1). A small change in the predicted probabilities might not change the discrete segmentation mask, resulting in a zero gradient, or it might flip a pixel, causing a large jump in the loss, hence the instability. To mitigate this, a soft version of the Jaccard Loss is often used [264]. Instead of calculating the intersection and union based on discrete segmentation masks, the Soft Jaccard Loss uses the predicted probabilities directly. For example, if $p_{\hat{y}_i}$ represents the predicted probability of a pixel i belonging to the foreground and \bar{y}_i represents the ground truth (1 for foreground, 0 for background) of a given pixel i . The Soft Jaccard Loss can be formulated as

$$\text{Soft Jaccard Loss} = 1 - \frac{\sum_i^N (p_{\hat{y}_i} \bar{y}_i) + \epsilon}{\sum_i^N (p_{\hat{y}_i} + \bar{y}_i - p_{\hat{y}_i} \bar{y}_i) + \epsilon}, \quad (3.2)$$

where ϵ is a small constant (e.g., 10^{-7}) added to both the numerator and denominator to ensure numerical stability, preventing division by zero and N is the total number of pixels. This formulation allows for gradients to flow smoothly even when predicted probabilities are not exactly 0 or 1, leading to more stable training. The Soft Jaccard Loss approximates the discrete Jaccard Loss but provides a smoother optimization landscape.

Another common choice is Cross-Entropy Loss, a standard loss function in classification tasks, adaptable to pixel-wise segmentation. Cross-Entropy measures the dissimilarity between the predicted probability distribution for each pixel and its true label. For binary segmentation (fingertip vs. background), Binary Cross-Entropy (BCE) is used, while Categorical Cross-Entropy (CCE) is applied to multi-class problems

$$\text{BCE} = -\frac{1}{N} \sum_{i=1}^N (\bar{y}_i \log(p_{\hat{y}_i}) + (1 - \bar{y}_i) \log(1 - p_{\hat{y}_i})), \quad (3.3)$$

$$\text{CCE} = -\sum_c \sum_{i=1}^N \bar{y}_i^{(c)} \log(p_{\hat{y}_i}^{(c)}). \quad (3.4)$$

Note that c is the class index for C classes and $p_{\hat{y}_i}^{(c)}$ indicates the predicted probability of pixel i belonging to class c , and $\bar{y}_i^{(c)}$ the ground truth label of pixel i belonging to class c . While Cross-Entropy is generally well-behaved and stable during training, it can be dominated by the majority class (background) in imbalanced scenarios, potentially leading to poor performance on the minority class (fingertip).

To address the class imbalance issue, Weighted Cross-Entropy Loss can be employed. This is a modification of the standard cross-entropy, assigning different weights to each class. By giving a higher weight to the minority class (fingertip), the loss function is encouraged to pay more attention to its correct classification. While effective in mitigating class imbalance, Weighted Cross-Entropy requires careful tuning of the class weights, which can be dataset-dependent.

Dice Loss offers another approach, related in concept to IoU. It is based on the Dice coefficient, which is calculated as $2 \cdot (\text{Area of Overlap}) / (\text{Total number of pixels in both masks})$. The Dice Loss is then defined as $1 - \text{Dice coefficient}$. Similar to IoU, Dice Loss measures the overlap between sets, making it naturally suited for segmentation tasks. It often demonstrates good performance in imbalanced situations. However, like Jaccard Loss, it can sometimes exhibit unstable gradients, especially when the foreground region is very small.

Focal Loss was specifically designed to tackle class imbalance [138]. It modifies the cross-entropy loss by introducing a modulating factor γ . This factor down-weights the contribution of well-classified examples, effectively focusing the training process on hard examples – those that are misclassified or classified with low confidence. This focused approach makes Focal Loss very effective in handling class imbalance, and it often outperforms weighted cross-entropy. A drawback is the introduction of a hyperparameter (the focusing parameter) that requires tuning.

$$\text{FL} = -\sum_c \sum_{i=1}^N (1 - p_{\hat{y}_i}^{(c)})^\gamma \bar{y}_i^{(c)} \log(p_{\hat{y}_i}^{(c)}) \quad (3.5)$$

Tversky Loss generalizes both Dice Loss and Jaccard Loss [221], as can be seen in the following

$$\text{Tversky} = \frac{|\hat{y} \cap \bar{y}|}{|\hat{y} \cap \bar{y}| + \alpha |\hat{y} \setminus \bar{y}| + \beta |\bar{y} \setminus \hat{y}|} \quad (3.6)$$

It introduces two parameters, α and β , to control the penalties assigned to False Positives and False Negatives, respectively. This provides greater flexibility in adjusting the balance between precision and recall, according to the specific needs of the application. However, using Tversky loss needs tuning the α and the β parameters.

Sometimes, a combination of loss functions is used. Combining Dice Loss and Cross-Entropy Loss, for example, can leverage the strengths of both approaches. The choice of the most

appropriate metric and loss function depends on the specifics of the dataset and the desired balance between correctly identifying the fingertip and minimizing false positives. In fingerprint segmentation, where accurate delineation of the fingertip region is critical, metrics like IoU and losses such as Dice, Jaccard, or Focal Loss are frequently favored for their ability to address class imbalance and emphasize accurate foreground segmentation.

3.1.3 Proposed Segmentation Algorithm

This thesis explores three distinct deep-learning-based segmentation algorithms, each designed to address specific challenges in contactless fingerprint image analysis. These algorithms progressively build upon each other, improving in capabilities and addressing limitations found in previous versions. We will progress from models requiring pre-cropped single-finger images to a model capable of segmenting all four fingertips directly from a hand image.

3.1.3.1 Custom U-Net Based Architecture

The first model, introduced in our prior work, *Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints using Pose Correction & Unwarping* [217], represents a custom U-Net-based architecture. This model was designed for segmenting the fingertip region from a pre-cropped, single-finger image.

The architecture, depicted in Figure 3.2, is based on the foundational U-Net design, known for its effectiveness in biomedical image segmentation [206]. A pre-trained MobileNet V2 serves as the encoder [222], using transfer learning to extract low-level features. The decoder uses a series of transposed convolutional layers to upsample the feature maps, progressively reconstructing the segmentation mask. Critically, instead of traditional residual connections, we concatenate the output of the previous layer with the input of the current layer, a design choice that enhances feature propagation. Leaky ReLU activation functions are used throughout the decoder [272]. Batch normalization layers are incorporated to stabilize training and mitigate the vanishing gradient problem [105]. The final layer uses a sigmoid activation function to produce a binary segmentation mask.

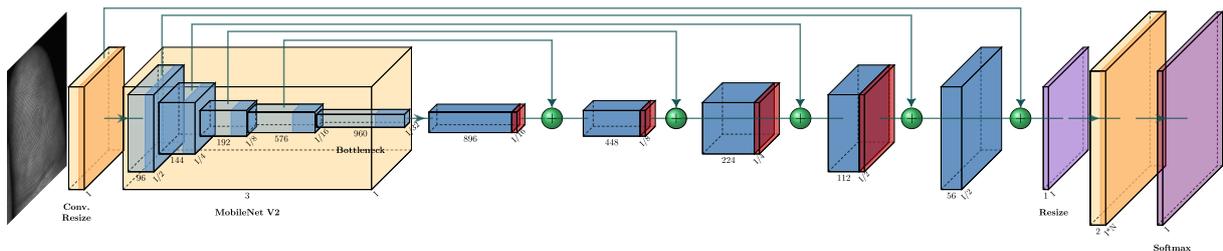


Fig. 3.2: Architecture of the custom U-Net model for fingertip segmentation

3.1.3.2 FingerUNeSt++

Building upon the U-Net-based approach, we developed FingerUNeSt++, an enhanced architecture that integrates elements from both ResNeSt and UNet++ [289]. This model, like its predecessor, operates on pre-cropped single-finger images. Its architecture is depicted in Figure 3.3. The model combines a ResNeSt-50d encoder with a simplified UNet++-like decoder. ResNeSt introduces split-attention blocks, enabling the network to focus on different feature subsets simultaneously. This improves the model's ability to capture diverse feature representations, important for dealing with the variability present in contactless fingerprint images. The ResNeSt-50d

variant offers a good balance between model depth and computational efficiency. The decoder employs a simplified UNet++ structure. We utilize nearest-neighbor upscaling, standard ReLU activations (instead of Leaky ReLU), and omit dropout layers. These simplifications streamline the architecture, reduce computational complexity, and improve training time, while retaining the benefits of the architecture. Following the UNet++ design, dense skip connections are used. These connections facilitate the fusion of multi-scale features by combining high-level semantic information from the decoder with low-level spatial information from the encoder.

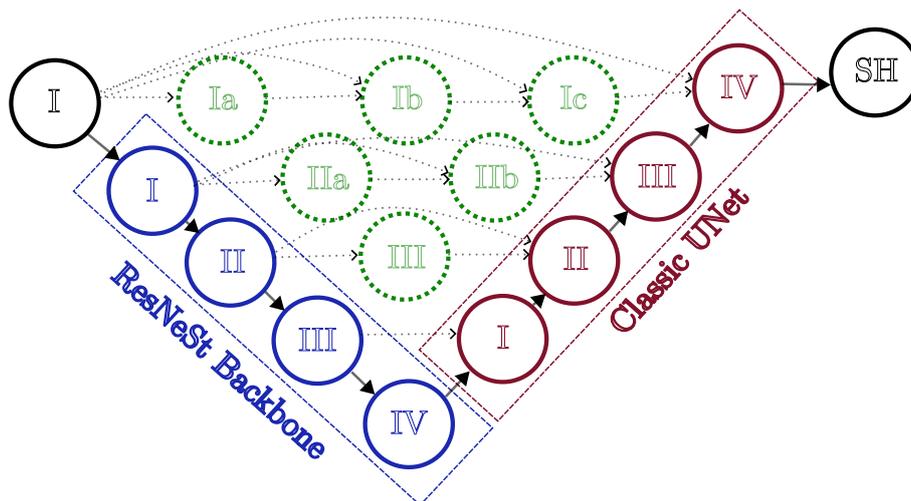


Fig. 3.3: Architecture of the FingerUNeSt++ model for fingertip segmentation

3.1.3.3 TipSegNet

The third model, TipSegNet, introduced in *TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging* [216], represents a significant advancement. Unlike the previous models, TipSegNet is designed to segment all four fingertips (excluding the thumb) directly from a whole-hand image, eliminating the need for prior finger detection and cropping. TipSegNet combines a ResNeXt-101 backbone [271] with a Feature Pyramid Network (FPN) decoder [137]. The ResNeXt-101 backbone provides feature extraction. ResNeXt introduces the concept of cardinality, increasing the model’s capacity by using groups of transformations. This allows for the capture of a richer and more diverse set of features, essential for distinguishing fingertips from complex backgrounds. The FPN decoder generates a multi-scale feature representation. It constructs a pyramid of feature maps, where each level represents the image at a different scale. This multi-scale approach ensures that the model can accurately segment fingertips regardless of their size or pose within the hand image. Lateral connections between the encoder and decoder combine high-resolution features from the encoder with semantically strong features from the decoder, further refining the segmentation.

To validate the design choices of TipSegNet, two key ablation studies were performed: First, the backbone ablation. The ResNeXt-101 backbone was compared with smaller ResNet architectures (ResNet-34, ResNet-50, and ResNet-101) to assess the impact of model complexity on segmentation performance. While all tested models achieved high accuracy (near 0.999) and mIoU (around 0.98), the larger ResNeXt-101 did provide a marginal improvement. This suggests that, for this specific task, performance is approaching saturation, and the benefits of increased model complexity are subtle. The study highlights the trade-off between computational cost and performance gains, suggesting that smaller models might be suitable for resource-constrained

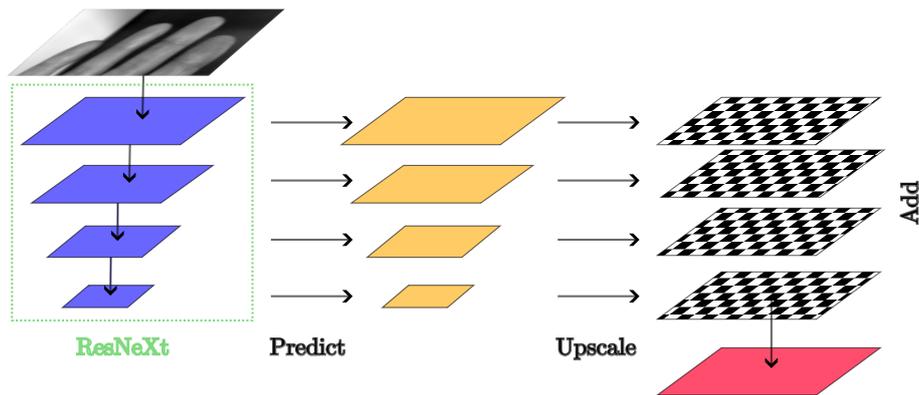


Fig. 3.4: TipSegNet model architecture.

environments. Second, the data augmentation ablation. The impact of the extensive data augmentation strategy was evaluated by training the model with: (a) full augmentation, (b) minimal augmentation (reduced strength), and (c) no augmentation. The results showed that while the prediction quality metrics (accuracy, mIoU) were very similar across all three scenarios, the training loss differed significantly. The model trained without augmentation reached the lowest training loss, indicating that it had learned the training data very well, potentially to the point of overfitting. The fully augmented model had a higher training loss, suggesting that the augmentations made the task more challenging. This is a desirable outcome as it suggests improved generalization ability to unseen data.

3.1.4 Experimental Evaluation and Results

3.1.4.1 Training Process

All three models were implemented and trained using the PyTorch 2 framework. A consistent training methodology was employed, with some variations in the optimization parameters. All models were trained using the Jaccard Loss function, also known as the Intersection over Union (IoU) loss. The Custom U-Net based architecture from [217] and FingerUNeSt++ [215] use the Stochastic Gradient Descent (SGD) optimizer [204]. To enhance model generalizability, and to mitigate overfitting, an extensive data augmentation strategy was applied during training. This strategy included a combination of geometric transformations and intensity adjustments: resizing and cropping, horizontal flipping, random rotations, perspective changes, Gaussian blur, solarization, posterization, and histogram equalization. These augmentations were applied randomly and with varying probabilities, exposing the model to a wide range of variations in finger pose, image quality, and simulated lighting conditions.

3.1.4.2 Model Complexity Comparison

Table 3.1 provides a comparison of the three models, as well as other SOTA models, in terms of trainable parameters, floating-point operations (FLOPs) and application scope. The Input column describes the expected model input, which can be either a cropped finger (Cropped Finger), resulting for a previous detection stage or the image of the whole hand, which removed the necessity of an object detection step before the segmentation. The output (Output) column describes whether the model predicts two classes, background and fingertip, or has an individual class for each finger's fingertip. This means that the model can differentiate between for example the index finger's tip and the middle finger's tip. The comparison in this table shows model

Model	Parameters	FLOPs	Input	Output
Squeeze U-Net [217]	2.5×10^6		Cropped Finger	Fingertip
EfficientUNet++ [217]	6.3×10^6		Cropped Finger	Fingertip
Custom U-Net [217]	1×10^7		Cropped Finger	Fingertip
FingerUNeSt++ [215]	5.1×10^7	5×10^{12}	Cropped Finger	Fingertip
DeepLabV3+ [116, 190]	6×10^6	1×10^{10}	Whole Hand	Fingertip
TipSegNet [216]	8×10^8	1×10^{12}	Whole Hand	Labeled Fingertip

Tab. 3.1: Overview of model complexity for the segmentation architectures evaluated. This table details the number of trainable parameters, Floating-Point Operations (FLOPs), input image type (Cropped Finger or Whole Hand), and output (Fingertip mask or Labeled Fingertip mask). Comparisons include models developed in this thesis (*Custom U-Net*, *FingerUNeSt++*, *TipSegNet*) and selected State-of-the-Art (SOTA) models.

complexity via their parameter counts and computational cost. It should be viewed in combination with the model efficacy comparison, which will be presented in section 3.1.4.4.

3.1.4.3 Dataset

In the paper *Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints using Pose Correction & Unwarping* [217], which evaluates Squeeze U-Net, EfficientUNet++ and a Custom U-Net, we leveraged the annotated dataset from [116] and extended it with new manual annotations based on the dataset captured in the work of Weissenfeld et al. in [266]. The annotations were conducted partially on the hand level, meaning the input image was the whole hand, and partially on the fingertip level. From the hand based annotations, we extracted the fingertips. Similarly, in *FingerUNeSt++: Improving Fingertip Segmentation in Contactless Fingerprint Imaging using Deep Learning* [220], we also work with cropped images of single fingertips to fit the application as a follow up step after the detection. The dataset contains 5,828 manually annotated images, of which 1457 images were used for validation and 1822 images for testing. The split was done on the user level, meaning no user appears in multiple sets (training, validation or test set).

The *TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging* paper [216] utilizes a combination of the dataset annotated for the two papers above and a substantial in-house dataset. The in-house dataset comprises of 2037 labeled hand images recorded using a smartphone, capturing hands against diverse background scenarios. This in-house dataset was combined with 220 manually annotated hand images from the dataset annotated in [217]. For training, 1788 images were used, with separate sets of 224 images each for testing and validation. Figure 3.5 provides an example of the training data used, along with added augmentations and annotations.

3.1.4.4 Model Efficacy Comparison

Table 3.2 presents a comparative analysis of various models’ performance in fingertip and finger segmentation tasks, utilizing Mean Intersection over Union (mIoU) as the primary evaluation metric, supplemented by accuracy where available. The table underscores the diversity of approaches, ranging from classical image processing techniques like Otsu thresholding and Gaussian Mixture models to advanced deep learning architectures such as EfficientUNet++, Squeeze U-Net, custom U-Net variations, FingerUNeSt++, DeepLabV3+, Mask R-CNN, HRNet,

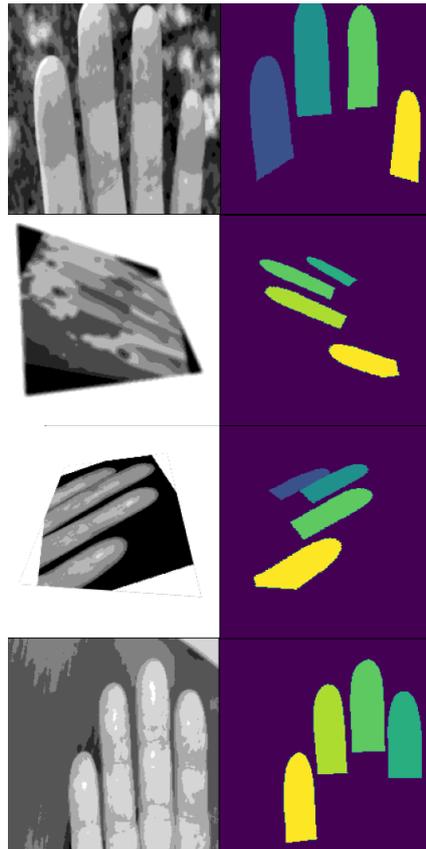


Fig. 3.5: Example of training data of the TipSegNet model, with augmentations used during the training.

Segnet, and TipSegNet. A critical aspect of this comparison lies in the distinction between the input data format (cropped single finger versus whole hand images) and the specific output objectives (segmenting individual fingertips or entire fingers, with or without finger identity labeling).

The models are broadly categorized based on their input scope: those designed for cropped finger images and those processing whole hand images. Furthermore, the output is differentiated between a binary mask isolating the fingertip or finger, and a labeled mask that additionally identifies the specific finger to which the segmented region belongs. Models developed within the scope of this thesis are denoted in italics.

Focusing initially on models designed for cropped single-finger inputs, traditional methods like Otsu's thresholding [180] achieve a respectable baseline mIoU of 0.85. However, as illustrated in Figure 3.6, the resulting masks often suffer from noise, incomplete coverage (particularly near the finger base), and inaccuracies within the fingertip region itself. Early deep learning models applied to this task, such as EfficientUNet++ [217], showed varied results, with this specific implementation yielding a lower mIoU (0.50) despite high accuracy, potentially indicating issues with boundary precision. Other architectures like Squeeze U-Net [217] and the *Custom U-Net* [217] developed during preliminary thesis work demonstrated significant improvements, reaching mIoU values of 0.86 and 0.91, respectively. Notably, the *FingerUNeSt++* model [215], also developed in this thesis, achieved near-perfect performance with an mIoU and Accuracy of 0.99. The qualitative difference is evident when comparing the noisy Otsu outputs (Figure 3.6) with

Model	mIoU	Accuracy	Input	Output
EfficientUNet++ [217]	0.50	0.88	Cropped Finger	Fingertip
Otsu [180]	0.85	0.85	Cropped Finger	Fingertip
Squeeze U-Net [217]	0.86	0.96	Cropped Finger	Fingertip
<i>Custom U-Net</i> [217]	0.91	0.98	Cropped Finger	Fingertip
<i>FingerUNeSt++</i> [215]	0.99	0.99	Cropped Finger	Fingertip
Gaussian Mixture [116]	0.31	-	Whole Hand	Finger
Color Histogram [116]	0.38	-	Whole Hand	Finger
Otsu [190]	0.92	-	Whole Hand	Finger
DeepLabV3+ [190]	0.95	-	Whole Hand	Fingertip
Mask R-CNN [116]	0.96	-	Whole Hand	Fingertip
HRNet [116]	0.85	-	Whole Hand	Labeled Fingertip
Segnet [116]	0.90	-	Whole Hand	Labeled Fingertip
DeepLabV3+ [116]	0.93	-	Whole Hand	Labeled Fingertip
<i>TipSegNet</i> [216]	0.99	1.00	Whole Hand	Labeled Fingertip

Tab. 3.2: Comparison of model efficacy across different architectures and input types. The mIoU (Mean Intersection over Union) is the primary metric, with accuracy provided where available. Models are sorted according to their mIoU performance inside each group. Note the distinction between models operating on cropped single fingers and those processing whole-hand images, as well as the output of either a single mask for each fingertip (Fingertip) or a labeled mask (Labeled Fingertip) that indicates to which finger this fingertip belongs to. The models written in *italic* were created in the context of this thesis.

the clean and precise masks generated by *FingerUNeSt++* (Figure 3.7), which closely resemble the ideal segmentation target shown in Figures 3.1a and 3.1b.

Transitioning to the more challenging task of segmenting fingertips from whole-hand images, the limitations of simpler methods become more apparent. Gaussian Mixture and Color Histogram approaches [116] yield very low mIoU scores (0.31 and 0.38), indicating their inadequacy for reliably locating and segmenting finger regions in complex backgrounds. While Otsu’s method, potentially with specific pre-processing or application context [190], performs better (mIoU 0.92), deep learning models again demonstrate superior capabilities for producing unlabeled fingertip masks. DeepLabV3+ [190] and Mask R-CNN [116] achieve high mIoU scores of 0.95 and 0.96, respectively, showing their effectiveness in identifying fingertip regions within the full hand context.

The most complex task involves not only segmenting but also correctly labeling each fingertip (e.g., distinguishing the index finger mask from the middle finger mask). For this Labeled Fingertip output category, several deep learning architectures like HRNet, DeepLabV3+, and Segnet [116] provide strong results, with mIoU scores ranging from 0.85 to 0.93. However, the *TipSegNet* model [216], developed as part of this thesis, sets a new benchmark, achieving an outstanding mIoU of 0.99 and a perfect Accuracy of 1.00. Figure 3.8 visually confirms this high performance, showing a random example from the test set, where *TipSegNet* accurately segments and labels all visible fingertips, closely matching the ground truth. Despite this exceptional performance, Figure 3.9 shows that even state-of-the-art models can have minor imperfections or struggle in particularly challenging cases, highlighting the complexities of image segmentation.

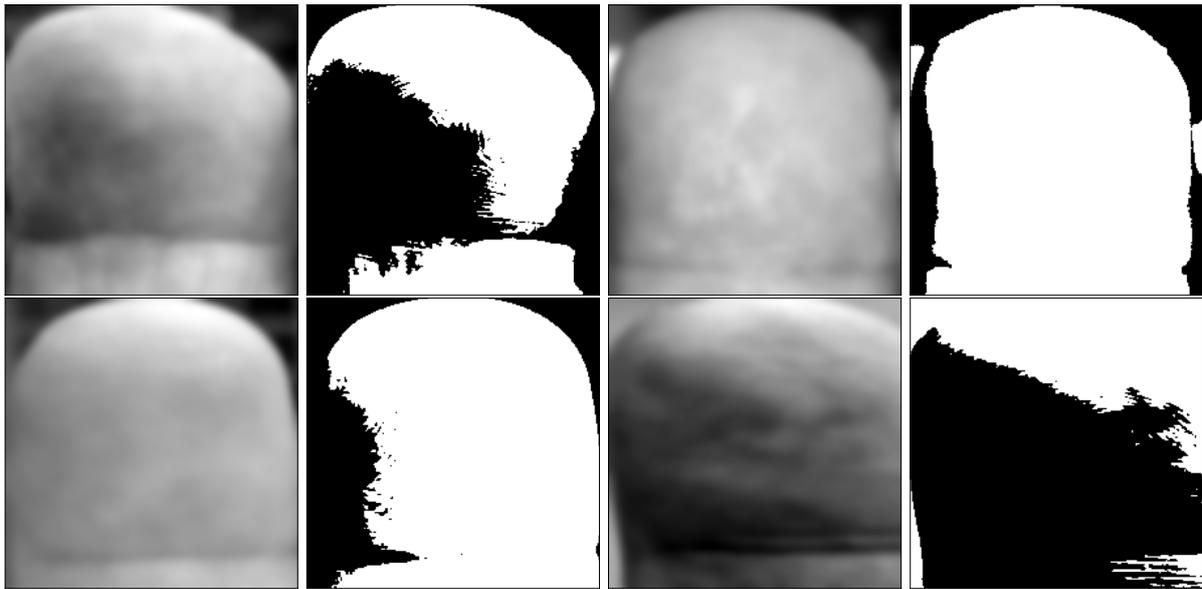


Fig. 3.6: Qualitative results of Otsu’s thresholding method applied to cropped fingertip images. The left image in each pair is the input, and the right is the generated binary mask. Note the presence of noise and incomplete segmentation, especially at the bottom edges.

3.1.5 Discussion

The experimental evaluation presented in the preceding sections highlights performance of the developed contactless fingertip segmentation models within this thesis. Moving beyond traditional methods like Otsu thresholding, Gaussian mixture or color histograms, which struggle with the inherent complexities of contactless imaging (as confirmed in Table 3.2 and Figure 3.6), the proposed architectures demonstrate substantial improvements in accuracy and generalizability.

The development of the segmentation algorithms in this thesis followed a deliberate progression, addressing increasingly complex challenges. The initial Custom U-Net (see Section 3.1.3.1) established a strong baseline for single-finger segmentation. However, to push the performance to near-perfect levels required for subsequent geometric modeling, the more advanced FingerUNeSt++ architecture (see Section 3.1.3.2) was developed, incorporating the feature-richness of ResNeSt. Both of these models, however, relied on a pre-cropped finger input. To create a truly practical end-to-end system and remove this dependency, the final and most advanced model, TipSegNet (see Section 3.1.3.3), was designed. Its ability to perform multi-class segmentation directly from a whole-hand image represents a significant step towards a fully automated pipeline.

3.1.5.1 Architectural Progression and Performance

These contributions address the challenges outlined in Section 3.1.1. The deep learning models, particularly *FingerUNeSt++* and *TipSegNet*, use hierarchical feature extraction capabilities inherent in their backbones (ResNeSt-50d and ResNeXt-101, respectively). This enables them to identify fingertip regions even in complex and variable backgrounds, which is the main limitation for simpler thresholding or color-based methods [116]. The extensive data augmentation strategies employed during training further enable the models to generalize across variations in lighting, skin tone, focus, and minor pose deviations, which are common pitfalls in contactless acquisition [116, 190].

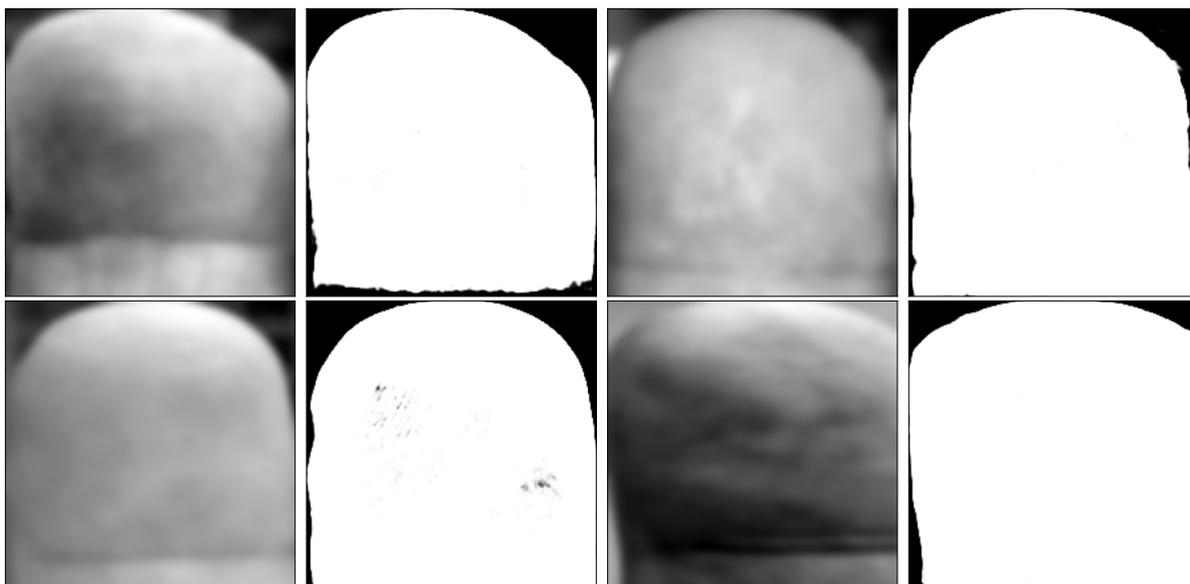


Fig. 3.7: Qualitative segmentation results from the *FingerUNet++* model on cropped fingertip images. The left image in each pair is the input, and the right is the predicted binary mask. The outputs are significantly cleaner and more accurate compared to the Otsu baseline (Figure 3.6).

Initial work of this thesis focused on refining segmentation for pre-cropped single-finger inputs, leading to *FingerUNet++* [215]. By combining the feature richness of ResNeSt with the dense connectivity of UNet++, this model achieved near-perfect scores (0.99 mIoU and Accuracy, Table 3.2), significantly outperforming both the baseline and earlier U-Net variants like the Custom U-Net developed in [217] (Figure 3.7).

As a next step, we approached the challenge of direct multi-finger segmentation from whole-hand images, resulting in *TipSegNet* [216]. This model eliminates the need for a cascaded approach often seen in related work [116, 217]. The combination of a ResNeXt backbone and a Feature Pyramid Network (FPN) proved highly effective. The FPN decoder explicitly addresses scale variations inherent in viewing multiple fingers at different distances and poses within a single hand image. *TipSegNet*'s state-of-the-art performance (0.99 mIoU, 1.00 Accuracy) showed that a multi-scale approach for complex scene understanding can be highly performant, setting a new benchmark for labeled fingertip segmentation from whole hands (Table 3.2 and Figure 3.8).

The ablation studies performed for *TipSegNet* provided valuable insights. While larger backbones like ResNeXt-101 offered marginal gains over smaller ResNets, the near-saturation of performance metrics suggests that for this task and dataset, the complexity might be approaching a point of diminishing returns. While SOTA performance was achieved, more computationally efficient backbones could offer a practical alternative in resource-constrained scenarios. Similarly, the data augmentation study confirmed its importance for generalization, making the learning task harder (higher training loss) but ultimately leading to a more robust model less prone to overfitting.

3.1.5.2 Implications and Limitations

Accurate segmentation is important for the reliability of entire contactless fingerprint recognition systems. As demonstrated in [217], precise fingertip masks are critical inputs for downstream

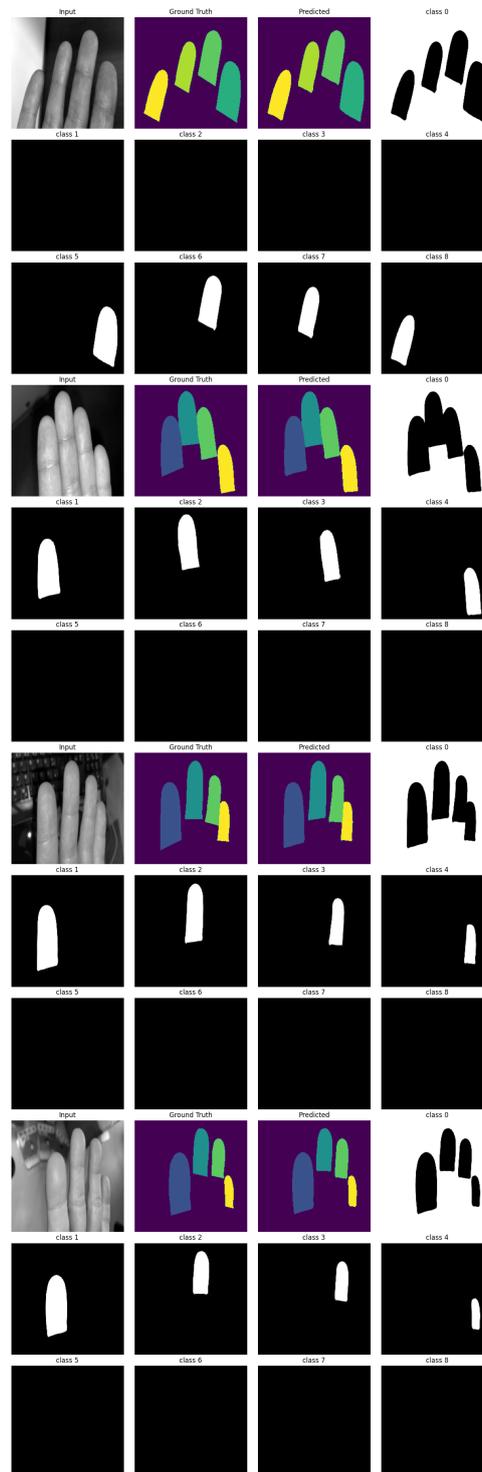


Fig. 3.8: Example of multi-finger segmentation and labeling by the *TipSegNet* model. From left to right: Input whole-hand image, Ground Truth segmentation (colored labels), Predicted segmentation (colored labels), and Background mask (class 0). Subsequent smaller images show individual predicted masks for each finger class. The prediction closely matches the ground truth.

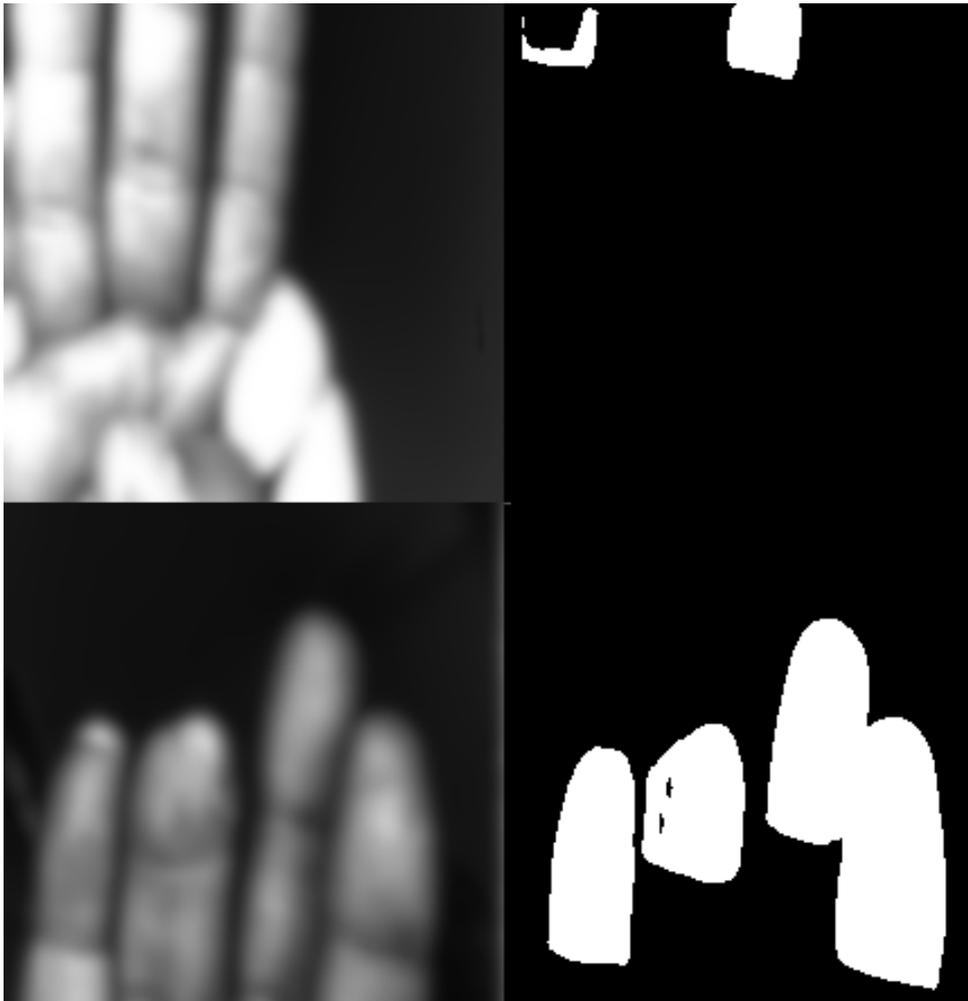


Fig. 3.9: Examples illustrating failure cases of the *TipSegNet* model. While overall performance is high, slight deviations from the ground truth can occur in the predicted masks, particularly around rare finger deformities or when the fingertips are largely covered or outside the frame. The input images are blurred for this visualization for privacy reasons.

tasks like pose correction and unwarping, which aim to bridge the gap between contactless and contact-based domains for improved interoperability. The high-fidelity masks generated by *FingerUNeSt++* and *TipSegNet* provide a significantly more reliable foundation for these subsequent steps compared to traditional methods or less accurate deep learning models.

While achieving SOTA performance, the models, especially *TipSegNet* with its 800M parameters (Table 3.1), demand considerable computational resources. This could pose challenges for deployment on low-power devices like smartphones or embedded systems, where models like Squeeze U-Net [217] might be preferred despite lower accuracy. Inference speed, while adequate on high-end GPUs (Table 2 in [215]), needs consideration for real-time applications on edge devices. Furthermore, despite extensive augmentation, generalization to truly unseen environments, extreme lighting conditions, or populations vastly different from the training data may still present challenges. Even *TipSegNet* exhibits minor imperfections, particularly at boundaries or in difficult cases (Figure 3.9), and its current implementation excludes the thumb.

3.2 Fingerprint Core Localization

For this section of the thesis, we transition from the segmentation of the fingertip vs background to finding a specific landmark point in the friction ridge pattern of the fingerprint. This landmark point is called the core and the core's location is frequently utilized in fingerprint analysis. Applications include template alignment prior to comparison [158], classification of the fingerprint pattern type [115], and serving as a reference point for estimating and correcting pose variations in contactless fingerprint systems [217, 247].

The core is commonly defined either as the point of maximum ridge curvature or the topmost point of the innermost recurring ridge [158]. This section addresses a fundamental question: How accurately can we detect fingerprint cores, and what are the inherent limitations of core-based alignment systems? We tackle this through two complementary approaches: first, developing a high-performance deep learning detector to achieve state-of-the-art detection accuracy [217], and second, conducting the first large-scale empirical analysis of core position variability to understand the theoretical limits of any core-based system that assume centrality of the core's location relative to the fingerprint pattern [220].

3.2.1 Challenges of Accurate Core Detection

Reliably locating the fingerprint core presents several challenges. General image quality degradation is a major factor. Sensor noise, low contrast between ridges and valleys, smudging artifacts (particularly in contact-based prints), or blur and lower resolution (common in contactless captures) can significantly obscure the fine ridge structures required for accurate core detection.

Contactless fingerprint imaging introduces additional complications unique to its capture modality. Specular reflections from the skin surface can create bright spots that interfere with pattern analysis. Uneven illumination across the finger can lead to inconsistent contrast, making ridge delineation difficult. Furthermore, perspective distortions arising from variations in finger pose (roll, pitch, yaw) relative to the sensor can alter the apparent location and shape of the core region within the captured image.

Beyond image quality and capture artifacts, the definition and even the existence of the core itself introduce ambiguity. As mentioned, different definitions based on ridge curvature [241] or the geometry of the innermost loop [9] (this is the definition provided in the ISO/IEC 19794-1:2011) are employed. Critically, cores are definitionally absent in certain fingerprint patterns, such as the arch type [5]. Other patterns, like whorls or double loops, can have two distinct cores [5]. This requires a clear strategy within any core-dependent algorithm for handling cases with no core or multiple cores, perhaps by selecting a canonical core or using alternative reference points.

The consequences of inaccurate core detection can impact downstream fingerprint processing tasks. Errors in core localization during the alignment stage lead to misregistered templates, resulting in reduced comparison scores and an increase in false rejections (missed matches). In pose correction frameworks that heavily rely on the core as a primary reference point [217, 247], localization errors directly propagate into the geometric transformation calculations, potentially leading to incorrect pose estimates and hindering interoperability efforts.

3.2.2 Related Works

Traditional algorithms for fingerprint core detection have primarily focused on analyzing characteristics of the ridge flow pattern. One category of methods analyzes the fingerprint orientation field, identifying cores by detecting specific rotational patterns using techniques like the Poincaré Index [17] or by applying complex filtering methods [118]. Another category focuses on ridge curvature,

aiming to locate the core at points where curvature is maximal [241]. Other approaches attempt to directly identify specific ridge structures according to standard definitions, such as locating the topmost point of the innermost recurving loop as specified in ISO/IEC 19794-1:2011 [9]. It is also common practice, particularly in commercial Automated Fingerprint Identification Systems (AFIS) like the Innovatrics IDKit used in our centrality analysis [220], to employ sophisticated, often proprietary (black-box), methods for core detection. These traditional methods often lack generalization over different modalities, being sensitive to image noise, variations in quality, and the ambiguities in core definitions.

One research direction frames core detection as a keypoint regression problem, where a neural network is trained to directly predict the (x, y) coordinates of the core [250]. An alternative research strategy, adopted in this thesis [217] and detailed below, treats core detection as a segmentation or heatmap prediction task. In this formulation, the network outputs a 2D probability map where the intensity at each pixel indicates the likelihood of the core's presence. This heatmap approach can better represent and handle the uncertainty associated with the core's precise location in the case of no or multiple cores. Architectures like the U-Net [206] have proven particularly well-suited for such segmentation-based localization tasks [217].

While methods for detecting the core are relatively well-studied, the systematic analysis of the core's statistical distribution and its typical location relative to the overall finger shape (its centrality) has received limited attention. The work presented in [220], discussed later in this section, provides, to our knowledge, the first large-scale empirical investigation into core centrality across substantial datasets of both rolled and plain fingerprint captures.

3.2.3 A Deep Learning Approach for Fingerprint Core Detection

Addressing the need for core detection, this thesis developed and evaluated a U-Net based deep learning model, leveraging the heatmap prediction approach also used for segmentation in [217].

3.2.3.1 Proposed Heatmap-Based Detection Algorithm

To implement the core detection, we utilized the same custom U-Net-based architecture previously described for fingertip segmentation (Section 3.1.3.1, Figure 3.2). This architecture features a MobileNetV2 [222] encoder backbone combined with a U-Net decoder structure incorporating skip connections (via concatenation), batch normalization [105], and leaky ReLU activations [272]. As mentioned, the core detection task was framed as a pixel-wise segmentation problem. Instead of outputting coordinates, the model is trained to output a 2D probability map, often called a heatmap. The intensity at each pixel in this heatmap represents the likelihood of the core's presence. This approach provides a richer output than simple coordinates and implicitly encodes spatial uncertainty about the core's exact position.

3.2.3.2 Annotation Strategy for Core Location Training

A key aspect of training this model is the annotation strategy for the ground truth heatmaps (conceptually shown in Figure 3.10). Rather than labeling a single ground truth pixel for the core location, we employed a target representation centered on the core but spatially distributed. Specifically, the ground truth was rendered as a circular region with a radial probability fall-off, following a 2D Gaussian profile. This soft annotation approach was chosen for several reasons. It provides a degree of tolerance for ambiguities and minor inaccuracies in annotating the exact core location. It also yields smoother gradients during the backpropagation compared to sparse single-pixel labels, which can contribute to more stable and effective model training. Furthermore,

this probabilistic target better represents the positional uncertainty often associated with the core concept itself.

The ground truth core locations used to center these heatmap annotations were primarily based on the ISO definition (innermost loop’s topmost point) where applicable [217, 220]. Importantly, for fingerprint pattern types lacking a true core, such as arches, a consistent landmark was still needed for potential use in subsequent pose correction steps. Therefore, during dataset creation for these arch-type cases, a reference point visually placed in the center of the fingerprint pattern was manually annotated to serve as the target location for the heatmap [217, 220]. This ensures the model learns to predict a consistent reference point even in the absence of a formally defined core. Of course, if the core detection model is designed to be used for classification of fingerprint pattern types, this step should be left out.

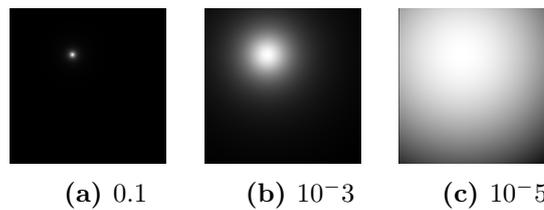


Fig. 3.10: Conceptual illustration of annotated core position for different fall-off rates, similar to Figure 3 in [217]. (a) Fast fall-off (e.g., 0.1); (b) Medium fall-off (e.g., 10^{-3}); (c) Slow fall-off (e.g., 10^{-5}).

3.2.3.3 Implementation and Training Details

The model benefited from transfer learning, initializing the MobileNetV2 encoder with weights pre-trained on a large image dataset [222]. The network was trained using a binary cross-entropy loss function, suitable for the probabilistic heatmap output (predicting the likelihood of the core’s presence vs absence at each pixel). Optimization was performed using Stochastic Gradient Descent (SGD) [204]. As implemented, this core detection model operates on pre-cropped single-finger images, meaning it relies on an upstream process (like the segmentation model described in Section 3.1) to first detect and isolate the fingertip region from the broader image context.

3.2.3.4 Performance Evaluation of the Core Detector

To assess the effectiveness of the developed core detection model, a rigorous evaluation was conducted [217]. A dedicated dataset was specifically created for this purpose, using a subset of the contactless fingerprint data collected in [266]. Annotations involved a semi-automatic process: initial core locations for the training set were generated using the Innovatrics IDKit commercial comparison algorithm [103], followed by manual review and correction to improve accuracy. To ensure the highest quality ground truth for evaluation, the validation and test set annotations were performed entirely manually. The dataset was carefully divided into user-disjoint training, validation, and testing splits to prevent data leakage and allow for a realistic assessment of generalization performance.

The primary metric for evaluating localization accuracy was the Average Euclidean Distance, measured in pixels, between the predicted core location (derived from the maximum likelihood pixel in the output heatmap) and the manually validated ground truth annotation. Additionally, the average number of distinct connected regions detected in the heatmap output was monitored to assess whether the model was prone to producing spurious, non-core detections [217].

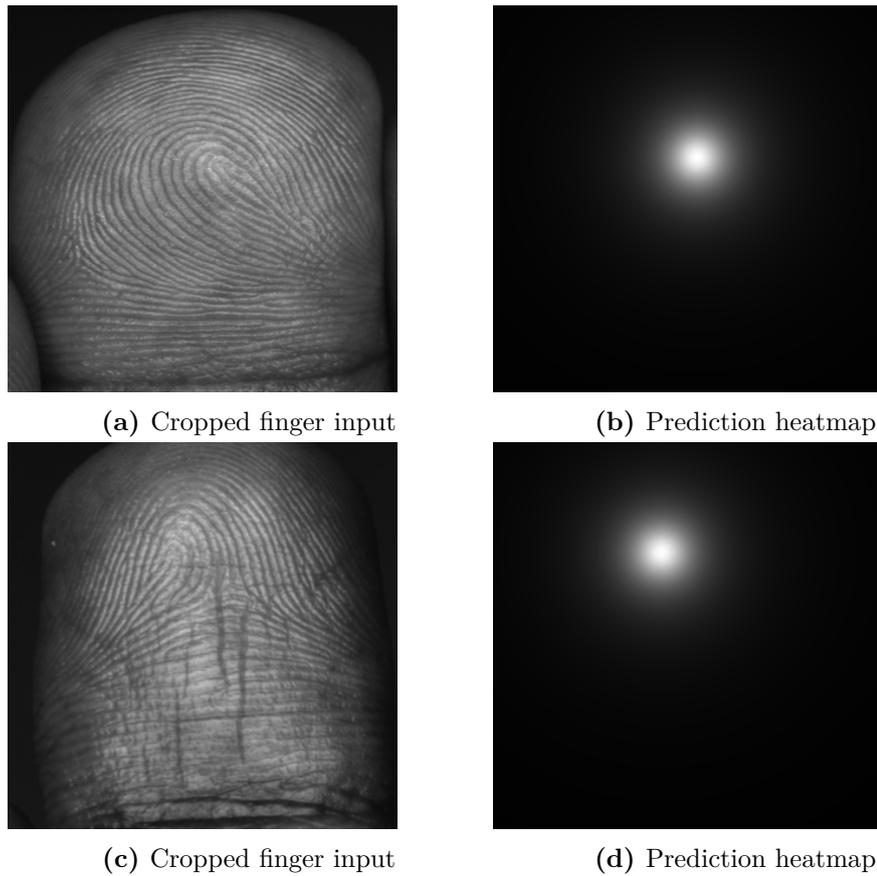


Fig. 3.11: Exemplary core position predictions by the Custom U-Net model on images from the test set.

The performance of our custom U-Net architecture was benchmarked against two contemporary, efficient U-Net variants adapted for the same core detection task: Squeeze U-Net [21] and EfficientUNet++ [232]. Furthermore, an ablation study was performed specifically on our custom U-Net model to investigate the impact of the ground truth heatmap’s radial fall-off rate on detection accuracy. This involved retraining and evaluating the model with three different fall-off profiles, corresponding to rates of 0.1, 10^{-3} , and 10^{-5} [217].

3.2.3.5 Results

Model	Avg. Distance [px]	Inf. Speed [s]
<i>Custom U-Net</i> [217]	9.49	0.16
EfficientUNet++ [232]	11.03	0.12
Squeeze U-Net [21]	9.63	0.10

Tab. 3.3: Average pixel distance (Avg. Distance) and inference speed (Inf. Speed) is seconds of the tested models running on a GTX 3090 with a fall-off rate of 10^{-5} . For both metrics, lower is better. The *Custom U-Net* is the model developed in the context of this thesis.

The quantitative evaluation results, presented in detail in Table 3.3 and in publication [217, Table 3], confirmed the effectiveness of the proposed custom U-Net architecture. It achieved the lowest average Euclidean pixel distance error, measuring 9.49 pixels, outperforming both Squeeze U-Net (9.63 pixels) and EfficientUNet++ (11.03 pixels) on this task. This accuracy was attained while maintaining competitive inference speeds compared to the other models. Figure 3.11 shows the models performance on a random example from the test set.

The ablation study focusing on the annotation fall-off rate provided a clear finding: the slowest fall-off rate tested (10^{-5}), corresponding to the smoothest and broadest target heatmap, resulted in the best detection accuracy (9.49 pixels vs. 23.46 pixels for 10^{-3} and 38.47 pixels for 0.1).

The results suggests that providing a less sharply localized, more distributed target signal during training is beneficial for learning accurate core localization with this heatmap prediction approach. Encouragingly, all tested models consistently learned to predict a single, connected region corresponding to the core location, indicating they were not prone to generating spurious detections.

3.2.4 Empirical Analysis of Fingerprint Core Centrality and Distribution

In addition to developing a detection model, the second approach involved a large-scale empirical investigation into the position and distribution of the fingerprint core within the finger area. This analysis, detailed in *Centrality of the Fingerprint Core Location* [220], provides insights into the properties of this landmark.

3.2.4.1 Datasets and Methodology for Centrality Analysis

This analysis utilized a larger and more diverse data collection than the detector evaluation. Six datasets were combined: the publicly available NIST Special Databases SD300a, SD302a, and SD302b [73, 74], the Neurotechnology dataset, the PolyU contact-based dataset [136], and the in-house AIT dataset collected in [266]. Some of the datasets provide a finger position description for each of the recorded fingerprints, which we denote in the FGP scheme introduced by NIST [168, p. 18], depicted in 3.4.

Finger	Right	Left
Thumb	1	6
Index	2	7
Middle	3	8
Ring	4	9
Little	5	10
Plain Thumb	11	12
Unknown	0	0

Tab. 3.4: FGP values mapped to finger names.

This aggregation of datasets provided a substantial number of samples for both rolled fingerprint impressions (30,602 samples) and plain (flat) fingerprint impressions (9,568 samples), allowing for comparisons between these common capture types. Table 3.5 provides an overview of the number of fingers present in each dataset.

The methodology involved several processing steps applied to each fingerprint image. First, the fingerprint region was segmented from the background using NIST’s publicly available nfseg

FGP	AIT	NIST 300a	NIST 302a	NIST 302b	+
1	1083	815	935	521	3354
2	875	801	981	542	3199
3	528	819	981	605	2933
4	523	832	984	632	2971
5	941	803	952	616	3312
6	1077	797	889	505	3268
7	512	806	963	575	2856
8	522	813	942	609	2886
9	525	816	993	630	2964
10	525	791	940	603	2859
+	7111	8193	9560	5829	30602

(a) Number of rolled fingerprints.

FGP	NIST 300a	NIST 302b	Neurotechnology	PolyU	+
0	0	0	840	2776	3616
1	0	226	0	0	226
2	704	228	0	0	932
3	662	255	0	0	917
4	768	270	0	0	1038
5	653	244	0	0	897
11	779	11	0	0	790
6	0	236	0	0	236
7	771	236	0	0	907
8	650	234	0	0	884
9	779	257	0	0	936
10	675	231	0	0	906
12	748	151	0	0	899
+	7189	2579	840	2776	9568

(b) Number of plain fingerprints.

Tab. 3.5: Finger-wise overview of dataset statistics for rolled and plain fingerprints used in the empirical study.

tool, which also rotates the segmented fingerprint to align its primary axis of symmetry with the vertical axis. Second, the core point(s) were detected using the commercial Innovatrics IDKit software [103]. If multiple cores were detected, their mean position was used. Third, based on the segmented finger area and the detected core position, several metrics were calculated to quantify the core’s location relative to the finger’s geometric center. These metrics, formally defined in [220], are important for understanding core position variability:

- **Core Offset** (o_{core}^x, o_{core}^y): This is the raw distance, in pixels, of the detected core point from the geometric center of the cropped (segmented) fingerprint image. o_{core}^x represents the horizontal offset, and o_{core}^y represents the vertical offset.
- **Relative Core Offset** (ro_{core}^x, ro_{core}^y): To make comparisons independent of image size or finger width/height, the raw offsets are normalized. The relative core offset is calculated

by dividing the raw offset by the half-width ($W_{cropped}/2$) or half-height ($H_{cropped}/2$) of the cropped fingerprint image. For the horizontal component

$$ro_{core}^x = \frac{o_{core}^x}{W_{cropped}/2} = \frac{2 \cdot o_{core}^x}{W_{cropped}} \quad (3.7)$$

The calculation for ro_{core}^y is analogous using the image height $H_{cropped}$. These relative offsets range from -1 (core at the left/top edge) to +1 (core at the right/bottom edge), with 0 indicating perfect geometric centering.

- **Bias (b_{core}):** This metric specifically quantifies any systematic horizontal shift in the core's position, potentially introduced by procedural factors like rolling. It is calculated as the mean of the horizontal relative core offsets (ro_{core}^x) across all samples for a specific finger or group

$$b_{core}(f) = \frac{1}{N_f} \sum_{i=1}^{N_f} ro_{core}^x; i, \quad (3.8)$$

where f denotes the specific finger position (e.g., right index), N_f is the number of samples for that finger, and i indexes the samples. A non-zero bias suggests a systematic tendency for the core to be offset horizontally.

- **Average Relative Offset (ARO, $aro_{core}^x, aro_{core}^y$):** This measures the average magnitude of the core's deviation from the geometric center, without correcting for any systematic bias. It is calculated as the mean of the absolute values of the relative core offsets for a given finger f ,

$$aro_{core}^x(f) = \frac{1}{N_f} \sum_{i=1}^{N_f} |ro_{core}^x; i| \quad (3.9)$$

The calculation for aro_{core}^y is analogous. ARO provides a measure of the overall expected deviation from the center in practical scenarios where bias might be present.

- **Corrected Average Relative Offset (CARO, $caro_{core}^x$):** This metric measures the average magnitude of the core's deviation after accounting for the systematic horizontal bias (b_{core}). It isolates the random or non-systematic component of the core's horizontal scatter around its mean (biased) position

$$caro_{core}^x(f) = \frac{1}{N_f} \sum_{i=1}^{N_f} |ro_{core}^x; i - b_{core}(f)| \quad (3.10)$$

CARO helps understand the core's positional variability independent of procedural biases. (Note: A vertical bias correction was deemed unnecessary as no significant systematic vertical bias was found comparable to the horizontal rolling bias).

3.2.4.2 Key Findings from the Centrality Analysis

The analysis yielded several findings regarding core location properties. A statistically significant horizontal bias (b_{core}) in core position was observed for rolled fingerprints, as can be seen in Table 3.6. Critically, the direction of this bias differed systematically between left-hand fingers compared to right-hand fingers. This systematic left-right difference, contrasted with the finding that such consistent bias was largely absent in plain fingerprint impressions (Table 5 of [220]), strongly suggesting that the bias originated primarily from the mechanics of the rolling procedure itself

(likely related to finger rotation direction or pressure application) rather than from anatomical asymmetry.

FGP	Bias [%]	FGP	Bias [%]
1	2.90 (4.10)	6	0.42 (-0.78)
2	0.67 (1.14)	7	0.33 (-0.17)
3	3.27 (4.31)	8	-1.11 (-1.92)
4	5.17 (5.33)	9	-3.62 (-3.77)
5	4.89 (5.67)	10	-0.50 (-0.78)

Tab. 3.6: Rolled fingerprint bias for averaged core x-Positions. Values in parenthesis show the bias for core x-positions when multiple cores inside a given fingerprint are not averaged over.

Furthermore, even after computationally correcting for this identified rolling bias, the analysis revealed that the fingerprint core does not, on average, reside precisely at the geometric center of the segmented fingerprint area. The corrected average relative offset in the horizontal direction ($caro_{core}^x$) was found to range between 5.7% ($\pm 5.2\%$) and 7.6% ($\pm 6.9\%$) of the half-finger-width, varying slightly depending on the specific finger. This can be seen in Table 3.7. A significant average deviation from the center was also consistently observed in the vertical direction (aro_{core}^y), ranging from approximately 9% ($\pm 6\%$) to 12% ($\pm 8\%$) of the half-finger-height for the rolled print samples analyzed. This empirically quantifies the typical extent to which the core is non-central.

FGP	CARO (ARO) [%]	FGP	CARO (ARO) [%]
1	6.7 ± 5.6 (7.4 ± 5.9)	6	7.2 ± 6.2 (7.7 ± 6.4)
2	7.6 ± 6.9 (8.5 ± 7.1)	7	7.0 ± 6.0 (7.7 ± 6.3)
3	6.6 ± 5.7 (7.3 ± 6.0)	8	7.0 ± 5.7 (7.4 ± 6.0)
4	5.7 ± 5.2 (6.6 ± 5.6)	9	6.3 ± 5.5 (6.9 ± 5.8)
5	6.4 ± 5.3 (7.2 ± 5.5)	10	6.6 ± 5.4 (7.2 ± 5.7)

Tab. 3.7: CARO (ARO) scores for rolled fingerprints

This procedure is often described as the parametric bootstrap test in literature [81, 121, 242].

Investigating the statistical nature of this positional variability, standard tests for normality (specifically, the Anderson-Darling test) strongly rejected the hypothesis that core positions follow a simple Normal (Gaussian) distribution, particularly for the dataset of rolled fingerprints. To identify a better fitting statistical model, further analysis using the Bayesian Information Criterion (BIC) [227] and Monte Carlo Goodness-of-Fit tests was conducted. The BIC combines the complexity of a model with its performance into one score and we used it to limit the search space of possible distributions. We started with a set of 110 distributions and selected a set of the 10 best matching distributions for each finger. For the next step, we used the Monte Carlo Goodness-of-Fit test, which is also often described as parametric bootstrap test in literature [81, 121, 242], to give a final ranking. Traditionally, goodness-of-fit tests use pre-calculated critical values based on fixed significance levels. The Generalized Monte Carlo goodness-of-fit procedure however performs Monte Carlo trials adapted to their particular data and this works as follows. First, the estimation of unknown parameters in the specified distribution family is

conducted using maximum likelihood estimation. The estimated parameters define the null-hypothesized distribution, representing the distribution from which the data were assumed to be sampled under the null hypothesis. We calculate the Anderson-Darling statistic for the given finger-wise ro_{core} . Subsequently, numerous new samples, each comprising of the same number of observations as the original data, were drawn from the null-hypothesized distribution. For each drawn sample, unknown parameters were re-estimated, and the corresponding Anderson-Darling statistic was calculated between the sample and its fitted distribution. These computed statistic values collectively formed the Monte Carlo null distribution, distinct from the aforementioned null-hypothesized distribution. The p-value of the goodness-of-fit test is determined as the proportion of statistic values in the Monte Carlo null distribution that are as extreme as or more extreme than the statistic value computed for the provided data. Specifically, the p-value is calculated as the ratio of the number of statistic values in the Monte Carlo null distribution greater than or equal to the statistic value of the data plus one to the total number of elements in the Monte Carlo null distribution also plus one. [135]

This analysis indicated that the *Non-Central Fischer (NCF) distribution* provided the best statistical fit for describing the core's positional distribution in most of the scenarios examined, especially for rolled fingerprints. The results for all fingers can be seen in Table 3.8.

-	Plain				Rolled			
	x		y		x		y	
	Dist.	p [%]	Dist.	p [%]	Dist.	p [%]	Dist.	p [%]
0	Dag	70	NCF	78	-	-	-	-
1	LNo	67	NCF	20	Bur	11	NCF	95
2	LNo	56	Bur	60	NCF	32	NCF	71
3	Bur	47	NCF	80	NCF	37	Bur	37
4	NCF	94	NCF	55	NCF	85	Bur	29
5	NCF	50	NCF	73	NCF	84	Bur	42
6	NCF	84	NCF	44	NCF	65	Bur	36
7	NCF	56	Bur	55	NCF	73	NCF	74
8	NCF	28	NCF	54	NCF	33	NCF	30
9	NCF	67	Bur	31	NCF	60	NCF	78
10	Nor	75	NCF	67	Bur	41	NCF	61
11	NCF	33	Bur	43	-	-	-	-
12	Nor	72	NCF	23	-	-	-	-

Tab. 3.8: Best fitting distributions - Non-Central Fischer (NCF), Dagum (Dag), Lognorm (LNo), Burr (Bur), Norm (Nor)

Finally, the relationship between core position and standardized fingerprint image quality was explored by calculating the Spearman rank correlation between the relative core offsets (ro_{core}^x, ro_{core}^y) and the NFIQ 2 quality score [246]. For rolled fingerprints, this revealed a weak but statistically significant positive correlation (correlation coefficient ≈ 0.23) between the vertical core offset (ro_{core}^y) and the NFIQ 2 score. This finding suggests that the NFIQ 2 metric exhibits a slight preference for rolled prints where the core is located lower down (further from the fingertip, closer to the base) relative to the geometric center. In contrast, no consistent or statistically significant correlation was found between the horizontal core offset (ro_{core}^x) and NFIQ 2 for rolled prints. The correlations observed for plain prints were generally weaker and often lacked statistical significance, as depicted in Figure 3.12.

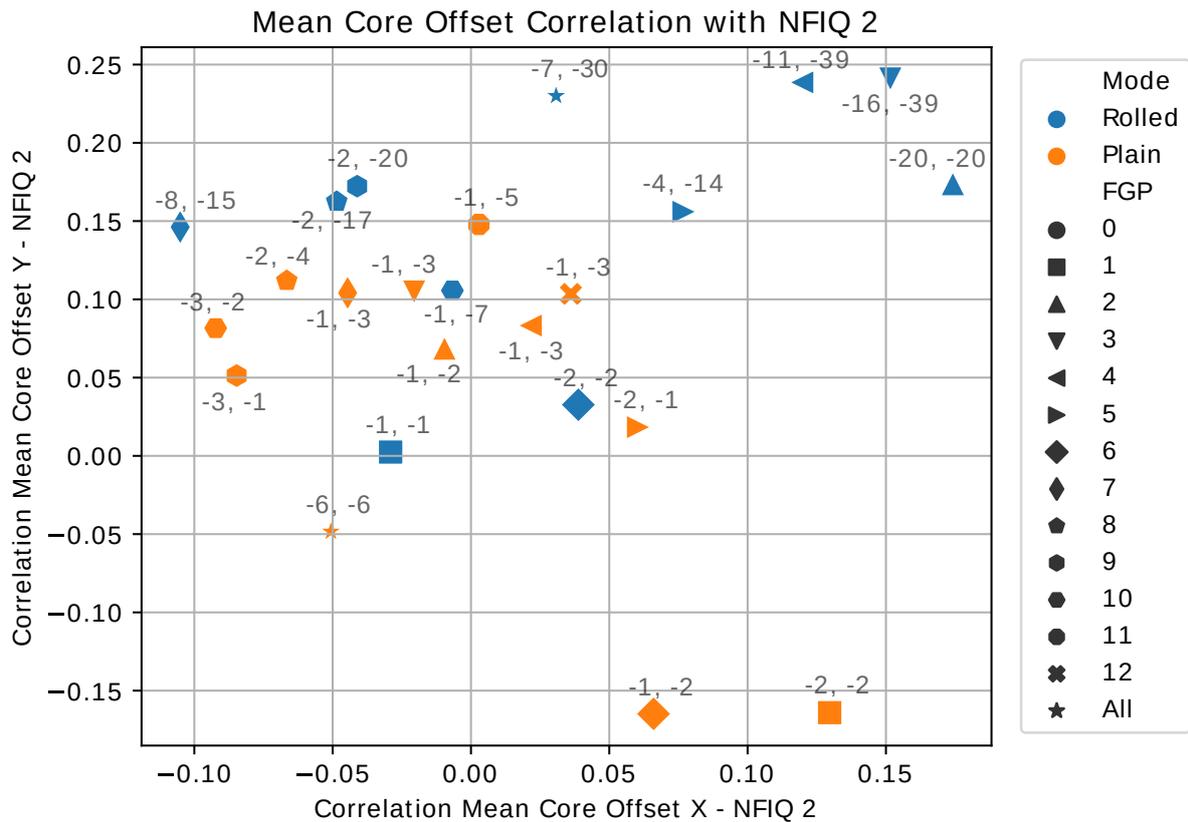


Fig. 3.12: Correlations between the Mean Core Offset in X and Y direction and the NFIQ 2 score for rolled (blue) and plain (orange) fingerprint recordings, for each finger position (see markers in legend).

3.2.5 Discussion

The deep learning model developed for core detection demonstrated high accuracy, achieving a low average localization error (9.49 pixels) [217]. The effectiveness of framing the problem as heatmap segmentation using a U-Net architecture, combined with a soft annotation strategy (slow fall-off Gaussian heatmap), proved successful in training a precise detector.

Moreover, the core centrality analysis [220] provides context by investigating the nature of the feature being detected. The key takeaway is that the fingerprint core is demonstrably not located precisely at the geometric center of the finger area, nor is its position perfectly fixed. It deviates systematically by approximately 6-8% horizontally (after bias correction) and 9-12% vertically from the center. This empirical quantification challenges the common simplifying assumption of perfect core centrality often made in algorithms and establishes a baseline level of positional variability that any system relying on the core must contend with.

The strong evidence indicating that the horizontal bias observed in rolled prints is primarily an artifact of the rolling procedure itself has significant implications for system interoperability. Algorithms designed to process fingerprints from different modalities (rolled, plain, contactless) may need to explicitly account for these distinct, modality-induced systematic offsets in core position to achieve accurate cross-comparisons.

Identifying the Non-Central Fischer (NCF) distribution as a more accurate statistical model for core position moves beyond simple assumptions of normality. This finding can be directly applied

to enhance the realism of synthetic fingerprint generation by sampling core positions according to this distribution, potentially leading to more better training and testing data quality.

The finding that the NFIQ 2 quality metric shows a slight preference for lower-positioned cores specifically in rolled prints requires further research. It remains unclear whether this correlation reflects genuine fingerprint quality characteristics often found in that region of rolled prints, whether it represents a subtle characteristic or bias within the NFIQ 2 algorithm itself when evaluating rolled captures or whether there is no causal reason behind the correlation. The lack of a similar strong correlation for horizontal position suggests NFIQ 2 is less sensitive to side-to-side variations within the typical range observed.

3.2.5.1 Impact on Research Questions

The findings from the core detection development and centrality analysis directly inform several key research questions posed in this thesis:

RQ3 (How can core point detection in contactless images be improved to enable better alignment with contact-based databases?): This work directly addresses RQ3 by presenting a high-performance deep learning heatmap prediction method for core detection [217]. The achieved low localization error (9.49 pixels) demonstrates a significant improvement in core detection capabilities. However, the centrality analysis [220] simultaneously reveals a fundamental limitation: the target landmark (the core) itself exhibits positional variability (approx. 6-8% horizontally and 9-12% vertically from the geometric center, even after bias correction). This uncertainty in the core's true location places a practical upper bound on the alignment accuracy achievable based solely on detecting this single point, even with a perfect detector. Thus, while core detection is improved, perfect alignment using only the core remains challenging due to the core's natural variance.

RQ2 (What techniques can effectively correct pose variations in contactless fingerprints to enhance comparison accuracy?): The empirical quantification of core deviation and the identification of rolling-induced bias are highly relevant to RQ2. Any pose correction technique that relies on the core as a primary reference point must contend with its variability. Effective techniques will need to be robust to this level of positional uncertainty (approx. 6-12% of finger dimensions) or be augmented with additional, potentially more stable, landmarks. Furthermore, the finding of a significant, modality-specific (rolling-induced) bias in core position suggests that pose correction frameworks aiming for interoperability across capture methods (rolled, plain, contactless) may require modality-dependent adjustments or calibrations to achieve optimal performance.

RQ1 (How can segmentation of fingerprints from contactless hand images be made more accurate and robust across different environments?): While this section focuses on core localization, its success is coupled to RQ1. The developed core detection model [217] operates on pre-cropped single-finger images, underscoring its reliance on accurate upstream segmentation (as addressed by work related to RQ1, e.g., Section 3.1). High-quality segmentation is therefore a critical enabler for precise core detection, which in turn is foundational for effective pose correction (RQ2).

RQ4 (What are the problems of established fingerprints quality assessment metrics and what can be change to improve their performance on contactless fingerprint images?): The core centrality analysis [220] provided an interesting insight related to RQ4 by examining the correlation between core position and NFIQ 2 scores [246]. The observation of a weak but significant positive correlation between vertical core offset (core lower on the finger) and NFIQ 2 scores for rolled prints suggests that established quality metrics might have subtle interactions with landmark positions or capture-induced characteristics. This shows the need to

understand how such metrics behave across different modalities and processing steps, which is important for developing or adapting quality assessment specifically for contactless fingerprints.

RQ6 (Can we produce synthetic 3D fingers and how do they perform compared to real fingers? What manufacturing techniques can be used?): The empirical finding that the Non-Central Fischer (NCF) distribution accurately models core positional variability [220] directly contributes to RQ6. This statistical model can be integrated into synthetic fingerprint generation algorithms to create more realistic datasets where core points are not simplistically centered but are distributed according to observed natural patterns. This enhanced realism in synthetic data is useful for training and testing various fingerprint processing algorithms, including core detectors and pose correction methods.

3.2.5.2 Limitations

Certain limitations pertaining to these studies should be acknowledged. The core detection model developed and evaluated in [217] currently requires a pre-cropped single-finger image as input, relying on an upstream detection or segmentation process. The large-scale centrality analysis [220], while leveraging large amounts of data, utilized a commercial software (IDKit) for the initial core detection step. Although validation on manually checked subsets showed high accuracy, the possibility of minor errors introduced by the automated detector in the full dataset cannot be entirely excluded. Finally, as with any empirical study, the specific characteristics (e.g., demographics, acquisition protocols) of the datasets used might influence the generalizability of the centrality and distribution findings.

3.2.5.3 Future Research Directions

The insights gained from this work suggest several avenues for future research. Developing end-to-end deep learning architectures capable of performing simultaneous finger detection, segmentation, and core localization directly from full hand images would simplify preprocessing pipelines. Integrating the validated Non-Central Fischer (NCF) distribution model into synthetic fingerprint generation algorithms could produce datasets with more realistic core position variability for training and testing purposes. Conducting targeted studies analyzing core location stability and distribution specifically across different types of contactless sensors and under various challenging contactless capture conditions (e.g., varying distances, angles, motion blur during mobile capture) is needed. Undertaking a deeper investigation into the observed correlation between NFIQ 2 quality scores and vertical core position in rolled prints could clarify its origin and implications for fingerprint quality assessment methodologies. Finally, given the demonstrated variability of the core, exploring the potential of using additional, possibly more positionally stable, landmarks within the fingerprint (such as friction ridge creases, delta points when present, or distinctive minutiae configurations) either in conjunction with or as alternatives to the core holds promise for developing more accurate pose estimation techniques, especially important for contactless fingerprint applications.

3.3 Pose Correction and Unwarping for Contactless Fingerprints

While contactless fingerprint capture offers advantages like hygiene [176] and avoids pressure distortions common in contact-based methods [237], it introduces significant geometric variations. Fingers can be presented to the sensor at various angles, and the natural curvature of the fingertip leads to a non-flat representation in the 2D image. These geometric inconsistencies pose challenges for accurate fingerprint comparison, especially when comparing contactless images to

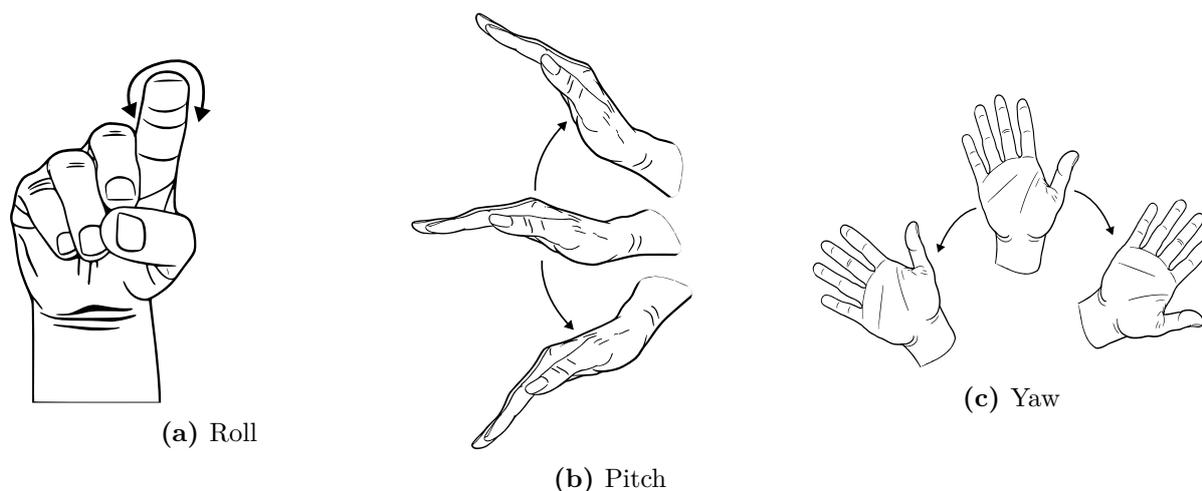


Fig. 3.13: Visualization of the three different hand rotations.

the large existing databases of contact-based prints [193, 217]. To achieve reliable interoperability, these variations must be addressed. This section introduces algorithms developed to correct for the finger's pose (orientation) relative to the sensor and to computationally flatten (unwarp) the fingerprint image, thereby reducing geometric differences between capture methods.

3.3.1 Challenges of Pose Variations in Contactless Fingerprints

A primary challenge in contactless fingerprint acquisition is the variability in finger presentation pose [217, Section 1, 2.3]. Unlike contact-based systems where the finger is pressed onto a platen, contactless methods allow the user to present their finger with freedom in orientation. This freedom results in variations along three rotational axes:

- **Roll:** Rotation of the finger around its own long axis, shown in Figure 3.13a.
- **Pitch:** Up-or-down tilt of the fingertip relative to the sensor, shown in Figure 3.13b.
- **Yaw:** Left-or-right tilt of the fingertip parallel to the sensor, shown in Figure 3.13c.

These variations in roll, pitch, and yaw lead to perspective changes and therefore geometric distortions in the captured 2D image when compared to a standard frontal view [217, 247].

Such geometric distortions directly impact the performance of fingerprint recognition systems. Feature extraction algorithms may incorrectly estimate the location and orientation of minutiae due to the distorted perspective [217, Section 1]. More critically, comparison algorithms struggle to align fingerprints captured with different poses. Misalignment leads to fewer correctly matched minutiae pairs and consequently, lower comparison scores, increasing the likelihood of false rejections (failing to match prints from the same finger) [136]. This problem is even more challenging when attempting to compare contactless fingerprints against contact-based databases [193]. The inherent geometric differences between a perspective contactless capture and a potentially pressure-distorted contact-based capture are hinder system interoperability [217, Section 1]. Therefore, correcting these pose variations is a necessary step towards building reliable contactless systems that can effectively utilize existing fingerprint data [217, Section 1, 2.3].

3.3.2 Related Works

Initial attempts to address pose variations, particularly lateral rotation (roll), focused on transforming extracted features rather than the image itself. Tan and Kumar [247] proposed an approach based on modeling the finger as an ellipsoid. By estimating the lateral rotation angle from the detected core point's position relative to the finger edges, they derived geometric transformations to correct the locations and angles of extracted minutiae points. While providing a geometric framework, this method's limitation lies in its application solely to minutiae, preventing further image-based enhancements after correction [217, Section 2.3] or the use of texture-based template comparison approaches.

Subsequent research shifted towards correcting the pose of the entire fingerprint image, allowing for subsequent processing like unwarping. Dabouei et al. [50] introduced a deep learning approach, training a network to learn a robust thin-plate spline transformation that directly warped the contactless fingerprint image to correct for pose distortions. This learned approach demonstrated effectiveness but lacked an explicit underlying geometric model or a clear mathematical mapping function relating the input pose to the output correction [217, Section 2.3].

The work detailed in this publication [217], upon which this part of the thesis is based, combines geometric modeling and deep learning. It utilizes deep learning models for prerequisite steps like fingertip segmentation and core point localization [217, Section 2.3.1, 2.3.2], which provide the inputs (finger contour and reference point) for the pose correction calculations. The correction itself employs classical geometric transformations applied to the entire image. It involves a two-stage process: first, correcting the in-plane horizontal rotation (yaw) based on the finger's axis derived from the segmentation mask, and second, correcting the out-of-plane lateral rotation (roll) using the detected core point and the assumed elliptical cross-section model, similar to the geometric foundation laid by Tan and Kumar [247] but applied to the image intensities [217, Section 2.3.3].

Beyond correcting the overall finger orientation relative to the sensor, the inherent 3D curvature of the fingertip surface needs to be addressed to simulate a flattened, contact-based impression. This process, termed unwarping or unwrapping, aims to computationally flatten the captured fingerprint texture. Söllinger and Uhl [237] systematically investigated various parametric unwarping models. Based on simplified assumptions of the finger's 3D shape, such as a cylinder or an ellipsoid, they developed circular, elliptic, and bidirectional unwarping methods, where bidirectionality implies unwarping of the curvature also along the longitudinal direction of the fingertip, to transform the curved surface into a 2D plane. These methods provide computationally efficient ways to reduce the geometric discrepancy caused by finger curvature. The pipeline presented in [217] incorporates these unwarping techniques as a subsequent step applied after the proposed pose correction, demonstrating that addressing pose and curvature sequentially can significantly improve comparison performance [217, Section 2.4, 4.3].

The effectiveness of many pose correction and unwarping algorithms, particularly those relying on geometric models like the one presented here [217], depend on the accuracy of preceding steps. Fingertip segmentation is used to define the finger's boundaries and derive its axis or contour [217, Section 2.3.1, 5.1], often achieved using deep learning models like U-Net and its variants [206, 215, 217]. See section 3.1 for more details. Similarly, detection of landmarks like the fingerprint core is required when used as a reference point for rotation estimation [217, 247, Section 2.3.2, 3.2]. This topic is presented in more details in 3.2. Therefore, advancements in segmentation and landmark detection directly benefit the development of more accurate pose correction and unwarping techniques.

3.3.3 Proposed Pose Correction Algorithm

Our pose correction approach aims to computationally rectify the three-dimensional orientation of the captured fingertip, transforming the image to an approximate standard frontal view. This aims to improve alignment for subsequent template comparison [217, Section 2.3]. Unlike methods that only adjust the coordinates of extracted minutiae points [247], our algorithm operates on the entire fingerprint image intensity values. This transformation allows us to apply further processing steps, such as surface unwarping, to the pose-corrected image [217, Section 1.1, 2.3]. The process relies fundamentally on accurate initial fingertip segmentation and core point localization, and consists of two main rotation correction steps addressing in-plane (horizontal) and out-of-plane (lateral) rotations.

3.3.3.1 Horizontal Rotation Correction

The first step corrects the rotation within the image plane (yaw angle, or horizontal rotation). This addresses fingers presented tilted left or right in the image frame. This step requires an accurate segmentation mask delineating the fingertip from the background. Instead of fitting a simple rectangular bounding box to the mask, which can be sensitive to noise or irregular finger shapes [189], we determine the finger’s central longitudinal axis differently.

From the segmentation mask, we extract the boundary contour points. For each horizontal row of pixels intersecting the mask, we identify the midpoint between the leftmost and rightmost contour points in that row. A linear function (representing the finger’s central axis) is then fitted to these midpoints using least squares [217, Section 2.3.1]. The slope of this fitted line indicates the finger’s orientation in the image plane. Let this slope correspond to an angle θ_{roll} . The entire image is then rotated by $-\theta_{roll}$ around the image center such that the finger’s central axis aligns vertically with the image axis [217, Section 2.3.1]. This method is less susceptible to outliers or minor imperfections in the segmentation boundary compared to the standard rotated bounding box technique.

3.3.3.2 Lateral Rotation Correction

The second step addresses the out-of-plane rotation (roll, collectively termed lateral rotation). This correction aims to transform the image as if the finger was viewed directly from the front, without addressing the distortions introduced from the pitch.

To mathematically model and correct for lateral rotation, we need an assumption about the finger’s 3D geometry. Following the approach presented by Tan and Kumar [247], we model the fingertip surface as a 3D shape whose cross-sections perpendicular to its longitudinal axis are ellipses [217, Section 2.3.3]. Let an ellipse in a frontal view cross-section (in a local y-z coordinate system, where y is width and z is depth) be parameterized by angle θ as

$$y = a \cos(\theta) \tag{3.11}$$

$$z = b \sin(\theta), \tag{3.12}$$

where a and b are the semi-major (half-width) and semi-minor (half-depth) axes, respectively. Furthermore, based on empirical observations and the work of Tan and Kumar, we assume a constant ratio k between the major and minor axes. In our practical implementation, this major-to-minor axis ratio $k = a/b$ is fixed at 1.2 [217, Section 2.3.3], while we show in the following derivation that the equations can be solved without fixing the axis ratio.

Using the ellipsoid model, we can define the effect of a lateral rotation (roll by the view angle) α around an axis parallel to the finger’s length (approximated as rotation around the z-axis in

the local y - z plane). According to Tan and Kumar [247, Eq. 2], the original coordinates (y, z) are transformed into the observed coordinates (y', z') in the sensor's view plane (after horizontal rotation correction) as follows

$$y' = y \cos(\alpha) - z \sin(\alpha) = a \cos(\theta) \cos(\alpha) - b \sin(\theta) \sin(\alpha), \quad (3.13)$$

$$z' = y \sin(\alpha) + z \cos(\alpha) = a \cos(\theta) \sin(\alpha) + b \sin(\theta) \cos(\alpha) \quad (3.14)$$

Here, y' represents the observed horizontal coordinate within the slice in the captured image, while z' represents the depth component which is not directly observed. Tan and Kumar further showed [247, Eq. 3, 4] that the expression for the observed coordinate y' can be rewritten using sum-to-product identities in the form of a phase-shifted cosine wave

$$y' = A' \cos(\theta + B'), \quad (3.15)$$

where the observed amplitude A' and phase shift B' are functions of the original ellipse parameters and the rotation angle α

$$A' = \sqrt{(a \cos(\alpha))^2 + (b \sin(\alpha))^2}, \quad (3.16)$$

$$B' = \arctan\left(\frac{b \sin(\alpha)}{a \cos(\alpha)}\right) = \arctan\left(\frac{1}{k} \tan(\alpha)\right) \quad (3.17)$$

The observed amplitude A' corresponds to half the width of the projected ellipse visible in the captured image slice. Also, the equation for the depth component, although not directly visible, can be written as

$$z' = C' \cos(\theta - D'), \quad (3.18)$$

where the amplitude C' of the theoretically observable depth and phase shift D' are functions of again the original ellipse parameters and the rotation angle α

$$C' = \sqrt{(a \sin(\alpha))^2 + (b \cos(\alpha))^2}, \quad (3.19)$$

$$D' = \arctan\left(\frac{b \cos(\alpha)}{a \sin(\alpha)}\right) \quad (3.20)$$

The correction furthermore requires a stable reference point on the finger surface. We utilize the fingerprint core, as described in section 3.2 as this reference [217, Section 2.3.2, 2.3.3, Figure 4]. The core's position relative to the finger edges in the horizontally-rotated image provides the key information to estimate the lateral viewing angle.

Let the horizontally rotated image be considered. For a given horizontal slice passing through the core, let P'_l and P'_r be the coordinates of the left and right finger edges (derived from the segmentation mask), and let P'_c be the coordinate of the detected core. Note that the points P'_l and P'_r are not special in any way in the transformed image, because they are no longer edge points in this view.

The lateral rotation angle α (defined as the negative viewing angle relative to the frontal view) influences the apparent position of the core between the edges due to perspective projection of the curved surface.

Tan and Kumar [247] showed that the lateral rotation angle α can be related to the ratio of distances between the edges and the core. Let the measured distances be $d_{lc} = \|P'_c - P'_l\|$ and $d_{cr} = \|P'_r - P'_c\|$. We define the ratio R as

$$R := \frac{d_{lc}}{d_{cr}} = \frac{P'_l P'_c}{P'_c P'_r}, \quad (3.21)$$

where we know that the edge points are at positions A' and $-A'$ of the observed, rotated ellipse. Furthermore, we know that the core sits at $\theta = -\frac{\pi}{2}$ for the ellipsis in the non-rotated view. We can calculate the core position in the observed, rotated ellipse by utilizing equation 3.15, leading to

$$y_{core} = 0, \quad (3.22)$$

$$y'_{core} = A' \cos\left(-\frac{\pi}{2} + B'\right) \quad (3.23)$$

The distances $d_{lc} = \|P'_c - P'_l\|$ and $d_{cr} = \|P'_r - P'_c\|$ can then be expressed as

$$d_{lc} = A' + A' \cos\left(-\frac{\pi}{2} + B'\right), \quad (3.24)$$

$$d_{cr} = A' - A' \cos\left(-\frac{\pi}{2} + B'\right) \quad (3.25)$$

$$\begin{aligned} R &= \frac{A' + A' \cos\left(-\frac{\pi}{2} + B'\right)}{A' - A' \cos\left(-\frac{\pi}{2} + B'\right)} \\ R &= \frac{1 + 1 \cos\left(-\frac{\pi}{2} + B'\right)}{1 - 1 \cos\left(-\frac{\pi}{2} + B'\right)} \\ R &= \frac{1 + 1 \sin(B')}{1 - 1 \sin(B')} \end{aligned} \quad (3.26)$$

This ratio R can be directly measured from the horizontally rotated image and the detected core position. The relationship between R and the viewing angle α , under the elliptical cross-section assumption, is derived through the geometry of projecting the ellipse onto the image plane. If B' represents the phase angle in the projected ellipse description, which is related to α via the ellipse parameters a and b as $B' = \arctan\left(\frac{b}{a} \tan(\alpha)\right)$ (assuming standard ellipse parameterization context), we substitute this into Equation 3.26. By substituting the expression for B' into 3.26, we get a direct relationship between the measurable ratio R and the desired angle α :

$$R = \frac{1 + \sin\left(\arctan\left(\frac{b}{a} \tan(\alpha)\right)\right)}{1 - \sin\left(\arctan\left(\frac{b}{a} \tan(\alpha)\right)\right)} \quad (3.27)$$

We now wish to solve Equation 3.27 for the angle α .

Let $S = \sin\left(\arctan\left(\frac{b}{a} \tan(\alpha)\right)\right)$. The equation becomes

$$R = \frac{1 + S}{1 - S} \quad (3.28)$$

Assuming $R \neq -1$ (which is guaranteed as R is a ratio of distances, hence $R \geq 0$), we can solve for S

$$\begin{aligned}
R(1 - S) &= 1 + S \\
R - RS &= 1 + S \\
R - 1 &= S + RS \\
R - 1 &= S(1 + R) \\
S &= \frac{R - 1}{R + 1}
\end{aligned} \tag{3.29}$$

Now, let's express S directly in terms of $\tan(\alpha)$. Let $u = \arctan\left(\frac{b}{a}\tan(\alpha)\right)$. By definition, $\tan(u) = \frac{b}{a}\tan(\alpha)$, and $u \in (-\pi/2, \pi/2)$. We want to find $S = \sin(u)$. We use the identity $\sin^2(u) + \cos^2(u) = 1$. Dividing by $\cos^2(u)$ (assuming $\cos(u) \neq 0$, which is true for $u \in (-\pi/2, \pi/2)$)

$$\begin{aligned}
\tan^2(u) + 1 &= \sec^2(u) = \frac{1}{\cos^2(u)} \\
\implies \cos^2(u) &= \frac{1}{1 + \tan^2(u)} \\
\implies \sin^2(u) &= 1 - \cos^2(u) = 1 - \frac{1}{1 + \tan^2(u)} = \frac{\tan^2(u)}{1 + \tan^2(u)}
\end{aligned}$$

Taking the square root, $\sin(u) = \pm \frac{\tan(u)}{\sqrt{1 + \tan^2(u)}}$. Since $u \in (-\pi/2, \pi/2)$, $\sin(u)$ and $\tan(u)$ have the same sign. Therefore, we must take the positive root

$$\sin(u) = \frac{\tan(u)}{\sqrt{1 + \tan^2(u)}}$$

Substitute $\tan(u) = \frac{b}{a}\tan(\alpha)$

$$\begin{aligned}
S &= \frac{\frac{b}{a}\tan(\alpha)}{\sqrt{1 + \left(\frac{b}{a}\tan(\alpha)\right)^2}} \\
&= \frac{\frac{b}{a}\tan(\alpha)}{\sqrt{1 + \frac{b^2}{a^2}\tan^2(\alpha)}} \\
&= \frac{\frac{b}{a}\tan(\alpha)}{\sqrt{\frac{a^2 + b^2\tan^2(\alpha)}{a^2}}} \\
&= \frac{\frac{b}{a}\tan(\alpha)}{\frac{\sqrt{a^2 + b^2\tan^2(\alpha)}}{|a|}} \quad (\text{Assuming } a \neq 0) \\
&= \frac{b\tan(\alpha)}{a} \frac{|a|}{\sqrt{a^2 + b^2\tan^2(\alpha)}} \\
&= \frac{\text{sgn}(a) b\tan(\alpha)}{\sqrt{a^2 + b^2\tan^2(\alpha)}}
\end{aligned} \tag{3.30}$$

Here, $\text{sgn}(a) = |a|/a$ is the sign function. We assume $a, b \neq 0$.

Equating the two expressions for S (Equations 3.29 and 3.30)

$$\frac{R-1}{R+1} = \frac{\text{sgn}(a) b \tan(\alpha)}{\sqrt{a^2 + b^2 \tan^2(\alpha)}} \quad (3.31)$$

To solve for $\tan(\alpha)$, we square both sides. Note that squaring introduces extraneous solutions, which we must consider later.

$$\begin{aligned} \left(\frac{R-1}{R+1}\right)^2 &= \left(\frac{\text{sgn}(a) b \tan(\alpha)}{\sqrt{a^2 + b^2 \tan^2(\alpha)}}\right)^2 \\ \frac{(R-1)^2}{(R+1)^2} &= \frac{(\text{sgn}(a))^2 b^2 \tan^2(\alpha)}{a^2 + b^2 \tan^2(\alpha)} \\ \frac{(R-1)^2}{(R+1)^2} &= \frac{b^2 \tan^2(\alpha)}{a^2 + b^2 \tan^2(\alpha)} \quad (\text{since } (\text{sgn}(a))^2 = 1 \text{ for } a \neq 0) \end{aligned}$$

Let $T = \tan^2(\alpha)$. Assume $b \neq 0$ and $R \geq 0$. If $R = 0$, the LHS is 1, requiring $S = -1$, which implies $y = -\pi/2$, meaning $\frac{b}{a} \tan(\alpha) \rightarrow -\infty$. The denominator below ($4b^2R$) becomes zero, so we assume $R > 0$.

$$\begin{aligned} (R-1)^2(a^2 + b^2T) &= (R+1)^2b^2T \\ a^2(R-1)^2 + b^2T(R-1)^2 &= b^2T(R+1)^2 \\ a^2(R-1)^2 &= b^2T[(R+1)^2 - (R-1)^2] \\ a^2(R-1)^2 &= b^2T[(R^2 + 2R + 1) - (R^2 - 2R + 1)] \\ a^2(R-1)^2 &= b^2T[4R] \\ T = \tan^2(\alpha) &= \frac{a^2(R-1)^2}{4b^2R} \end{aligned}$$

Taking the square root

$$\tan(\alpha) = \pm \sqrt{\frac{a^2(R-1)^2}{4b^2R}} = \pm \frac{|a||R-1|}{2|b|\sqrt{R}}$$

Since a represents the major axis and b the minor axis of the elliptical finger model, we know that they are positive.

$$\tan(\alpha) = \pm \frac{a(R-1)}{2b\sqrt{R}} \quad (3.32)$$

In order to figure out which sign is correct, we must check against the equation before squaring (Equation 3.31).

Case 1: $\tan(\alpha) = \frac{a(R-1)}{2b\sqrt{R}}$. Substitute into the RHS of Equation 3.31

$$\begin{aligned}
\text{RHS} &= \frac{\text{sgn}(a) b \left(\frac{a(R-1)}{2b\sqrt{R}} \right)}{\sqrt{a^2 + b^2 \left(\frac{a(R-1)}{2b\sqrt{R}} \right)^2}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{\sqrt{a^2 + \frac{b^2 a^2 (R-1)^2}{4b^2 R}}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{\sqrt{a^2 \left(1 + \frac{(R-1)^2}{4R} \right)}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{|a| \sqrt{\frac{4R+R^2-2R+1}{4R}}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{|a| \sqrt{\frac{R^2+2R+1}{4R}}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{|a| \sqrt{\frac{(R+1)^2}{4R}}} \\
&= \frac{\text{sgn}(a) a(R-1)/(2\sqrt{R})}{|a| |R+1|/(2\sqrt{R})} \\
&= \frac{\text{sgn}(a) a(R-1)}{|a|(R+1)} \quad (\text{Since } R > 0, R+1 > 0, |R+1| = R+1) \\
&= \frac{|a|(R-1)}{|a|(R+1)} \quad (\text{Since } \text{sgn}(a)a = |a|) \\
&= \frac{R-1}{R+1}
\end{aligned}$$

This matches the LHS of Equation 3.31. So, $\tan(\alpha) = \frac{a(R-1)}{2b\sqrt{R}}$ is a valid solution.

Case 2: $\tan(\alpha) = -\frac{a(R-1)}{2b\sqrt{R}} = \frac{a(1-R)}{2b\sqrt{R}}$. Substituting this into the RHS of Equation 3.31: The denominator $\sqrt{a^2 + b^2 \tan^2(\alpha)}$ remains the same because $\tan^2(\alpha)$ is unchanged.

$$\begin{aligned}
\text{RHS} &= \frac{\text{sgn}(a) b \left(\frac{a(1-R)}{2b\sqrt{R}} \right)}{\sqrt{a^2 + b^2 \tan^2(\alpha)}} \\
&= \frac{\text{sgn}(a) a(1-R)/(2\sqrt{R})}{|a|(R+1)/(2\sqrt{R})} \\
&= \frac{\text{sgn}(a) a(1-R)}{|a|(R+1)} \\
&= \frac{|a|(1-R)}{|a|(R+1)} \\
&= \frac{1-R}{R+1}
\end{aligned}$$

This only matches the LHS $\frac{R-1}{R+1}$ if $\frac{1-R}{R+1} = \frac{R-1}{R+1}$, which implies $1 - R = R - 1$, or $2 = 2R$, meaning $R = 1$. If $R = 1$, both formulas give $\tan(\alpha) = 0$, indicating that we observed no rotation because the core is found at the center.

The structure of the original equation $R = (1+S)/(1-S)$ has a symmetry, which is the reason we found two solutions. If we replace R with $1/R$, we get

$$1/R = \frac{1+S'}{1-S'}$$

which solves to

$$\begin{aligned} S' &= \frac{1/R - 1}{1/R + 1} \\ &= \frac{1 - R}{1 + R} \\ &= -S \end{aligned}$$

So, if (R, α) is a solution pair leading to S , then $(1/R, \alpha')$ is a solution pair leading to $S' = -S$. We have $\sin(\arctan(\frac{b}{a} \tan(\alpha'))) = -\sin(\arctan(\frac{b}{a} \tan(\alpha)))$. Since $\sin(-u) = -\sin(u)$ and \arctan maps to $(-\pi/2, \pi/2)$, this implies

$$\begin{aligned} \arctan\left(\frac{b}{a} \tan(\alpha')\right) &= \\ -\arctan\left(\frac{b}{a} \tan(\alpha)\right) &= \\ \arctan\left(-\frac{b}{a} \tan(\alpha)\right) & \end{aligned}$$

Taking \tan of both sides gives $\frac{b}{a} \tan(\alpha') = -\frac{b}{a} \tan(\alpha)$, so $\tan(\alpha') = -\tan(\alpha)$. Therefore, the solution corresponding to $1/R$ has the negative tangent of the solution corresponding to R . The first solution form $\tan(\alpha) = \frac{a(1-R)}{2b\sqrt{R}}$ solves the equation for $1/R$. The second solution form $\tan(\alpha) = \frac{a(R-1)}{2b\sqrt{R}}$ solves the equation for R .

Thus, the two solutions found arise naturally from this symmetry. Depending on whether the measured ratio is R or $1/R$ (which might depend on labeling conventions), one of the two formulas applies. Both can be presented as potential solutions stemming from the equation structure.

The solutions for α are then given by

$$\alpha = \arctan\left(\frac{a - aR}{2b\sqrt{R}}\right) + n\pi \quad (\text{Corresponds to } 1/R \text{ case, or } \tan(\alpha')) \quad (3.33)$$

and

$$\alpha = \arctan\left(\frac{aR - a}{2b\sqrt{R}}\right) + n\pi \quad (\text{Corresponds to } R \text{ case, or } \tan(\alpha)) \quad (3.34)$$

for any integer n . The \arctan function returns the principal value in $(-\pi/2, \pi/2)$. The addition of $n\pi$ accounts for the periodicity of the tangent function. Note however that for our usecase of calculating the lateral rotation angle based on the core offset, it does not make sense to consider angles outside the range of $(-\pi/2, \pi/2)$, therefore, n can be set to zero.

Once α is estimated, we use the elliptical model again to transform the pixel intensities to the frontal view. For each pixel (x', y') in the observed image, we calculate its corresponding 3D coordinates on the ellipsoid based on the calculated lateral rotation angle α . We then apply the

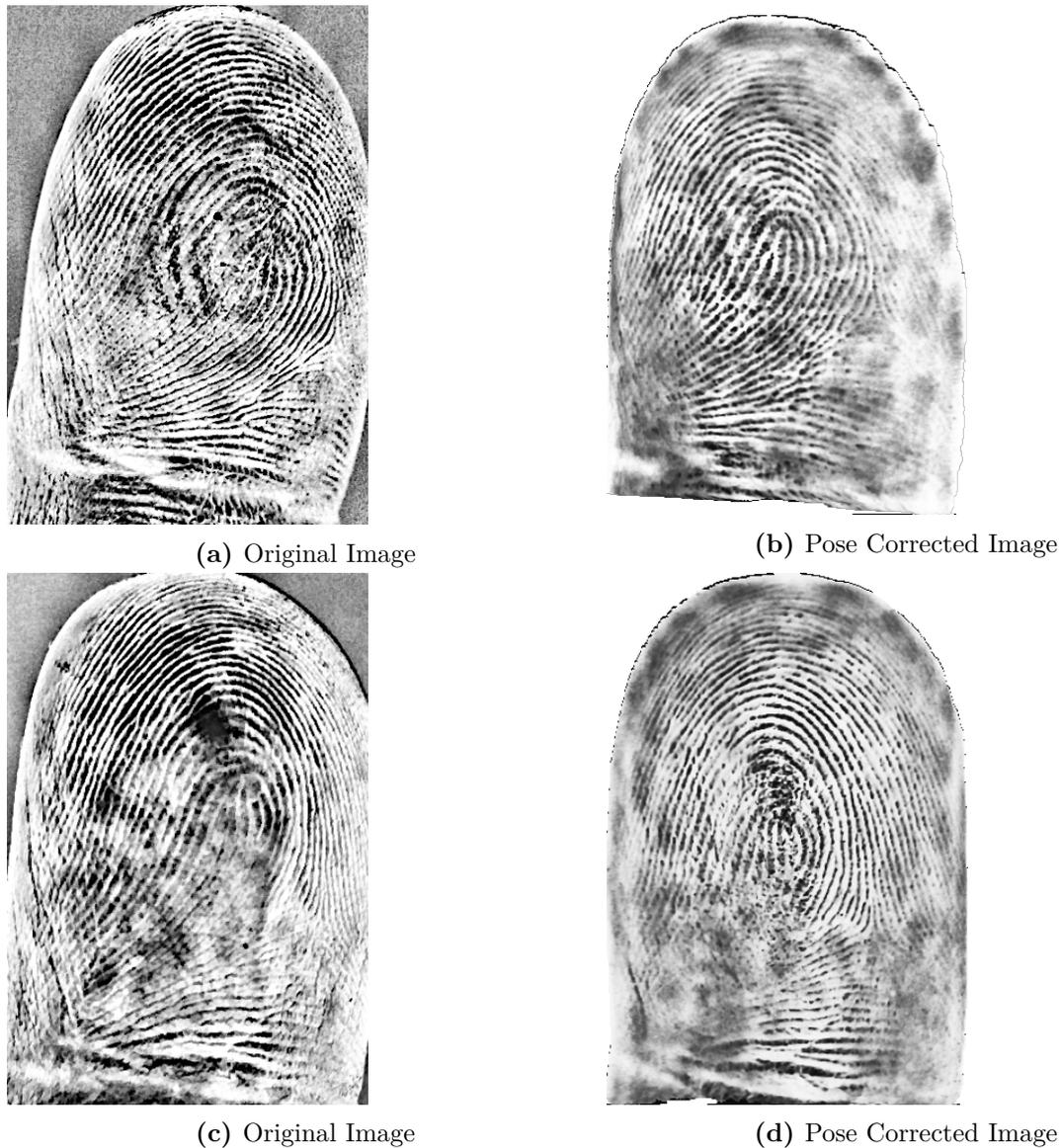


Fig. 3.14: Effect of pose correction on two example recordings. Index finger in 3.14a and 3.14b and middle finger in 3.14c and 3.14d.

inverse rotation by $-\alpha$ to find the corresponding 3D point on the rotated ellipsoid model, which corresponds to an axes parallel ellipsoid. The parameters of the ellipsoid at each slice (major axis a , minor axis $b = a/1.2$) are determined using the width of the finger contour in the observed image at that slice. We then project the rotated 3D point back onto the original sensor plane to find the target pixel coordinates (x, y) in the input (horizontally-rotated) image. The intensity value at (x, y) (using interpolation if necessary) are set to be the source pixel's (x', y') value. This process effectively unrotates the finger surface texture onto the frontal plane [217, Section 2.3.3]. Regions of the target frontal view that correspond to parts of the ellipsoid invisible in the original rotated view (occluded surfaces) are filled with a background value. The result is a pose-corrected image, aligned as if captured from a direct frontal viewpoint (exemplary shown in Figure 3.14), ready for the subsequent unwarping step.

3.3.4 Proposed Unwarping Algorithm

Even after correcting the overall orientation of the finger using the pose correction algorithm described previously (Section 3.3.3), the resulting image still represents a perspective projection of the inherently curved fingertip surface, because it was taken by a camera with a finite focus length. To further improve interoperability between contactless captures and flattened contact-based impressions, a subsequent unwarping step is required [217, Section 2.4]. This step aims to computationally flatten the fingerprint texture, simulating an orthographic projection of the 3D fingerprint.

In our work presented in [217], we integrated and evaluated several parametric unwarping algorithms, originally proposed and detailed by Söllinger and Uhl in their publication *Optimizing contactless to contact-based fingerprint comparison using simple parametric warping models* [237]. These algorithms operate by assuming a simplified 3D model of the finger's shape and then mathematically mapping the captured texture onto a flat 2D plane.

The specific methods explored, following the work of Söllinger and Uhl [237], include:

- **Circular Unwarping:** This approach models the finger as a cylinder, leading to a circular cross-section. The captured image texture is then projected onto the unfolded surface of this cylinder. As investigated in [217, Section 2.4], two variants of this method were considered: one using a fixed cylinder radius determined from the overall image width, and another employing an adaptive radius. The adaptive approach uses the fingertip contour derived from the segmentation mask (obtained as described in Section 3.3.3.1) to estimate the finger width, and thus the cylinder radius, locally for each part of the finger.
- **Elliptic Unwarping:** Recognizing that a simple cylinder might not capture the finger's shape, this method assumes an elliptical cross-section, consistent with the model used in the lateral pose correction step (Section 3.3.3.2). The texture is then unwrapped from the surface of this assumed ellipsoid [237].
- **Bidirectional Unwarping:** This method aims to better account for the rounded fingertip. As described by Söllinger and Uhl [237], it combines two circular unwarping operations: one along the main finger axis and another perpendicular to it [217, Section 2.4].

These unwarping algorithms, developed by Söllinger and Uhl [237], were applied as a post-processing step after our proposed pose correction within the pipeline evaluated in [217, Section 4.3]. The goal was to assess their combined effect on reducing geometric distortions and improving the interoperability between contactless and contact-based fingerprint templates. The visual impact of applying these different unwarping techniques to a contactless fingerprint image is illustrated in Figure 5 of the paper [217, Figure 5].

3.3.5 Experimental Evaluation and Results

To assess the efficacy of the proposed pose correction and unwarping pipeline, we focused on its impact on contactless to contact-based fingerprint comparison performance [217]. The overall processing pipeline, consisting of the live preprocessing steps and the postprocessing steps including pose correction and unwarping evaluated in this section, is illustrated in Figure 3.15.

3.3.5.1 Dataset and Demographics

The evaluation utilized a dataset comprising both contactless and contact-based fingerprint recordings obtained from a law enforcement agency's operational data [217, Section 3.2.3]. Due

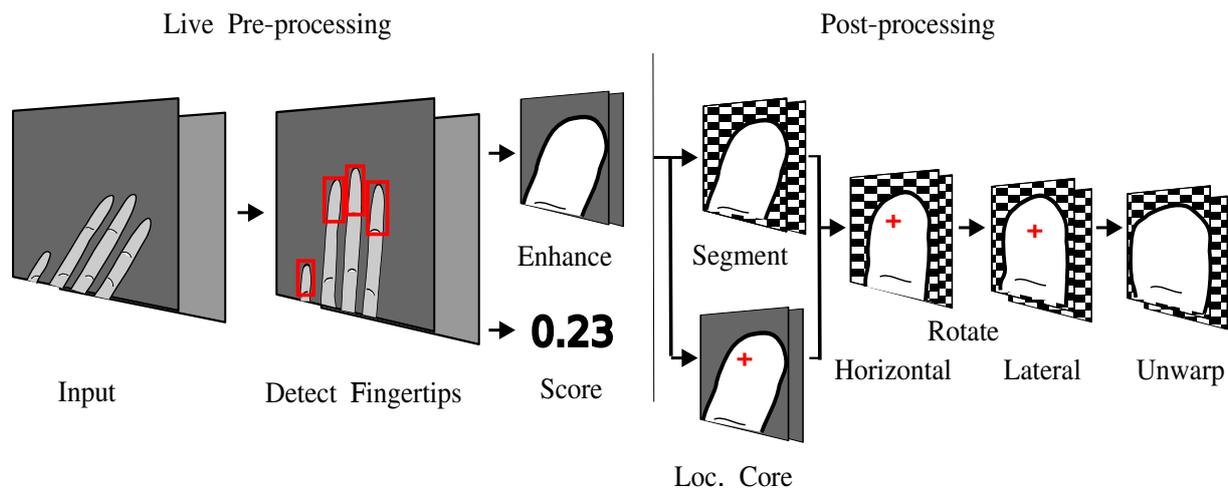


Fig. 3.15: The preprocessing and postprocessing pipeline evaluated. The input is the raw grayscale image from the contactless sensor. Live preprocessing runs in real-time and involves detecting fingertips, enhancing the cropped images, and scoring their sharpness. The results are saved. Postprocessing, performed later, involves segmenting the finger contour, locating the finger’s core, using these to perform horizontal and lateral rotation corrections, and finally unwarping the pose-corrected image before template comparison. Adapted from Figure 1 in [217].

to time and computational constraints, a representative random subset of 78 participants from the larger dataset collected in [266] was selected for this evaluation, encompassing a total of 37,162 contactless fingerprint images [217, Section 1.1, 3.2.3].

Contactless images were captured using the sensor prototype detailed in [217, Section 2.1]. Contact-based counterparts consisted of rolled fingerprint images acquired using a commercially available Idemia TP 5300 scanner at a resolution of 1000 DPI [217, Section 3.2.3]. Before analysis, the dataset underwent a cleaning process to correct finger identification errors introduced during acquisition, involving the correction of mappings for 102 finger instances and the exclusion of 25 individual fingers where a definitive assignment between contactless and contact-based recordings could not be established [217, Section 3.2.6].

The participants had diverse demographics, with gender and origin information not available for only three individuals. Of the remaining 75 participants, 42 were identified as male and 33 as female. Geographic origin was varied: Asia was the most represented continent with 50 participants (24 male, 26 female) from eight different countries (predominantly Syria (36), Afghanistan (5), and China (3)). Africa was represented by 10 participants (all male) from nine countries (primarily Somalia (4) and Nigeria (2)). Europe contributed 5 participants (all male) from five countries (including Ukraine (3) and Türkiye (2)) [217, Section 3.2.4, Figure 7]. The age and country of origin distributions for these 75 participants are visualized in Figure 3.16.

A quality analysis using the NIST Fingerprint Image Quality 2 (NFIQ 2) metric [246] was performed on both the contact-based and the preprocessed contactless images (after live preprocessing including enhancement [116] but before pose correction/unwarping). It is important to remember that NFIQ 2 was designed and trained specifically for contact-based, plain (flat) fingerprint images [246], whereas in this analysis, it was applied to contact-based rolled impressions and contactless images. Therefore, the application of NFIQ 2 here falls outside its intended scope. Given the mismatch between the NFIQ 2 design and the image types analyzed, these quality scores should be considered only as a rough indication of potential sample utility, rather

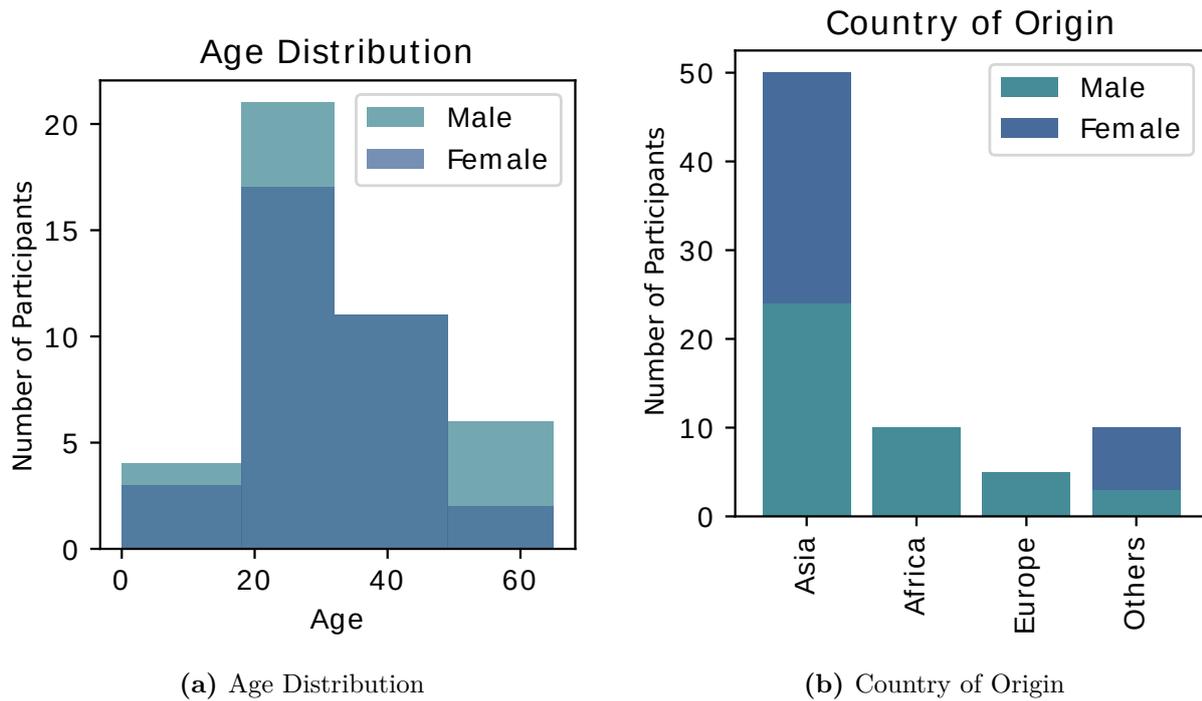


Fig. 3.16: Age and country of origin distribution, separated by gender, for the 75 participants (out of 78 total) with available demographic information used in the contactless to contact-based template comparison dataset. Adapted from Figure 7 in [217].

than a direct measure. Furthermore, as noted in [189], NFIQ 2 scores are not directly indicative of comparison performance (which can be measured by EER or similar metrics), but predict the sample utility.

The results, detailed in 3.9, showed that the contact-based rolled prints (Idemia TP 5300) had a higher average NFIQ 2 score (50.0 ± 18.9) compared to the contactless images (32.3 ± 13.5). The contact-based images also exhibited a wider score range ([2, 93]) and larger interquartile range (IQR [37, 63]) compared to the contactless images (range [0, 91], IQR [22, 42]). When compared to scores from publicly available contact-based datasets reported in [189], the contact-based portion of this dataset demonstrated relatively high quality (50 ± 18 vs. e.g., MCYT dp: 38 ± 15 , FVC06: 36 ± 9), while the contactless scores were comparatively low (32 ± 14) [217, Section 3.2.5]. It is important to note that NFIQ 2 calculation failed for 3,464 contactless images, primarily due to the recognized fingerprint area being too small [217, Section 3.2.5].

3.3.5.2 Evaluation Methodology

The primary metric used to evaluate the impact of the proposed pipeline was the Equal Error Rate (EER). EER represents the point where the False Acceptance Rate (FAR) equals the False Rejection Rate (FRR) and is a standard measure for biometric system performance, with lower values indicating better accuracy [217, Section 1.1, 3.1].

For each finger of each participant in a recording session, the top five sharpest contactless fingerprint images were selected based on the sharpness score calculated during the live preprocessing stage (Section 2.2 in [217]). Templates were generated for these five selected contactless images and for all available contact-based rolled print templates using the commercial IDKit software from Innovatrics [103]. Each of the five contactless templates was then compared against

Recording Mode	All	Thumb	Index	Middle	Ring	Little
Contact-based	50.0 ± 18.9	55.4 ± 13.4	54.6 ± 18.2	50.6 ± 16.5	46.0 ± 18.8	43.5 ± 23.3
[<i>Min</i> , <i>Max</i>]	[2, 93]	[13, 92]	[10, 93]	[3, 85]	[3, 93]	[2, 90]
[$q_{0.25}$, $q_{0.75}$]	[37, 63]	[48, 63]	[42, 67]	[39, 62]	[34, 60]	[25, 62]
[$q_{0.10}$, $q_{0.90}$]	[24, 74]	[39, 72]	[28, 76]	[31, 73]	[21, 70]	[12, 75]
Contactless	32.3 ± 13.5	38.8 ± 10.1	31.7 ± 13.5	35.7 ± 12.8	30.9 ± 13.2	23.6 ± 12.5
[<i>Min</i> , <i>Max</i>]	[0, 91]	[0, 82]	[0, 80]	[0, 81]	[0, 86]	[0, 91]
[$q_{0.25}$, $q_{0.75}$]	[22, 42]	[33, 46]	[22, 41]	[26, 45]	[20, 41]	[14, 32]
[$q_{0.10}$, $q_{0.90}$]	[14, 49]	[25, 51]	[14, 50]	[18, 52]	[14, 48]	[8, 40]

Tab. 3.9: Mean NIFQ 2 scores with standard deviation for the contact-based recordings collected using the Idemia TP 5300 sensor as well as the contactless recordings after the pre-processing stage. Taken from [217] and extend to show more extensive metrics like minimum and maximum ($[Min, Max]$) and different quartile ranges ($[q_{start}, q_{end}]$).

all contact-based templates of the corresponding finger from the same participant (generating genuine scores) and against templates from the other participants (generating imposter scores). This approach accounts for variations in recording quality by focusing on the sharpest images and reduces the total number of comparisons needed compared to using all captured images [217, Section 3.1].

To assess the variability of the EER results, a random subsampling strategy was employed. For each EER calculation, 80% of the genuine scores and 80% of the imposter scores were randomly selected. This subsampling and EER calculation process was repeated 100 times. The final reported EER values represent the mean and standard deviation across these 100 repetitions [217, Section 3.1, 4.3].

3.3.5.3 Comparison Performance Results (EER)

The experimental results, detailed in [217, Table 5] and shown in Table 3.10, demonstrate the impact of the different processing steps on the contactless-to-contact-based comparison EER. The baseline performance, using only the standard preprocessing pipeline (including Kauba et al. enhancement [116]) without any subsequent pose correction or unwarping, yielded an overall mean EER of 1.57 ± 0.11 % across all fingers (Row B). Applying only the proposed pose correction algorithm (Section 3.3.3) resulted in a slight increase in the overall EER to 1.61 ± 0.10 %. However, examining individual fingers revealed improvements for the Thumb and Ring fingers, offset by degradations for others, indicating a finger-specific impact (Row PC). Applying only the unwarping methods (introduced by [237]) without prior pose correction showed that Elliptic unwarping (El) provided the largest average EER reduction compared to the baseline, achieving an overall EER of 1.35 ± 0.10 %, representing a 14.0% relative improvement (Row El). The core evaluation involved combining pose correction with the various unwarping techniques. When applied after pose correction, Bidirectional unwarping (Bi) yielded the largest average improvement over the pose-correction-only result, reaching an EER of 1.47 ± 0.08 % (a 6.4% relative improvement compared to pose correction only) (Row PC+Bi). Elliptic unwarping after pose correction (PC+El) achieved a similar overall EER of 1.48 ± 0.11 % (Row PC+El). The results in 3.10 showed that the optimal enhancement strategy (including whether to apply pose correction and which unwarping method to use) varied significantly depending on the specific finger. By selecting the best performing combination for each finger individually based on the lowest EER in Table 5, a finger-wise combined enhancement pipeline was constructed. This involved:

Enhancement	All [%]	Thumb [%]	Index [%]	Middle [%]	Ring [%]	Little [%]
B	1.57 ± 0.11	2.38 ± 0.58	1.12 ± 0.13	1.28 ± 0.21	<i>0.85 ± 0.23</i>	2.97 ± 0.33
CF	1.46 ± 0.08	<i>0.47 ± 0.02</i>	1.39 ± 0.14	0.89 ± 0.20	1.12 ± 0.19	3.04 ± 0.34
CA	3.30 ± 0.19	1.51 ± 0.20	1.69 ± 0.57	2.28 ± 0.35	<i>0.85 ± 0.50</i>	8.67 ± 0.84
El	1.35 ± 0.10	0.82 ± 0.22	0.94 ± 0.16	<i>0.66 ± 0.15</i>	0.98 ± 0.15	3.43 ± 0.34
Bi	1.37 ± 0.10	1.52 ± 0.22	1.33 ± 0.31	<i>1.00 ± 0.18</i>	1.31 ± 0.17	1.77 ± 0.15
PC	1.61 ± 0.10	<i>0.78 ± 0.40</i>	2.03 ± 0.39	1.68 ± 0.37	0.94 ± 0.22	2.04 ± 0.21
PC + CF	1.62 ± 0.08	<i>0.68 ± 0.05</i>	2.00 ± 0.19	0.80 ± 0.19	1.45 ± 0.21	2.15 ± 0.13
PC + CA	2.27 ± 0.14	<i>0.52 ± 0.06</i>	3.02 ± 0.19	0.63 ± 0.10	2.16 ± 0.28	3.12 ± 0.14
PC + El	1.48 ± 0.11	<i>0.41 ± 0.16</i>	1.45 ± 0.26	0.89 ± 0.22	1.06 ± 0.16	2.60 ± 0.15
PC + Bi	1.47 ± 0.08	<i>0.57 ± 0.16</i>	1.76 ± 0.29	0.79 ± 0.10	1.96 ± 0.18	1.98 ± 0.20

Tab. 3.10: Equal error rates (EERs) given in percentages for contactless to contact-based template comparison. Depicted are mean and standard deviation of EERs for 100 random sub-samples containing 80% of the template comparison scores. The baseline was calculated for the enhanced images after the pre-processing pipeline without any further post-processing steps. Results for the Pose correction (PC) run were calculated by applying pose correction as a post-processing step to the results of the baseline (B). For the next runs, additional unwarping was added: Circular unwarping with fixed finger width (CF), circular unwarping with an adaptive finger width (CA), elliptic unwarping with a fixed finger width (El) and finally bidirectional unwarping with a fixed finger width (Bi). Bold entries highlight the best score for this column and italicized entries the best score for the row. Ties are broken by smaller standard deviation. Taken from [217].

- **Thumb:** Pose Correction + Elliptic Unwarping (PC+El)
- **Index:** Elliptic Unwarping only (El)
- **Middle:** Elliptic Unwarping only (El)
- **Ring:** Baseline preprocessing only (None)
- **Little:** Bidirectional Unwarping only (Bi)

This finger-wise approach, detailed in [217, Table 6], achieved the lowest overall EER of 0.99 ± 0.07 %. This represents an absolute reduction of 0.58 percentage points and a relative decrease in EER of 36.9% compared to the baseline performance [217, Section 4.3, 6].

3.3.5.4 Image Utility Results (NFIQ 2)

To complement the EER analysis, the impact of the enhancement pipeline on image quality was assessed using NFIQ 2 scores [217, Section 4.4]. Exemplarily, the NFIQ 2 scores were calculated for the contactless images after applying both pose correction and bidirectional unwarping (PC+Bi), one of the effective combined methods. The results are presented in 3.11.

Compared to the NFIQ 2 scores of the baseline preprocessed contactless images (see Table 3.9), the combined enhancement (PC+Bi) led to an increase in the average NFIQ 2 score across all fingers from 32.3 ± 13.5 to 33.5 ± 10.9 . This constitutes a relative increase of 3.72%. Furthermore, the standard deviation decreased from 13.5 to 10.9, indicating a reduction in quality score volatility by 19.26%.

Analysis per finger showed improvements in average NFIQ 2 scores for the Thumb (+7.22%), Index (+4.1%), Ring (+1.29%), and Little (+15.25%) fingers. Only the Middle finger showed a

-	All	Thumb	Index	Middle	Ring	Little
Mean \pm Std.	33.5 \pm 10.9	41.6 \pm 7.0	33.0 \pm 10.8	35.0 \pm 10.2	31.3 \pm 10.8	27.2 \pm 9.8
[<i>Min</i> , <i>Max</i>]	[0, 71]	[10, 71]	[1, 71]	[2, 64]	[1, 60]	[0, 69]
[$q_{0.25}$, $q_{0.75}$]	[26, 42]	[38, 46]	[27, 41]	[29, 42]	[24, 39]	[20, 34]
[$q_{0.10}$, $q_{0.90}$]	[18, 46]	[32, 49]	[17, 45]	[20, 46]	[16, 44]	[14, 40]

Tab. 3.11: NFIQ 2 scores for the contactless recordings after the post-processing stage with bidirectional unwarping, taken from [217].

slight decrease (-1.96%). Importantly, the variability (standard deviation) of NFIQ 2 scores was reduced for all fingers: Thumb (-30.69%), Index (-20.00%), Middle (-20.31%), Ring (-18.18%), and Little (-21.60%). This suggests that the pose correction and unwarping steps generally not only slightly improve the average measurable feature quality but also make the quality more consistent across recordings.

3.3.6 Discussion

The experimental results confirm that addressing geometric variations through pose correction and surface unwarping significantly enhances the interoperability between contactless and contact-based fingerprint systems. The developed pipeline, combining pose correction steps with unwarping techniques, achieved a substantial relative reduction in Equal Error Rate (EER) of 36.9% compared to a baseline using only standard preprocessing. This finding directly addresses the core challenge outlined in Section 3.3.1, demonstrating a way to mitigate the geometric inconsistencies inherent in contactless capture.

While the proposed pose correction method improves performance, its accuracy is fundamentally limited by the inherent positional variability of the fingerprint core, which this thesis empirically quantifies in Section 3.2 to be as high as 12% of the finger height. This finding underscores the need for future pose correction algorithms to incorporate more stable landmarks or be robust to this level of core deviation.

3.3.6.1 Effectiveness of Pipeline Components

The effectiveness of the overall pipeline comes from the combination of all of its components. The initial horizontal rotation correction (yaw correction), based on fitting the finger’s central axis from the segmentation contour, proved more reliable than simpler bounding box methods, by being less susceptible to noise or irregularities in the finger shape.

The subsequent lateral rotation correction, which utilizes the detected core point and an assumed ellipsoid finger model, builds upon the geometric principles established by Tan and Kumar [247]. However, the results indicate that applying pose correction alone yields mixed results, highlighting the complexity of the distortions and the potential limitations of relying solely on the core point. Integrating the findings from the core centrality analysis (Section 3.2.5) we see that the fingerprint core shows positional variability (6-8% horizontally, 9-12% vertically relative to the geometric center) and modality-induced bias (especially from rolling). Therefore, even with a perfect core detector, the reference point for lateral correction is not fixed or centered, inherently limiting the achievable accuracy of any correction solely based upon it. This variability likely contributes to why pose correction alone did not uniformly improve EER across all fingers, as seen in Table 3.10.

The significant EER improvements observed when applying unwarping (particularly elliptic unwarping) even without prior pose correction (Table 3.10) demonstrate the impact of finger curvature on comparison performance. When combined with pose correction, unwarping provides further gains, effectively addressing both orientation and shape discrepancies.

3.3.6.2 Finger-Specificity and Optimization

No single combination of pose correction and unwarping proved universally best across all five fingers. For instance, the ring finger performed best with no additional steps beyond baseline preprocessing, while the thumb benefited most from combining pose correction with elliptic unwarping, and the index/middle fingers favored elliptic unwarping alone (Table 3.10). This variability likely comes from differences in the shape, size, typical presentation pose, and potentially the core location characteristics of individual fingers. The success of the finger-wise combined approach, yielding the overall best EER (0.99 ± 0.07 %), shows the importance of adapting enhancement strategies to individual finger characteristics for optimal system performance.

3.3.6.3 Impact on Image Quality

While the primary goal was improving comparison accuracy (in the form of EER scores), the analysis also considered the impact on image utility using NFIQ 2. The results suggested a modest improvement in average NFIQ 2 scores and a more notable reduction in score variability after applying pose correction and bidirectional unwarping. However, as emphasized in Section 3.3.5.1, using NFIQ 2 (designed for contact-based plain prints) for contactless and rolled images provides only a rough indication. The observed slight increase in average score and decrease in variability suggest the geometric corrections might lead to features that are somewhat more consistent or better align with NFIQ 2's underlying quality definitions, but definitive conclusions require contactless-specific quality metrics, such as those explored in Section 3.4.

3.3.6.4 Strengths and Limitations

The proposed pipeline offers several strengths. Correcting the pose of the entire image, rather than just minutiae coordinates [247], enables the application of subsequent image-based enhancements like unwarping, leading to greater overall improvement. The improved method for horizontal rotation (yaw) correction enhances reliability compared to simpler approaches. The study explicitly demonstrates the significant gains achievable by optimizing the processing chain on a per-finger basis.

However, limitations must be acknowledged. The pipeline's performance is dependent on the accuracy of the upstream segmentation and core localization modules. Errors in these initial steps will propagate and worsen the effectiveness of the pose correction [217, Section 5.1]. The reliance on a fixed elliptical model (with $k=1.2$) for lateral correction and unwarping is a simplification. Real finger shapes exhibit greater complexity and variation. Furthermore, the computational cost of the full pipeline, including potentially complex deep learning models for segmentation and core detection, needs consideration for resource-constrained applications [217, Section 5.1].

3.3.6.5 Impact on Research Questions

The development and evaluation of the pose correction and unwarping pipeline presented in this section provide direct insights into several of the research questions introduced at the beginning of this thesis.

Firstly, this work directly addresses **RQ2 (What techniques can effectively correct pose variations in contactless fingerprints to enhance comparison accuracy?)** The results demonstrate that a hybrid approach, combining deep learning for segmentation and core localization with geometrically-grounded transformations for rotation correction, is effective. Moreover, the study shows that addressing both orientation (pose correction) *and* surface curvature (unwarping) yields the most substantial improvements in comparison accuracy (given as EER improvements). Furthermore, the finding that the optimal combination of techniques is finger-specific shows that effective pose variation correction necessitates tailored approaches. The significant 36.9% relative EER reduction achieved by the finger-wise optimized pipeline provides a strong answer regarding the efficacy of these combined techniques.

Secondly, the findings provide context for **RQ3 (How can core point detection in contactless images be improved to enable better alignment with contact-based databases?)** While Section 3.2 detailed the development of an accurate core detector, the centrality analysis within that section, and its implications discussed here, reveal a fundamental limitation. The positional variability and modality-induced bias of the fingerprint core itself mean that even a perfect core detector cannot enable perfect alignment if the core is the only reference point. This implies that while improving core detection accuracy (RQ3) is valuable, achieving further enhancements in alignment for pose correction likely requires incorporating additional reference points or developing methods robust to this inherent core uncertainty.

Thirdly, this section highlights the interdependencies between different processing stages, linking back to **RQ1 (How can segmentation of fingerprints from contactless hand images be made more accurate and robust?)** The success of the proposed pose correction, particularly the horizontal rotation component relying on the finger contour, is dependent on the quality of the input segmentation mask.

Finally, the results touch upon **RQ4 (What are the problems of established fingerprints quality assessment metrics [...] on contactless fingerprint images?)** The application of NFIQ 2 to the enhanced images yielded ambiguous results, with only minor average score improvements despite significant EER reductions. This reinforces the notion, central to RQ4, that established metrics designed for contact-based prints (like NFIQ 2) may not capture the quality improvements relevant to contactless comparison after geometric corrections.

3.4 Contactless Fingerprint Quality Assessment

The performance and reliability of any biometric system are linked to the quality of the acquired samples. This holds particularly true for contactless fingerprint recognition, where the acquisition process presents unique challenges compared to traditional contact-based methods [194]. While contactless capture avoids issues like latent prints and non-linear distortions from pressure [87, 237], it introduces variability due to factors like inconsistent illumination, distance variations leading to focus or scaling issues, specular reflections, and unconstrained finger pose [116, 194]. These factors can degrade the captured image, making subsequent feature extraction and comparison less reliable, impacting the overall system accuracy and interoperability, especially with legacy contact-based databases [133, 217]. Therefore, a method for assessing the quality of contactless fingerprint samples is important for operational systems to ensure only sufficiently useful samples are processed or to provide feedback for recapture [194].

3.4.1 Challenges of Quality Assessment for Contactless Fingerprints

The concept of biometric sample quality is formally defined in ISO/IEC 29794-1 and it is based on three concepts: Character, relating to the inherent properties of the source (e.g., a scarred

finger possesses poor character); Fidelity, reflecting the degree of similarity between the sample and its source (e.g., a low-resolution capture results in low fidelity); and Utility, representing the predicted positive influence of the sample on the biometric system's performance, combining both character and fidelity. Quality assessment algorithms aim to map a biometric sample to a numerical score, ideally reflecting its utility, where higher scores predict better recognition performance [89, 194].

While contact-based fingerprint quality assessment is well-established and largely standardized around the NIST Fingerprint Image Quality 2 (NFIQ 2) algorithm [246], which is the reference implementation for ISO/IEC 29794-4, the field of contactless fingerprint quality assessment is lacking such a standard. The distinct challenges of contactless acquisition, such as variable illumination, focus issues, and pose distortions (as discussed in Section 3.3), mean that quality metrics developed for contact-based prints may not directly translate or perform optimally [192, 194]. For instance, sharpness becomes a much more important factor in contactless scenarios due to potential motion blur and narrow depth of field, whereas pressure consistency is irrelevant [116, 194].

3.4.2 Existing Approaches and Limitations

Prior research into contactless fingerprint quality assessment focused on general image quality metrics or features adapted from contact-based analysis. Parziale and Chen [182] proposed a ridge coherence-based measure derived from local gradient analysis. Labati et al. [57] investigated a large set of features including ROI characteristics, Fourier features, and Gabor features, using machine learning classifiers (Neural Networks, kNN) to aggregate them into a quality score. Li et al. [132] developed metrics for smartphone-captured images using spatial and frequency domain features combined with an SVM. Liu et al. [144] evaluated generic image attributes like contrast, sharpness, and luminance for contactless modalities including fingerprints.

These studies often suffered from similar limitations [194]. Many evaluations were conducted on private datasets, hindering reproducibility and comparability. Furthermore, the evaluation methodologies were sometimes limited, often relying on simply dividing data into quality bins and observing error rates rather than using standardized metrics like EDC curves (discussed below). A common pattern was the emphasis on sharpness and contrast measures, potentially neglecting other aspects of fingerprint utility related to character, such as the clarity and continuity of the ridge pattern itself [194]. Wild et al. [269] and a preliminary study by Priesnitz et al. [192] evaluated the direct applicability of contact-based metrics like NFIQ 1.0 and NFIQ 2.0 on contactless data, finding they showed the intended behaviour (assigning lower scores to lower-utility samples) but performed sub-optimally compared to their performance on contact-based data, indicating a need for adaptation or a new approach.

3.4.3 The NFIQ 2 Framework as a Basis

Given the limitations of existing contactless-specific metrics and the established nature of NFIQ 2 for contact-based fingerprints, Jannis Priesnitz proposed adapting the NFIQ 2 framework for mobile contactless fingerprint images, which we explored in [194]. Compared to creating a new framework, NFIQ 2 offers a big advantage: it uses a set of 74 hand-crafted quality features, formally standardized in ISO/IEC 29794-4, which cover various aspects like contrast, ridge clarity, minutiae count, and ROI size. Many of these features, particularly those related to fidelity, such as sharpness, local clarity, and orientation certainty, are highly relevant to the challenges encountered in contactless acquisition [194]. Furthermore, NFIQ 2 uses a random forest classifier to map these features to a final quality score between 0 and 100, where the decision trees inside

the random forest allows for better interpretability of the model results than deep learning models. Furthermore, since the random forest classifier is a machine learning component, it can be re-trained on data specific to a particular capture modality. This re-weights the importance of the individual features to optimize the quality prediction for that specific use case and modality [194, 246]. The hypothesis behind our work in [194] was that the NFIQ 2 feature set is suitable for contactless fingerprints.

3.4.4 The MCLFIQ Algorithm: Adaptation and Training

3.4.4.1 Sample Pre-processing Requirements

A requirement for most quality assessment algorithms is consistent sample presentation and therefore a pre-processing that unifies the input. Raw images captured by contactless devices (typically color images from smartphone cameras) are not directly suitable for NFIQ 2 feature extraction [194]. The framework conditions defined in [194], are also common in many contactless processing workflows [87, 98, 112, 262] and they include:

- **Rotation:** The segmented fingerprint image must be rotated to an upright position.
- **Cropping:** The image should contain only the fingerprint area, typically cropped approximately at the first finger knuckle.
- **Grayscale Conversion:** The image must be converted to grayscale.
- **Ridge Pattern Emphasis:** Techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE) [185] should be applied to enhance the ridge-valley contrast.
- **Normalization:** The ridge-line frequency should be normalized to a standard range (e.g., 8-12 pixels, aligning with approximately 500 PPI contact-based scans).
- **Background Separation:** The fingerprint ROI must be precisely segmented from the background, with the background typically set to white.

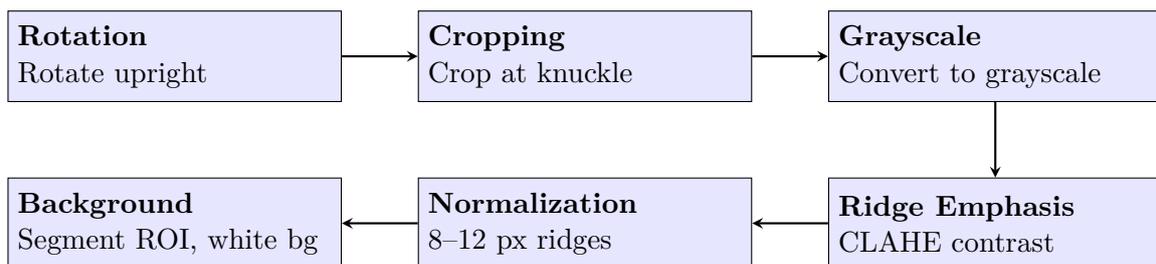


Fig. 3.17: Fingerprint image preprocessing pipeline with horizontal wrapping.

3.4.4.2 Training with Synthetic Data

Training the random forest classifier for MCLFIQ required a large dataset of contactless fingerprint images labeled with ground truth quality information (e.g., high utility vs. low utility). However, large-scale, publicly available, and appropriately labeled datasets of mobile contactless fingerprints are scarce [194]. Manually labeling existing datasets is time-consuming, requires expertise, and is impractical for the scale needed for machine learning model training. Furthermore, using

available real-world datasets for training would reduce the data available for independent testing and evaluation [191].

To overcome this data scarcity, Priesnitz et al. [194] used synthetically generated contactless fingerprint data for training the MCLFIQ random forest classifier. Priesnitz et al. utilized the SynCoLFinGer generator [191], which is capable of creating synthetic contactless fingerprint images with varying levels of quality. SynCoLFinGer models various aspects of contactless capture, including ridge patterns (using SFinGe [33]), deformations, distortions, noise, subject characteristics (like skin color), and environmental influences. This generation process allows for explicit control over the simulated quality level, enabling the automatic generation of ground truth quality labels (high/low) alongside the synthetic images [191, 194]. By training on a large dataset (e.g., 40,000 samples as used in [194]) of diverse synthetic contactless fingerprints with reliable labels, the MCLFIQ random forest could learn the relationship between the NFIQ 2 features and the expected utility for this specific modality, without using real-world data needed for evaluation later on. Figure 3.18 shows examples of synthetic data generated for training.

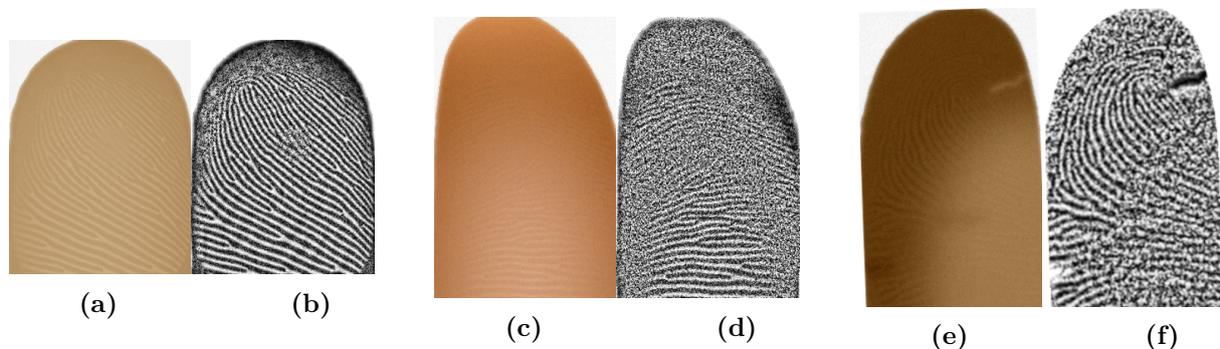


Fig. 3.18: Example images from the synthetic training database generated by SynCoLFinGer [191] for training the MCLFIQ model, showing presets for high (3.18a and 3.18b) and low quality samples (3.18c, 3.18d, 3.18e and 3.18f) and their appearance after pre-processing. Adapted from [194, Fig. 3].

3.4.5 Experimental Evaluation and Results

The effectiveness of the proposed MCLFIQ algorithm was evaluated by comparing its predictive performance against several baseline methods on real-world contactless and contact-based fingerprint datasets.

3.4.5.1 Evaluation Methodology: Performance Metrics

The primary metric for evaluating the predictive power of the quality assessment algorithms is the Error-vs-Discard Characteristic (EDC) curve, following the methodology outlined in ISO/IEC DIS 29794-1 and originally proposed by Grother and Tabassi [89]. An EDC curve plots the biometric system's error rate (typically the False Non-Match Rate, FNMR) as a function of the fraction of lowest-quality comparison scores being discarded. A more detailed explanation is given in the following paragraph.

For a given quality algorithm, quality scores (q_i) are computed for both samples in a genuine mated pair (i). A combined quality score for the pair is determined (typically) using the minimum function, $q_i = \min(q_i^{(1)}, q_i^{(2)})$ [89]. Comparisons are then iteratively discarded based on this combined quality score, starting with the lowest quality pairs. The FNMR is recomputed at each

discard fraction. If the quality score correlates with comparison performance, we would expect a decrease of FNMR, since we remove samples with a low utility. And a steeper decrease in the EDC curve indicates that the quality algorithm is effectively identifying low-utility samples whose removal leads to a reduction in the error rate [194]. To summarize the EDC curve, the Partial Area Under the Curve (EDC PAUC) is often calculated, typically up to a discard fraction of 20% (0.2), providing a single scalar value where lower values indicate better predictive performance [194].

In addition to EDC curves, system performance is also visualized using Detection Error Tradeoff (DET) curves. A DET curve plots the FNMR against the False Match Rate (FMR) across a range of decision thresholds for the biometric comparison scores. DET curves provide a view of the trade-off between the two types of errors for a given system. Performance is considered better for curves that lie closer to the origin (lower error rates for both FMR and FNMR). The Equal Error Rate (EER) is the point on the DET curve where FMR equals FNMR, often used as a single summary statistic for system performance. The DET provides an estimate of the inherent dataset utility.

3.4.5.2 Datasets and Setup

The evaluation utilized three distinct mobile contactless fingerprint databases: the AIT database [116], the HDA database [189], and the ISPFID v1 [223]. These datasets represent captures using different smartphone models under various environmental conditions (indoor, outdoor, constrained, unconstrained). Additionally, two contact-based datasets from the FVC2006 competition [35] (DB2 and DB3) and the contact-based subset of the AIT database were included for a counter-experiment. All contactless images underwent the standardized pre-processing pipeline described earlier before quality assessment and recognition [194].

Three different fingerprint recognition workflows were employed to assess the comparison score: a Commercial-Off-The-Shelf (COTS) system (Innovatrics IDKit [103]), an open-source system combining FingerNet [250] feature extraction with SourceAFIS [258] minutiae comparison, and the NIST NBIS framework [265], consisting of MindTCT for minutiae extraction and Bozorth3 for minutiae comparison.

The baseline quality assessment algorithms used for comparison were: the original NFIQ 2.2 [246], a sharpness-based metric developed by AIT specifically for contactless images [116], and the general-purpose no-reference image quality assessment algorithm BRISQUE [162], which was also re-trained on the same SynCoLFinGer dataset as MCLFIQ [194]. The corresponding DET curves for contact-based and contactless evaluations are presented in Figure 3.19a and Figure 3.19b, showing the overall recognition performance achieved on these datasets using the different comparison algorithms.

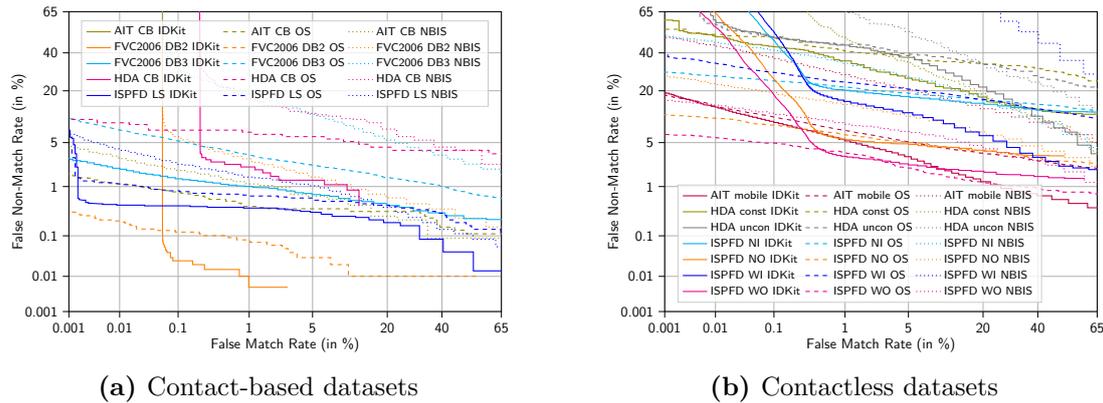


Fig. 3.19: Detection Error Tradeoff (DET) curves showing FNMR vs. FMR using the different recognition systems (IDKit, OS, NBIS). Adapted from [194, Fig. 9].

3.4.5.3 Performance Analysis

The EDC curves for the contactless datasets, presented in [194, Fig. 10] and shown Figure 3.20 and the corresponding EDC PAUC values [194, Tables VI-VIII] versus the EDC curves for the contact based datasets, as presented in the same sections in [194] and shown in Figure 3.21, demonstrated the superior performance of MCLFIQ on the contactless datasets compared to the baseline methods. Across all three contactless databases and all three recognition systems, MCLFIQ consistently yielded lower EDC PAUC values on average, indicating it was more effective at predicting sample utility [194], which can also be seen when comparing the EDC PAUC values depicted in Figure 3.22. While the AIT sharpness metric performed well on some specific database/comparison algorithm combinations, and NFIQ 2.2 showed reasonable performance, MCLFIQ exhibited greater robustness and overall better predictive power. The retrained BRISQUE generally performed poorly on the contactless data. Figure 3.22 provides a visual summary of the average performance.

The counter-experiment on the contact-based datasets, including FVC2006 DB2 and DB3 [194, Fig. 11, Tables IX-XI], results shown in Figure 3.21, confirmed that the original NFIQ 2.2 outperformed MCLFIQ on these datasets. This validates that NFIQ 2.2 is well-tuned for contact-based data, whereas MCLFIQ is specifically optimized for contactless challenges, demonstrating the necessity of modality-specific adaptation.

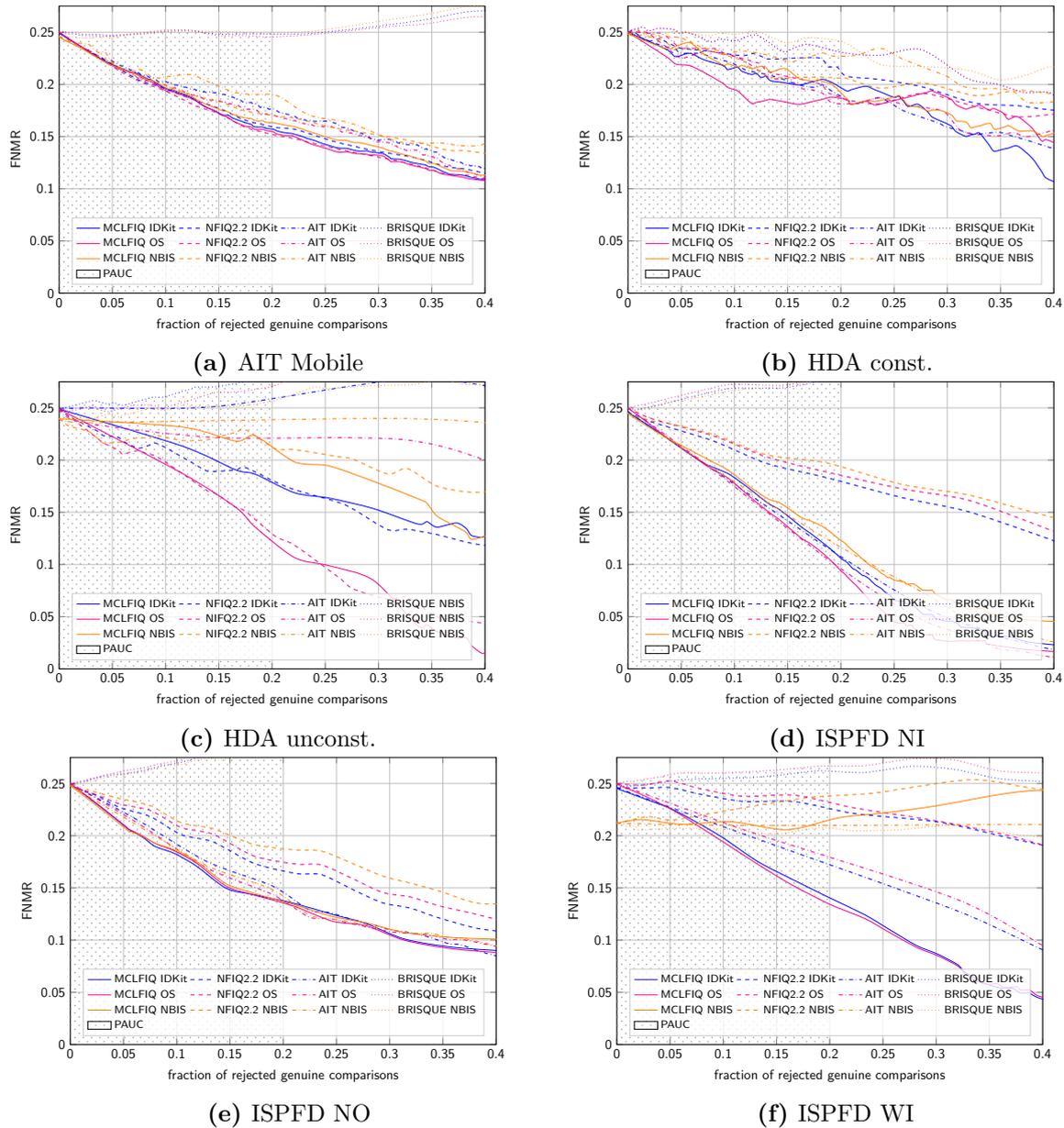


Fig. 3.20: Error-vs-Discard Characteristic (EDC) curves for various contactless datasets (AIT Mobile, HDA constrained/unconstrained, ISPFD variations). The FNMR is plotted against the fraction of discarded lowest-quality genuine comparison scores. Curves showing a steeper decrease indicate better quality prediction. Adapted from [194, Fig. 10].

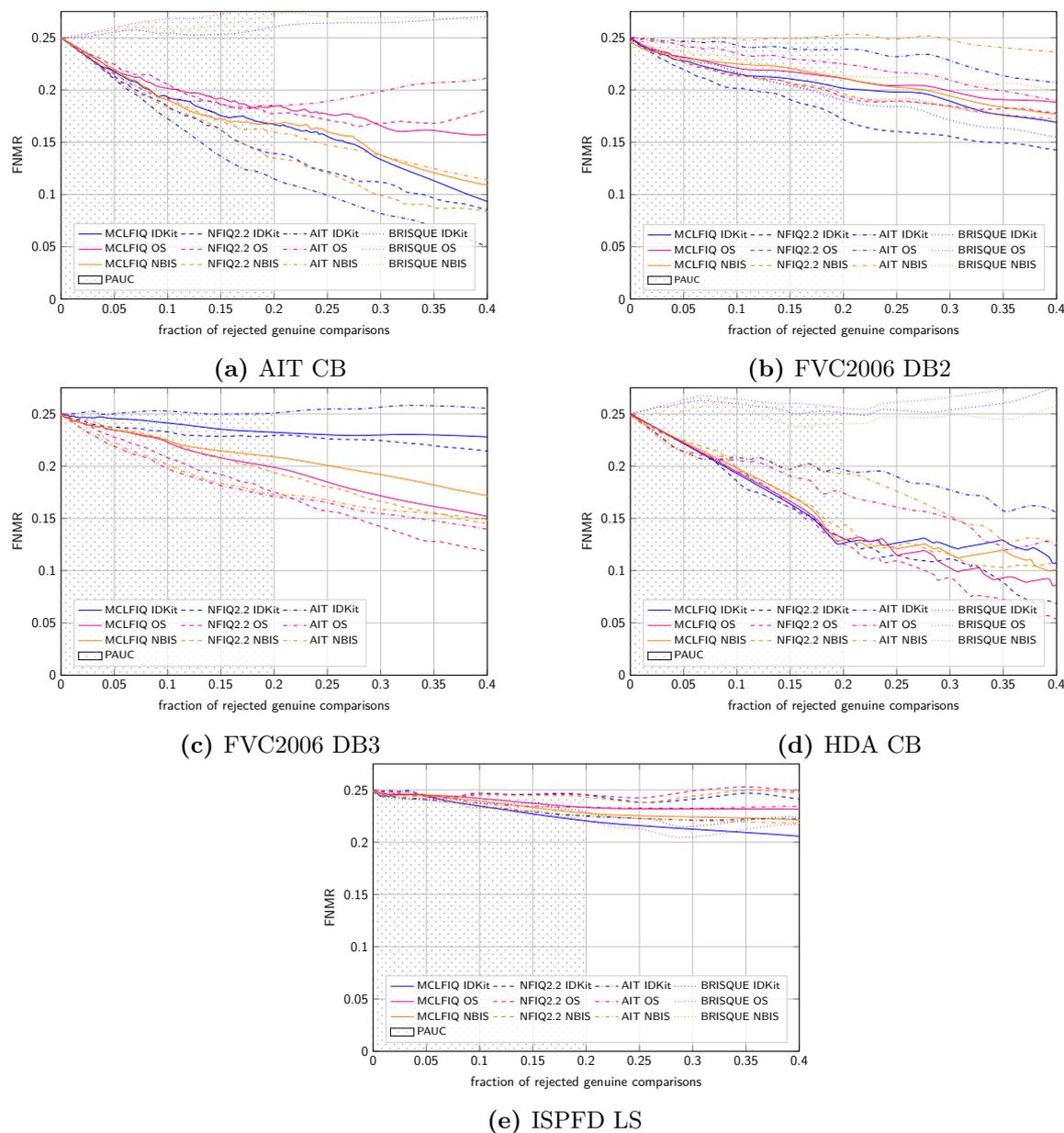


Fig. 3.21: Error-vs-Discard Characteristic (EDC) curves for various contact-based datasets (AIT CB, FVC2006 DB2/DB3, HDA CB, ISPF D LS). As expected, the original NFIQ 2.2 (red lines) generally shows better or comparable performance to MCLFIQ (blue lines) on these datasets, confirming its optimization for contact-based capture. Adapted from [194, Fig. 11].

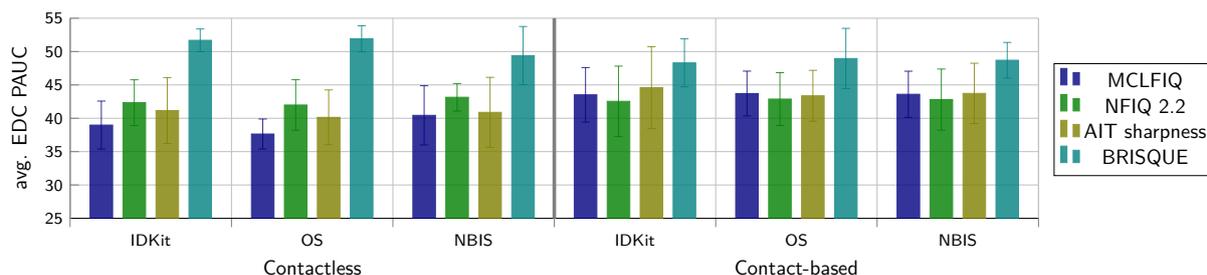


Fig. 3.22: Average EDC PAUC values (lower is better) across contactless and contact-based datasets using different recognition workflows (IDKit, OS, NBIS), comparing MCLFIQ against baseline quality assessment algorithms (NFIQ 2.2, AIT Sharpness, BRISQUE). MCLFIQ demonstrates lower average PAUC and generally lower standard deviation on contactless data, indicating superior predictive power and robustness for this modality. NFIQ 2.2 performs best on contact-based data. Adapted from [194, Fig. 12].

MCLFIQ		NFIQ 2.2	
Feature name	Feature Importance (%)	Feature name	Feature Importance (%)
ROI Relative Orientation Map Coherence Sum	38.56	Frequency Domain Analysis Standard Deviation	6.72
Orientation Certainty Level Mean	21.29	FingerJet FX OSE COM Minutiae Count	4.40
Orientation Certainty Level Bin 0	7.6	FingerJet FX OSE OCL Minutiae Quality	3.96
ROI Orientation Map Coherence Sum	7.45	Ridge Valley Uniformity_Mean	3.32
FingerJet FX OSE OCL Minutiae Quality	5.46	Frequency Domain Analysis Mean	2.97
Orientation Certainty Level Bin 7	5.27	FingerJet FX OSE Total Minutiae Count	2.75
Frequency Domain Analysis Bin 0	2.13	Ridge Valley Uniformity Standard Deviation	2.43
Orientation Certainty Level Bin 8	1.56	Local Clarity Score Bin 7	2.42
Orientation Certainty Level Bin 6	1.29	Local Clarity Score Bin 8	2.39
Frequency Domain Analysis Mean	1.07	Frequency Domain Analysis Bin 9	2.28
Sum	91.68	Sum	33.65

Tab. 3.12: Feature importance comparison between MCLFIQ and NFIQ 2.2 for the top 10 features, taken from [194].

3.4.5.4 Feature Importance Analysis

An analysis of the feature importance within the trained random forest models gives an explanation of the functioning of the quality scores. The result of the 10 most important features are shown in Table 3.12. For MCLFIQ, the most important features were those related to image sharpness and local ridge clarity, such as *ROI Relative Orientation Map Coherence Sum* and *Orientation Certainty Level Mean*. These two features alone accounted for over 50% of the feature importance in the MCLFIQ model. This aligns with the understanding that sharpness and focus are critical challenges in mobile contactless capture. In contrast, the feature importance in the original NFIQ 2.2 model was more uniformly distributed, with features related to minutiae count and quality playing a more significant role alongside frequency domain analysis. This difference highlights how the retraining process adapted the model to prioritize features most indicative of quality degradation specific to the contactless modality (primarily fidelity issues), whereas NFIQ 2.2

balances character and fidelity aspects relevant to contact-based capture [194]. This difference in feature importance distribution also led to a smaller model size for MCLFIQ (approx. 296 KB) compared to NFIQ 2.2 (approx. 53 MB) due to shallower decision trees in the random forest [194].

3.4.6 Discussion

The work presented in [194] and summarized here demonstrates the potential and necessity of adapting quality assessment frameworks for the contactless fingerprint recognition. The development and evaluation of the Mobile Contactless Fingerprint Image Quality (MCLFIQ) algorithm confirm that while the feature set of NFIQ 2 [246] captures relevant quality aspects, modality-specific retraining of the underlying classifier can improve performance in the contactless domain.

3.4.6.1 Effectiveness of the Adapted Framework

The strength of the MCLFIQ approach is its ability to use the standardized feature set of NFIQ 2 while adapting the predictive model to the specific degradations common in contactless captures. The experimental results (Figures 3.20, 3.21, and 3.22) show that MCLFIQ outperforms the original NFIQ 2.2 and other baseline metrics (like sharpness or retrained BRISQUE) in predicting the utility of contactless samples across various datasets and recognition systems [194].

Furthermore, the successful use of synthetic data generated by SynCoLFinGer [191] validates synthetic data generation as a data alternative, enabling the development of larger models without compromising real data needed for testing [194]. The feature importance analysis (Table 3.12) shows that MCLFIQ prioritizes features related to image sharpness and local ridge clarity (fidelity aspects), such as *ROI Relative Orientation Map Coherence Sum* and *Orientation Certainty Level Mean*. This contrasts with NFIQ 2.2, where feature importance is more distributed, reflecting the shift in focus towards the primary quality challenges encountered in mobile contactless acquisition [194].

3.4.6.2 Impact on Research Questions

This investigation into contactless fingerprint quality assessment directly addresses key research questions posed earlier in this thesis.

Most centrally, it tackles **RQ4 (What are the problems of established fingerprints quality assessment metrics [...] on contactless fingerprint images? What can be changed to improve their performance?)** The evaluation confirms that established metrics like NFIQ 2.2, while effective for contact-based prints, show suboptimal performance when directly applied to contactless data (Figure 3.20 vs. 3.21). The problem identified is the mismatch in the weighting of quality features relative to their impact on utility in the contactless domain. The solution demonstrated is the adaptation through retraining the classifier component (MCLFIQ). By retraining the random forest on modality-specific (synthetic) data, the feature weights are adjusted (Table 3.12), leading to improved predictive performance for contactless samples.

The work also implicitly relates to other questions. Accurate quality assessment is important for the overall system pipeline. For instance, quality scores from MCLFIQ can inform decisions about sample usability before computationally expensive steps like **pose correction (RQ2)** are attempted. A low-quality score might indicate severe blur or illumination issues that could hinder the accurate segmentation (**RQ1**) or core localization (**RQ3**) required for pose correction algorithms, suggesting recapture might be necessary.

3.4.6.3 Limitations

Despite the success of MCLFIQ, certain limitations should be noted. Firstly, the performance of MCLFIQ, like NFIQ 2, is dependent on a standardized pre-processing pipeline (Section 3.4.4.1). Deviations from this pipeline could negatively impact feature extraction and thus the resulting quality score. Secondly, while the NFIQ 2 feature set proved largely applicable, it is based on hand-crafted features. It remains possible that deep learning-based feature extractors could capture more subtle or complex quality indicators specific to contactless modalities that are missed by the current set. Thirdly, the evaluation relied on publicly available contactless datasets which are still limited in size and diversity compared to contact-based benchmarks. Further validation on larger, more varied datasets would strengthen the conclusions regarding MCLFIQ's generalizability [194]. Finally, it's important to reiterate that the core development and training of the MCLFIQ algorithm were primarily conducted by Jannis Priesnitz, with the author of this thesis contributing significantly to the extensive evaluation phase documented in [194].

3.5 Fingerprint Mosaicking Artifact Detection

Fingerprint mosaicking is the process of combining multiple, potentially overlapping, partial fingerprint images to form a single, larger, and more complete fingerprint representation. This is particularly relevant for generating images equivalent to traditional rolled fingerprints, which capture the entire nail-to-nail surface area. Mosaicking techniques are applied in both contact-based and increasingly in contactless fingerprint acquisition scenarios [49, 109, 142]. While mosaicking aims to enhance the captured data, the process itself can introduce errors, known as mosaicking artifacts. These artifacts can manifest as structural inconsistencies or geometric distortions, degrading the quality and biometric utility of the resulting image [142, 218]. This section presents the challenges of combining individual fingerprint recordings leading to these artifacts, reviews related work, details a novel detection algorithm based on self-supervised learning, presents evaluation results, and discusses the overall findings and implications.

3.5.1 Challenges of Fingerprint Mosaicking

Errors introduced during the fingerprint mosaicking process can compromise the integrity of the final composite image. These errors generally fall into two categories based on their visual characteristics and impact on ridge structures, termed soft and hard artifacts in [218].

Soft artifacts primarily involve geometric distortions, such as stretching, compression, or warping, that might occur during the alignment and blending of partial images. While these distortions alter the overall fingerprint shape or relative minutiae positions, they typically maintain the continuity of individual ridge lines across the boundaries where the original images were joined. Assessing the severity of soft artifacts can be difficult without comparing the mosaicked image to a ground truth or reference print, as some degree of natural elastic deformation is inherent in fingerprint capture, especially with contact-based rolling [218].

Hard artifacts on the other hand represent more severe errors characterized by visible structural discontinuities in the ridge-valley pattern. These typically occur at the seams or overlapping regions between the stitched partial images. Examples include ridges that terminate instead of connecting to their corresponding counterparts in the adjacent patch, noticeable misalignments where ridge patterns do not smoothly transition across seams, or even the duplication or deletion of ridge segments due to poor registration or blending [218]. Such hard artifacts corrupt the underlying biometric information encoded in the ridge pattern. They can create spurious minutiae or displace genuine ones, increasing the likelihood of recognition errors, including false matches

(incorrect identifications) and false non-matches (failure to identify a true match) [45, 142]. In some cases, stitching algorithms might attempt to hide these discontinuities using post-processing techniques like localized blurring or alpha blending along the seams. While this might improve the visual appearance, it does not correct the underlying structural error and can make artifact detection more complex [218]. Established fingerprint quality metrics, like NFIQ 2 [246], primarily evaluate factors such as ridge clarity, contrast, and minutiae count, and are not designed to detect these kinds of structural flaws arising from the mosaicking process. Therefore, dedicated approaches are required to identify hard mosaicking artifacts.

3.5.2 Related Work

The creation of composite images from multiple views, known as image stitching or mosaicking, is a well-established technique in computer vision with applications from panoramic photography to medical imaging and remote sensing [245]. Research in this area is done for both creating seamless mosaics as well as techniques for evaluating their quality or detecting errors.

Fingerprint Mosaicking Techniques: In the context of fingerprints, early work by Jain and Ross demonstrated the feasibility of mosaicking and emphasized the critical role of accurate alignment (registration) of the partial images [109]. Subsequent research explored various capture and processing strategies. Choi et al. investigated mosaicking based on sequences captured during finger rolling and sliding motions to acquire a larger surface area [42]. Ross et al. compared the merits of aligning images based on overall image information (intensity-based) versus aligning based on extracted features like minutiae (feature-based) [210]. Liu et al. specifically addressed the challenges of mosaicking images acquired using contactless sensors, which often have a greater pose variability [142]. Algorithms based on phase correlation [23] and minutiae-free approaches [38] were also proposed for specific alignment scenarios. More recently, deep learning has emerged as a powerful tool, with Cui et al. developing an end-to-end network for dense registration, achieving highly accurate alignment for fingerprint mosaicking [49]. These techniques aim to minimize artifacts during creation, but errors can still occur due to factors like non-linear distortion, acquisition inconsistencies, or algorithmic limitations.

Artifact and Error Detection in Image Stitching: Detecting errors in stitched images is important for quality control. General image stitching literature addresses common artifacts like ghosting (misalignment causing semi-transparent duplicates), blurring, visible seams, and geometric inconsistencies [28, 245, 283]. Reference-free quality assessment methods evaluate the stitched image intrinsically, without comparing it to a ground truth. General-purpose no-reference image quality metrics, such as BRISQUE [162], assess perceptual quality based on statistical features but are not able to detect specific structural stitching errors. Fingerprint-specific quality metrics like NFIQ 2 [246] and the adapted MCLFIQ [194] (discussed in Section 3.4) focus on features indicative of sample utility, such as ridge clarity and minutiae quality, rather than explicitly identifying mosaicking-induced structural flaws. Metrics analyzing local sharpness variations [261] have the potential to detect blurring which was introduced along seams to mask discontinuities, but they do not address the underlying misalignment [218]. Methods developed for detecting errors in panoramic video stitching, such as the pairwise inconsistency assessment by Nabil et al. [166], operate without a final perfect reference but focus on inconsistencies between adjacent input frames, which is analogous to detecting errors along seams in a live recording, but this may not capture all types of patch misalignments within a fingerprint context and is also not suitable for image recordings. The core limitation of these reference-free approaches is often their lack of specificity for the unique, localized structural discontinuities, misaligned or broken ridges, that define hard mosaicking artifacts in fingerprint patterns.

Reference-based methods, on the other hand, assess quality by comparing the stitched image against a known reference or ground truth. General metrics like the Structural Similarity Index (SSIM) [174] provide an overall measure of similarity but may lack sensitivity to localized defects. In fields like 3D view synthesis assessment, metrics evaluate geometric accuracy and visual realism against the original 3D model or multiple source views [14, 47]. Within fingerprint biometrics, reference-based evaluation concepts exist, such as NIST’s definition of Geometric Accuracy for assessing acquisition devices [133] and spectral validation metrics proposed by Libert et al. [134], which inherently rely on comparing acquired data to known targets or references. While potentially accurate for detecting both soft geometric distortions and hard structural errors, the major drawback of reference-based methods in the context of operational mosaicking is the general unavailability of a pristine, artifact-free ground truth image corresponding to the mosaicked sample for real-world usescases. One possible solution is to test sensors in a laboratory environment with known ground truths and repeatable experiments. For this, standardized finger phantoms are required, see section 3.6 for more. This however brings the risk that the laboratory setting differs from the real-world environment, resulting in possible missed issues. Consequently, a reference-free approach specifically designed to detect the characteristic structural flaws of hard mosaicking artifacts in fingerprints was needed, motivating the work presented in [218]. The conceptual distinction between soft (geometric/appearance) and hard (structural) artifacts aligns with discussions of geometric versus photometric inconsistencies in general stitching literature [245], however, the specific way these manifest in complex fingerprint ridge patterns necessitates a specialized detection strategy.

3.5.3 Proposed Mosaicking Artifact Detection Algorithm

To overcome the limitations of existing methods and the need for manual labeling, a novel mosaicking artifact detection algorithm was developed using a self-supervised deep learning strategy [[ruzicka_mosaickingpaper_2025](#)]. The core principle involves training a neural network to differentiate between authentic fingerprint ridge structures and the structural anomalies introduced by mosaicking errors, using artificially generated artifacts for supervision. The full framework, including the artifact insertion code as well as the training and inference code can be found publicly available under the Mozilla Public License Version 2.0 here: [218].

3.5.3.1 Self-Supervised Training Strategy

The self-supervised training strategy builds on a data generation pipeline that creates labeled training data from unlabeled, artifact-free fingerprint images [218]. This process begins by taking a known, artifact free fingerprint image. Then, artificial hard mosaicking artifacts are programmatically introduced. Two distinct types of structural discontinuities are simulated. The first type involves simulating patch-based misalignments, where a randomly sized rectangular region within the image is selected and its pixels are shifted by a small, random offset (constrained to 2-7% of the image dimension in the study). This replicates localized errors that can occur when stitching together segments. The second type simulates seam-based errors by shifting all pixels along a randomly chosen horizontal or vertical line by a small offset. This mimics discontinuities along the boundary between two stitched parts. As a result, we can record the exact locations and extent of these introduced artifacts during generation. This information is used to automatically create corresponding pixel-level ground truth segmentation masks, where artifact regions are labeled positively. The original image areas serve as negative examples. This automatic labeling process enables training on large datasets without manual work. To ensure the model learns to detect the structural characteristics of the artifacts rather than superficial

image features, standard data augmentation techniques (like rotation, perspective shifts, blur, etc.) are applied to the base images before the artificial artifacts are introduced [218]. Figure 3.23 illustrates this artifact simulation process.

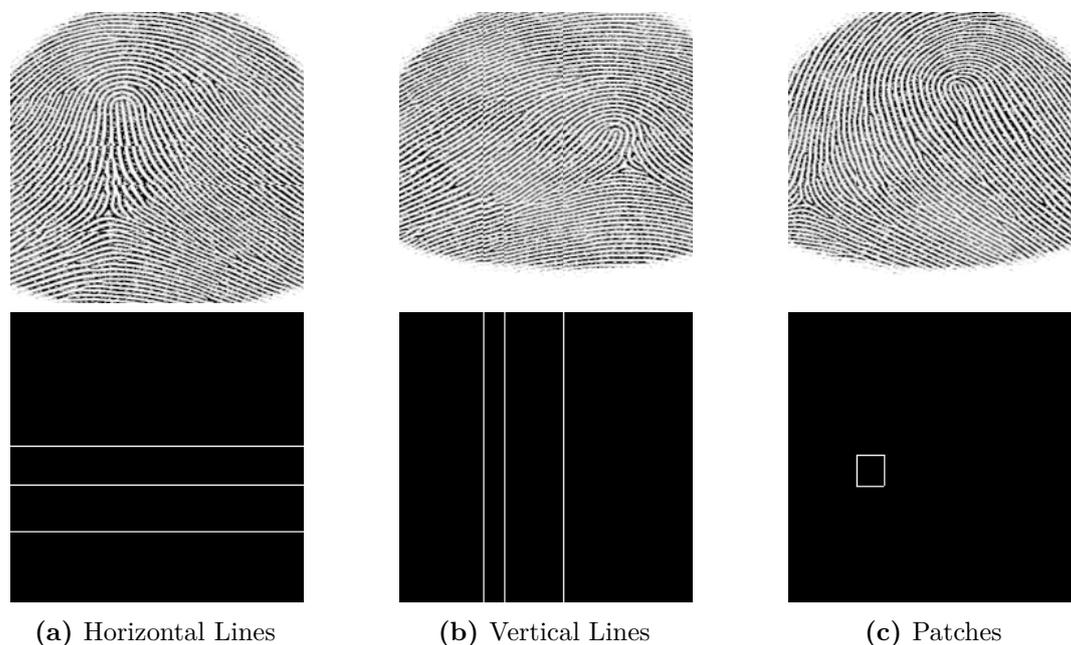


Fig. 3.23: Illustration of the self-supervised artifact generation. Top row shows synthetic fingerprint images with artificially introduced hard mosaicking artifacts: (a) Horizontal line shifts, (b) Vertical line shifts, (c) Patch shifts. Bottom row shows the corresponding automatically generated ground truth segmentation masks (artifact locations in white) used for training the detector. Figure adapted from [218].

3.5.3.2 Model Architecture and Implementation

The artifact detector uses the UNet++ segmentation architecture idea [289]. This architecture is implemented using a ResNeSt-50d encoder [281], which is initialized with weights pre-trained on the ImageNet dataset [55] for feature extraction. The UNet++ design utilizes nested and dense skip connections, allowing for feature fusion across multiple scales, which is beneficial for localizing the potentially small and varied artifact regions [218, Fig. 2]. The network takes a fingerprint image as input. Its output is a binary mask of the same dimensions, where pixels identified as belonging to a hard mosaicking artifact are predicted. The model training optimizes a Jaccard loss (also known as Intersection-over-Union loss), commonly used for segmentation tasks, via the Stochastic Gradient Descent (SGD) optimizer [218]. Figure 3.24 provides an abstract overview of the architecture.

3.5.3.3 Mosaicking Artifact Score (MAS)

To enable automated analysis and provide a summary of the severity of artifacts in an image, the Mosaicking Artifact Score (MAS) was defined [218]. This score consolidates the information from the output segmentation mask into a single value. It calculates a weighted sum of the areas of all detected artifact regions (both patches and lines). The weights are chosen to highlight the

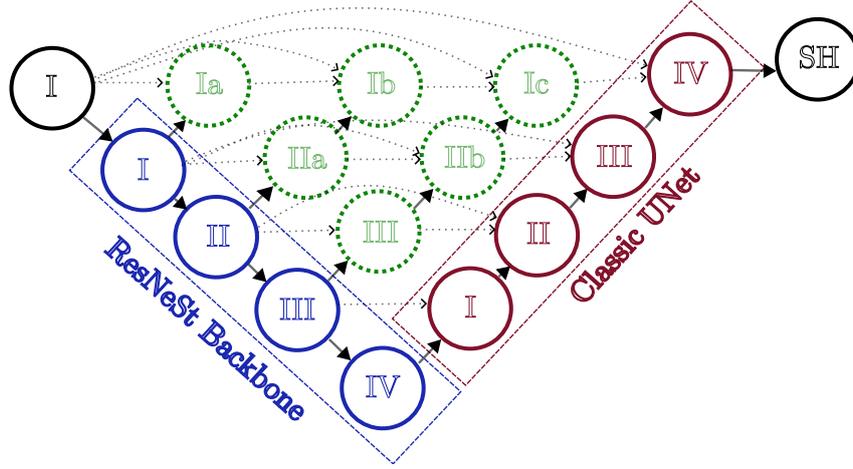


Fig. 3.24: Abstract visualization of the UNet++ based model architecture used for artifact detection, combining a ResNeSt encoder (blue, labeled ResNeSt Backbone) with a UNet++ decoder (green/red, including intermediate layers like Ia, Ib, Ic, IIa, IIb, IIIa, and final decoder stages I, II, III, IV leading to the Segmentation Head SH). Arrows indicate data flow, dotted lines represent skip connections. Adapted from [218, Fig. 2].

severity and also the nature of each artifact type, with patch artifacts receiving a significantly higher weight than line artifacts in the study (details in [218, Eq. 1]).

$$\begin{aligned}
 \text{MAS} &:= \left(\sum_i^n b_{\text{patch}} + w_{\text{patch}} \times h_{\text{patch}} + \right. \\
 &c \times \left(\sum_i^m s_{\text{height}} \times w_{\text{line}} + \sum_i^o s_{\text{width}} \times h_{\text{line}} \right) \left. \right) \times \\
 &\frac{100}{s_{\text{width}} \times s_{\text{height}}}, \tag{3.35} \\
 b_{\text{patch}} &:= b \times \frac{s_{\text{width}} \times s_{\text{height}}}{100},
 \end{aligned}$$

where n is the number of detected patches where misalignment occurred, b a constant that weights the importance of multiple patches and is set to 5 in this setting, $w_{\text{patch}}, h_{\text{patch}}$ stand for the width and the height of the detected patch, c is a constant that balances the importance of the single line artifact and is set to 0.025 in our case, m is the number of detected purely vertical line artifacts, s_{height} stands for the height of the segmentation mask, w_{line} is the width of the detected line artifact, h_{line} is the height of the detected line artifact, o depicts the number of detected horizontal line artifacts and s_{width} stands for the width of the segmentation mask.

A MAS of 0 signifies an artifact-free prediction. Increasing scores indicate progressively more severe or widespread artifacts. A threshold can be set based on the patch weight (e.g., $\text{MAS} \geq 5$ in the study) to make a binary decision on whether significant artifacts are present [218].

3.5.4 Experimental Evaluation and Results

A set of experiments was conducted to evaluate the proposed artifact detection method’s accuracy, its robustness against various image degradations, and the impact of the artifacts it detects on fingerprint recognition performance [218].

3.5.4.1 Datasets and Setup

Two main versions of the detection model were trained using the self-supervised strategy. The contactless (CL) version of the model was trained on a large dataset of artifact-free single-shot contactless fingerprint images taken from [266]. The pressed (PR) fingerprint version of the model was trained on a dataset of artifact-free contact-based pressed prints. Model performance was evaluated using standard image segmentation metrics (Intersection-over-Union (IoU), F1-score, F2-score, Accuracy, Recall) on held-out test portions of their respective training datasets. Generalization ability was assessed by testing these models on entirely different datasets and modalities, including the challenging NIST SD300a dataset containing scanned slap and rolled ink fingerprints [73], an additional pressed print dataset (PRD-2), and a dataset of rolled prints (ROD-1) corresponding to the same subjects as in PRD-1 [218]. The Mean Score Difference between the predicted MAS and the ground truth MAS (based on the known locations of simulated artifacts during testing) was also calculated as a measure of score accuracy.

The robustness ablation study tested the models’ specificity by evaluating their response to images containing common quality defects that are not mosaicking artifacts. One hundred synthetic fingerprint base images (from SFinGe [33]) were altered using methods described in [191] to simulate three levels of severity (low, medium, high) for various degradations: general skin defects, ink variations (smudging/lightness), additive sensor noise, scars, and wounds. Examples are shown in Figure 3.25. The MAS was calculated for each of these altered images to determine if the detector would produce false positives triggered by these unrelated quality issues [218].

The impact of mosaicking artifacts on fingerprint recognition accuracy was quantified by introducing controlled artificial artifacts (small offsets of 1-2% image dimension, large offsets of 2-7%) into a set of clean images. The Equal Error Rate (EER) was then computed by matching these artifact-containing images and comparing the resulting comparison scores with the scores of the original artifact-free counterparts. This analysis employed three distinct fingerprint recognition systems: an open-source pipeline (FingerNet [250] for minutiae extraction combined with SourceAFIS [258] for template comparison), the standard NIST NBIS tools (MindTCT for minutiae extraction and Bozorth3 for template comparison) [265], and a high-performing commercial comparison algorithm (Innovatrics IDKit for the whole ABIS pipeline) [103].

3.5.4.2 Detection Accuracy and Generalization

The experimental results showed that both the CL and PR trained models achieved excellent performance in detecting the simulated artifacts on test data matching their training modality, with IoU scores around 0.98 [218, Table II]. Table 3.13 summarizes key metrics. Both models demonstrated substantial generalization ability as well. The CL model, despite being trained only on contactless images, successfully detected artifacts in scanned slap ink prints (NIST SD300a slap, IoU 0.959) and in rolled ink prints (NIST SD300a rolled, IoU 0.908). The PR model also generalized well to other pressed and rolled datasets. Figure 3.26 provides qualitative examples of the detector output on challenging NIST SD300a slap and rolled images. The Mean Score Difference between the predicted and ground truth MAS was consistently low (typically $\ll b$) across diverse test sets, indicating that the quantitative score accurately reflected the severity of the simulated artifacts [218].

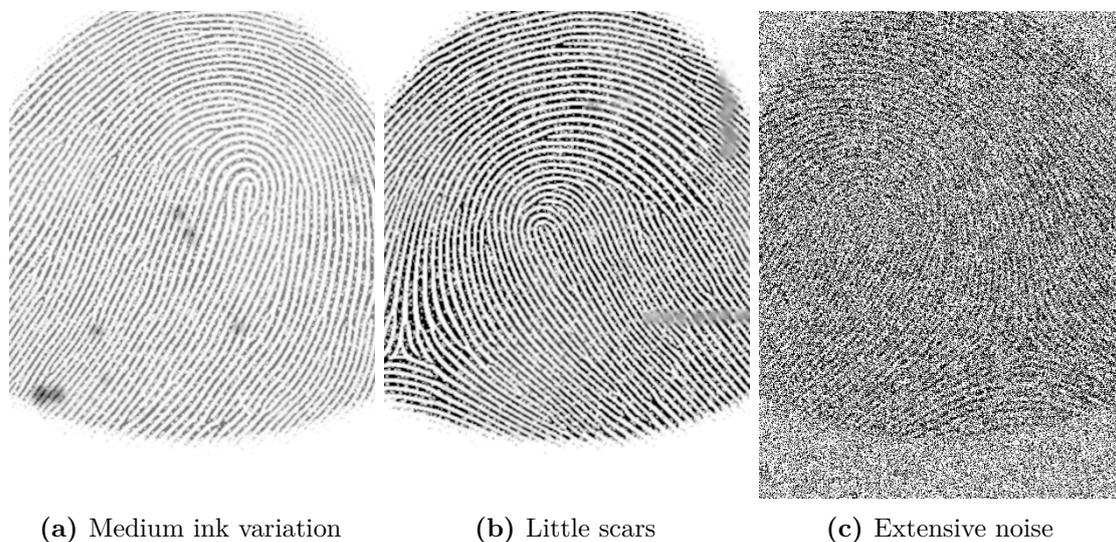
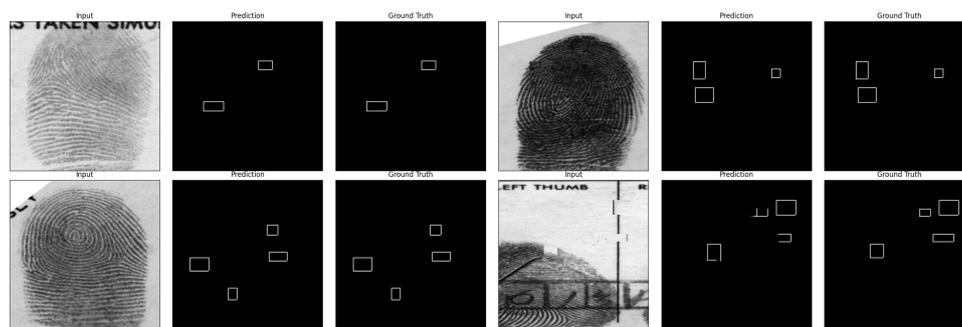


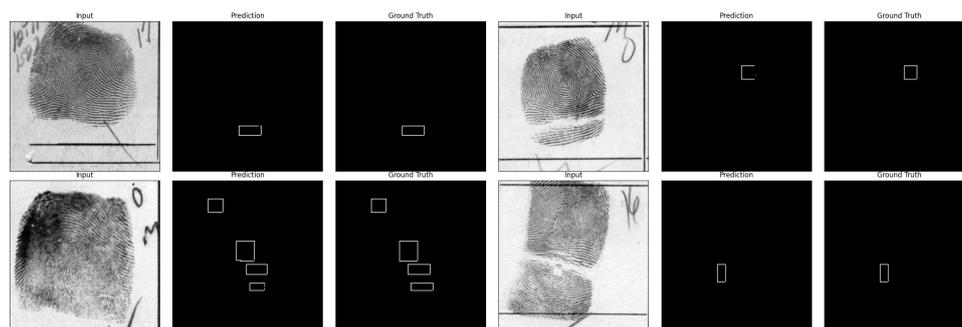
Fig. 3.25: Examples of synthetic fingerprint alterations used in the robustness ablation study, simulating common image quality issues unrelated to mosaicking. These include (a) medium ink variations, (b) minor scarring, and (c) extensive sensor noise. Images generated using methods from [191] applied to SFinGe [33] bases. Adapted from [218, Fig. 4].

	Dataset	IoU	F1	F2	Accuracy	Recall	Mean Score Dif.
CL	Weissenfeld et al. [266]	0.982	0.991	0.990	1.000	0.989	0.264
	NIST 300a slap	0.959	0.979	0.975	1.000	0.972	0.483
	NIST 300a rolled	0.908	0.952	0.940	1.000	0.932	1.061
PR	PRD-1-Test	0.977	0.988	0.987	1.000	0.986	0.355
	PRD-2	0.978	0.989	0.988	1.000	0.987	0.351
	ROD-1	0.931	0.964	0.957	1.000	0.952	0.815

Tab. 3.13: Summary of model performance metrics (IoU, F1, F2, Accuracy, Recall, Mean Score Difference) for the Contactless (CL) and Pressed (PR) trained models across various test datasets. Performance is high on native modalities and shows good generalization. Adapted from [218, Table II].



(a) NIST SD300a Slap Examples



(b) NIST SD300a Rolled Examples

Fig. 3.26: Examples of the artifact detector output (CL model) on challenging images from the NIST SD300a dataset. Each triplet shows: Input Image (left), Predicted Artifact Mask (center, white areas indicate detected artifacts), Ground Truth Mask (right, based on simulated artifacts). The model effectively identifies simulated artifacts even on these significantly different ink-based modalities. Adapted from [218, Fig. 3].

3.5.4.3 Robustness Ablation Study

The robustness tests shown in Table 3.14 confirmed the models' high specificity (taken from [218, Table III]). When evaluating images synthetically altered to include common quality defects like scars, noise, or ink variations, the MAS scores produced by both models remained consistently near zero across almost all conditions and intensities. This demonstrates that the models are robust against these factors and do not incorrectly flag general quality degradation as mosaicking artifacts. Only medium/high levels of additive noise led to a minimal number of false positives (1-3 images out of 100 slightly exceeding the MAS threshold of 5), but the average scores remained low, confirming the detectors' focus on structural discontinuities typical of mosaicking errors. Table 3.14 shows the mosaicking score for the different alteration types.

3.5.4.4 Impact on EER

The analysis clearly quantified the negative impact of mosaicking artifacts on fingerprint recognition accuracy [218, Table IV], as shown in Figure 3.15. Introducing even small artifacts (simulating 1-2% positional offsets) resulted in the EER more than doubling for both the open-source SourceAFIS system and the commercial Innovatrics IDKit system compared to matching artifact-free images. Larger artifacts (2-7% offsets) caused further substantial increases in EER for the SourceAFIS and Idkit recognition systems. This experiment provides strong evidence that

		No	Skin			Ink			Noise			Scar			Wounds		
			↓	~	↑	~	↑	↓	~	↑	↓	~	↑	↓	~	↑	
CL Model	max	1.12	1.13	0.81	1.19	1.09	1.07	0.44	5.18	0.56	0.93	1.14	1.31	1.11	1.08	1.35	
	median	0.01	0.03	0.01	0.02	0.03	0.00	0.00	0.00	0.00	0.02	0.04	0.02	0.02	0.03	0.08	
	mean	0.15	0.15	0.10	0.14	0.15	0.11	0.03	0.09	0.05	0.15	0.15	0.14	0.14	0.18	0.23	
	std	0.26	0.25	0.18	0.24	0.25	0.24	0.08	0.56	0.12	0.26	0.24	0.26	0.24	0.26	0.32	
PR Model	max	0.00	0.00	0.00	0.00	0.00	0.01	0.00	10.41	0.13	0.00	0.00	0.00	0.00	0.00	0.21	
	median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	std	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.27	0.01	0.00	0.00	0.00	0.00	0.00	0.03	

Tab. 3.14: Robustness ablation results of the model trained on contact-less images (CL Model) and contact based images (PR Model). Columns indicate no modification (No), skin damage (Skin), ink problems (Ink), added image noise (Noise), added scars (Scar) and added wounds (Wounds). The arrows indicate low (↓), medium (~) and high (↑) intensity of the image modification. The table presents the maximum value (max), median value, mean value and the standard deviation (std) of the mosaick artifact score.

the failure to detect and mitigate mosaicking artifacts can severely compromise the reliability of fingerprint comparison systems in practice.

Artifacts	SourceAFIS [%]	Bozorth3 [%]	Idkit [%]
None	0.43	3.97	0.38
Small Offset	0.91	5.41	0.88
Large Offset	0.99	4.82	0.88

Tab. 3.15: Equal Error Rates (EERs) showing the significant increase in comparison errors when small or large mosaicking artifacts are present, compared to artifact-free images (None). Results are shown for three different fingerprint recognition systems (SourceAFIS, Bozorth3/MindTCT, IDKit). Adapted from [218, Table IV].

3.5.5 Discussion

3.5.5.1 Effectiveness of the Self-Supervised Approach

Firstly, we showed that self-supervised learning can train a model to detect complex, structurally defined artifacts like hard fingerprint mosaicking errors. The automatic generation of artifacts and corresponding ground truth labels from unlabeled data makes the method highly scalable and avoids the intensive manual labor required for annotating such defects manually. The resulting models showed high accuracy and strong generalization capabilities across different fingerprint capture modalities (contactless, pressed, rolled, slap) and data sources (live scans, ink prints). This suggests the models learned to identify the fundamental structural discontinuities associated with hard mosaicking, rather than overfitting to the characteristics of a specific dataset [218]. The proposed Mosaicking Artifact Score (MAS) provides an interpretable quantitative measure of artifact severity, suitable for automated quality assessment pipelines. Its low false-positive rate and robustness against common non-mosaicking image degradations further shows its potential for real-world use [218].

3.5.5.2 Impact on Research Questions

This research directly addresses several questions posed within this thesis. Primarily, it offers a concrete solution to **RQ5 (How can we detect errors in the mosaicking process of contactless fingerprint sensor systems?)**. The validated self-supervised method and the MAS metric provide an automated tool for this purpose.

It also significantly informs **RQ4 (What are the problems of established fingerprints quality assessment metrics and what can be change to improve their performance on contactless fingerprint images?)**. This work identifies mosaicking artifacts as a distinct quality issue concerning structural integrity, which is not explicitly targeted by general metrics like NFIQ 2 or the fidelity-oriented MCLFIQ. The proposed detector and score represent an addition to fingerprint quality assessment, enabling evaluation of the mosaicking process outcome itself. This is especially important for systems using mosaicking to create rolled print equivalents from contactless captures.

Furthermore, the detection of these artifacts impacts other processing stages. For **RQ2 (What techniques can effectively correct pose variations in contactless fingerprints to enhance comparison accuracy?)** and **RQ3 (How can core point detection in contactless images be improved to enable better alignment with contact-based databases?)**, preventing images with severe mosaicking artifacts from being processed is important, as the corrupted ridge structure could lead to errors in segmentation, landmark localization, and subsequent geometric transformations. For **RQ7 (How can we create a privacy preserving comparison approach that circumvents the risk of exposing biometric data to other parties?)**, ensuring the structural soundness of fingerprint templates before they enter secure computation protocols enhances the overall reliability and trustworthiness of the comparison result.

3.5.5.3 Limitations

Certain limitations should be acknowledged regarding the current work [ruzicka_toward_2025]. The training relies on simulated artifacts. While designed based on observed errors, these simulations might not encompass the full diversity of artifacts produced by every possible real-world mosaicking algorithm. The detection mechanism is specifically focused on hard structural discontinuities. Reliably detecting softer geometric distortions (soft artifacts) without access to a reference image remains an open research problem. While good generalization was observed, performance might potentially degrade when encountering images from vastly different sensor technologies or mosaicking software than those represented in the datasets used for training the base models. Lastly, the computational demands of the deep learning model might require further optimization techniques (like model pruning or quantization) for efficient deployment in highly resource-constrained environments, such as directly within fingerprint sensors for real-time analysis.

3.5.5.4 Future Research Directions

A next step would be to collect and annotate a dataset containing real-world mosaicking artifacts generated by various operational systems. Fine-tuning the self-supervised models on real data could enhance detection performance, especially for abnormal artifact types. Integrating the artifact detection model directly into fingerprint capture devices or acquisition software could enable immediate feedback to the operator, facilitating re-acquisition of faulty captures. Developing techniques to detect soft geometric distortions without a reference, perhaps using anomaly detection, would be a valuable extension. Combining the specialized Mosaicking

Artifact Score with general-purpose quality metrics like NFIQ 2 or MCLFIQ could lead to a more comprehensive, multi-dimensional quality assessment framework, providing a richer understanding of sample utility. Adapting and evaluating the artifact detection methodology for emerging 3D fingerprint capture and mosaicking techniques is another important area. Finally, given the demonstrated impact of these artifacts, advocating for the inclusion of specific mosaicking artifact checks within future fingerprint image quality standards (e.g., future versions of ISO/IEC 29794) should be considered to promote better practices across the biometrics community.

3.6 Synthetic Finger Phantoms

Ensuring the reliability and performance of fingerprint recognition systems requires sensor evaluation and calibration methodologies. Traditionally, this involves human sampling, where real users interact with sensors [179]. However, this approach has limitations: participant selection bias, lack of repeatability, and the absence of a definitive ground truth for comparison [219]. Furthermore, increasing data privacy regulations¹ [90, 235, 243] restrict the collection and usage of real biometric data. While digitally synthesized fingerprint images [33, 63, 155, 191, 203, 259] address privacy concerns and offer control over generated patterns, they cannot fully replicate the interaction between a physical finger and a sensor, including factors like skin deformation, moisture, and pressure application for contact-based sensors and perspective distortions, changes in lighting and motion blur in contactless sensors.

To provide a standardized, repeatable, and privacy-compliant evaluation method, we developed physical synthetic fingerprint targets, also called phantoms. These phantoms aim to simulate the relevant physical and geometric properties of human fingers, allowing for controlled testing against a known ground truth. However, creating realistic phantoms presents its own set of challenges [6, 62]. This section explores these challenges, details various manufacturing methodologies for phantom creation based on [219], presents the evaluation of their properties and performance, and discusses the implications for fingerprint sensor standardization and evaluation.

3.6.1 Related Work

The creation of synthetic fingerprints, both digital and physical, became more and more important to overcome limitations associated with real fingerprint data. Research spans generating realistic 2D images to fabricating physical 3D targets for sensor evaluation.

3.6.1.1 Synthetic Fingerprint Image Generation

The scarcity of large, diverse, and publicly available fingerprint datasets motivated the biometric community to start generating synthetic fingerprint images algorithmically. Early work by [33] introduced SFinGe, a method that models fingerprint ridge patterns based on mathematical principles, and additionally considers various factors that transform a raw ridge pattern into a realistic fingerprint. This tool enables the creation of large databases for algorithm training and evaluation.

More recently, deep learning techniques, particularly Generative Adversarial Networks (GANs), have entered the field. PrintsGAN by [63] demonstrated the ability to generate realistic fingerprint patches and full images with associated multiple impressions, validated through crowd-sourced realism evaluation. SynFi, proposed by [203], utilized a combination of GANs and Super-Resolution to achieve high-fidelity synthetic prints, resulting in samples which are computational indistinguishable from real fingerprints. The Clarkson Fingerprint Generator also employs

¹<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2016:119:FULL>

GANs with a multi-resolution progressive growth approach [10]. Other approaches leverage variations like pix2pix networks trained on minutiae patterns [155] or focus on specific aspects like generating Level 3 features [259]. Addressing the specific need for contactless fingerprint data, Jannis et al. developed SynCoLFinGer [191], which extends SFinGe by modeling contactless acquisition characteristics, including environmental influences and subject properties. However, these methods focus on generating 2D image data, lacking the possibility of physical sensor interaction testing.

3.6.1.2 Physical Fingerprint Phantom Generation

The creation of physical fingerprint phantoms is important for standardized sensor calibration and evaluation [6]. This area is related with the field of Presentation Attack Detection (PAD), where researchers create physical replicas (Presentation Attack Instruments - PAIs) to test system vulnerabilities [79, 160]. Materials commonly used for PAIs, such as gelatin, silicone, latex, and glues [79, 240], are also candidates for phantom creation due to their ability to mimic skin properties. However, the goal differs: phantoms require high fidelity to a known ground truth and long-term stability for repeatable measurements, whereas PAIs aim primarily to deceive the sensor.

Several manufacturing techniques have been explored for creating structures with fingerprint-like details. 3D printing offers high design flexibility. Early work by [6] involved projecting 2D patterns onto generic 3D finger shapes printed using materials mimicking skin hardness and elasticity. [62] further developed universal wearable 3D targets suitable for optical and capacitive sensors. While flexible, achieving high resolution for fine ridge details and avoiding printing artifacts remain challenging depending on the printer technology (e.g., SLA, MSLA, PolyJet) [6, 219]. Laser engraving enables precise ablation of patterns onto various materials. It has been effectively used in PAD research to create textured surfaces [122] and, as explored in [219], can be applied directly to silicone or elastomer plates or used to create high-fidelity molds (e.g., in aluminum). The precision allows for fine detail replication, but direct engraving on flat plates doesn't inherently produce a 3D finger shape, and mold creation adds complexity. Molding and casting are often used in conjunction with other techniques. Masters created via CNC machining or 3D printing can be used to create negative molds, which are then cast with materials like silicone or gelatin [219, 226]. Schultz et al. [226] demonstrated a method using solvent-softened polycarbonate to capture real fingerprint details, followed by casting with PDMS, offering high fidelity but limited scalability and lack of a controllable synthetic ground truth. The choice of casting material significantly impacts elasticity, durability, and shrinkage [219]. CNC machining provides high precision for creating master targets, typically from metals like aluminum, as detailed in [219]. While capable of producing very accurate and durable masters for subsequent molding [225], it is generally the most expensive and time-consuming method, potentially limited by tool size for extremely fine fingerprint features. Despite these advancements, creating standardized, durable, and realistic physical phantoms that accurately replicate both the 3D geometry and material properties of diverse human fingers across different manufacturing methods remains an active area of research, motivating the comprehensive comparison presented in [219] and detailed in this section.

3.6.2 Challenges of Synthetic Fingerprints and Fingers

One major challenge is replicating the three-dimensional ridge structure of a fingerprint and assuring the characteristic material properties match those of human skin at the same time [177, 256]. Human skin has a unique elasticity, hardness, and surface texture [2, 51, 229, 279], properties

that critically influence how a finger interacts with different types of sensors, particularly during dynamic processes like rolling or slap acquisitions. Identifying materials that not only mimic these biomechanical properties accurately but are also compatible with high-resolution fabrication processes remains difficult [46].

Beyond material science, the manufacturing process needs to be precise and repeatable to ensure that multiple phantoms are consistent and reliable evaluation tools [39]. Each fabrication technique, whether it be laser engraving, 3D printing, or CNC machining, presents a unique profile of advantages and disadvantages concerning accuracy, cost, scalability, and the potential introduction of process-specific artifacts [111, 219]. For example, certain 3D printing methods might compromise fine ridge detail or surface finish, while molding techniques could potentially introduce defects like shrinkage or air bubbles, affecting the phantom's fidelity.

The design of the phantom must also account for its intended use across various sensor modalities, including optical, capacitive, ultrasonic, and thermal sensors, as well as different acquisition types like rolled prints, slap prints, or contactless captures. A phantom optimized for rolled acquisition on an optical scanner might prove unsuitable for slap capture on a capacitive sensor due to differing requirements in material conductivity, or surface characteristics [219].

Finally, the long-term stability and durability of the phantom materials is important for practical application as standards. The chosen materials must withstand repeated use and resist degradation processes such as hardening, shrinkage, or surface wear over time to ensure consistent and reliable evaluation results [219]. Materials like gelatin, while initially offering appealing skin-like properties, are known to be susceptible to dehydration and deformation, limiting their lifespan as reliable phantoms [46].

3.6.3 Manufacturing Methodologies

To systematically investigate potential solutions for phantom creation, several distinct manufacturing methodologies were implemented and evaluated [219]. These methods can be broadly categorized into direct engraving techniques and processes involving the fabrication of intermediate molds.

3.6.3.1 Laser Engraving

Laser engraving technology, known for its ability to precisely ablate materials and create fine patterns [30, 183], was explored in multiple configurations. One approach involved the direct engraving of synthetic fingerprint patterns onto flat substrates composed of specialized materials. Specifically, plates of Laserline EPDM elastomer, a material typically used in security printing, were engraved at a resolution of 5080 dpi with a target penetration depth of 100 μm [219, Fig. 2]. Two variants (CSX 5K, 0.95 mm thick, 78 Shore A; CSC, 1.42 mm thick, 76.5 Shore A) were tested. For testing, these flat plates were wrapped around cylindrical wooden holders (20 mm diameter), though their inherent stiffness, especially the thicker variant, posed challenges for achieving good contact during simulated rolling. Additionally, flat plates of silicone, including commercially available tattoo training skin (Gospire, 1.8 mm thick) and an in-house formulation based on Dragon Skin 10 Fast (1.8 mm thick), were directly engraved [219, Fig. 3]. These silicone plates provided more flexibility than the elastomer but remained fundamentally flat targets. Examples of these directly engraved phantoms are shown in Figure 3.27. The cost for direct laser engraving was estimated at approximately 200€ per phantom [219].

A novel alternative involved using laser engraving to create an intermediate mold with a 3D shape. The negative (inverse) of a fingerprint pattern was engraved into the inner concave surface of an aluminum half-pipe (EN-AW 6063 alloy, 23 mm outer diameter, 20 mm inner

diameter) [219, Fig. 4a]. This mold could subsequently be filled with a casting material, such as silicone, to produce a phantom with a more realistic cylindrical finger geometry. To complete the finger shape, a 3D-printed plug was designed to form the rounded fingertip during the casting process [219, Fig. 4b]. This indirect method, while enabling the creation of 3D shapes with high-resolution patterns, adds process steps (mold creation, casting, curing) and risks introducing artifacts if the underlying aluminum half-pipe geometry is inconsistent or errors in any of the processing steps are made. The cost was significantly higher, estimated at 8000€ for initial setup and 1000€ per subsequent mold/phantom [219].

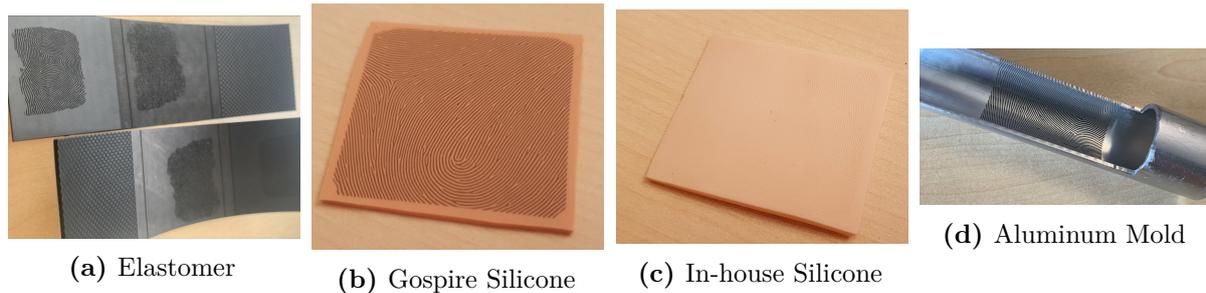


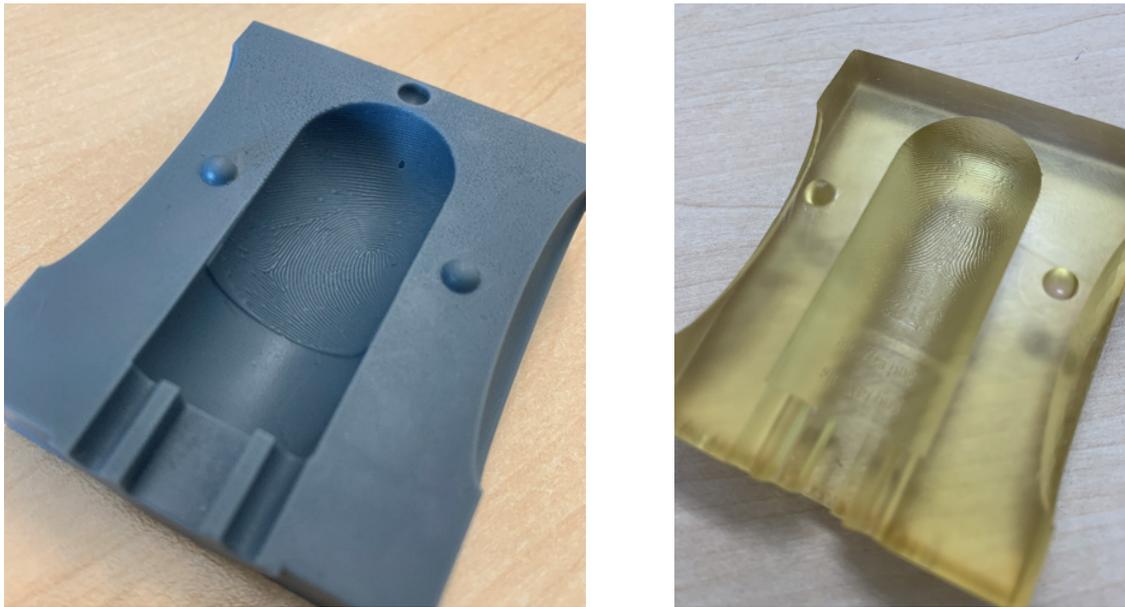
Fig. 3.27: Examples of Laser-Engraved Phantoms. (a) Direct engraving on elastomer. (b) Direct engraving on in-house silicone. (c) Direct engraving on Gopspire silicone. (d) Aluminum half-pipe mold. Images taken from [219, Fig. 10].

3.6.3.2 3D Printing of Resin Molds

For the 3D printing methodology, negative molds of fingerprints were fabricated using stereolithography (SLA and MSLA) techniques [239]. This approach offers significant design flexibility [49]. The first step of this approach was the development of a process to project 2D synthetic fingerprint patterns onto a generic 3D finger model. For our use-case, we chose a cylinder shape capped with a spherical tip. For this, we utilized the ABF++ algorithm [228] within the Blender modeling software. This enabled the creation of molds that would produce phantoms with both realistic overall shape and arbitrarily defined, ground-truth fingerprint patterns. The molds were designed as two-part structures to facilitate easier removal of the cast phantom and potentially allow for different fingerprint patterns on opposing sides of the same phantom [219, Fig. 5]. To ensure smooth demolding and capture fine details, the fingerprint texture was upscaled and the depth projection of the fingerprint texture was smoothed via blurring before projection onto the 3D model. This rounds the fingerprint ridges, leading to easier removal of the phantom from the mold [219].

Two stereolithography-based printing techniques were compared: an in-house, consumer-grade Masked Stereolithography (MSLA) printer (Elegoo Saturn 2-8K, costing approximately 400€ for setup and under 5€ for materials per mold) and a professional SLA printing service (Alpine3D GmbH, costing approximately 3000€ for setup and 50€ per mold). Initial trials with the MSLA printer revealed minor artifacts in the molds, such as small holes and inconsistencies in ridge depth near the edges (Figure 3.28a). Design refinements, including improved mold half connections and ensuring uniform ridge depth, were implemented for the professionally printed SLA molds (Figure 3.28b), which were largely artifact-free and additionally transparent [219], to avoid reactions of the color pigments in the mold with the casting material.

The resulting molds were then filled with casting materials. Kryolan Gelafix, a gelatinous substance used for special effects, was tested for its skin-like appearance (Figure 3.29b). It required heating to liquefy before casting and cured relatively quickly. Silicone, using Dragon Skin



(a) Elegoo Saturn 2-8K

(b) Alpine3D GmbH

Fig. 3.28: Examples of 3D printed mold. (a) In-house printer, (b) printing service.

variants (10 Fast or 20), required mixing two components, degassing in a vacuum chamber, and a longer curing time (4-24 hours), but offered much greater stability and durability (Figure 3.29a). Table 3.16 provides a comparison of relevant Dragon Skin properties against those of human skin, showing Dragon Skin 20 closely matches the lower end of human skin hardness (Shore A 20). Material costs for filling were low (under 10€ for Gelafix, under 5€ for Silicone per phantom) [219].

Tab. 3.16: Dragon Skin vs. Human Skin Properties. Adapted from [219, Table 1].

	Dragon Skin 10 Fast [58]	Dragon Skin 20 [58]	Human Skin
Density [kg/m ³]	1070	1080	1250 [130]
Pot Life [min.]	8	25	-
Cure Time	75 min	4 h	-
Hardness [Shore A]	10	20	20–41 [60, 67]
Shrinkage [m/m]	<0.001	<0.001	-

3.6.3.3 CNC Machining

Computer Numerical Control (CNC) machining was used to produce highly precise master targets directly from aluminum round bars (6061 alloy) [225]. A 5-axis milling machine created master cylinders featuring various surface patterns, including checkerboards, vertical/horizontal Ronchi lines, concentric rings, and a synthetic fingerprint pattern, to test machining precision and consistency [219, Fig. 7]. Micro-tools were selected based on the smallest feature size (e.g., 0.2 mm for the fingerprint). Surface roughness could be controlled, with options for sandblasting (achieving Rq of 2.5 μm) versus leaving the milled finish (Rq 1.2 μm) (Figure 3.30a and Figure 3.30c). These high-precision metal masters then served as the basis for a two-stage silicone molding process: first, a flexible negative mold was cast from the master using Dragon Skin 10

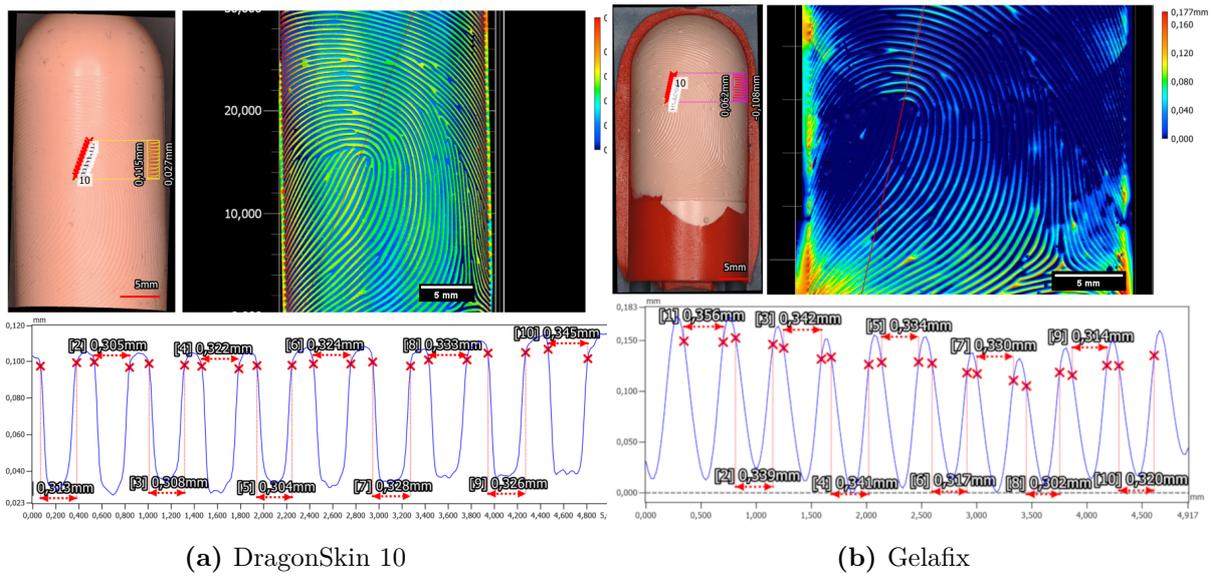


Fig. 3.29: Width measurements of casted phantoms made from 3D printed mold. (a) DragonSkin 10 A silicone, (b) Gelafix.

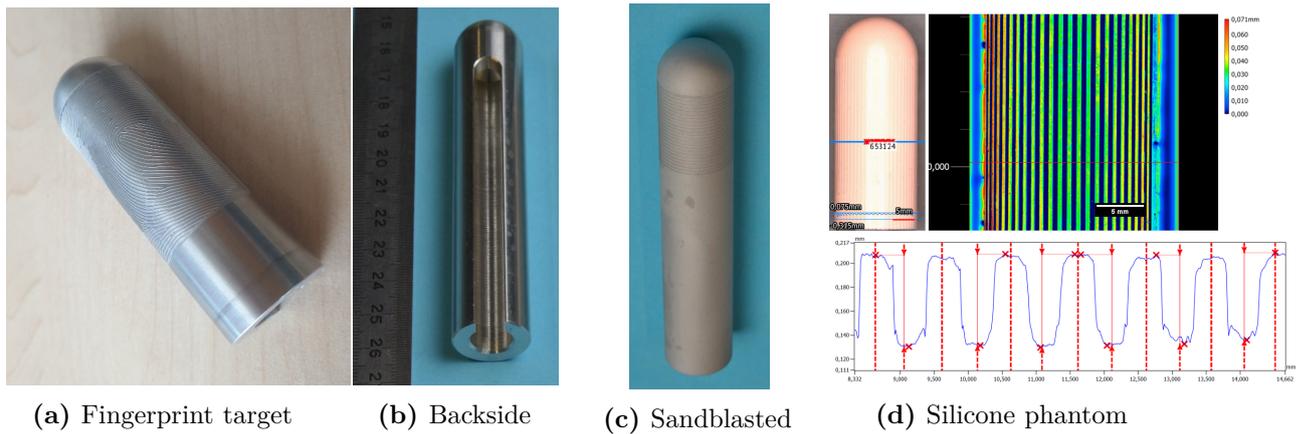


Fig. 3.30: Examples of CNC-Machined masters with the fingerprint pattern 3.30a, the backside showing the slit for inserting the target holder 3.30b, the sandblasted target with increased surface roughness 3.30c and measurements of the phantom with Ronchi lines after the two-stage silicone molding.

Fast (after applying a release agent), and second, this negative mold was used to cast the final positive phantom using the same silicone (Figure 3.30d). While capable of yielding extremely accurate master forms, CNC machining is inherently costly (approximately 1000€ per subsequent master after initial setup) and labor-intensive, particularly for the initial setup, programming, and the two-stage casting process. It also encountered limitations, as the intended checkerboard pattern could not be fully realized due to the challenges of milling such fine features (0.1 mm) without tool breakage [219].

3.6.4 Experimental Evaluation and Results

The manufactured phantoms were evaluated to assess their quality, fidelity, consistency, and suitability for different sensor types and recording modalities [219]. Measurements involved profilometry (Keyence VR-5200) to determine ridge/valley dimensions and surface characteristics, microscopy (Vision Engineering Lynx) for visual inspection, and fingerprint acquisition using a Greenbit DactyScan 84c optical scanner. Subsequent analysis included NFIQ 2 quality assessment and template comparison using commercial (Idkit) and open-source (NBIS, FinSource) ABIS to quantify utility, fidelity and variability.

3.6.4.1 Generation Results and Material Properties

Profilometry analysis confirmed the successful transfer of microstructures but revealed important differences between the methods [219]. Direct laser engraving on elastomer showed significant deviation from the target depth (mean $115.6 \mu\text{m} \pm 1.2 \mu\text{m}$ measured vs. $100 \mu\text{m}$ target) and exhibited unnaturally large valley width variations (mean $304.0 \mu\text{m} \pm 19.1 \mu\text{m}$) along with pixelated edges due to the initial 500 dpi ground truth resolution. Direct engraving on silicone resulted in more consistent ridge shapes but produced shallower valley depths (Gospire: $86.8 \mu\text{m} \pm 2.7 \mu\text{m}$; In-house: $63.8 \mu\text{m} \pm 2.0 \mu\text{m}$) compared to the target, the Gospire silicone also showed surface burning effects from the laser. Phantoms cast from the laser-engraved aluminum half-pipe mold achieved accurate ridge depths ($89.9 \mu\text{m} \pm 7.0 \mu\text{m}$) but inherited visible line artifacts caused by offsets during the mold engraving process.

Phantoms produced using 3D-printed molds showed characteristics dependent on the printer and filling material. Gelafix fillings initially provided good ridge definition (mean depth $146.0 \mu\text{m} \pm 7.6 \mu\text{m}$ for ES2 mold) but suffered severe degradation: diameter shrinkage exceeded 17% after 6 months, accompanied by significant hardening and a reduction in ridge depth to $118.1 \mu\text{m} \pm 5.9 \mu\text{m}$. Silicone fillings proved highly stable (<1% shrinkage after 6 months). Phantoms from the professionally printed mold exhibited minimal artifacts and were designed with a more realistic, shallower ridge depths (mean $38 \mu\text{m} \pm 5 \mu\text{m}$) compared to those from the in-house ES2 mold (mean $128 \mu\text{m} \pm 9 \mu\text{m}$). The in-house ES2 mold suffered from small holes and edge inconsistencies as a result of the printing process. Both molds showed the pointed ridge profiles typical of the printing/casting process. Phantoms derived from CNC-machined aluminum masters accurately reproduced the target ridge depth ($75.2 \mu\text{m} \pm 2.0 \mu\text{m}$) and demonstrated consistent ridge and valley widths (valley width $320.8 \mu\text{m} \pm 10.6 \mu\text{m}$). Table 3.17 provides a consolidated comparison of the measured dimensions and observed artifacts for each method.

Methodology	Price	Manual Labor	Artifacts	Valley Width [μm]	Ridge Depth [μm]
CNC Alum + Silicone	setup + $\approx 1000\text{€}$	High	Minimal	530 ± 13	86 ± 4
Laser Alum HP + Silicone	$\approx 8000\text{€} + 1000\text{€}$	Medium	Offset of ridge lines	364 ± 20	90 ± 7
Laser Elastomer	$\approx 200\text{€}$	Low	Varying ridge thickness	304 ± 19	115 ± 1
Laser Silicone Gospire	$\approx 200\text{€}$	Low	Minimal	340 ± 13	87 ± 3
Laser Silicone In-house	$\approx 200\text{€}$	Medium	Small offset of ridge lines	303 ± 8	64 ± 2
Print ES2 + Gelafix	$\approx 400\text{€} + <10\text{€}$	Medium	Air bubbles + small holes in mold	330 ± 18	146 ± 8
Print ES2 + Silicone	$\approx 400\text{€} + <5\text{€}$	Medium	Small holes in mold	370 ± 37	128 ± 9
Print ALP + Silicone	$\approx 3000\text{€} + 50\text{€}$	Medium	Minimal	265 ± 18	38 ± 5

Tab. 3.17: Overview of Manufacturing Methodologies. Adapted from [219, Table 3].

3.6.4.2 End-to-End Fidelity and Intra-Class Variability

The utility of the phantoms was investigated by assessing the quality and consistency of the fingerprints recorded from them. NFIQ 2 scores [246] were calculated for both the initial 2D synthetic fingerprints and the rolled impressions acquired from the corresponding phantoms using the Greenbit scanner [219, Table 4]. It is important to note that the usage of NFIQ 2 for rolled impressions is not in its intended scope, since it was developed for flat recordings, however, it is still the best freely available tool to estimate the utility of the rolled fingerprint samples. Most 2D synthetic images scored in the 44-59 range, considered medium to good quality [44]. The rolled recordings generally maintained similar scores (40-61 range), indicating that the basic ridge structure quality was preserved during manufacturing and recording for most methods. However, a notable drop in score was observed for phantoms made using the in-house ES2 3D printer (average scores of 36 for Gelafix, 37 for Silicone), suggesting artifacts introduced by this specific printing process negatively impacted sample utility.

End-to-end fidelity, important for understanding how accurately a phantom represents its intended digital counterpart in a recognition scenario, was assessed by comparing templates generated from the 2D synthetic images against templates from their corresponding rolled phantom recordings. In order to archive a fair comparison and eliminate scale variations introduced in the design process or with the age related shrinking, a preprocessing ablation study was conducted. Here, we investigated the scaling values of the recorded images, since scaling issues can drastically impact ABIS comparison performance.

For this ablation study, we scaled the rolled phantom recording in the x-direction and also in the y-direction and calculated template comparison scores with the corresponding synthetic fingerprint image for each scaling value.

We selected the current scaling values using a grid search algorithm. For this, we sampled 50 points for scaling along the x-direction and 50 points for scaling along the y-direction, resulting in 2500 possible scaling combinations per image.

The starting points of the grid and the end points of the grid were chosen such that the template comparison score maxima sat roughly in the middle of the grid, ensuring we found the right optima. We then conducted this search over all samples recorded from this phantom.

Next, we selected the 10 best scaling values per phantom and used their weighted average over all recordings of all phantoms for a given synthetic fingerprint and manufacturing method. There, the template comparison scores served as weights in the averaging process. The result of this process were the final scaling values used in the downstream tasks, shown in Table 3.18. Also, the quality of the identified scaling value could be measured by the variability of this score, given by the standard deviation.

To calculate the template comparison scores, we used the state-of-the-art commercially available ABIS IDkit from Innovatrics [103].

We applied the found scaling values to the recordings to ensure optimal performance in the following end-to-end fidelity study. For this study, we utilized three distinct fingerprint identification systems: the commercial Innovatrics Idkit, the NIST Biometric Image Software (NBIS) suite (Mindtct for minutiae extraction and Bozorth3 for minutiae comparison), and a combined open-source approach using FingerNet [250] for minutiae extraction and SourceAFIS [258] for minutiae comparison (referred to as FinSource) [219, Table 6]. High template comparison scores indicate high end-to-end fidelity.

The results depicted in Table 3.19 show significant performance differences. Phantoms created via direct laser engraving on silicone substrates (both Gospire and in-house produced) demonstrated high fidelity, with average scores across the three systems being 771 and 689, respectively. They achieved perfect scores (1000) with the Idkit comparison algorithm. Similarly, phantoms

Methodology	x	y	σ_{score}
CNC Alum	0.57	0.58	0.0
Laser Alum HP Silicone	0.91	1.03	0.0
Laser Elastomer	0.99	1.02	8.3
Laser Silicone Gospire	0.93	1.04	0.0
Laser Silicone Inouse	0.86	1.00	0.0
Print ES2 Gelafix	0.79	0.83	13.9
Print ES2 Silicone	0.68	0.80	13.7
Print Alpine Silicone	0.64	0.88	20.7

Tab. 3.18: Optimal scaling values for the recorded rolled fingerprint samples in the horizontal direction (x) and the vertical direction (y). Adapted from [219].

cast from the laser-engraved aluminum half-pipe mold (Laser Alum HP + Silicone, average 593) and those derived from CNC-machined aluminum masters (CNC Alum + Silicone, average 481) also showed high fidelity, again reaching maximum scores with Idkit. These methods, involving precise master creation, effectively transferred the fingerprint pattern.

In contrast, phantoms originating from 3D-printed resin molds generally yielded lower fidelity scores. However, it should be noted that the phantoms originating from the 3D-printed resin models are the only phantoms where parts of the fingerprint pattern was projected onto the rounded fingertip area. This introduces additional challenges in the template comparison stages. First, there is a reduced exposure of the fingerprint surface area to the sensor, since the front-most part of the tip cannot be rolled simultaneously with the flatter parts of the fingertip. Also, those parts require a more complicated mosaicking process, potentially introducing artifacts, leading to misplacement of the minutiae. Finally, those rounded areas introduce geometric distortions that have to be accounted for.

The professionally printed Alpine3D (ALP) molds (average 388 for silicone fillings across different fingerprint types) outperformed those from the in-house ES2 molds (average 275 for Gelafix, 295 for Silicone). This highlights the impact of printer precision and mold quality on the final phantom fidelity. Gelafix as a filling material consistently underperformed compared to silicone. The laser-engraved elastomer phantom exhibited very poor fidelity (average score 93), primarily due to its stiffness, which made it difficult to capture a complete and undeformed impression.

It is also noteworthy that comparison algorithm choice influenced the perceived fidelity for certain phantoms. For example, the Laser Alum HP + Silicone phantom achieved a perfect score with Idkit but a significantly lower score of 104 with NBIS, likely because the line artifacts present in this phantom type were handled differently by the feature extraction and comparison algorithms of the NBIS system. Table 3.19 summarizes these average fidelity scores.

Intra-class variability evaluates the consistency of recordings obtained from the same phantom type across multiple acquisitions. High comparison scores between different recordings of the same phantom indicate good repeatability and low variability. The same three comparison systems were used for this assessment [219, Table 7, 8].

The results largely mirrored the fidelity findings. Phantoms created using laser-engraved aluminum molds (Laser Alum HP + Silicone, average variability score 624) and CNC-machined masters (CNC Alum + Silicone, average 569) exhibited high consistency, indicating that these methods produce phantoms that yield repeatable recordings. 3D-printed phantoms demonstrated more varied consistency. Those from the refined Alpine3D (ALP) process with silicone filling

Methodology	Avg	Idkit	Nbis	FinSource
CNC Alum	481	1000	158	285
Laser Alum HP Silicone	593	1000	104	674
Laser Elastomer ¹	93	112	44	123
Laser Silicone Gospire	771	1000	186	1126
Laser Silicone In-house	689	1000	158	908
Print ES2 Gelafix ²	275	570	54	201
Print ES2 Silicone ²	295	615	70	200
Print ALP Silicone ²	388	829	92	243
-Arch	421	904	53	304
-Tented Arch	373	803	71	245
-Left Loop	465	977	130	289
-Right Loop	430	906	137	246
-Whorl	252	553	72	132

¹ The elastomer phantom was not rolled but placed on the acquisition area, and the wooden cylindrical finger model was rolled with pressure over the elastomer phantom to simulate the rolling motion.

² The phantoms created from the 3D-printed molds are the only phantoms with a fingerprint pattern that is partially placed on the rounded fingertip area.

Tab. 3.19: End-to-end fidelity for phantoms measured via template comparison scores of Idkit (Idkit), Mindtct + Bozorth3 (Nbis), and FingerNet + SourceAFIS (FinSource) comparison algorithms and their average (Avg). The finger types (-Arch, -Tented Arch, ...) are sub-categories for the phantoms printed using the SLA printer of Alpine3D GmbH.

showed reasonable consistency (average 538 across fingerprint types). However, phantoms from the in-house ES2 printer, particularly with Gelafix filling (average 363), had considerably lower consistency, suggesting that recordings from these phantoms were less reliable. The laser-engraved elastomer phantom also showed poor repeatability (average 537), aligning with the difficulties in its handling.

Table 3.20 (derived from Table 7 of [219]) summarizes the average intra-class variability for several key methodologies. It is important to note that the direct laser-engraved silicone phantoms (Gospire and In-house) are not shown because they are of a simpler 2D form factor.

When comparing different high-performing fabrication methods that produced the same underlying fingerprint pattern [219, Table 8], the laser-based methods (direct silicone, aluminum mold) and the CNC-based method grouped together, exhibiting high interoperability in terms of consistency. The 3D-printed methods formed a distinct cluster with generally lower, though still useful for some types, consistency scores. This highlights that while different precise methods can create consistent phantoms, the nuances of 3D printing can lead to greater variability in the final recorded impressions.

3.6.5 Discussion

The evaluation presented above and in more detail in [219] shows that selecting an optimal approach requires consideration of manufacturing methodology, material properties, cost, labor,

Tab. 3.20: Intra-Class Variability Scores. Adapted from [219, Table 7].

Methodology	Average Variability Score
CNC Alum + Silicone	569
Laser Alum HP + Silicone	624
Laser Elastomer	537
Print ES2 + Gelafix	363
Print ES2 + Silicone	514
Print ALP + Silicone (Avg over types)	538

and the resulting phantom’s characteristics and performance in the context of the intended application – be it rolled, slap, or contactless sensor evaluation – and the desired level of fidelity versus resource constraints.

3.6.5.1 Methodology Comparison

Direct laser engraving onto flat silicone plates stands out for its balance of cost-effectiveness (approx. €200 per phantom), low labor requirement, minimal artifact generation (especially with Gospire silicone), and high end-to-end fidelity and recording consistency. Its main drawback is the inherent flatness, requiring mounting for realistic rolled print simulation. In contrast, creating a 3D phantom via casting silicone into a laser-engraved aluminum half-pipe mold achieves comparable high fidelity and consistency while providing the desired 3D geometry. However, this comes at a significantly higher cost (setup >€8k, unit >€1k), involves more labor (two casting steps), and is susceptible to line artifacts originating from the mold engraving process.

CNC machining of aluminum masters, followed by a two-stage silicone casting process, offers the best precision for the master target, translating to good fidelity and consistency in the final phantom. This method allows for fine control over surface properties like roughness. Yet, it represents the most expensive (approx. €1k per subsequent master) and labor-intensive approach, and faced limitations with highly detailed patterns like the checkerboard, where tool breakage occurred.

3D printing of resin molds provides the greatest design flexibility and is the most economical option, particularly when using in-house MSLA printers (setup €400, material <€5 per mold). Professional SLA printing offers higher precision and artifact-free molds but at a higher cost (setup €3k, unit €50). While offering versatility in shape and pattern integration, the resulting phantoms generally exhibited lower fidelity and consistency compared to the top laser-engraved or CNC-derived methods. The quality is highly contingent on careful mold design (e.g., avoiding liquid traps, ensuring smooth edges), printer resolution, and the choice of printing service (ALP outperformed ES2).

3.6.5.2 Material Suitability

Silicone, specifically the Dragon Skin platinum-cure variants tested, proved to be the superior material for creating durable and realistic phantoms. Its advantages include extremely low shrinkage (<0.1%), long-term stability (minimal hardening over months), and elasticity (Shore A 10-20) that reasonably approximates the lower range of human skin hardness, facilitating realistic deformation during rolling. Kryolan Gelafix, while visually similar to skin initially and potentially useful for specific PAD tests, is unsuitable for standardized, reusable phantoms due to its rapid and severe shrinkage (>17% in 6 months) and hardening. The Laserline EPDM

elastomer, designed for printing plates, was found to be too stiff for effective interaction with contact-based fingerprint sensors, especially for simulating rolled acquisition.

3.6.5.3 Application-Specific Recommendations

For simulating rolled fingerprint recordings, where capturing the deformation of a 3D shape is critical, methods yielding elastic 3D phantoms are necessary. CNC machining or laser engraving of aluminum molds, followed by silicone casting, provide the highest fidelity and consistency but are resource-intensive. 3D printing of high-quality (e.g., professional SLA) molds filled with silicone offers a more accessible and flexible alternative, potentially sacrificing some fidelity for cost savings and design freedom. They also have the benefit of allowing arbitrary finger geometries that deviate from a cylindrical approximation, while CNC machining or laser engraving are constrained to cylindrical shapes.

For simulating slap fingerprint recordings, the requirements are different. They primarily need a flat, detailed surface. Direct laser engraving onto flat silicone plates (like Gopire or a stable in-house formulation) is highly recommended due to its simplicity, high fidelity, low artifact rate, and appropriate geometry.

For evaluating contactless fingerprint sensors, where 3D shape and visual appearance (including color) are often important, 3D-printed molds filled with appropriately pigmented silicone are the most advantageous. This approach allows the creation of phantoms with realistic finger geometry and controlled surface characteristics.

3.6.5.4 Impact on Research Questions

This investigation into phantom manufacturing directly addresses **RQ6 (Can we produce synthetic 3D fingers and how do they perform compared to real fingers? What manufacturing techniques can be used?)**. The results definitively show that yes, synthetic 3D fingers can be produced using laser engraving of molds, 3D printing of molds, or CNC machining of masters, followed by casting, primarily with silicone. Direct engraving is suitable for flat (slap) simulations. Performance, measured by fidelity and consistency, can be very high, approaching the limits of current comparison systems for the best methods (CNC/Laser molds). While material properties like elasticity approximate human skin, perfect replication remains a challenge. The phantoms successfully replicate known ridge patterns, providing a ground truth missing in human testing. Laser engraving, 3D printing, and CNC machining (all with casting) are confirmed as viable techniques, with specific trade-offs detailed above.

This work also establishes the necessary tools and understanding to advance other research questions. The created phantoms serve as standardized targets. They provide a fixed geometry and pattern ground truth essential for rigorously evaluating pose correction algorithms (**RQ2**) and core/landmark detectors (**RQ3**) without facing human variability. They enable controlled studies on the influence of physical characteristics and acquisition conditions on quality metrics (**RQ4**). Furthermore, these physical phantoms can generate realistic image sequences (e.g., through controlled rolling or sweeping) necessary for testing and validating fingerprint mosaicking algorithms and artifact detectors (**RQ5**). The ability to reliably create physical instances of digitally defined fingerprints provides a powerful methodology for validating algorithms and sensors against known references.

3.6.5.5 Limitations

While this study provides a broad comparison, some limitations persist. The exploration of materials, while including representative examples, could be expanded, particularly investigating

multi-layer composites to better mimic skin structure. The long-term stability assessment, though demonstrating silicone's superiority over Gelafix, was limited in duration to under a year. Fully replicating the complex, non-homogeneous biomechanics of human skin remains a challenge. The performance evaluation relied predominantly on a single high-quality optical contact-based sensor. This means that results might differ for capacitive, ultrasonic, or various contactless sensor technologies. Cost estimations provided are indicative and depend heavily on specific vendors and production volumes. Lastly, while phantoms reduce variability compared to human subjects, manual handling during recording still introduces minor inconsistencies. Therefore, robotic acquisition protocols would be required to ensure full repeatability.

3.7 Privacy Preserving Fingerprint Comparison

The increasing reliance on biometric data, such as fingerprints, for identification and authentication presents a privacy paradox. While offering unique, permanent, and difficult-to-forge characteristics that are advantageous over traditional methods like passwords or tokens, biometric data, once compromised, cannot be revoked or changed [152]. This immutability, coupled with the inherent personal nature of biometric identifiers, raises concerns regarding data collection, storage, potential misuse, and the risk of mass surveillance [233]. Standard practice often involves local processing and storage of fingerprint data on dedicated secure hardware (e.g., secure elements in smartphones) to mitigate these risks. However, this localized approach inherently limits the application of fingerprint biometrics to primarily local authentication scenarios. This stops more advanced use cases that require remote comparison or data pooling [152].

One use-case requiring remote comparison is for example privacy preserving border security at international checkpoints, such as airports. The envisioned scenario involves comparing a traveler's scanned fingerprint not only against their passport data but also against national and international No-Fly Lists or watchlists collaboratively pooled by multiple law enforcement agencies. Such a multi-national system requires privacy guarantees: fingerprints of regular travelers must remain confidential and should never be disclosed or stored centrally, while the contents of the sensitive watchlists, potentially from different jurisdictions, also need to be protected during the comparison process. Only in the event of a confirmed match should relevant, limited information be revealed to initiate further action. Traditional approaches, where raw biometric data is transmitted or compared in the clear, are inadequate for such sensitive, multi-party applications.

Multiparty Computation (MPC) offers a cryptographic solution to these challenges [66]. MPC enables a set of parties to jointly compute a function over their private inputs without revealing these inputs to each other or any other party involved in the computation. In the context of fingerprint comparison, MPC allows for the comparison of an encrypted (e.g., secret-shared) fingerprint template against a database of similarly encrypted templates, with only the match result (e.g., a match/no-match decision or a similarity score) being revealed. This approach can preserve the privacy of both the individual providing the fingerprint and the sensitive data within the reference database. However, MPC introduces a significant computational overhead, which has historically been a major hindrance to its practical application for complex algorithms like minutiae-based fingerprint comparison [147]. While simpler feature vector comparisons (e.g., Euclidean distance) are relatively efficient in MPC [12], state-of-the-art fingerprint recognition predominantly relies on more intricate minutiae-based comparison algorithm due to their superior accuracy and explainability [151]. Recent advancements in MPC protocols and frameworks have, however, started to make real-world deployment feasible even for these more demanding tasks. This section details research focused on implementing and optimizing a minutiae-based

fingerprint comparison algorithm using MPC, demonstrating its practical feasibility for real-time applications while preserving privacy [152].

3.7.1 Related Works

The challenge of performing fingerprint comparison while preserving the privacy of the biometric data has been an active area of research for several years. The inherent sensitivity of biometric templates, which are irreplaceable if compromised, requires solutions that go beyond traditional data encryption, especially when the comparison is performed by untrusted parties or on untrusted platforms [108]. Early efforts in biometric template protection focused on creating cancelable templates [197] or biometric cryptosystems [199, 257].

Cancelable biometrics aim to transform the original biometric template into a new, intentionally distorted version in a non-invertible way [196]. If a transformed template is compromised, it can be cancelled and a new one re-issued by applying a different transformation. While offering revocability, challenges remain in ensuring that the transformed templates maintain high discriminability and are robust against various attacks [198]. Several approaches have been proposed for fingerprint minutiae, often involving transformations in the minutiae space, such as random projections or permutations.

Biometric cryptosystems, including fuzzy extractors and secure sketches, aim to bind a cryptographic key with the biometric data or derive a key from it [27, 56]. For fingerprints, these methods often struggle with the intra-user variability (distortions between different acquisitions of the same finger) and the lack of a fixed-length representation for minutiae sets. Successfully extracting a stable key from noisy biometric data without leaking significant information about the template itself is a primary challenge [199].

More directly related to the MPC approach discussed in this thesis are methods that perform comparison in the encrypted domain. Homomorphic Encryption (HE) schemes allow computations to be performed directly on encrypted data without decrypting it first. Several works have explored applying HE to biometric comparison, including fingerprints [12, 274]. However, minutiae-based fingerprint comparison involves complex operations (e.g., geometric comparisons, alignment, scoring functions) that are often inefficient or difficult to implement with current HE schemes, particularly fully homomorphic encryption (FHE), which supports arbitrary computations but comes with significant performance overhead [25]. Partial HE schemes (e.g., Paillier) have been used for simpler distance calculations or components of the comparison process, but a full, efficient minutiae comparison algorithm remains a challenge.

Secure Multiparty Computation (MPC) has emerged as a more flexible alternative for complex biometric comparison. Unlike HE where one party typically encrypts data for another to compute on, MPC allows multiple parties to jointly compute a function on their private inputs. Early work by [123] demonstrated secure two-party computation for edit distance, relevant to string-based biometric representations. For fingerprints, research has focused on securely implementing various components of the minutiae comparison pipeline. For instance, [24] explored secure computation of fingerprint alignment and comparison scores. [102] proposed a privacy-preserving identification system using secure sum and comparison protocols for minutiae data. More recent works, like those by [145] and [187], have proposed protocols for secure minutiae-based fingerprint comparison, often breaking down the comparison algorithm into MPC-friendly sub-protocols. These approaches typically involve secure comparisons of minutiae coordinates and orientations, and aggregation of comparison scores. The work by [147] benchmarked different MPC frameworks for cryptographic building blocks relevant to biometric comparison, highlighting the trade-offs between security models and performance. The challenges in adapting complex algorithms like SourceAFIS, with its specific edge-based representation and tree-crawling mechanism, for MPC,

as detailed in [152], involve careful algorithmic re-engineering to balance privacy, accuracy, and computational efficiency. These efforts often involve hybrid approaches where some non-sensitive pre-processing or data structuring is done in plaintext, while the core comparison logic involving sensitive template data is performed within the MPC framework.

3.7.2 Fingerprint Comparison

At the core of this work is a minutiae-based fingerprint comparison algorithm. Minutiae, which are specific points in a fingerprint such as ridge endings or bifurcations, form the primary features used for comparison. The specific algorithm adapted for MPC implementation is SourceAFIS, an open-source, commercially-used comparison algorithm developed by Robert Važan [258]. The SourceAFIS comparison process can be divided into two stages: feature extraction and template comparison.

Feature Extraction in SourceAFIS: During feature extraction, SourceAFIS processes an input fingerprint image to identify and characterize minutiae. Each minutia is stored in a template with its type (ending or bifurcation), its (x, y) position, and its direction (angle). Since these raw minutiae features are not inherently rotation- or translation-invariant, the template generation process also involves computing and storing edges. An edge represents a connection between a reference minutia and one of its neighboring minutiae. For each reference minutia, a fixed number of edges are calculated and described by their length and two angles relative to the reference minutia and the neighbor. These edge parameters are designed to be invariant to rotation and translation of the overall fingerprint, which is needed for stable comparison [152]. This feature extraction phase is typically performed locally (e.g., on the sensor side or a trusted local device) and does not require MPC, as it operates on a single fingerprint image before any comparison is made with other potentially sensitive data. The result of this operation is a fingerprint template.

Template Comparison in SourceAFIS: The template comparison stage aims to determine if two fingerprint templates, referred to as the probe (the newly acquired fingerprint) and the candidate (a template from the database), originate from the same finger. The SourceAFIS algorithm accomplishes this through a three-phase process for each probe-candidate pair: (i) enumerate, (ii) crawl, and (iii) score [152, Algorithm 1].

In the *enumerate* phase, the algorithm seeks potential root pairs of matching edges between the probe and candidate templates. It iterates through the minutiae of the candidate, forming edges to other minutiae. For each candidate edge, its hash value is computed and looked up in a pre-computed edge hash table of the probe template. If a probe edge with a similar hash exists, it is considered a potential match, forming a root pair. This hash-based lookup is an optimization in the original plaintext algorithm to speed up the search for similar edges.

Starting with a root pair, the *crawl* phase attempts to build a spanning tree of mated minutiae. The algorithm iteratively compares the surrounding edges of the current minutiae pair in the probe and candidate templates. If a pair of surrounding edges is deemed sufficiently similar, the minutiae at the ends of these edges are added to the spanning tree, and the process continues from these newly added minutiae. This step considers the spatial arrangement of minutiae by evaluating edge neighborhoods, ensuring that not only individual edges match but also their local context (Figure 3.31, adapted from [152, Fig. 1]).

Finally, in the *score* phase, a match probability score is computed for the generated spanning tree. This score quantifies the similarity between the probe and candidate based on factors like the number of mated minutiae in the tree, the number of supporting edges, and the similarity of these features. The entire process of enumerate, crawl, and score is repeated for multiple potential root pairs, and the highest probability score obtained is returned as the overall similarity score



Fig. 3.31: Illustration of a Spanning Tree for a Matching Fingerprint Pair. Minutiae are marked in red, the initial root pair in blue, the pairing tree in green, and supporting edges in yellow. Adapted from [152, Fig. 1].

for the probe-candidate pair. This score is then typically compared against a threshold to make a match or no-match decision.

3.7.3 Privacy Preserving Comparison with MPC

The primary objective of this research was to adapt the SourceAFIS minutiae comparison algorithm for execution within an MPC framework, specifically MP-SPDZ [119], to enable privacy-preserving fingerprint comparisons. As the feature extraction step operates on a single fingerprint locally, it remains unchanged and is performed outside the MPC environment. The core challenge lies in re-engineering the comparison stages (enumerate, crawl, score) for efficient and secure computation on encrypted data.

3.7.3.1 Minutiae Extractors and Template Compatibility

An important aspect of the research was to ensure the MPC-adapted comparison algorithm could work with minutiae extracted by different algorithms, not just the native SourceAFIS extractor. This enhances interoperability and allows leveraging potentially more accurate or specialized extractors. The study evaluated four distinct minutiae extraction methods: the original SourceAFIS extractor [258], Mindtct from the NBIS suite [265], FingerNet, a hybrid deep learning approach [250], and the commercial Idkit from Innovatrics [103].

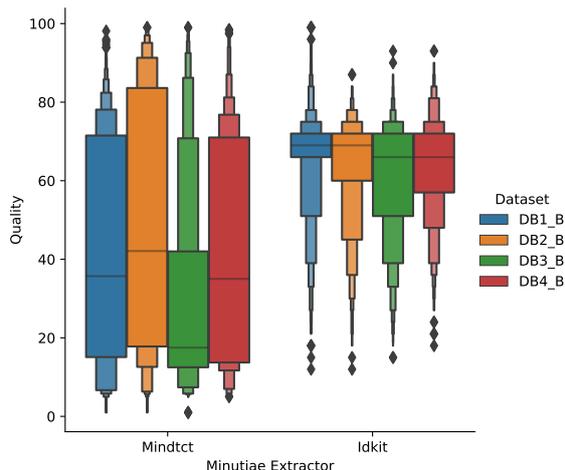
Minutiae found by these extractors were fed into the SourceAFIS template generation logic (which computes the invariant edges) before being input to the MPC comparison protocol. The performance was evaluated based on the Equal Error Rate (EER) and the number of minutiae extracted, using the FVC2000 databases [154]. The results showed that FingerNet, when paired with the SourceAFIS comparison logic, significantly outperformed the other combinations in terms of EER (FingerNet: 0.03 ± 0.02 EER vs. SourceAFIS: 0.12 ± 0.05 EER), while extracting a comparable number of minutiae (Table 3.21, adapted from [152, Table 1]). Mindtct extracted significantly more minutiae but yielded a higher EER, suggesting a higher incidence of false positives or less discriminative features. Idkit also provided competitive results. This confirmed that the comparison algorithm is adaptable and can benefit from advancements in minutiae extraction. Note however that the sample size of the dataset is small in comparison to other studies that compare comparison performance of different algorithms, indicating that the performance

ranking between the various extractors has to be taken cautiously, which is also reflected in the high standard deviations of the EER. Also note that Idkit has its own comparison logic, which when coupled with their template creation process beats the tested open source solutions.

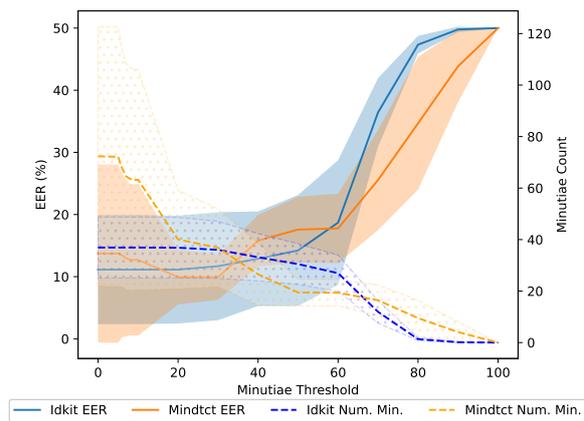
Further analysis involving minutiae quality scores (Figure 3.32, adapted from [152, Fig. 2, 3]) showed that filtering minutiae based on quality can improve EER, especially for extractors like Mindtct that produce many low-quality minutiae. This implies that a pre-filtering step based on locally computed quality scores could reduce template size and potentially speed up MPC computations without sacrificing accuracy.

Tab. 3.21: Comparison of Minutiae Extractors in terms of Equal Error Rate (EER) when used with the SourceAFIS comparison logic (plaintext) and the average number of extracted minutiae per fingerprint. Adapted from [152, Table 1].

Extractor	EER (\pm std)	Avg. Num. Minutiae (\pm std)
SourceAFIS	0.12 ± 0.05	35.96 ± 20.2
Mindtct	0.14 ± 0.14	72.34 ± 50.34
FingerNet	0.03 ± 0.02	36.21 ± 9.75
Idkit	0.11 ± 0.09	36.91 ± 11.81



(a) Minutiae Quality Distribution



(b) EER vs. Minutiae Count after Quality Thresh.

Fig. 3.32: (a) Distribution of minutiae quality scores (normalized) estimated by Idkit (blue) and Mindtct (orange/red) across FVC2000 datasets. (b) EER and average minutiae count after applying a quality threshold for Idkit (blue) and Mindtct (orange/red). Adapted from [152, Fig. 2, 3].

3.7.3.2 MPC Algorithm Optimizations

A naive translation of the SourceAFIS comparison algorithm into MPC would be computationally prohibitive due to its reliance on MPC-unfriendly operations like data-dependent loops, extensive hash table lookups, and re-computations. Therefore, significant algorithmic refactoring was necessary [152].

For the *enumerate* step, the original hash-based lookup for candidate edges cannot be efficiently replicated in MPC as it reveals information. Instead, the MPC version must compare every

candidate edge with every probe edge. This $N \times M$ comparison, where N and M are the number of edges in the probe and candidate respectively, is computationally intensive. Algorithm 2 in [152] outlines this MPC-friendly enumeration. The output is a mated array indicating which edge pairs match based on length and angle differences within defined tolerances. This mated array, containing only permuted edge indices, can then be revealed and used in plaintext for subsequent steps, forming an important optimization.

The *crawl* step (Algorithm 3 in [152]) was also significantly modified. In the original algorithm, edges are sorted by length and processed sequentially. This data-dependent sorting is inefficient in MPC. The optimized MPC approach omits this re-sorting. While this leads to different spanning trees being generated compared to the original (Figure 3.33, adapted from [152, Fig. 4]), tests showed that average error rates were not significantly impacted, though individual comparison outcomes could differ. The core crawl logic, which builds the tree using the pre-computed (now revealed) root pair array, can largely be performed in plaintext, significantly reducing MPC overhead.

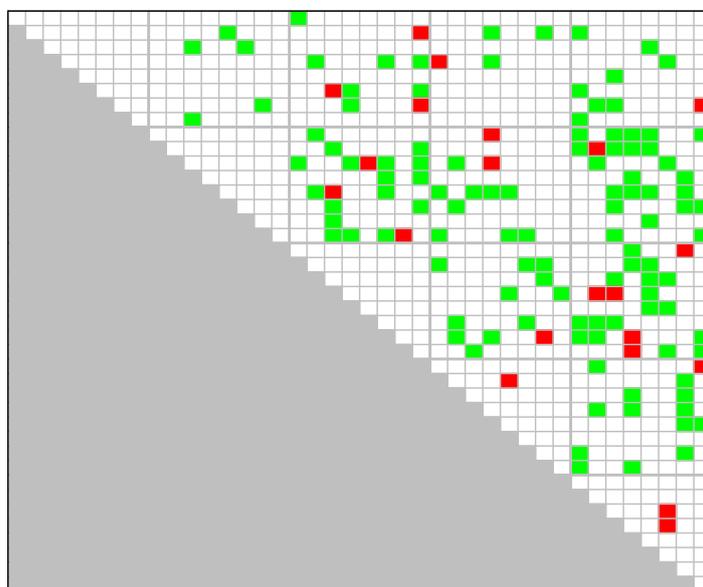


Fig. 3.33: Visual comparison of comparison decisions for fingerprint pairs from FVC2000 DB3 using the original SourceAFIS (evaluating at its EER threshold) versus the MPC-adapted version (evaluating at its EER threshold). White indicates identical match/non-match decisions. Green indicates the MPC version correctly matches where the original did not. Red indicates the MPC version fails to match where the original did. Adapted from [152, Fig. 4].

For the *scoring* step, operations that do not involve sensitive data (e.g., counting minutiae in the tree, number of supporting edges) are performed in plaintext. Only comparisons involving sensitive features like minutia type or precise edge geometry remain within the MPC protocol (Listing 2 in [152]). The study also explored replacing floating-point arithmetic with fixed-point arithmetic for MPC, as true floating-point operations are costly. It was found that even with only one decimal place of precision for fixed-point numbers, the EER did not significantly degrade compared to using floating-point numbers, due to the algorithm's reliance on comparisons with integer constants (e.g., degrees for angles) [152, Table 4].

3.7.3.3 Performance Evaluation with MP-SPDZ

The optimized algorithm was implemented and benchmarked using the MP-SPDZ framework, employing a 16-core Intel Xeon CPU. Initial tests with MPyC, another Python-based MPC framework, yielded execution times of around 7 minutes for the enumerate step, which was too slow for the target real-time application. MP-SPDZ, particularly with its support for eda-bits (efficient data-agnostic bits for comparisons), offered significant speed-ups [64].

For the *enumerate* step in MP-SPDZ, with 3 parties and no network latency simulated, finding roots for templates with an average of 39 minutiae (and 9 edges per minutia, capped at 100 minutiae) took approximately 12.25 seconds using the Shamir protocol (Table 3.22a, adapted from [152, Table 8]). The *scoring* step for a single tree of average size (6 nodes, see Table 3.22c, adapted from [152, Table 7]) took approximately 0.02 seconds (Table 3.22b, adapted from [152, Table 9]). Given that an average fingerprint comparison involves around 250 root pairs (Table 3.22d, adapted from [152, Table 10]) requiring a crawl and score iteration each, the total sequential execution time was estimated at $12.25s + (250 * 0.02s) = 17.25$ seconds. However, as the scoring of different root pairs is independent, these operations can be parallelized. If parallelized, the dominant factor remains the 12.25-second enumeration step, bringing the real-time goal within reach, especially with further optimizations like minutiae pre-filtering.

(a) Time for root finding (s) [152, Table 8]

Minutiae	<i>rep-field</i> (s)	<i>shamir</i> (s)
10	1.04	3.32
20	3.25	11.78
30	7.03	26.17
40	12.25	46.57
50	18.94	72.09

(b) Time for scoring trees (s) [152, Table 9]

Size	<i>rep-field</i> (s)	<i>shamir</i> (s)
1	0.01	0.02
2	0.02	0.02
5	0.02	0.04
10	0.03	0.06
20	0.05	0.10
50	0.03	0.08
100	0.04	0.14

(c) Tree size per fingerprint [152, Table 7]

	Matches			Non-Matches		
	Min.	Max.	Avg.	Min.	Max.	Avg.
DB1	5	45	7.4	1	7	2.3
DB2	4	77	6.3	2	7	2.8
DB3	5	100	5.5	2	9	2.9
DB4	4	33	6.2	1	7	3.0
Total	4	100	6.4	1	9	2.8

(d) Root pairs per fingerprint [152, Table 10]

	Matches			Non-Matches		
	Min.	Max.	Avg.	Min.	Max.	Avg.
DB1	8	899	194	0	428	65
DB2	29	1887	367	2	803	153
DB3	112	4007	1176	25	2484	590
DB4	16	619	178	3	321	58
Total	8	4007	479	0	2484	217

Tab. 3.22: MP-SPDZ performance benchmarks for key components of the SourceAFIS algorithm, using the Shamir or Replicated Field (*rep-field*) protocol with 3 parties and no network latency. (a) Time for root finding (*enumerate* step). (b) Time for scoring trees of different sizes. (c) Average tree size per fingerprint comparison. (d) Number of root pairs found per fingerprint comparison. Adapted from [152, Tables 7-10].

The accuracy of the MPC-adapted plaintext implementation (using fixed-point arithmetic and the modified crawl) was comparable to the original SourceAFIS, with EERs of 0.04 on FVC2000 DB1 using the respective native minutiae extractors [152, Table 2]. Combining the MPC-adapted

logic with the FingerNet extractor yielded an EER of 0.01 on DB1, demonstrating high accuracy (Table 3.21 using MPC logic would show this). Further optimizations, such as filtering root pairs based on tree size (e.g., discarding trees with <4 mated minutiae or >9 , which rarely occur in genuine or impostor matches respectively, Table 11 in [152]) and more aggressive minutiae quality filtering, could further reduce computation time for the scoring phase.

3.7.4 Discussion

The research presented in [152] successfully demonstrates the feasibility of implementing a state-of-the-art, minutiae-based fingerprint comparison algorithm (SourceAFIS) using Multiparty Computation, achieving comparison times that approach real-world applicability for scenarios like airport security checks.

3.7.4.1 Effectiveness of Privacy-Preserving Fingerprint Comparison

The study shows that a fully-fledged privacy-preserving minutiae-based comparison algorithm can be realized with practical efficiency. The achieved EERs, especially when combined with advanced minutiae extractors like FingerNet (resulting in EERs as low as 0.01-0.03 on FVC2000 datasets), are competitive with plaintext state-of-the-art systems. This indicates that the transition to MPC does not necessitate a significant compromise on comparison accuracy, a critical factor for deployment in security-sensitive applications. The estimated execution time of around 17 seconds for a full comparison (reducible to 12 seconds with parallel scoring, dominated by the enumerate step) using MP-SPDZ is a substantial improvement over previous MPC attempts for complex comparison algorithms and is within a range that could be acceptable for certain less time-critical identity verification scenarios, or further improved for time-critical ones.

3.7.4.2 Impact on Research Questions

This work directly addresses **RQ7 (How can we create a privacy preserving comparison approach that circumvents the risk of exposing biometric data to other parties?)**. The research provides a concrete answer by:

1. Demonstrating a successful implementation of a complex minutiae-based fingerprint comparison algorithm (SourceAFIS) using MPC (MP-SPDZ framework).
2. Identifying and implementing algorithmic optimizations (e.g., modified enumerate and crawl steps, selective encryption of sensitive features, fixed-point arithmetic) that make MPC performance practical.
3. Showing that this privacy-preserving approach maintains high comparison accuracy, comparable to its plaintext counterpart, especially when paired with modern minutiae extractors.

The findings suggest that MPC is a viable technology for secure remote fingerprint comparison, enabling new use cases like confidential watchlist checking without compromising the privacy of individuals or the sensitive data held by participating organizations.

The research also implicitly touches upon other questions by highlighting the importance of the preceding steps. For instance, the accuracy of segmentation (**RQ1**) and landmark detection (**RQ3**) directly impacts the quality of minutiae extracted, which in turn affects the performance of any comparison algorithm, including an MPC-based one. Similarly, if mosaicking (**RQ5**) is used to generate templates for an MPC system, ensuring the structural integrity of those templates via artifact detection becomes important for reliable privacy-preserving fingerprint comparison.

3.7.4.3 Limitations

Several limitations should be acknowledged. The current performance benchmarks were conducted with simulated no-network-latency between the MPC parties. Real-world network latency will increase execution times. The formal security model and proof for the implemented MPC protocol, quantifying exact leakage from revealed intermediate values (like the edge match array), are yet to be fully developed, though efforts were made to minimize such leakage. The current optimizations focused on a 3-party computation scenario using Shamir secret sharing, performance and suitability might differ for other MPC protocols or numbers of parties. While the comparison algorithm is compatible with various minutiae extractors, the feature extraction itself is still performed in plaintext locally, meaning the privacy of the raw fingerprint image relies on the security of the capture device and local processing environment.

3.7.4.4 Future Research Directions

Future work should focus on several key areas. Further optimization of the MPC protocols, especially the enumerate step (e.g., exploring MPC-friendly near-neighbor search algorithms or locality-sensitive hashing), is essential to reduce execution times closer to traditional plaintext comparison algorithms. A comprehensive formal security analysis of the complete MPC-adapted comparison protocol is needed to quantify any information leakage. Extending the implementation and benchmarking to real-world distributed environments with network latency will provide more realistic performance figures. Investigating the integration of minutiae quality information more directly into the MPC comparison logic could offer further speed-ups by allowing secure pruning of less reliable features during the encrypted comparison. Finally, exploring end-to-end MPC solutions that also encompass parts of the feature extraction process, if feasible, would enhance overall system privacy.

Chapter 4

Conclusion

4.1 Conclusion

This thesis has presented a series of contributions to the field of contactless fingerprint recognition, addressing the complex challenges that arise from its unconstrained acquisition process. Unlike contact-based methods, contactless capture introduces variations in pose, illumination, and scale, which require specialized algorithms to correct for and ensure interoperability with legacy systems. To systematically address these issues, this research was guided by a set of core questions spanning the entire recognition pipeline. These questions can be grouped into three main areas:

- *Image Normalization and Interoperability:* RQ1: How can segmentation of fingerprints from contactless hand images be made more accurate and robust across different environments? RQ2: What techniques can effectively correct pose variations in contactless fingerprints to enhance comparison accuracy? and RQ3: How can core point detection in contactless images be improved to enable better alignment with contact-based databases?
- *Data Assurance and Quality:* RQ4: What are the problems of established fingerprints quality assessment metrics and what can be changed to improve their performance on contactless fingerprint images? and RQ5: How can we detect errors in the mosaicking process of contactless fingerprint sensor systems?
- *Validation and Secure Deployment:* RQ6: Can we produce realistic, synthetic 3D fingers and how do they perform compared to real fingers? What manufacturing techniques can be used? and RQ7: How can we create a privacy preserving comparison approach that circumvents the risk of exposing biometric data to other parties?

This chapter summarizes the key findings and contributions of this thesis in response to these questions, demonstrates their integration into a holistic pipeline, and discusses the limitations of the proposed methods.

4.1.1 Image Normalization and Interoperability

This research directly addresses the challenges of image normalization (RQ1, RQ2, RQ3) by developing a series of algorithms that transform raw contactless captures into standardized, interoperable fingerprints. To address the foundational challenge of isolating fingertips from varied backgrounds (RQ1), this work developed progressively capable deep learning models. The initial work on single-finger segmentation led to a custom U-Net architecture [217] and later progressed to *FingerUNeSt++*, a model combining a ResNeSt encoder with a UNet++-like decoder, which achieved a near-perfect mean Intersection-over-Union (mIoU) of 0.99 on its test set, significantly outperforming traditional methods and baseline deep learning architectures [215]. Building on this, *TipSegNet* was developed to perform multi-class segmentation directly from a whole-hand image, eliminating the need for a preceding finger detection step. By leveraging a

ResNeXt-101 backbone with a Feature Pyramid Network (FPN), *TipSegNet* achieved state-of-the-art performance with an mIoU of 0.99 and an accuracy of 1.00, setting a new benchmark for labeled, multi-finger segmentation from a single capture [216]. The significance of these models lies in providing the masks that are a prerequisite for all subsequent normalization steps. An inaccurate segmentation mask would introduce cascading errors, rendering any pose correction or unwarping ineffective.

To improve interoperability with contact-based systems (RQ2, RQ3), a novel pipeline was developed to correct geometric distortions. It combines segmentation-driven horizontal rotation correction with an ellipsoid-based model for lateral rotation correction, using the fingerprint core as a reference point [217]. When combined with parametric unwarping techniques [237], this pipeline demonstrated a substantial enhancement in comparison performance. On a challenging operational dataset, a finger-wise optimized application of the pipeline reduced the Equal Error Rate (EER) for contactless-to-contact-based comparison by a relative 36.9%, from a baseline of 1.57% to 0.99% [217]. Additional work also empirically quantified the inherent positional variability of the fingerprint core, revealing systematic modality-induced biases and a fundamental limit on the accuracy of any alignment method relying solely on this landmark [220], suggesting that future interoperability efforts must either accept this variability or incorporate additional reference features. Moreover, the core’s position was found to be best modeled by a finger dependent distribution. The one distribution however that fit best for most finger was the Non-Central Fischer (NCF) distribution. This provides insights into synthetic fingerprint generation, which could lead to more realistic core distribution in non-machine learning generators.

4.1.2 Data Assurance and Quality

This thesis introduces two novel contributions to data assurance and quality, addressing RQ4 and RQ5. In response to the lack of dedicated quality assessment metrics for mobile contactless fingerprints (RQ4), this work contributed to the evaluation of *MCLFIQ*, an adaptation of the established NFIQ 2 framework [246]. By retraining the framework’s random forest classifier on modality-specific synthetic data, *MCLFIQ* demonstrated superior predictive power for contactless image utility compared to the original NFIQ 2.2 and other baseline metrics across multiple datasets and recognition systems. Feature importance analysis confirmed that the retrained model correctly prioritized fidelity-related aspects like sharpness and local ridge clarity, which are critical for the contactless modality [194]. As a result, we saw that the feature set of NFIQ 2 is still largely relevant, also for other modalities.

To address a gap in standard quality assessment, this work also developed a novel, self-supervised framework to detect structural errors introduced during fingerprint mosaicking (RQ5) [218]. By programmatically introducing realistic structural discontinuities into artifact-free images, a deep learning model (*UNet++* with a ResNeSt encoder) was trained to detect these hard artifacts without requiring any manual annotation. The resulting detector demonstrated high accuracy (IoU > 0.9) and strong generalization across diverse modalities (contactless, rolled, slap) and was robust against non-mosaicking quality defects like scars or noise. Its performance was tested by the Bundesamt für Sicherheit in der Informationstechnik (Germany) by retraining a model on their data and evaluating it on both publicly available data, as well as on institutional data. Both the AIT and the BSI models had excellent false-positive rates and very high accuracy rates, implying that they can be rolled out to operate on large datasets without excessive flagging of intact data. Additionally, a mosaicking artifact score was developed to infer artifact severity for a given image, enabling the usability of the framework in an automated way on large datasets. This mosaicking artifact framework provides the first dedicated, reference-free tool for verifying

the structural integrity of mosaicked fingerprints and its code is available under the Mozilla Public License Version 2.0 [218].

4.1.3 Validation and Secure Deployment

To enable robust validation and secure deployment (RQ6, RQ7), this thesis provides both a physical methodology and a cryptographic framework. To create standardized and repeatable validation methods (RQ6), this research provides a comprehensive investigation into the manufacturing of synthetic physical fingerprint targets. A comprehensive investigation of manufacturing methodologies, including direct laser engraving, CNC machining, and 3D printing of molds, and materials was conducted [219]. This work established that high-fidelity, physically realistic, and stable 3D phantoms can be produced, particularly using silicone cast in precisely machined, laser-engraved or 3D printed molds. These phantoms successfully replicate a known ground-truth pattern and provide the necessary physical properties for realistic sensor interaction, serving as essential tools for the objective evaluation of sensor performance and processing algorithms like pose correction and mosaicking.

Finally, to enable the deployment of fingerprint recognition in privacy-sensitive, multi-party applications (RQ7), this research demonstrated the feasibility of a privacy-preserving minutiae-based comparison algorithm. A state-of-the-art open-source algorithm, SourceAFIS, was successfully adapted for execution within a Multiparty Computation (MPC) framework. Through significant algorithmic refactoring to optimize MPC-unfriendly operations, the implementation achieved a comparison time of approximately 17 seconds while maintaining high recognition accuracy, comparable in accuracy to its plaintext counterpart. This work makes complex, secure remote fingerprint comparison practical for real-world scenarios, representing a significant advancement over previous attempts [152]. The implication is that sensitive operations, such as checking a traveler’s fingerprint against a confidential multi-agency watchlist, can be performed without either party revealing their private data. This removes the data-privacy barrier to the adoption of biometrics in distributed, trust-less environments.

4.1.4 The Complete Pipeline: A Holistic View

The individual contributions of this thesis, while distinct, are designed to function as interconnected components of a contactless fingerprint recognition pipeline. To illustrate their workings, consider a plausible and challenging real-world use-case: a mobile, high-security identity verification system at a border control checkpoint. In this scenario, a border agent uses a standard smartphone to capture a traveler’s fingerprints and verify them against both the traveler’s digital identity document and a confidential, multi-agency watchlist.

The process begins with the acquisition phase, where the traveler presents their hand to the smartphone’s camera, perhaps performing a guided nail-to-nail rolling motion for one or more fingers to generate a rolled-equivalent impression. Immediately, the algorithms developed in this thesis begin processing the incoming video stream. The *TipSegNet* model (RQ1) operates directly on the whole-hand images, segmenting and labeling each fingertip [216]. For each segmented fingertip image, the *MCLFIQ* algorithm (RQ4) calculates a quality score, providing immediate feedback on factors like focus and illumination and rejecting frames that are not of sufficient utility for further processing [194].

Once a sufficient number of frames are acquired, they are passed to the mosaicking stage to create a complete, rolled-equivalent fingerprint image. At this point, the self-supervised artifact detector (RQ5) is employed to assess the structural integrity of the composite image [218]. If significant hard artifacts, such as misaligned or broken ridges resulting from an imperfect roll, are

detected, the mosaicked image is flagged as corrupt, and a re-capture is requested. This integrity check is important to prevent an erroneous template from causing a downstream comparison failure.

Assuming the resulting fingerprint image passes all quality and integrity checks, it is then normalized for interoperability. The deep learning-based core detector (RQ3) localizes the fingerprint core, which, along with the segmentation mask, serves as input to the pose correction and unwarping pipeline (RQ2) [217]. This step computationally corrects for finger orientation (roll, yaw) and flattens the curved surface texture, transforming the contactless image into a standardized format that is geometrically comparable to legacy contact-based templates stored in government databases. The empirical analysis of core centrality and modality-specific biases [220] informs this process, providing a quantitative understanding of the inherent geometric transformations required.

Finally, the normalized template enters the secure comparison phase. To verify the traveler against the sensitive, pooled watchlist without compromising the privacy of either the traveler or the watchlist data, the privacy-preserving MPC protocol is initiated (RQ7). The traveler's template and the agencies' watchlist templates are secret-shared, and the optimized MPC comparison algorithm compares them in the encrypted domain [152]. Only the final binary result (a match or no-match) is revealed to the border agent. Throughout this entire interaction, the traveler's raw biometric data is never exposed to the central databases, and the contents of the confidential watchlists remain protected.

The development, validation, and certification of this entire complex pipeline would be underpinned by the synthetic physical phantoms (RQ6). These phantoms provide the standardized, repeatable targets necessary to test each algorithmic component, from segmentation accuracy under different poses to the EER of the complete pipeline, against a known ground truth, ensuring the system's reliability and performance before deployment in such a critical security environment [219].

4.1.5 Limitations of the Proposed Methods

Despite the advancements presented, the proposed methods have inherent limitations that require discussion and guide future research. These limitations primarily stem from the interdependencies within the processing pipeline, the simplifications made in modeling complex real-world phenomena, and the practical boundaries of the developed security protocols.

A critical limitation is the effect of cascading errors throughout the multi-stage pipeline. The performance of downstream modules like the pose correction and unwarping pipeline is fundamentally bound by the accuracy of the initial segmentation and landmark detection steps [217]. A minor, seemingly insignificant error in delineating the finger's contour can propagate into an incorrect axis estimation, leading to a flawed rotation correction. This, in turn, can shift the apparent location of the core, compromising the lateral rotation estimate and ultimately degrading, rather than improving, the final template. This sensitivity shows that the pipeline's overall reliability is not the sum of its parts but is constrained by its weakest link.

Furthermore, several algorithms rely on simplified models of reality, which creates a potential reality gap. The self-supervised artifact detector, while robust, is trained exclusively on programmatically simulated hard artifacts. It may not generalize to the full, diverse range of structural errors produced by every proprietary mosaicking algorithm, particularly those that may produce novel or unforeseen discontinuity patterns [218]. Similarly, the pose correction algorithm's assumption of an elliptical finger cross-section is a simplification. It does not account for more complex and irregular 3D topology of real human fingers. This simplification places a ceiling on the achievable geometric accuracy of the unwarping process.

The approach to quality assessment, *MCLFIQ*, is also constrained by its foundational architecture. While retraining the classifier on modality-specific data proved effective, the algorithm is still limited to the expressive power of the 74 hand-crafted features defined in the NFIQ 2 standard [194, 246]. It is plausible that unique contactless quality degradations exist, for example, subtle optical distortions or complex motion blur patterns, that are not fully captured by this predefined feature set, limiting the predictive power of the metric.

Finally, the privacy-preserving comparison framework has important architectural boundaries. The security of the protocol begins with the minutiae template. The initial feature extraction from the raw fingerprint image is performed in plaintext on a local device. Thus, the privacy of the raw biometric data still depends entirely on the trustworthiness of the acquisition environment, which is outside the scope of the MPC protocol itself. Additionally, the optimized protocol reveals a minimal amount of intermediate information (the permuted edge match array) to achieve its near real-time performance. While this data is structured to minimize leakage, it is not formally proven to be zero-knowledge, presenting a trade-off between computational efficiency and perfect, provable privacy [152].

Bibliography

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. Mar. 16, 2016. DOI: 10.48550/arXiv.1603.04467. arXiv: 1603.04467[cs]. URL: <http://arxiv.org/abs/1603.04467> (visited on 03/07/2025).
- [2] I. M. Alarifi. “A comprehensive review on advancements of elastomers for engineering applications”. In: *Advanced Industrial and Engineering Polymer Research*. Biodegradable Polymers for Engineering Applications 6.4 (Oct. 1, 2023), pp. 451–464. ISSN: 2542-5048. DOI: 10.1016/j.aiepr.2023.05.001. URL: <https://www.sciencedirect.com/science/article/pii/S2542504823000350> (visited on 01/08/2024).
- [3] J. Almog, A. Hirshfeld, and J. T. Klug. “Reagents for the chemical development of latent fingerprints: synthesis and properties of some ninhydrin analogues”. In: *Journal of Forensic Sciences* 27.4 (Oct. 1982), pp. 912–917. ISSN: 0022-1198.
- [4] F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, J. Gonzalez-Rodriguez, H. Fronthaler, K. Kollreider, and J. Bigun. “A Comparative Study of Fingerprint Image-Quality Estimation Methods”. In: *IEEE Transactions on Information Forensics and Security* 2.4 (Dec. 2007). Conference Name: IEEE Transactions on Information Forensics and Security, pp. 734–743. ISSN: 1556-6021. DOI: 10.1109/TIFS.2007.908228. URL: <https://ieeexplore.ieee.org/document/4380297> (visited on 03/06/2025).
- [5] D. Ametefe, S. Seroja, D. mohd ali, and Z. Muhammad. “Fingerprint pattern classification using deep transfer learning and data augmentation”. In: *The Visual Computer* 39 (Apr. 7, 2022), pp. 1–14. DOI: 10.1007/s00371-022-02437-x.
- [6] S. S. Arora, A. K. Jain, and N. G. Paulter. “3D Whole Hand Targets: Evaluating Slap and Contactless Fingerprint Readers”. In: *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*. 2016 International Conference of the Biometrics Special Interest Group (BIOSIG). Darmstadt, Germany: IEEE, Sept. 2016, pp. 1–8. ISBN: 978-3-88579-654-1. DOI: 10.1109/BIOSIG.2016.7736900. URL: <http://ieeexplore.ieee.org/document/7736900/> (visited on 11/25/2021).
- [7] R. T. Azuma. “A Survey of Augmented Reality”. In: *Presence: Teleoperators and Virtual Environments* 6.4 (Aug. 1, 1997), pp. 355–385. DOI: 10.1162/pres.1997.6.4.355. URL: <https://doi.org/10.1162/pres.1997.6.4.355> (visited on 03/07/2025).
- [8] J. L. Ba, J. R. Kiros, and G. E. Hinton. *Layer Normalization*. July 21, 2016. DOI: 10.48550/arXiv.1607.06450. arXiv: 1607.06450[cs,stat]. URL: <http://arxiv.org/abs/1607.06450> (visited on 03/20/2023).

- [9] G. A. Bahgat, A. H. Khalil, N. S. Abdel Kader, and S. Mashali. “Fast and accurate algorithm for core point detection in fingerprint images”. In: *Egyptian Informatics Journal* 14.1 (Mar. 1, 2013), pp. 15–25. ISSN: 1110-8665. DOI: 10.1016/j.eij.2013.01.002. URL: <https://www.sciencedirect.com/science/article/pii/S1110866513000030> (visited on 11/03/2022).
- [10] K. Bahmani, R. Plesh, P. Johnson, S. Schuckers, and T. Swyka. “High Fidelity Fingerprint Generation: Quality, Uniqueness, And Privacy”. In: *2021 IEEE International Conference on Image Processing (ICIP)*. 2021 IEEE International Conference on Image Processing (ICIP). ISSN: 2381-8549. Sept. 2021, pp. 3018–3022. DOI: 10.1109/ICIP42928.2021.9506386. URL: https://ieeexplore.ieee.org/abstract/document/9506386?casa_token=-A05gKSliSsAAAAA:rNcScY2dXjvGSiVJmTC00vAVU1pAh5kcF40Fu3rSPtkuGxQCcfNeCnriM6dLzeYlGg-rUZJgNA (visited on 03/07/2024).
- [11] J. G. Barnes. “History”. In: *The Fingerprint Sourcebook*. 1st ed. NCJ 225320. National Institute of Justice, Aug. 2022, p. 18. URL: <https://www.ojp.gov/pdffiles1/nij/225321.pdf> (visited on 09/09/2024).
- [12] M. Barni, T. Bianchi, D. Catalano, M. Di Raimondo, R. Donida Labati, P. Failla, D. Fiore, R. Lazzeretti, V. Piuri, F. Scotti, and A. Piva. “Privacy-preserving fingercode authentication”. In: *Proceedings of the 12th ACM workshop on Multimedia and security. MM&Sec '10*. New York, NY, USA: Association for Computing Machinery, Sept. 9, 2010, pp. 231–240. ISBN: 978-1-4503-0286-9. DOI: 10.1145/1854229.1854270. URL: <https://dl.acm.org/doi/10.1145/1854229.1854270> (visited on 05/21/2025).
- [13] K. Bashir, M. Amin, M. Majid, F. A. Butt, J. A. Rather, W. A. Wani, W. A. Khanday, A. H. Malik, and M. A. Yattoo. “From invisible to visible: a concise review on conjugated polymer materials in latent fingerprint analysis”. In: *Journal of Polymer Research* 31.8 (Aug. 2024), p. 235. ISSN: 1022-9760, 1572-8935. DOI: 10.1007/s10965-024-04086-1. URL: <https://link.springer.com/10.1007/s10965-024-04086-1> (visited on 10/06/2025).
- [14] F. Battisti, E. Bosc, M. Carli, P. Le Callet, and S. Perugia. “Objective image quality assessment of 3D synthesized views”. In: *Signal Processing: Image Communication* 30 (Jan. 31, 2015), pp. 78–88. DOI: 10.1016/j.image.2014.10.005.
- [15] H. Bay, T. Tuytelaars, and L. Van Gool. “SURF: Speeded Up Robust Features”. In: *Computer Vision – ECCV 2006*. Ed. by A. Leonardis, H. Bischof, and A. Pinz. Berlin, Heidelberg: Springer, 2006, pp. 404–417. ISBN: 978-3-540-33833-8. DOI: 10.1007/11744023_32.
- [16] A. M. Bazen and S. H. Gerez. “Thin-plate Spline Modelling of Elastic Deformations in Fingerprints”. In: *Proceedings of the 3rd IEEE Benelux Signal Processing Symposium (SPS-2002)*. Katholieke Universiteit, Mar. 21, 2002, pp. 205–208. URL: <https://research.utwente.nl/en/publications/thin-plate-spline-modelling-of-elastic-deformations-in-fingerprin> (visited on 03/03/2025).
- [17] A. Bazen and S. Gerez. “Systematic methods for the computation of the directional fields and singular points of fingerprints”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (July 2002). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 905–919. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2002.1017618. URL: <https://ieeexplore.ieee.org/document/1017618> (visited on 03/06/2025).
- [18] A. M. Bazen and S. H. Gerez. “Segmentation of Fingerprint Images”. In: (2001).
- [19] A. M. Bazen, G. T. B. Verwaaijen, S. H. Gerez, and L. P. J. Veelenturf. “A Correlation-Based Fingerprint Verification System”. In: (2000).

- [20] A. M. Bazen and S. H. Gerez. “Fingerprint matching by thin-plate spline modelling of elastic deformations”. In: *Pattern Recognition* 36.8 (Aug. 2003), pp. 1859–1867. ISSN: 00313203. DOI: 10.1016/S0031-3203(03)00036-0. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0031320303000360> (visited on 10/05/2022).
- [21] N. Beheshti and L. Johnsson. “Squeeze U-Net: A Memory and Energy Efficient Image Segmentation Network”. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle, WA, USA: IEEE, June 2020, pp. 1495–1504. ISBN: 978-1-72819-360-1. DOI: 10.1109/CVPRW50498.2020.00190. URL: <https://ieeexplore.ieee.org/document/9150777/> (visited on 07/18/2023).
- [22] Y. Bengio, A. Courville, and P. Vincent. “Representation Learning: A Review and New Perspectives”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.8 (Aug. 2013). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1798–1828. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2013.50. URL: <https://ieeexplore.ieee.org/document/6472238> (visited on 03/10/2025).
- [23] S. H. Bhati and U. C. Pati. “Novel algorithm for fingerprint mosaicing using phase correlation method”. In: *2015 Global Conference on Communication Technologies (GCCT)*. 2015 Global Conference on Communication Technologies (GCCT). Apr. 2015, pp. 29–33. DOI: 10.1109/GCCT.2015.7342618. URL: <https://ieeexplore.ieee.org/document/7342618> (visited on 07/11/2024).
- [24] M. Blanton and P. Gasti. “Secure and Efficient Protocols for Iris and Fingerprint Identification”. In: *Computer Security – ESORICS 2011*. Ed. by V. Atluri and C. Diaz. Berlin, Heidelberg: Springer, 2011, pp. 190–209. ISBN: 978-3-642-23822-2. DOI: 10.1007/978-3-642-23822-2_11.
- [25] M. Blanton and D. Murphy. “Privacy Preserving Biometric Authentication for Fingerprints and Beyond”. In: *Proceedings of the Fourteenth ACM Conference on Data and Application Security and Privacy*. CODASPY ’24. New York, NY, USA: Association for Computing Machinery, June 19, 2024, pp. 367–378. ISBN: 979-8-4007-0421-5. DOI: 10.1145/3626232.3653269. URL: <https://dl.acm.org/doi/10.1145/3626232.3653269> (visited on 05/21/2025).
- [26] M. Bojarski, D. D. Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba. *End to End Learning for Self-Driving Cars*. Apr. 25, 2016. DOI: 10.48550/arXiv.1604.07316. arXiv: 1604.07316[cs]. URL: <http://arxiv.org/abs/1604.07316> (visited on 03/07/2025).
- [27] X. Boyen. “Reusable cryptographic fuzzy extractors”. In: *Proceedings of the 11th ACM conference on Computer and communications security*. CCS ’04. New York, NY, USA: Association for Computing Machinery, Oct. 25, 2004, pp. 82–91. ISBN: 978-1-58113-961-7. DOI: 10.1145/1030083.1030096. URL: <https://dl.acm.org/doi/10.1145/1030083.1030096> (visited on 05/21/2025).
- [28] M. Brown and D. G. Lowe. “Automatic Panoramic Image Stitching using Invariant Features”. In: *International Journal of Computer Vision* 74.1 (Aug. 2007), pp. 59–73. ISSN: 0920-5691, 1573-1405. DOI: 10.1007/s11263-006-0002-3. URL: <https://link.springer.com/10.1007/s11263-006-0002-3> (visited on 05/21/2025).

- [29] R. Brunelli and T. Poggiot. “Template matching: matched spatial filters and beyond”. In: *Pattern Recognition* 30.5 (May 1, 1997), pp. 751–768. ISSN: 0031-3203. DOI: 10.1016/S0031-3203(96)00104-5. URL: <https://www.sciencedirect.com/science/article/pii/S0031320396001045> (visited on 03/07/2025).
- [30] S. Brüning. *Latest developments in direct laser engraving*. Apr. 14, 2021. URL: https://www.researchgate.net/publication/351064598_Latest_developments_in_direct_laser_engraving.
- [31] J. Canny. “A Computational Approach To Edge Detection”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on PAMI-8* (Dec. 1, 1986), pp. 679–698. ISSN: 9780080515816. DOI: 10.1109/TPAMI.1986.4767851.
- [32] K. Cao, D.-L. Nguyen, C. Tymoszek, and A. K. Jain. *End-to-End Latent Fingerprint Search*. Dec. 26, 2018. DOI: 10.48550/arXiv.1812.10213. arXiv: 1812.10213. URL: <http://arxiv.org/abs/1812.10213> (visited on 11/19/2024).
- [33] R. Capelli. “Synthetic Fingerprint Generation”. In: D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. London: Springer London, 2009, pp. 271–302. ISBN: 978-1-84882-253-5. DOI: 10.1007/978-1-84882-254-2_6. URL: http://link.springer.com/10.1007/978-1-84882-254-2_6 (visited on 10/02/2023).
- [34] R. Cappelli. “Exploring the Power of Simplicity: A New State-of-the-Art in Fingerprint Orientation Field Estimation”. In: *IEEE Access* 12 (2024). Conference Name: IEEE Access, pp. 55998–56018. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2024.3389701. URL: <https://ieeexplore.ieee.org/document/10500816> (visited on 03/06/2025).
- [35] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni. “Fingerprint verification competition 2006”. In: *Biometric Technology Today* 15.7 (July 1, 2007), pp. 7–9. ISSN: 0969-4765. DOI: 10.1016/S0969-4765(07)70140-6. URL: <https://www.sciencedirect.com/science/article/pii/S0969476507701406> (visited on 07/21/2023).
- [36] K. Casas-Zamora. *ICTs in Elections Database*. International Institute for Democracy and Electoral Assistance. Feb. 18, 2025. URL: <https://www.idea.int/data-tools/data/question> (visited on 02/18/2025).
- [37] C. Champod, C. J. Lennard, P. Margot, and M. Stoilovic. *Fingerprints and Other Ridge Skin Impressions*. CRC Press, Apr. 27, 2004. 300 pp. ISBN: 978-0-203-48504-0.
- [38] J. Chen, H. Zhao, Z. Cao, W. Zhao, and L. Pang. “Successive minutia-free mosaicking for small-sized fingerprint recognition”. In: *IET Image Processing* 13.7 (2019). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-ipr.2018.5972>, pp. 1146–1151. ISSN: 1751-9667. DOI: 10.1049/iet-ipr.2018.5972. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1049/iet-ipr.2018.5972> (visited on 07/11/2024).
- [39] J. V. Chen, A. B. C. Dang, and A. Dang. “Comparing cost and print time estimates for six commercially-available 3D printers obtained through slicing software for clinically relevant anatomical models”. In: *3D Printing in Medicine* 7.1 (Jan. 6, 2021), p. 1. ISSN: 2365-6271. DOI: 10.1186/s41205-020-00091-4. URL: <https://doi.org/10.1186/s41205-020-00091-4> (visited on 01/08/2024).
- [40] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille. “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PP (June 2, 2016). DOI: 10.1109/TPAMI.2017.2699184.

- [41] W. Chen, L. Sui, Z. Xu, and Y. Lang. “Improved Zhang-Suen thinning algorithm in binary line drawing applications”. In: *2012 International Conference on Systems and Informatics (ICSAI2012)*. 2012 International Conference on Systems and Informatics (ICSAI2012). May 2012, pp. 1947–1950. DOI: 10.1109/ICSAI.2012.6223430. URL: <https://ieeexplore.ieee.org/document/6223430> (visited on 03/06/2025).
- [42] K. Choi, H.-s. Choi, and J. Kim. “Fingerprint Mosaicking by Rolling and Sliding”. In: *Audio- and Video-Based Biometric Person Authentication*. Ed. by T. Kanade, A. Jain, and N. K. Ratha. Berlin, Heidelberg: Springer, 2005, pp. 260–269. ISBN: 978-3-540-31638-1. DOI: 10.1007/11527923_27.
- [43] F. Chollet et al. *Keras*. Version 3. Github, 2015.
- [44] M. Cohen. *NFIQ-1-2 Observations*. 2015.
- [45] S. A. Cole. “More than Zero: Accounting for Error in Latent Fingerprint Identification”. In: *The Journal of Criminal Law and Criminology (1973-)* 95.3 (2005). Publisher: Northwestern University School of Law, pp. 985–1078. ISSN: 00914169. URL: <http://www.jstor.org/stable/3491332> (visited on 09/04/2024).
- [46] M. Controzzi, M. D’Alonzo, C. Peccia, C. M. Oddo, M. C. Carrozza, and C. Cipriani. “Bioinspired Fingertip for Anthropomorphic Robotic Hands”. In: *Applied Bionics and Biomechanics* 11.1 (2014), pp. 25–38. ISSN: 1176-2322, 1754-2103. DOI: 10.1155/2014/864573. URL: <http://www.hindawi.com/journals/abb/2014/864573/> (visited on 01/12/2024).
- [47] P.-H. Conze, P. Robert, and L. Morin. “Objective view synthesis quality assessment”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 8288 (Feb. 1, 2012), p. 53. DOI: 10.1117/12.908762.
- [48] C. Cortes and V. Vapnik. “Support-vector networks”. In: *Machine Learning* 20.3 (Sept. 1995), pp. 273–297. ISSN: 0885-6125, 1573-0565. DOI: 10.1007/BF00994018. URL: <http://link.springer.com/10.1007/BF00994018> (visited on 03/07/2025).
- [49] Z. Cui, J. Feng, and J. Zhou. “Dense Registration and Mosaicking of Fingerprints by Training an End-to-End Network”. In: *IEEE Transactions on Information Forensics and Security* 16 (2021), pp. 627–642. ISSN: 1556-6013, 1556-6021. DOI: 10.1109/TIFS.2020.3017926. arXiv: 2004.05972[cs]. URL: <http://arxiv.org/abs/2004.05972> (visited on 07/11/2024).
- [50] A. Dabouei, S. Soleymani, J. Dawson, and N. M. Nasrabadi. “Deep Contactless Fingerprint Unwarping”. In: *2019 International Conference on Biometrics (ICB)*. 2019 International Conference on Biometrics (ICB). Crete, Greece: IEEE, June 2019, pp. 1–8. ISBN: 978-1-72813-640-0. DOI: 10.1109/ICB45273.2019.8987292. URL: <https://ieeexplore.ieee.org/document/8987292/> (visited on 11/25/2021).
- [51] A. K. Dąbrowska, G.-M. Rotaru, S. Derler, F. Spano, M. Camenzind, S. Annaheim, R. Stämpfli, M. Schmid, and R. M. Rossi. “Materials used to simulate physical properties of human skin”. In: *Skin Research and Technology* 22.1 (2016). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/srt.12235>, pp. 3–14. ISSN: 1600-0846. DOI: 10.1111/srt.12235. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/srt.12235> (visited on 01/08/2024).

- [52] D. Das. “A Fingerprint Segmentation Scheme Based on Adaptive Threshold Estimation”. In: *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Oct. 2018, pp. 1–6. DOI: 10.1109/CISP-BMEI.2018.8633064. URL: <https://ieeexplore.ieee.org/abstract/document/8633064> (visited on 03/17/2025).
- [53] D. Das and S. Mukhopadhyay. “A Pixel Based Segmentation Scheme for Fingerprint Images”. In: *Information Systems Design and Intelligent Applications*. Ed. by J. K. Mandal, S. C. Satapathy, M. Kumar Sanyal, P. P. Sarkar, and A. Mukhopadhyay. New Delhi: Springer India, 2015, pp. 439–448. ISBN: 978-81-322-2247-7. DOI: 10.1007/978-81-322-2247-7_45.
- [54] J. Daugman. “High confidence visual recognition of persons by a test of statistical independence”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15.11 (Nov. 1993). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1148–1161. ISSN: 1939-3539. DOI: 10.1109/34.244676. URL: <https://ieeexplore.ieee.org/document/244676> (visited on 03/07/2025).
- [55] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. “ImageNet: A large-scale hierarchical image database”. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009 IEEE Conference on Computer Vision and Pattern Recognition. ISSN: 1063-6919. June 2009, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848. URL: <https://ieeexplore.ieee.org/document/5206848> (visited on 03/07/2025).
- [56] Y. Dodis, L. Reyzin, and A. Smith. “Fuzzy Extractors”. In: *Security with Noisy Data: On Private Biometrics, Secure Key Storage and Anti-Counterfeiting*. Ed. by P. Tuyls, B. Skoric, and T. Kevenaar. London: Springer, 2007, pp. 79–99. ISBN: 978-1-84628-984-2. DOI: 10.1007/978-1-84628-984-2_5. URL: https://doi.org/10.1007/978-1-84628-984-2_5 (visited on 05/21/2025).
- [57] R. Donida Labati, V. Piuri, and F. Scotti. “Neural-based quality measurement of fingerprint images in contactless biometric systems”. In: *The 2010 International Joint Conference on Neural Networks (IJCNN)*. The 2010 International Joint Conference on Neural Networks (IJCNN). ISSN: 2161-4407. July 2010, pp. 1–8. DOI: 10.1109/IJCNN.2010.5596694. URL: <https://ieeexplore.ieee.org/document/5596694> (visited on 05/21/2025).
- [58] *Dragon Skin™ 10 FAST Product Information*. Smooth-On, Inc. URL: <https://www.smooth-on.com/products/dragon-skin-10-fast/> (visited on 01/12/2024).
- [59] T. Duarte, J. P. Pimentão, P. Sousa, and S. Onofre. “Biometric access control systems: A review on technologies to improve their efficiency”. In: *2016 IEEE International Power Electronics and Motion Control Conference (PEMC)*. 2016 IEEE International Power Electronics and Motion Control Conference (PEMC). Sept. 2016, pp. 795–800. DOI: 10.1109/EPEPMC.2016.7752095. URL: <https://ieeexplore.ieee.org/abstract/document/7752095> (visited on 02/18/2025).
- [60] C. Edwards and R. Marks. “Evaluation of biomechanical properties of human skin”. In: *Clinics in Dermatology*. Bioengineering of the Skin 13.4 (July 1, 1995), pp. 375–380. ISSN: 0738-081X. DOI: 10.1016/0738-081X(95)00078-T. URL: <https://www.sciencedirect.com/science/article/pii/0738081X9500078T> (visited on 01/12/2024).

- [61] *ELECTRONIC BIOMETRIC TRANSMISSION SPECIFICATION (EBTS)*. 242-HQ-A6687913-SYSDOCU. Clarksburg: U.S. Department of Justice, FBI, Apr. 16, 2021, p. 298. URL: file:///home/ruzickal/Downloads/EBTS%20v11.0_Final_508.pdf (visited on 04/03/2024).
- [62] J. J. Engelsma, S. S. Arora, A. K. Jain, and N. G. Paulter Jr. “Universal 3D Wearable Fingerprint Targets: Advancing Fingerprint Reader Evaluations”. In: *arXiv:1705.07972 [cs]* (May 22, 2017). arXiv: 1705.07972. URL: <http://arxiv.org/abs/1705.07972> (visited on 11/25/2021).
- [63] J. J. Engelsma, S. A. Grosz, and A. K. Jain. “PrintsGAN: Synthetic Fingerprint Generator”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022), pp. 1–14. ISSN: 0162-8828, 2160-9292, 1939-3539. DOI: 10.1109/TPAMI.2022.3204591. URL: <https://ieeexplore.ieee.org/document/9893541/> (visited on 01/08/2024).
- [64] D. Escudero, S. Ghosh, M. Keller, R. Rachuri, and P. Scholl. “Improved primitives for MPC over mixed arithmetic-binary circuits”. In: *CRYPTO 2020*. CRYPTO 2020. tex.howpublished: Cryptology ePrint Archive, Paper 2020/338. 2020. URL: <https://eprint.iacr.org/2020/338>.
- [65] European Border and Coast Guard Agency. *Technology foresight on biometrics for the future of travel: executive summary*. LU: Publications Office, 2022. URL: <https://data.europa.eu/doi/10.2819/387848> (visited on 11/09/2023).
- [66] D. Evans, V. Kolesnikov, and M. Rosulek. “A Pragmatic Introduction to Secure Multi-Party Computation”. In: *Foundations and Trends® in Privacy and Security* 2.2 (Dec. 18, 2018). Publisher: Now Publishers, Inc., pp. 70–246. ISSN: 2474-1558, 2474-1566. DOI: 10.1561/33000000019. URL: <https://www.nowpublishers.com/article/Details/SEC-019> (visited on 05/21/2025).
- [67] V. Falanga and B. Bucalo. “Use of a durometer to assess skin hardness”. In: *Journal of the American Academy of Dermatology* 29.1 (July 1, 1993), pp. 47–51. ISSN: 0190-9622. DOI: 10.1016/0190-9622(93)70150-R. URL: <https://www.sciencedirect.com/science/article/pii/019096229370150R> (visited on 01/12/2024).
- [68] Federal Bureau of Investigation. *Technical Specifications Document for the Contactless Fingerprint Pilot Project Version 1.1*. Sept. 29, 2023. URL: https://fbibiospecs.fbi.gov/file-repository/contactless_pilot/contactless_fingerprint_pilot_technical_specifications_rev09292023-final-draft.pdf/view (visited on 06/24/2025).
- [69] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. “Object Detection with Discriminatively Trained Part-Based Models”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.9 (Sept. 2010). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1627–1645. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2009.167. URL: <https://ieeexplore.ieee.org/document/5255236> (visited on 03/07/2025).
- [70] J. Feng. “Combining minutiae descriptors for fingerprint matching”. In: *Pattern Recognition* 41.1 (Jan. 1, 2008), pp. 342–352. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2007.04.016. URL: <https://www.sciencedirect.com/science/article/pii/S003132030700204X> (visited on 03/07/2025).
- [71] J. Fierrez-Aguilar, L. Nanni, J. Ortega-Garcia, R. Cappelli, and D. Maltoni. “Combining Multiple Matchers for Fingerprint Verification: A Case Study in FVC2004”. In: *Image Analysis and Processing – ICIAP 2005*. Ed. by F. Roli and S. Vitulano. Berlin, Heidelberg: Springer, 2005, pp. 1035–1042. ISBN: 978-3-540-31866-8. DOI: 10.1007/11553595_127.

- [72] B. A. J. Fisher and D. R. Fisher. *Techniques of Crime Scene Investigation*. 9th ed. Boca Raton: CRC Press, July 5, 2022. 368 pp. ISBN: 978-0-429-27201-1. DOI: 10.4324/9780429272011.
- [73] G. Fiumara, P. Flanagan, J. Grantham, B. Bandini, K. Ko, and J. Libert. *NIST special database 300: uncompressed plain and rolled images from fingerprint cards*. NIST TN 1993. Gaithersburg, MD: National Institute of Standards and Technology, June 2018, NIST TN 1993. DOI: 10.6028/NIST.TN.1993. URL: <http://nvlpubs.nist.gov/nistpubs/TechnicalNotes/NIST.TN.1993.pdf> (visited on 06/07/2023).
- [74] G. Fiumara, M. Schwarz, J. Heising, J. Peterson, P. Flanagan, and K. Marshall. *NIST Special Database 302: Supplemental Release of Latent Annotations*. National Institute of Standards and Technology, Nov. 2, 2021. DOI: 10.6028/NIST.TN.2190. URL: <https://nvlpubs.nist.gov/nistpubs/TechnicalNotes/NIST.TN.2190.pdf> (visited on 06/07/2023).
- [75] D. Forsyth and J. Ponce. *Computer Vision: A Modern Approach: International Edition*. 2. Auflage. Harlow: Pearson Education, Limited, 2012. 1 p. ISBN: 978-1-292-01408-1.
- [76] Y. Freund and R. E. Schapire. “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”. In: *Journal of Computer and System Sciences* 55.1 (Aug. 1, 1997), pp. 119–139. ISSN: 0022-0000. DOI: 10.1006/jcss.1997.1504. URL: <https://www.sciencedirect.com/science/article/pii/S002200009791504X> (visited on 03/07/2025).
- [77] K. Fukushima. “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”. In: *Biological Cybernetics* 36.4 (Apr. 1, 1980), pp. 193–202. ISSN: 1432-0770. DOI: 10.1007/BF00344251. URL: <https://doi.org/10.1007/BF00344251> (visited on 03/10/2025).
- [78] K.-Y. Fung and X. Jing. “Measurement and Characterization of Acoustic Impedance by the Impulse Response of Harmonic Oscillators”. In: *10th AIAA/CEAS Aeroacoustics Conference*. Aeroacoustics Conferences. American Institute of Aeronautics and Astronautics, May 10, 2004. DOI: 10.2514/6.2004-2840. URL: <https://arc.aiaa.org/doi/10.2514/6.2004-2840> (visited on 03/03/2025).
- [79] J. Galbally, J. Fierrez, and R. Cappelli. “An Introduction to Fingerprint Presentation Attack Detection”. In: *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection*. Ed. by S. Marcel, M. S. Nixon, J. Fierrez, and N. Evans. Advances in Computer Vision and Pattern Recognition. Cham: Springer International Publishing, 2019, pp. 3–31. ISBN: 978-3-319-92627-8. DOI: 10.1007/978-3-319-92627-8_1. URL: https://doi.org/10.1007/978-3-319-92627-8_1 (visited on 03/07/2024).
- [80] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez. *A Review on Deep Learning Techniques Applied to Semantic Segmentation*. Apr. 22, 2017. DOI: 10.48550/arXiv.1704.06857. arXiv: 1704.06857 [cs]. URL: <http://arxiv.org/abs/1704.06857> (visited on 08/13/2024).
- [81] C. Genest and B. Rémillard. “Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models”. In: *Annales de l’I.H.P. Probabilités et statistiques* 44.6 (2008), pp. 1096–1127. ISSN: 1778-7017. DOI: 10.1214/07-AIHP148. (Visited on 06/16/2023).
- [82] S. Ghosh, N. Das, I. Das, and U. Maulik. *Understanding Deep Learning Techniques for Image Segmentation*. July 13, 2019. DOI: 10.48550/arXiv.1907.06119. arXiv: 1907.06119 [cs]. URL: <http://arxiv.org/abs/1907.06119> (visited on 08/13/2024).

- [83] R. Girshick. *Fast R-CNN*. Sept. 27, 2015. DOI: 10.48550/arXiv.1504.08083. arXiv: 1504.08083[cs]. URL: <http://arxiv.org/abs/1504.08083> (visited on 03/18/2025).
- [84] X. Glorot and Y. Bengio. “Understanding the difficulty of training deep feedforward neural networks”. In: *Appearing in Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)* (2010).
- [85] R. C. Gonzalez and R. E. Woods. *Digital image processing*. Fourth, global edition. New York, New York: Pearson Education, 2018. 1 p. ISBN: 978-1-292-22304-9.
- [86] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. Adaptive computation and machine learning. Cambridge, Mass: The MIT press, 2016. ISBN: 978-0-262-03561-3.
- [87] S. A. Grosz, J. J. Engelsma, E. Liu, and A. K. Jain. *C2CL: Contact to Contactless Fingerprint Matching*. Dec. 9, 2021. arXiv: 2104.02811[cs,eess]. URL: <http://arxiv.org/abs/2104.02811> (visited on 10/05/2022).
- [88] S. A. Grosz, J. J. Engelsma, R. Ranjan, N. Ramakrishnan, M. Aggarwal, G. G. Medioni, and A. K. Jain. *Minutiae-Guided Fingerprint Embeddings via Vision Transformers*. Oct. 26, 2022. DOI: 10.48550/arXiv.2210.13994. arXiv: 2210.13994[cs]. URL: <http://arxiv.org/abs/2210.13994> (visited on 03/06/2025).
- [89] P. Grother and E. Tabassi. “Performance of Biometric Quality Measures”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.4 (Apr. 2007), pp. 531–543. ISSN: 0162-8828, 2160-9292. DOI: 10.1109/TPAMI.2007.1019. URL: <http://ieeexplore.ieee.org/document/4107559/> (visited on 04/17/2025).
- [90] B. Günay. “Biometrische Daten aus der Perspektive der DSGVO”. In: *Datenschutz und Datensicherheit - DuD* 47.2 (Feb. 1, 2023), pp. 96–99. ISSN: 1862-2607. DOI: 10.1007/s11623-023-1724-x. URL: <https://doi.org/10.1007/s11623-023-1724-x> (visited on 04/03/2024).
- [91] H. Han and Y. Koshimoto. “Characteristics of thermal-type fingerprint sensor”. In: *Biometric Technology for Human Identification V*. Biometric Technology for Human Identification V. Vol. 6944. SPIE, Mar. 17, 2008, pp. 214–225. DOI: 10.1117/12.777024. URL: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/6944/69440P/Characteristics-of-thermal-type-fingerprint-sensor/10.1117/12.777024.full> (visited on 03/03/2025).
- [92] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. 2. edition, 17.printing. Cambridge: Cambridge Univ. Press, 2018. 655 pp. ISBN: 978-0-521-54051-3.
- [93] C. Hazirbas, L. Ma, C. Domokos, and D. Cremers. “FuseNet: Incorporating Depth into Semantic Segmentation via Fusion-Based CNN Architecture”. In: *Computer Vision – ACCV 2016*. Ed. by S.-H. Lai, V. Lepetit, K. Nishino, and Y. Sato. Cham: Springer International Publishing, 2017, pp. 213–228. ISBN: 978-3-319-54181-5. DOI: 10.1007/978-3-319-54181-5_14.
- [94] K. He, G. Gkioxari, P. Dollár, and R. Girshick. “Mask R-CNN”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017 IEEE International Conference on Computer Vision (ICCV). ISSN: 2380-7504. Oct. 2017, pp. 2980–2988. DOI: 10.1109/ICCV.2017.322. URL: <https://ieeexplore.ieee.org/document/8237584/?arnumber=8237584> (visited on 08/26/2024).
- [95] K. He, X. Zhang, S. Ren, and J. Sun. *Deep Residual Learning for Image Recognition*. Dec. 10, 2015. DOI: 10.48550/arXiv.1512.03385. arXiv: 1512.03385[cs]. URL: <http://arxiv.org/abs/1512.03385> (visited on 03/26/2023).

- [96] K. He, X. Zhang, S. Ren, and J. Sun. *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. Feb. 6, 2015. DOI: 10.48550/arXiv.1502.01852. arXiv: 1502.01852[cs]. URL: <http://arxiv.org/abs/1502.01852> (visited on 06/25/2025).
- [97] Y. He, J. Tian, X. Luo, and T. Zhang. “Image enhancement and minutiae matching in fingerprint verification”. In: *Pattern Recognition Letters* 24.9 (June 1, 2003), pp. 1349–1360. ISSN: 0167-8655. DOI: 10.1016/S0167-8655(02)00376-8. URL: <https://www.sciencedirect.com/science/article/pii/S0167865502003768> (visited on 03/07/2025).
- [98] B. Hiew, A. B. Teoh, and Y. Pang. “Touch-less Fingerprint Recognition System”. In: *2007 IEEE Workshop on Automatic Identification Advanced Technologies*. 2007 IEEE Workshop on Automatic Identification Advanced Technologies. June 2007, pp. 24–29. DOI: 10.1109/AUTOID.2007.380587. URL: <https://ieeexplore.ieee.org/document/4263208> (visited on 05/21/2025).
- [99] L. Hong, Y. Wan, and A. Jain. “Fingerprint image enhancement: algorithm and performance evaluation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.8 (Aug. 1998). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 777–789. ISSN: 1939-3539. DOI: 10.1109/34.709565. URL: <https://ieeexplore.ieee.org/document/709565> (visited on 03/03/2025).
- [100] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. Apr. 17, 2017. DOI: 10.48550/arXiv.1704.04861. arXiv: 1704.04861[cs]. URL: <http://arxiv.org/abs/1704.04861> (visited on 03/10/2025).
- [101] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. “Densely Connected Convolutional Networks”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). ISSN: 1063-6919. July 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243. URL: <https://ieeexplore.ieee.org/document/8099726> (visited on 03/10/2025).
- [102] Y. Huang, L. Malka, D. Evans, and J. Katz. “Efficient privacy-preserving biometric identification.” In: *Proceedings of the 18th Network and Distributed System Security Conference (NDSS 2011)*. Jan. 2011.
- [103] *IDKit SDK*. Innovatrics. URL: <https://support.innovatrics.com/support/solutions/folders/5000267908> (visited on 03/06/2023).
- [104] G. of India. *About Aadhaar*. About Aadhaar. Feb. 18, 2025. URL: https://uidai.gov.in/en/?option=com_content&view=article&id=14 (visited on 02/18/2025).
- [105] S. Ioffe and C. Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. Mar. 2, 2015. DOI: 10.48550/arXiv.1502.03167. arXiv: 1502.03167[cs]. URL: <http://arxiv.org/abs/1502.03167> (visited on 03/06/2023).
- [106] A. Jain, L. Hong, and R. Bolle. “On-line fingerprint verification”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19.4 (Apr. 1997). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 302–314. ISSN: 1939-3539. DOI: 10.1109/34.587996. URL: <https://ieeexplore.ieee.org/document/587996/?arnumber=587996> (visited on 11/19/2024).
- [107] A. Jain, S. Prabhakar, L. Hong, and S. Pankanti. “Filterbank-based fingerprint matching”. In: *IEEE Transactions on Image Processing* 9.5 (May 2000), pp. 846–859. ISSN: 10577149. DOI: 10.1109/83.841531. URL: <http://ieeexplore.ieee.org/document/841531/> (visited on 03/06/2025).

- [108] A. Jain, A. Ross, and S. Pankanti. “Biometrics: a tool for information security”. In: *IEEE Transactions on Information Forensics and Security* 1.2 (June 2006). Conference Name: IEEE Transactions on Information Forensics and Security, pp. 125–143. ISSN: 1556-6021. DOI: 10.1109/TIFS.2006.873653. URL: <https://ieeexplore.ieee.org/document/1634356> (visited on 03/03/2025).
- [109] A. Jain and A. Ross. “Fingerprint mosaicking”. In: *IEEE International Conference on Acoustics Speech and Signal Processing*. Proceedings of ICASSP ’02. Orlando, FL, USA: IEEE, May 2002, pp. IV–4064–IV–4067. ISBN: 978-0-7803-7402-7. DOI: 10.1109/ICASSP.2002.5745550. URL: <http://ieeexplore.ieee.org/document/5745550/> (visited on 11/25/2021).
- [110] A. K. Jain and J. Feng. “Latent Fingerprint Matching”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.1 (Jan. 2011). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 88–100. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2010.59. URL: <https://ieeexplore.ieee.org/abstract/document/5432204> (visited on 03/07/2025).
- [111] A. Jandyal, I. Chaturvedi, I. Wazir, A. Raina, and M. I. Ul Haq. “3D printing – A review of processes, materials and applications in industry 4.0”. In: *Sustainable Operations and Computers* 3 (Jan. 1, 2022), pp. 33–42. ISSN: 2666-4127. DOI: 10.1016/j.susc.2021.09.004. URL: <https://www.sciencedirect.com/science/article/pii/S2666412721000441> (visited on 01/08/2024).
- [112] B. Jawade, A. Agarwal, S. Setlur, and N. Ratha. “Multi Loss Fusion For Matching Smartphone Captured Contactless Finger Images”. In: *2021 IEEE International Workshop on Information Forensics and Security (WIFS)*. 2021 IEEE International Workshop on Information Forensics and Security (WIFS). ISSN: 2157-4774. Dec. 2021, pp. 1–6. DOI: 10.1109/WIFS53200.2021.9648393. URL: <https://ieeexplore.ieee.org/document/9648393> (visited on 05/21/2025).
- [113] T.-Y. Jea and V. Govindaraju. “A minutia-based partial fingerprint recognition system”. In: *Pattern Recognition* 38.10 (Oct. 1, 2005), pp. 1672–1684. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2005.03.016. URL: <https://www.sciencedirect.com/science/article/pii/S0031320305001445> (visited on 03/07/2025).
- [114] X. Jiang and W.-Y. Yau. “Fingerprint minutiae matching based on the local and global structures”. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000. Vol. 2. ISSN: 1051-4651. Sept. 2000, 1038–1041 vol.2. DOI: 10.1109/ICPR.2000.906252. URL: <https://ieeexplore.ieee.org/document/906252> (visited on 03/07/2025).
- [115] K. Karu and A. K. Jain. “Fingerprint classification”. In: *Pattern Recognition* 29.3 (Mar. 1, 1996), pp. 389–404. ISSN: 0031-3203. DOI: 10.1016/0031-3203(95)00106-9. URL: <https://www.sciencedirect.com/science/article/pii/0031320395001069> (visited on 06/07/2023).
- [116] C. Kauba, D. Söllinger, S. Kirchgasser, A. Weissenfeld, G. Fernández Domínguez, B. Strobl, and A. Uhl. “Towards Using Police Officers’ Business Smartphones for Contactless Fingerprint Acquisition and Enabling Fingerprint Comparison against Contact-Based Datasets”. In: *Sensors* 21.7 (2021). ISSN: 1424-8220. DOI: 10.3390/s21072248. URL: <https://www.mdpi.com/1424-8220/21/7/2248>.

- [117] R. Kaur, P. S. Sandhu, and A. Kamra. “A novel method for fingerprint feature extraction”. In: *2010 International Conference on Networking and Information Technology*. 2010 International Conference on Networking and Information Technology. ISSN: 2324-8203. June 2010, pp. 1–5. DOI: 10.1109/ICNIT.2010.5508569. URL: <https://ieeexplore.ieee.org/document/5508569> (visited on 03/06/2025).
- [118] H. Kekre and V. Bharadi. “Fingerprint Core Point Detection Algorithm using Orientation Field Based Multiple Features”. In: *International Journal of Computer Applications* 1 (Feb. 25, 2010). DOI: 10.5120/314-482.
- [119] M. Keller. “MP-SPDZ: A Versatile Framework for Multi-Party Computation”. In: *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. CCS '20: 2020 ACM SIGSAC Conference on Computer and Communications Security. Virtual Event USA: ACM, Oct. 30, 2020, pp. 1575–1590. ISBN: 978-1-4503-7089-9. DOI: 10.1145/3372297.3417872. URL: <https://dl.acm.org/doi/10.1145/3372297.3417872> (visited on 05/21/2025).
- [120] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. Jan. 30, 2017. DOI: 10.48550/arXiv.1412.6980. arXiv: 1412.6980[cs]. URL: <http://arxiv.org/abs/1412.6980> (visited on 03/10/2025).
- [121] I. Kojadinovic and J. Yan. “Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap”. In: *Canadian Journal of Statistics* 40 (Sept. 1, 2012). DOI: 10.1002/cjs.11135.
- [122] J. Kolberg, J. Priesnitz, C. Rathgeb, and C. Busch. “COLFISPOOF: A New Database for Contactless Fingerprint Presentation Attack Detection Research”. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE/CVF Winter Conference on Applications of Computer Vision. 2023, pp. 653–661. URL: https://openaccess.thecvf.com/content/WACV2023W/MAA-A/html/Kolberg_COLFISPOOF_A_New_Database_for_Contactless_Fingerprint_Presentation_Attack_Detection_WACVW_2023_paper.html (visited on 03/07/2024).
- [123] V. Kolesnikov and T. Schneider. “Improved Garbled Circuit: Free XOR Gates and Applications”. In: *Automata, Languages and Programming*. Ed. by L. Aceto, I. Damgård, L. A. Goldberg, M. M. Halldórsson, A. Ingólfssdóttir, and I. Walukiewicz. Vol. 5126. ISSN: 0302-9743, 1611-3349 Series Title: Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 486–498. ISBN: 978-3-540-70582-6. DOI: 10.1007/978-3-540-70583-3_40. URL: http://link.springer.com/10.1007/978-3-540-70583-3_40 (visited on 05/21/2025).
- [124] P. Krähenbühl and V. Koltun. *Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials*. Oct. 20, 2012. DOI: 10.48550/arXiv.1210.5644. arXiv: 1210.5644[cs]. URL: <http://arxiv.org/abs/1210.5644> (visited on 05/22/2025).
- [125] A. Krizhevsky, I. Sutskever, and G. E. Hinton. “ImageNet classification with deep convolutional neural networks”. In: *Communications of the ACM* 60.6 (May 24, 2017), pp. 84–90. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/3065386. URL: <https://dl.acm.org/doi/10.1145/3065386> (visited on 03/07/2025).
- [126] A. Kumar. “Introduction to Trends in Fingerprint Identification”. In: *Contactless 3D Fingerprint Identification*. Ed. by A. Kumar. Cham: Springer International Publishing, 2018, pp. 1–15. ISBN: 978-3-319-67681-4. DOI: 10.1007/978-3-319-67681-4_1. URL: https://doi.org/10.1007/978-3-319-67681-4_1 (visited on 02/18/2025).

- [127] A. Kumar and Y. Zhou. “Human identification using finger images”. In: *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society* 21.4 (Apr. 2012), pp. 2228–2244. ISSN: 1941-0042. DOI: 10.1109/TIP.2011.2171697.
- [128] S. Kumar, D. S. Bhagat, and G. S. Bumbrah. “1,2-Indanedione (IND) Reagent for The Detection of Latent Fingermarks: A Review”. In: *Arab Journal of Forensic Sciences & Forensic Medicine* 3.1 (June 1, 2021). Number: 1, pp. 77–93. ISSN: 1658-6794. DOI: 10.26735/NRYD9355. URL: <https://journals.nauss.edu.sa/index.php/AJFSFM/article/view/1466> (visited on 03/03/2025).
- [129] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (Nov. 1998). Conference Name: Proceedings of the IEEE, pp. 2278–2324. ISSN: 1558-2256. DOI: 10.1109/5.726791. URL: <https://ieeexplore.ieee.org/document/726791> (visited on 03/10/2025).
- [130] M. Leider. “On the Weight of the Skin**From the New York Skin and Cancer Unit, Department of Dermatology and Syphilology (Marion B. Sulzberger, M.D. Director) and the Post Graduate Medical School of the New York University-Bellevue Medical Center.” In: *Journal of Investigative Dermatology* 12.3 (Mar. 1, 1949), pp. 187–191. ISSN: 0022-202X. DOI: 10.1038/jid.1949.28. URL: <https://www.sciencedirect.com/science/article/pii/S0022202X15503679> (visited on 01/12/2024).
- [131] M. Lepley. *Test Procedures for Verifying Image Quality Requirements for Personal Identity Verification (PIV) Single Finger Capture Devices*. MITRE TECHNICAL REPORT MTR060170R5. Bedford: Center for Integrated Intelligence Systems, Dec. 2006, p. 59.
- [132] G. Li, B. Yang, M. A. Olsen, and C. Busch. “Quality Assessment for Fingerprints Collected by Smartphone Cameras”. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). OR, USA: IEEE, June 2013, pp. 146–153. ISBN: 978-0-7695-4990-3. DOI: 10.1109/CVPRW.2013.29. URL: <http://ieeexplore.ieee.org/document/6595867/> (visited on 06/25/2021).
- [133] J. Libert, J. Grantham, B. Bandini, S. Wood, M. Garris, K. Ko, F. Byers, and C. Watson. *Guidance for evaluating contactless fingerprint acquisition devices*. NIST SP 500-305. Gaithersburg, MD: National Institute of Standards and Technology, July 2018, NIST SP 500–305. DOI: 10.6028/NIST.SP.500-305. URL: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-305.pdf> (visited on 03/20/2023).
- [134] J. M. Libert, J. Grantham, and S. Orandi. *A 1D spectral image validation verification metric for fingerprints*. NIST IR 7599. Edition: 0. Gaithersburg, MD: National Institute of Standards and Technology, 2009, NIST IR 7599. DOI: 10.6028/NIST.IR.7599. URL: <https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7599.pdf> (visited on 07/11/2024).
- [135] H. W. Lilliefors. “On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown”. In: *Journal of the American Statistical Association* 62.318 (June 1967), pp. 399–402. ISSN: 0162-1459, 1537-274X. DOI: 10.1080/01621459.1967.10482916. URL: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1967.10482916> (visited on 06/16/2023).
- [136] C. Lin and A. Kumar. “Matching Contactless and Contact-Based Conventional Fingerprint Images for Biometrics Identification”. In: *IEEE Transactions on Image Processing* 27.4 (Apr. 2018), pp. 2008–2021. ISSN: 1057-7149, 1941-0042. DOI: 10.1109/TIP.2017.2788866. URL: <http://ieeexplore.ieee.org/document/8244291/> (visited on 11/25/2021).

- [137] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. *Feature Pyramid Networks for Object Detection*. Apr. 19, 2017. DOI: 10.48550/arXiv.1612.03144. arXiv: 1612.03144[cs]. URL: <http://arxiv.org/abs/1612.03144> (visited on 08/09/2024).
- [138] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar. “Focal Loss for Dense Object Detection”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, Oct. 2017, pp. 2999–3007. ISBN: 978-1-5386-1032-9. DOI: 10.1109/ICCV.2017.324. URL: <http://ieeexplore.ieee.org/document/8237586/> (visited on 11/19/2024).
- [139] T. Lindeberg. “Feature Detection with Automatic Scale Selection”. In: *International Journal of Computer Vision* 30.2 (Nov. 1, 1998), pp. 79–116. ISSN: 1573-1405. DOI: 10.1023/A:1008045108935. URL: <https://doi.org/10.1023/A:1008045108935> (visited on 03/07/2025).
- [140] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez. “A survey on deep learning in medical image analysis”. In: *Medical Image Analysis* 42 (Dec. 1, 2017), pp. 60–88. ISSN: 1361-8415. DOI: 10.1016/j.media.2017.07.005. URL: <https://www.sciencedirect.com/science/article/pii/S1361841517301135> (visited on 03/07/2025).
- [141] F. Liu and D. Zhang. “3D fingerprint reconstruction system using feature correspondences and prior estimated finger model”. In: *Pattern Recognition* 47.1 (Jan. 1, 2014), pp. 178–193. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2013.06.009. URL: <https://www.sciencedirect.com/science/article/pii/S0031320313002616> (visited on 10/22/2021).
- [142] F. Liu, D. Zhang, C. Song, and G. Lu. “Touchless Multiview Fingerprint Acquisition and Mosaicking”. In: *IEEE Transactions on Instrumentation and Measurement* 62.9 (Sept. 2013). Conference Name: IEEE Transactions on Instrumentation and Measurement, pp. 2492–2502. ISSN: 1557-9662. DOI: 10.1109/TIM.2013.2258248. URL: <https://ieeexplore.ieee.org/document/6562785> (visited on 07/11/2024).
- [143] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. “SSD: Single Shot MultiBox Detector”. In: vol. 9905. 2016, pp. 21–37. DOI: 10.1007/978-3-319-46448-0_2. arXiv: 1512.02325[cs]. URL: <http://arxiv.org/abs/1512.02325> (visited on 03/07/2025).
- [144] X. Liu. “Multi-modality quality assessment for unconstrained biometric samples”. PhD thesis. France: Université de Caen Normandie, June 22, 2018. 225 pp. URL: https://theses.hal.science/tel-02299278/file/sygal_fusion_28726-liu-xinwei.pdf.
- [145] Y. Liu, T. Zhou, Z. Yue, W. Liu, L. Han, Q. Li, and X. Yang. “Secure and efficient online fingerprint authentication scheme based on cloud computing”. In: *IEEE Transactions on Cloud Computing* PP (Aug. 2021), pp. 1–1. DOI: 10.1109/TCC.2021.3103546.
- [146] J. Long, E. Shelhamer, and T. Darrell. “Fully convolutional networks for semantic segmentation”. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). ISSN: 1063-6919. June 2015, pp. 3431–3440. DOI: 10.1109/CVPR.2015.7298965. URL: <https://ieeexplore.ieee.org/document/7298965> (visited on 03/10/2025).
- [147] T. Lorünser and F. Wohner. “Performance Comparison of Two Generic MPC-frameworks with Symmetric Ciphers.” in: *Proceedings of the 17th International Joint Conference on e-Business and Telecommunications*. 17th International Conference on Security and Cryptography. Lieusaint - Paris, France: SCITEPRESS - Science and Technology Publications, 2020, pp. 587–594. ISBN: 978-989-758-446-6. DOI: 10.5220/0009831705870594.

- URL: <https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0009831705870594> (visited on 05/21/2025).
- [148] D. G. Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. In: *International Journal of Computer Vision* 60.2 (Nov. 1, 2004), pp. 91–110. ISSN: 1573-1405. DOI: 10.1023/B:VISI.0000029664.99615.94. URL: <https://doi.org/10.1023/B:VISI.0000029664.99615.94> (visited on 03/07/2025).
- [149] A. V. Maceo. “Anatomy and Physiology of Adult Friction Ridge Skin”. In: *The Fingerprint Sourcebook*. 1st. NCJ 225320. National Institute of Justice, July 2011, p. 422. URL: <https://www.ojp.gov/library/publications/fingerprint-sourcebook>.
- [150] J. MacQueen. “Some methods for classification and analysis of multivariate observations”. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Vol. 5.1. University of California Press, Jan. 1, 1967, pp. 281–298. URL: <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992> (visited on 03/07/2025).
- [151] J. Mader and T. Lorünser. “Feasibility of Privacy Preserving Minutiae-Based Fingerprint Matching:” in: *Proceedings of the 10th International Conference on Information Systems Security and Privacy*. 10th International Conference on Information Systems Security and Privacy. Rome, Italy: SCITEPRESS - Science and Technology Publications, 2024, pp. 899–906. ISBN: 978-989-758-683-5. DOI: 10.5220/0012472300003648. URL: <https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0012472300003648> (visited on 02/18/2025).
- [152] J. Mader, F. Wohner, L. Ruzicka, and T. Loruenser. “Towards Real-Time Privacy-Preserving Minutiae Matching”. In: *Proceedings of the 23rd Workshop on Privacy in the Electronic Society*. WPES ’24. New York, NY, USA: Association for Computing Machinery, Nov. 21, 2024, pp. 98–108. ISBN: 979-8-4007-1239-5. DOI: 10.1145/3689943.3695049. URL: <https://dl.acm.org/doi/10.1145/3689943.3695049> (visited on 11/24/2024).
- [153] A. Maeva and F. Severin. “High resolution ultrasonic method for 3D fingerprint recognizable characteristics in biometrics identification”. In: *2009 IEEE International Ultrasonics Symposium*. 2009 IEEE International Ultrasonics Symposium. ISSN: 1948-5727. Sept. 2009, pp. 2260–2263. DOI: 10.1109/ULTSYM.2009.5441399. URL: <https://ieeexplore.ieee.org/document/5441399> (visited on 03/03/2025).
- [154] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain. “FVC2000: fingerprint verification competition”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.3 (Mar. 2002). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 402–412. ISSN: 1939-3539. DOI: 10.1109/34.990140. URL: <https://ieeexplore.ieee.org/document/990140> (visited on 03/03/2025).
- [155] A. Makrushin, V. S. Mannam, and J. Dittmann. “Privacy-Friendly Datasets of Synthetic Fingerprints for Evaluation of Biometric Algorithms”. In: *Applied Sciences* 13.18 (Jan. 2023). Number: 18 Publisher: Multidisciplinary Digital Publishing Institute, p. 10000. ISSN: 2076-3417. DOI: 10.3390/app131810000. URL: <https://www.mdpi.com/2076-3417/13/18/10000> (visited on 01/08/2024).

- [156] D. Maltoni and R. Cappelli. “Fingerprint Recognition”. In: *Handbook of Biometrics*. Ed. by A. K. Jain, P. Flynn, and A. A. Ross. Boston, MA: Springer US, 2008, pp. 23–42. ISBN: 978-0-387-71041-9. DOI: 10.1007/978-0-387-71041-9_2. URL: https://doi.org/10.1007/978-0-387-71041-9_2 (visited on 03/06/2025).
- [157] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. London: Springer London, 2009. ISBN: 978-1-84882-253-5. DOI: 10.1007/978-1-84882-254-2. URL: <http://link.springer.com/10.1007/978-1-84882-254-2> (visited on 11/19/2024).
- [158] L. Manhua, X. Jiang, and A. Kot. “Fingerprint Reference-Point Detection”. In: *EURASIP Journal on Advances in Signal Processing* 2005 (Mar. 15, 2005). DOI: 10.1155/ASP.2005.498.
- [159] S. Marcel, M. S. Nixon, J. Fierrez, and N. Evans, eds. *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection*. Advances in Computer Vision and Pattern Recognition. Cham: Springer International Publishing, 2019. ISBN: 978-3-319-92626-1. DOI: 10.1007/978-3-319-92627-8. URL: <http://link.springer.com/10.1007/978-3-319-92627-8> (visited on 04/03/2024).
- [160] M. Micheletto, G. Orrù, R. Casula, D. Yambay, G. L. Marcialis, and S. Schuckers. “Review of the Fingerprint Liveness Detection (LivDet) Competition Series: From 2009 to 2021”. In: *Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability Assessment*. Ed. by S. Marcel, J. Fierrez, and N. Evans. Advances in Computer Vision and Pattern Recognition. Singapore: Springer Nature, 2023, pp. 57–76. ISBN: 978-981-19528-8-3. DOI: 10.1007/978-981-19-5288-3_3. URL: https://doi.org/10.1007/978-981-19-5288-3_3 (visited on 03/07/2024).
- [161] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos. *Image Segmentation Using Deep Learning: A Survey*. Nov. 14, 2020. arXiv: 2001.05566 [cs]. URL: <http://arxiv.org/abs/2001.05566> (visited on 09/18/2023).
- [162] A. Mittal, A. K. Moorthy, and A. C. Bovik. “No-Reference Image Quality Assessment in the Spatial Domain”. In: *IEEE Transactions on Image Processing* 21.12 (Dec. 2012), pp. 4695–4708. ISSN: 1057-7149, 1941-0042. DOI: 10.1109/TIP.2012.2214050. URL: <http://ieeexplore.ieee.org/document/6272356/> (visited on 10/17/2023).
- [163] Y. Moon, H. Yeung, K. Chan, and S. Chan. “Template synthesis and image mosaicking for fingerprint registration: an experimental study”. In: *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. Vol. 5. ISSN: 1520-6149. May 2004, pp. V–409. DOI: 10.1109/ICASSP.2004.1327134. URL: <https://ieeexplore.ieee.org/document/1327134/?arnumber=1327134> (visited on 07/11/2024).
- [164] M. G. S. Murshed, K. Bahmani, S. Schuckers, and F. Hussain. *Deep Age-Invariant Fingerprint Segmentation System*. Mar. 6, 2023. DOI: 10.48550/arXiv.2303.03341. arXiv: 2303.03341 [cs]. URL: <http://arxiv.org/abs/2303.03341> (visited on 06/10/2024).
- [165] M. G. S. Murshed, R. Kline, K. Bahmani, F. Hussain, and S. Schuckers. *Deep Slap Fingerprint Segmentation for Juveniles and Adults*. May 3, 2022. DOI: 10.48550/arXiv.2110.04067. arXiv: 2110.04067 [cs, eess]. URL: <http://arxiv.org/abs/2110.04067> (visited on 10/02/2023).
- [166] S. Nabil, F. Devernay, and J. L. Crowley. “Error Detection in Panoramic Videos: a Pairwise Assessment within Stitching”. 2017. URL: <https://hal.science/hal-01849267> (visited on 07/11/2024).

- [167] V. Nair and G. E. Hinton. “Rectified Linear Units Improve Restricted Boltzmann Machines”. In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (2010).
- [168] National Institute of Standards and Technology. *American national standard for information systems*. NIST SP 500-245. Edition: 0. Gaithersburg, MD: National Institute of Standards and Technology, 2000, NIST SP 500–245. DOI: 10.6028/NIST.SP.500-245. URL: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication500-245.pdf> (visited on 06/07/2023).
- [169] S. Nayar and Y. Nakagawa. “Shape from focus”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.8 (Aug. 1994). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 824–831. ISSN: 1939-3539. DOI: 10.1109/34.308479. URL: <https://ieeexplore.ieee.org/document/308479> (visited on 03/03/2025).
- [170] M. Nazarkevych, V. Dordiak, V. Brytkovskyi, K. Pelekh, I. Pikh, and Y. Voznyi. “Comparison of the Effectiveness of Fingerprint Skeletal Methods”. In: *2019 XIth International Scientific and Practical Conference on Electronics and Information Technologies (ELIT)*. 2019 XIth International Scientific and Practical Conference on Electronics and Information Technologies (ELIT). Sept. 2019, pp. 19–23. DOI: 10.1109/ELIT.2019.8892293. URL: <https://ieeexplore.ieee.org/document/8892293> (visited on 03/06/2025).
- [171] A. Neubeck and L. Van Gool. “Efficient Non-Maximum Suppression”. In: *18th International Conference on Pattern Recognition (ICPR’06)*. 18th International Conference on Pattern Recognition (ICPR’06). Vol. 3. ISSN: 1051-4651. Aug. 2006, pp. 850–855. DOI: 10.1109/ICPR.2006.479. URL: <https://ieeexplore.ieee.org/document/1699659> (visited on 03/07/2025).
- [172] A. Y. Ng. “Feature selection, L1 vs. L2 regularization, and rotational invariance”. In: *Proceedings of the twenty-first international conference on Machine learning*. ICML ’04. New York, NY, USA: Association for Computing Machinery, July 4, 2004, p. 78. ISBN: 978-1-58113-838-2. DOI: 10.1145/1015330.1015435. URL: <https://dl.acm.org/doi/10.1145/1015330.1015435> (visited on 03/10/2025).
- [173] S. B. Nikam and S. Agarwal. “Local binary pattern and wavelet-based spoof fingerprint detection”. In: *International Journal of Biometrics* 1.2 (2008), p. 141. ISSN: 1755-8301, 1755-831X. DOI: 10.1504/IJBM.2008.020141. URL: <http://www.inderscience.com/link.php?id=20141> (visited on 03/07/2025).
- [174] J. Nilsson and T. Akenine-Möller. *Understanding SSIM*. June 29, 2020. DOI: 10.48550/arXiv.2006.13846. arXiv: 2006.13846[eess]. URL: <http://arxiv.org/abs/2006.13846> (visited on 05/21/2025).
- [175] M. S. Nixon and A. S. Aguado. *Feature extraction and image processing*. 2. ed., reprinted. Amsterdam: Acad. Press, 2010. 406 pp. ISBN: 978-0-12-372538-7.
- [176] K. Okerefor, I. Ekong, I. O. Markson, and K. Enwere. “Fingerprint Biometric System Hygiene and the Risk of COVID-19 Transmission”. In: *JMIR Biomedical Engineering* 5.1 (Sept. 8, 2020). Company: JMIR Biomedical Engineering Distributor: JMIR Biomedical Engineering Institution: JMIR Biomedical Engineering Label: JMIR Biomedical Engineering Publisher: JMIR Publications Inc., Toronto, Canada, e19623. DOI: 10.2196/19623. URL: <https://biomedeng.jmir.org/2020/1/e19623> (visited on 10/04/2022).

- [177] C. W. J. Oomens, M. van Vijven, and G. W. M. Peters. “Chapter 16 - Skin Mechanics”. In: *Biomechanics of Living Organs*. Ed. by Y. Payan and J. Ohayon. Vol. 1. Translational Epigenetics. Oxford: Academic Press, Jan. 1, 2017, pp. 347–357. DOI: 10.1016/B978-0-12-804009-6.00016-X. URL: <https://www.sciencedirect.com/science/article/pii/B978012804009600016X> (visited on 01/09/2024).
- [178] A. Opiah. *South Africa adopts biometrics for social protection, following India’s example / Biometric Update*. biometric update. Nov. 22, 2024. URL: <https://www.biometricupdate.com/202411/south-africa-adopts-biometrics-for-social-protection-following-indias-example> (visited on 02/18/2025).
- [179] S. Orandi, J. M. Libert, J. Grantham, K. Ko, B. Bandini, and C. I. Watson. *Specification for certification testing of contactless fingerprint acquisition devices, v1.0*. NIST SP 500-339. Gaithersburg, MD: National Institute of Standards and Technology (U.S.), Apr. 4, 2023, NIST SP 500–339. DOI: 10.6028/NIST.SP.500-339. URL: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.500-339.pdf> (visited on 01/09/2024).
- [180] N. Otsu. “A Threshold Selection Method from Gray-Level Histograms”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (Jan. 1979). Conference Name: IEEE Transactions on Systems, Man, and Cybernetics, pp. 62–66. ISSN: 2168-2909. DOI: 10.1109/TSMC.1979.4310076. URL: <https://ieeexplore.ieee.org/document/4310076> (visited on 03/06/2025).
- [181] S. J. Pan and Q. Yang. “A Survey on Transfer Learning”. In: *IEEE Transactions on Knowledge and Data Engineering* 22.10 (Oct. 2010). Conference Name: IEEE Transactions on Knowledge and Data Engineering, pp. 1345–1359. ISSN: 1558-2191. DOI: 10.1109/TKDE.2009.191. URL: <https://ieeexplore.ieee.org/document/5288526> (visited on 03/10/2025).
- [182] G. Parziale and Y. Chen. “Advanced Technologies for Touchless Fingerprint Recognition”. In: *Handbook of Remote Biometrics: for Surveillance and Security*. Ed. by M. Tistarelli, S. Z. Li, and R. Chellappa. London: Springer, 2009, pp. 83–109. ISBN: 978-1-84882-385-3. DOI: 10.1007/978-1-84882-385-3_4. URL: https://doi.org/10.1007/978-1-84882-385-3_4 (visited on 05/21/2025).
- [183] R. Paschotta. “Laser Material Processing - an encyclopedia article”. In: *RP Photonics Encyclopedia*. RP Photonics AG, 2021. DOI: 10.61835/wmq. URL: https://www.rp-photonics.com/laser_material_processing.html (visited on 01/08/2024).
- [184] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. In: *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc., 2019. URL: https://papers.nips.cc/paper_files/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html (visited on 03/10/2025).
- [185] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld. “Adaptive histogram equalization and its variations”. In: *Computer Vision, Graphics, and Image Processing* 39.3 (Sept. 1, 1987), pp. 355–368. ISSN: 0734-189X. DOI: 10.1016/S0734-189X(87)80186-X. URL: <https://www.sciencedirect.com/science/article/pii/S0734189X8780186X> (visited on 03/06/2025).
- [186] W. Pool. “Use of autoencoders for fingerprint encoding and comparison”. In: (2021).

- [187] G. Pradel and C. Mitchell. “Privacy-Preserving Biometric Matching Using Homomorphic Encryption”. In: *2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*. 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). Shenyang, China: IEEE, Oct. 2021, pp. 494–505. ISBN: 978-1-6654-1658-0. DOI: 10.1109/TrustCom53373.2021.00079. URL: <https://ieeexplore.ieee.org/document/9724306/> (visited on 05/21/2025).
- [188] J. Prewitt. “Object enhancement and extraction”. In: *Picture processing and Psychopictorics* (1970). URL: <https://web.eecs.utk.edu/~hqi/ece472-572/reference/edge-Prewitt70.pdf> (visited on 03/07/2025).
- [189] J. Priesnitz, R. Huesmann, C. Rathgeb, N. Buchmann, and C. Busch. “Mobile Contactless Fingerprint Recognition: Implementation, Performance and Usability Aspects”. In: *Sensors* 22.3 (Jan. 2022). Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, p. 792. ISSN: 1424-8220. DOI: 10.3390/s22030792. URL: <https://www.mdpi.com/1424-8220/22/3/792> (visited on 09/09/2022).
- [190] J. Priesnitz, C. Rathgeb, N. Buchmann, and C. Busch. “Deep Learning-Based Semantic Segmentation for Touchless Fingerprint Recognition”. In: *Pattern Recognition. ICPR International Workshops and Challenges*. Ed. by A. Del Bimbo, R. Cucchiara, S. Sclaroff, G. M. Farinella, T. Mei, M. Bertini, H. J. Escalante, and R. Vezzani. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 154–168. ISBN: 978-3-030-68793-9. DOI: 10.1007/978-3-030-68793-9_11.
- [191] J. Priesnitz, C. Rathgeb, N. Buchmann, and C. Busch. “SynCoLFinGer: Synthetic contactless fingerprint generator”. In: *Pattern Recognition Letters* 157 (May 1, 2022), pp. 127–134. ISSN: 0167-8655. DOI: 10.1016/j.patrec.2022.04.003. URL: <https://www.sciencedirect.com/science/article/pii/S0167865522000915> (visited on 01/08/2024).
- [192] J. Priesnitz, C. Rathgeb, N. Buchmann, and C. Busch. “Touchless Fingerprint Sample Quality: Prerequisites for the Applicability of NFIQ2.0”. In: *2020 International Conference of the Biometrics Special Interest Group (BIOSIG)*. 2020 International Conference of the Biometrics Special Interest Group (BIOSIG). ISSN: 1617-5468. Sept. 2020, pp. 1–5.
- [193] J. Priesnitz, C. Rathgeb, N. Buchmann, C. Busch, and M. Margraf. “An overview of touchless 2D fingerprint recognition”. In: *EURASIP Journal on Image and Video Processing* 2021.1 (Dec. 2021), p. 8. ISSN: 1687-5281. DOI: 10.1186/s13640-021-00548-4. URL: <https://jivp-urasipjournals.springeropen.com/articles/10.1186/s13640-021-00548-4> (visited on 11/25/2021).
- [194] J. Priesnitz, A. Weißenfeld, L. Ruzicka, C. Rathgeb, B. Strobl, R. Lessmann, and C. Busch. “MCLFIQ: Mobile Contactless Fingerprint Image Quality”. In: *IEEE Transactions on Biometrics, Behavior, and Identity Science* 6.2 (Apr. 2024), pp. 272–287. ISSN: 2637-6407. DOI: 10.1109/TBIOM.2024.3377686. URL: <https://ieeexplore.ieee.org/document/10473152> (visited on 07/07/2025).
- [195] S. J. D. Prince. *Computer Vision: Models, Learning, and Inference*. 1st ed. Cambridge University Press, June 18, 2012. ISBN: 978-0-511-99650-4. DOI: 10.1017/CB09780511996504. URL: <https://www.cambridge.org/core/product/identifier/9780511996504/type/book> (visited on 03/07/2025).

- [196] N. Ratha, R. Bolle, V. Pandit, and V. Vaish. “Robust fingerprint authentication using local structural similarity”. In: *Proceedings Fifth IEEE Workshop on Applications of Computer Vision*. Fifth IEEE Workshop on Applications of Computer Vision. WACV 2000. Palm Springs, CA, USA: IEEE Comput. Soc, 2000, pp. 29–34. ISBN: 978-0-7695-0813-9. DOI: 10.1109/WACV.2000.895399. URL: <http://ieeexplore.ieee.org/document/895399/> (visited on 03/03/2025).
- [197] N. K. Ratha, S. Chikkerur, J. H. Connell, and R. M. Bolle. “Generating Cancelable Fingerprint Templates”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.4 (Apr. 2007), pp. 561–572. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2007.1004. URL: <https://ieeexplore.ieee.org/abstract/document/4107561> (visited on 07/02/2025).
- [198] C. Rathgeb, R. Tolosana, R. Vera-Rodriguez, and C. Busch, eds. *Handbook of Digital Face Manipulation and Detection: From DeepFakes to Morphing Attacks*. Advances in Computer Vision and Pattern Recognition. Cham: Springer International Publishing, 2022. ISBN: 978-3-030-87663-0. DOI: 10.1007/978-3-030-87664-7. URL: <https://link.springer.com/10.1007/978-3-030-87664-7> (visited on 10/24/2024).
- [199] C. Rathgeb and A. Uhl. “A survey on biometric cryptosystems and cancelable biometrics”. In: *EURASIP Journal on Information Security* 2011.1 (Sept. 23, 2011), p. 3. ISSN: 1687-417X. DOI: 10.1186/1687-417X-2011-3. URL: <https://doi.org/10.1186/1687-417X-2011-3> (visited on 05/21/2025).
- [200] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. *You Only Look Once: Unified, Real-Time Object Detection*. May 9, 2016. DOI: 10.48550/arXiv.1506.02640. arXiv: 1506.02640[cs]. URL: <http://arxiv.org/abs/1506.02640> (visited on 03/07/2025).
- [201] M. Rejman-Green, K. Brzozowski, T. Mansfield, R. Sanchez-Reillo, P. Waggett, and G. Whitaker. *The application of biometrics in critical infrastructures operations: guidance for security managers*. EUR 27191 EN. UK: Publications Office, 2015, p. 34. URL: <https://data.europa.eu/doi/10.2788/761739> (visited on 02/18/2025).
- [202] S. Ren, K. He, R. Girshick, and J. Sun. *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. Jan. 6, 2016. DOI: 10.48550/arXiv.1506.01497. arXiv: 1506.01497[cs]. URL: <http://arxiv.org/abs/1506.01497> (visited on 03/07/2025).
- [203] M. S. Riazi, S. M. Chavoshian, and F. Koushanfar. *SynFi: Automatic Synthetic Fingerprint Generation*. Feb. 16, 2020. arXiv: 2002.08900[cs,eess]. URL: <http://arxiv.org/abs/2002.08900> (visited on 01/08/2024).
- [204] H. Robbins and S. Monro. “A Stochastic Approximation Method”. In: *The Annals of Mathematical Statistics* 22.3 (1951). Publisher: Institute of Mathematical Statistics, pp. 400–407. ISSN: 0003-4851. URL: <https://www.jstor.org/stable/2236626> (visited on 03/10/2025).
- [205] T. Rohwedder, D. Osorio-Roig, C. Rathgeb, and C. Busch. *Benchmarking fixed-length Fingerprint Representations across different Embedding Sizes and Sensor Types*. July 17, 2023. arXiv: 2307.08615[cs]. URL: <http://arxiv.org/abs/2307.08615> (visited on 07/27/2023).
- [206] O. Ronneberger, P. Fischer, and T. Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: 10.1007/978-3-319-24574-4_28.

- [207] F. Rosenblatt. “The perceptron: A probabilistic model for information storage and organization in the brain.” In: *Psychological Review* 65.6 (1958), pp. 386–408. ISSN: 1939-1471, 0033-295X. DOI: 10.1037/h0042519. URL: <https://doi.apa.org/doi/10.1037/h0042519> (visited on 03/10/2025).
- [208] A. Ross, S. Dass, and A. Jain. “A deformable model for fingerprint matching”. In: *Pattern Recognition* 38.1 (Jan. 1, 2005), pp. 95–103. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2003.12.021. URL: <https://www.sciencedirect.com/science/article/pii/S0031320304002444> (visited on 03/07/2025).
- [209] A. Ross, A. Jain, and J. Reisman. “A hybrid fingerprint matcher”. In: *Pattern Recognition* 36.7 (July 1, 2003), pp. 1661–1673. ISSN: 0031-3203. DOI: 10.1016/S0031-3203(02)00349-7. URL: <https://www.sciencedirect.com/science/article/pii/S0031320302003497> (visited on 03/07/2025).
- [210] A. Ross, S. Shah, and J. Shah. “Image versus feature mosaicing: A case study in fingerprints”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 6202 (Apr. 17, 2006). DOI: 10.1117/12.666278.
- [211] R. K. Rowe, K. A. Nixon, and P. W. Butler. “Multispectral Fingerprint Image Acquisition”. In: *Advances in Biometrics: Sensors, Algorithms and Systems*. Ed. by N. K. Ratha and V. Govindaraju. London: Springer, 2008, pp. 3–23. ISBN: 978-1-84628-921-7. DOI: 10.1007/978-1-84628-921-7_1. URL: https://doi.org/10.1007/978-1-84628-921-7_1 (visited on 03/03/2025).
- [212] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. “ORB: An efficient alternative to SIFT or SURF”. In: *2011 International Conference on Computer Vision*. 2011 International Conference on Computer Vision. ISSN: 2380-7504. Nov. 2011, pp. 2564–2571. DOI: 10.1109/ICCV.2011.6126544. URL: <https://ieeexplore.ieee.org/document/6126544> (visited on 03/07/2025).
- [213] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. “Learning representations by back-propagating errors”. In: *Nature* 323.6088 (Oct. 1986), pp. 533–536. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/323533a0. URL: <https://www.nature.com/articles/323533a0> (visited on 03/10/2025).
- [214] A. P. Russo. “Method and system for fingerprint template matching”. U.S. pat. 7299360B2. P. B. AB. Nov. 20, 2007. URL: <https://patents.google.com/patent/US7299360B2/en> (visited on 07/23/2025).
- [215] L. Ruzicka, B. Kohn, and C. Heitzinger. *FingerUNeSt++: Improving Fingertip Segmentation in Contactless Fingerprint Imaging using Deep Learning*. IET Biometrics, Nov. 2024.
- [216] L. Ruzicka, B. Kohn, and C. Heitzinger. “TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging”. In: *Sensors* 25.6 (Mar. 14, 2025), p. 1824. ISSN: 1424-8220. DOI: 10.3390/s25061824. URL: <https://www.mdpi.com/1424-8220/25/6/1824> (visited on 04/15/2025).
- [217] L. Ruzicka, D. Söllinger, B. Kohn, C. Heitzinger, A. Uhl, and B. Strobl. “Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints Using Pose Correction and Unwarping”. In: *IET Biometrics* 2023 (Dec. 18, 2023). Ed. by V. Conti. Publisher: Hindawi, p. 7519499. ISSN: 2047-4938. DOI: 10.1049/2023/7519499. URL: <https://doi.org/10.1049/2023/7519499>.

- [218] L. Ruzicka, A. Spence, S. Bergmann, G. Nolden, B. Kohn, and C. Heitzinger. *Towards Fingerprint Mosaicking Artifact Detection: A Self-Supervised Deep Learning Approach*. Jan. 9, 2025. DOI: 10.48550/arXiv.2501.05034. arXiv: 2501.05034[cs]. URL: <http://arxiv.org/abs/2501.05034> (visited on 06/25/2025).
- [219] L. Ruzicka, B. Strobl, S. Bergmann, G. Nolden, T. Michalsky, C. Domscheit, J. Priesnitz, F. Blümel, B. Kohn, and C. Heitzinger. “Toward Synthetic Physical Fingerprint Targets”. In: *Sensors* 24.9 (Jan. 2024). Number: 9 Publisher: Multidisciplinary Digital Publishing Institute, p. 2847. ISSN: 1424-8220. DOI: 10.3390/s24092847. URL: <https://www.mdpi.com/1424-8220/24/9/2847> (visited on 04/15/2025).
- [220] L. Ruzicka, B. Strobl, B. Kohn, and C. Heitzinger. “Centrality of the Fingerprint Core Location”. In: *Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies BIOSTEC*. Biostec. Vol. Volume 1: Biosignals, Biodevices, Bioinformatics, Bioimaging. Rome, Italy: Scitepress, 2024, pp. 713–720. ISBN: 978-989-758-688-0. DOI: 10.5220/0012309300003657. URL: <https://www.scitepress.org/PublicationsDetail.aspx?ID=k/DomVEyngU=&t=1>.
- [221] S. S. M. Salehi, D. Erdogmus, and A. Gholipour. *Tversky loss function for image segmentation using 3D fully convolutional deep networks*. June 18, 2017. DOI: 10.48550/arXiv.1706.05721. arXiv: 1706.05721[cs]. URL: <http://arxiv.org/abs/1706.05721> (visited on 03/18/2025).
- [222] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen. *MobileNetV2: Inverted Residuals and Linear Bottlenecks*. Mar. 21, 2019. DOI: 10.48550/arXiv.1801.04381. arXiv: 1801.04381[cs]. URL: <http://arxiv.org/abs/1801.04381> (visited on 03/02/2023).
- [223] A. Sankaran, A. Malhotra, A. Mittal, M. Vatsa, and R. Singh. “On smartphone camera based fingerphoto authentication”. In: *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). Sept. 2015, pp. 1–7. DOI: 10.1109/BTAS.2015.7358782. URL: <https://ieeexplore.ieee.org/document/7358782> (visited on 04/22/2025).
- [224] A. M. Saxe, J. L. McClelland, and S. Ganguli. *Exact solutions to the nonlinear dynamics of learning in deep linear neural networks*. Feb. 19, 2014. DOI: 10.48550/arXiv.1312.6120. arXiv: 1312.6120[cs]. URL: <http://arxiv.org/abs/1312.6120> (visited on 06/25/2025).
- [225] T. Schmitz, M. Davies, B. Dutterer, and J. Ziegert. “The application of high-speed CNC machining to prototype production”. In: *International Journal of Machine Tools and Manufacture* 41.8 (June 1, 2001), pp. 1209–1228. ISSN: 0890-6955. DOI: 10.1016/S0890-6955(01)00005-0. URL: <https://www.sciencedirect.com/science/article/pii/S0890695501000050> (visited on 01/11/2024).
- [226] C. W. Schultz, J. X. H. Wong, and H.-Z. Yu. “Fabrication of 3D Fingerprint Phantoms via Unconventional Polycarbonate Molding”. In: *Scientific Reports* 8.1 (June 25, 2018). Number: 1 Publisher: Nature Publishing Group, p. 9613. ISSN: 2045-2322. DOI: 10.1038/s41598-018-27885-1. URL: <https://www.nature.com/articles/s41598-018-27885-1> (visited on 01/09/2024).

- [227] G. Schwarz. “Estimating the Dimension of a Model”. In: *The Annals of Statistics* 6.2 (Mar. 1978). Publisher: Institute of Mathematical Statistics, pp. 461–464. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/aos/1176344136. URL: <https://projecteuclid.org/journals/annals-of-statistics/volume-6/issue-2/Estimating-the-Dimension-of-a-Model/10.1214/aos/1176344136.full> (visited on 06/21/2023).
- [228] A. Sheffer, B. Lévy, M. Mogilnitsky, and A. Bogomyakov. “ABF++: fast and robust angle based flattening”. In: *ACM Transactions on Graphics* 24.2 (Apr. 2005), pp. 311–330. ISSN: 0730-0301, 1557-7368. DOI: 10.1145/1061347.1061354. URL: <https://dl.acm.org/doi/10.1145/1061347.1061354> (visited on 01/29/2024).
- [229] A. Shit, N. N. Giang, and S. Y. Park. “Visible light-responsive mechanically and electronically controllable conductive carbon dot-hydrogel-based pressure-strain sensor for wireless monitoring of antifouling performance”. In: *Composites Science and Technology* 218 (Feb. 8, 2022), p. 109212. ISSN: 0266-3538. DOI: 10.1016/j.compscitech.2021.109212. URL: <https://www.sciencedirect.com/science/article/pii/S0266353821005686> (visited on 04/02/2024).
- [230] C. Shorten and T. M. Khoshgoftaar. “A survey on Image Data Augmentation for Deep Learning”. In: *Journal of Big Data* 6.1 (July 6, 2019), p. 60. ISSN: 2196-1115. DOI: 10.1186/s40537-019-0197-0. URL: <https://doi.org/10.1186/s40537-019-0197-0> (visited on 03/10/2025).
- [231] B. Siciliano and O. Khatib, eds. *Springer Handbook of Robotics*. Springer Handbooks. Cham: Springer International Publishing, 2016. ISBN: 978-3-319-32550-7. DOI: 10.1007/978-3-319-32552-1. URL: <https://link.springer.com/10.1007/978-3-319-32552-1> (visited on 03/07/2025).
- [232] J. Silva, M. Nobre Menezes, T. Rodrigues, B. Silva, F. Pinto, and A. Oliveira. *Encoder-Decoder Architectures for Clinically Relevant Coronary Artery Segmentation*. June 21, 2021.
- [233] K. Simoens, P. Tuyls, and B. Preneel. “Privacy Weaknesses in Biometric Sketches”. In: *2009 30th IEEE Symposium on Security and Privacy*. 2009 30th IEEE Symposium on Security and Privacy. ISSN: 2375-1207. May 2009, pp. 188–203. DOI: 10.1109/SP.2009.24. URL: <https://ieeexplore.ieee.org/document/5207645/> (visited on 05/21/2025).
- [234] K. Simonyan and A. Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Apr. 10, 2015. DOI: 10.48550/arXiv.1409.1556. arXiv: 1409.1556[cs]. URL: <http://arxiv.org/abs/1409.1556> (visited on 03/10/2025).
- [235] J. M. Siracuse. “The Future of Our Fingerprints: The Importance of Instituting Biometric Data Protections in Pennsylvania”. In: *Duquesne Law Review* 59.2 (2021).
- [236] I. Sobel and G. Feldman. *An Isotropic 3x3 Image Gradient Operator*. Aug. 20, 2015. DOI: 10.13140/RG.2.1.1912.4965.
- [237] D. Sollinger and A. Uhl. “Optimizing contactless to contact-based fingerprint comparison using simple parametric warping models”. In: *2021 IEEE International Joint Conference on Biometrics (IJCB)*. 2021 IEEE International Joint Conference on Biometrics (IJCB). Shenzhen, China: IEEE, Aug. 4, 2021, pp. 1–7. ISBN: 978-1-66543-780-6. DOI: 10.1109/IJCB52358.2021.9484364. URL: <https://ieeexplore.ieee.org/document/9484364/> (visited on 11/25/2021).

- [238] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis and Machine Vision*. Boston, MA: Springer US, 1993. ISBN: 978-0-412-45570-4. DOI: 10.1007/978-1-4899-3216-7. URL: <http://link.springer.com/10.1007/978-1-4899-3216-7> (visited on 03/07/2025).
- [239] R. Sood and S. K. Pradhan. “Design and development of a low-cost open-source 3D printer and its single response optimization using polylactic acid (PLA) material”. In: *Materials Today: Proceedings*. International Conference on Materials and Manufacturing Methods – 2019 27 (Jan. 1, 2020), pp. 2981–2991. ISSN: 2214-7853. DOI: 10.1016/j.matpr.2020.04.905. URL: <https://www.sciencedirect.com/science/article/pii/S2214785320338992> (visited on 01/08/2024).
- [240] C. Sousedik and C. Busch. “Presentation attack detection methods for fingerprint recognition systems: a survey”. In: *IET Biometrics* 3.4 (2014). eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1049/bmt.2013.0020>, pp. 219–233. ISSN: 2047-4946. DOI: 10.1049/iet-bmt.2013.0020. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1049/iet-bmt.2013.0020> (visited on 03/12/2024).
- [241] V. S. Srinivasan and N. N. Murthy. “Detection of singular points in fingerprint images”. In: *Pattern Recognition* 25.2 (Feb. 1, 1992), pp. 139–153. ISSN: 0031-3203. DOI: 10.1016/0031-3203(92)90096-2. URL: <https://www.sciencedirect.com/science/article/pii/0031320392900962> (visited on 06/06/2023).
- [242] W. Stute, W. G. Manteiga, and M. P. Quindimil. “Bootstrap based goodness-of-fit-tests”. In: *Metrika* 40.1 (Dec. 1, 1993), pp. 243–256. ISSN: 1435-926X. DOI: 10.1007/BF02613687. URL: <https://doi.org/10.1007/BF02613687> (visited on 06/16/2023).
- [243] G. Sydow, L. Bienemann, D. Ennöckl, H. Greve, M. Helfrich, A. Hense, and R. Mantz. *Europäische Datenschutzgrundverordnung: Handkommentar*. 2. Auflage. NomosKommentar. Baden-Baden: Nomos, 2018. 1638 pp. ISBN: 978-3-8487-4892-1.
- [244] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. “Going deeper with convolutions”. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). ISSN: 1063-6919. June 2015, pp. 1–9. DOI: 10.1109/CVPR.2015.7298594. URL: <https://ieeexplore.ieee.org/document/7298594> (visited on 03/10/2025).
- [245] R. Szeliski. *Computer Vision: Algorithms and Applications*. Texts in Computer Science. Cham: Springer International Publishing, 2022. ISBN: 978-3-030-34371-2. DOI: 10.1007/978-3-030-34372-9. URL: <https://link.springer.com/10.1007/978-3-030-34372-9> (visited on 03/07/2025).
- [246] E. Tabassi, M. Olsen, O. Bausinger, C. Busch, A. Figlarz, G. Fiumara, O. Henniger, J. Merkle, T. Ruhland, C. Schiel, and M. Schwaiger. “NIST Fingerprint Image Quality 2”. In: *NIST* (July 13, 2021). Last Modified: 2022-11-29T09:11-05:00 Publisher: Elham Tabassi, Martin Olsen, Oliver Bausinger, Christoph Busch, Andrew Figlarz, Gregory Fiumara, Olaf Henniger, Johannes Merkle, Timo Ruhland, Christopher Schiel, Michael Schwaiger. URL: <https://www.nist.gov/publications/nist-fingerprint-image-quality-2> (visited on 10/03/2023).
- [247] H. Tan and A. Kumar. “Towards More Accurate Contactless Fingerprint Minutiae Extraction and Pose-Invariant Matching”. In: *IEEE Transactions on Information Forensics and Security* 15 (2020). Conference Name: IEEE Transactions on Information Forensics and Security, pp. 3924–3937. ISSN: 1556-6021. DOI: 10.1109/TIFS.2020.3001732.

- [248] M. Tan and Q. V. Le. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. Sept. 11, 2020. DOI: 10.48550/arXiv.1905.11946. arXiv: 1905.11946[cs, stat]. URL: <http://arxiv.org/abs/1905.11946> (visited on 07/21/2023).
- [249] X. Tan and B. Bhanu. “Fingerprint matching by genetic algorithms”. In: *Pattern Recognition* 39.3 (Mar. 1, 2006), pp. 465–477. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2005.09.005. URL: <https://www.sciencedirect.com/science/article/pii/S0031320305003444> (visited on 03/07/2025).
- [250] Y. Tang, F. Gao, J. Feng, and Y. Liu. “FingerNet: An unified deep network for fingerprint minutiae extraction”. In: *2017 IEEE International Joint Conference on Biometrics (IJCB)*. 2017 IEEE International Joint Conference on Biometrics (IJCB). Denver, CO: IEEE, Oct. 2017, pp. 108–116. ISBN: 978-1-5386-1124-1. DOI: 10.1109/BTAS.2017.8272688. URL: <http://ieeexplore.ieee.org/document/8272688/> (visited on 04/08/2024).
- [251] X. Tao, X. Chen, X. Yang, and J. Tian. “Fingerprint Recognition with Identical Twin Fingerprints”. In: *PLoS ONE* 7.4 (Apr. 27, 2012), e35704. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0035704. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3338710/> (visited on 02/18/2025).
- [252] S. Tarar and E. Kumar. “Fingerprint Mosaicking Algorithm to Improve the Performance of Fingerprint Matching System”. In: *Computer Science and Information Technology 2.3* (Mar. 2014), pp. 142–151. ISSN: 2331-6063, 2331-6071. DOI: 10.13189/csit.2014.020304. URL: http://www.hrpub.org/journals/article_info.php?aid=1330 (visited on 07/11/2024).
- [253] M. Tartagni and R. Guerrieri. “A 390 dpi live fingerprint imager based on feedback capacitive sensing scheme”. In: *1997 IEEE International Solids-State Circuits Conference. Digest of Technical Papers*. 1997 IEEE International Solids-State Circuits Conference. Digest of Technical Papers. ISSN: 0193-6530. Feb. 1997, pp. 200–201. DOI: 10.1109/ISSCC.1997.585332. URL: <https://ieeexplore.ieee.org/document/585332> (visited on 03/03/2025).
- [254] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin. “Deep learning on image denoising: An overview”. In: *Neural Networks* 131 (Nov. 1, 2020), pp. 251–275. ISSN: 0893-6080. DOI: 10.1016/j.neunet.2020.07.025. URL: <https://www.sciencedirect.com/science/article/pii/S0893608020302665> (visited on 03/07/2025).
- [255] F. Tomaz, T. Candeias, and H. Shahbazkia. “Fast and accurate skin segmentation in color images”. In: *First Canadian Conference on Computer and Robot Vision, 2004. Proceedings*. First Canadian Conference on Computer and Robot Vision, 2004. Proceedings. May 2004, pp. 180–187. DOI: 10.1109/CCCRV.2004.1301442. URL: <https://ieeexplore.ieee.org/document/1301442> (visited on 10/02/2023).
- [256] V. Tuchin, A. Bashkatov, E. Genina, V. Kochubey, V. Lychagov, S. Portnov, N. Trunina, D. Miller, S. Cho, H. Oh, B. Shim, M. Kim, J. Oh, H. Eum, Y. Ku, D. Kim, and Y. Yang. “Finger tissue model and blood perfused skin tissue phantom”. In: *Proceedings of SPIE - The International Society for Optical Engineering* 7898 (Feb. 1, 2011). DOI: 10.1117/12.881604.
- [257] U. Uludag, S. Pankanti, S. Prabhakar, and A. Jain. “Biometric cryptosystems: issues and challenges”. In: *Proceedings of the IEEE* 92.6 (June 2004), pp. 948–960. ISSN: 1558-2256. DOI: 10.1109/JPROC.2004.827372. URL: <https://ieeexplore.ieee.org/abstract/document/1299169> (visited on 07/02/2025).

- [258] Vazan. *SourceAFIS fingerprint matcher*. URL: <https://sourceafis.machinezoo.com/> (visited on 01/22/2024).
- [259] A. B. Vieira Wyzykowski, M. P. Segundo, and R. De Paula Lemes. “Level Three Synthetic Fingerprint Generation”. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. 2020 25th International Conference on Pattern Recognition (ICPR). Milan, Italy: IEEE, Jan. 10, 2021, pp. 9250–9257. ISBN: 978-1-72818-808-9. DOI: 10.1109/ICPR48806.2021.9412304. URL: <https://ieeexplore.ieee.org/document/9412304/> (visited on 01/08/2024).
- [260] P. Viola and M. J. Jones. “Robust Real-Time Face Detection”. In: *International Journal of Computer Vision* 57.2 (May 1, 2004), pp. 137–154. ISSN: 1573-1405. DOI: 10.1023/B:VISI.0000013087.49260.fb. URL: <https://doi.org/10.1023/B:VISI.0000013087.49260.fb> (visited on 03/07/2025).
- [261] C. Vu, T. Phan, and D. Chandler. “S-3: A Spectral and Spatial Measure of Local Perceived Sharpness in Natural Images”. In: *Image Processing, IEEE Transactions on* 21 (Apr. 1, 2012), pp. 934–945. DOI: 10.1109/TIP.2011.2169974.
- [262] Y. Wang, J. Hu, and F. Han. “Enhanced gradient-based algorithm for the estimation of fingerprint orientation fields”. In: *Applied Mathematics and Computation*. Special Issue on Intelligent Computing Theory and Methodology 185.2 (Feb. 15, 2007), pp. 823–833. ISSN: 0096-3003. DOI: 10.1016/j.amc.2006.06.082. URL: <https://www.sciencedirect.com/science/article/pii/S0096300306007648> (visited on 03/06/2025).
- [263] Y. Wang, J. Hu, and D. Phillips. “A Fingerprint Orientation Model Based on 2D Fourier Expansion (FOMFE) and Its Application to Singular-Point Detection and Fingerprint Indexing”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.4 (Apr. 2007). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 573–585. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2007.1003. URL: <https://ieeexplore.ieee.org/document/4107562> (visited on 03/03/2025).
- [264] Z. Wang, X. Ning, and M. B. Blaschko. *Jaccard Metric Losses: Optimizing the Jaccard Index with Soft Labels*. Mar. 20, 2024. arXiv: 2302.05666[cs]. URL: <http://arxiv.org/abs/2302.05666> (visited on 06/14/2024).
- [265] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko. *User’s guide to NIST biometric image software (NBIS)*. NIST IR 7392. Edition: 0. Gaithersburg, MD: National Institute of Standards and Technology, 2007, NIST IR 7392. DOI: 10.6028/NIST.IR.7392. URL: <https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7392.pdf> (visited on 04/08/2024).
- [266] A. Weissenfeld, R. Schmid, B. Kohn, B. Strobl, and G. F. Domínguez. “Case study of the acquisition of contactless fingerprints in a real police setting”. In: *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*. 2022 International Conference of the Biometrics Special Interest Group (BIOSIG). ISSN: 1617-5468. Sept. 2022, pp. 1–5. DOI: 10.1109/BIOSIG55365.2022.9897060.
- [267] K. Wertheim. “Embryology and Morphology of Friction Ridge Skin”. In: *The Fingerprint Sourcebook*. 1st. NCJ 225320. National Institute of Justice, July 2011. URL: <https://www.ojp.gov/library/publications/fingerprint-sourcebook>.
- [268] L. Wieclaw. “Fingerprint Orientation Field Enhancement”. In: *Computer Recognition Systems 4*. Ed. by R. Burduk, M. Kurzyński, M. Woźniak, and A. Żołnierok. Berlin, Heidelberg: Springer, 2011, pp. 33–40. ISBN: 978-3-642-20320-6. DOI: 10.1007/978-3-642-20320-6_4.

- [269] P. Wild, F. Daubner, H. Penz, and G. F. Dominguez. “Comparative Test of Smartphone Finger Photo vs. Touch-based Cross-sensor Fingerprint Recognition”. In: *2019 7th International Workshop on Biometrics and Forensics (IWBF)*. 2019 7th International Workshop on Biometrics and Forensics (IWBF). Cancun, Mexico: IEEE, May 2019, pp. 1–6. ISBN: 978-1-72810-622-9. DOI: 10.1109/IWBF.2019.8739191. URL: <https://ieeexplore.ieee.org/document/8739191/> (visited on 03/01/2023).
- [270] A. J. Willis and L. Myers. “A cost-effective fingerprint recognition system for use with low-quality prints and damaged fingertips”. In: *Pattern Recognition* 34.2 (Feb. 1, 2001), pp. 255–270. ISSN: 0031-3203. DOI: 10.1016/S0031-3203(00)00003-0. URL: <https://www.sciencedirect.com/science/article/pii/S0031320300000030> (visited on 03/03/2025).
- [271] S. Xie, R. Girshick, P. Dollar, Z. Tu, and K. He. “Aggregated Residual Transformations for Deep Neural Networks”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: IEEE, July 2017, pp. 5987–5995. ISBN: 978-1-5386-0457-1. DOI: 10.1109/CVPR.2017.634. URL: <http://ieeexplore.ieee.org/document/8100117/> (visited on 08/12/2024).
- [272] B. Xu, N. Wang, T. Chen, and M. Li. *Empirical Evaluation of Rectified Activations in Convolutional Network*. Nov. 27, 2015. DOI: 10.48550/arXiv.1505.00853. arXiv: 1505.00853[cs,stat]. URL: <http://arxiv.org/abs/1505.00853> (visited on 03/06/2023).
- [273] J. Yang, Y.-W. Oh, S. Siracosit, H. Park, J. Lee, S. G. Kim, H. Kim, K. Lee, G.-J. Jeon, J.-H. Ahn, S.-H. Jung, I.-S. Kang, and S.-H. K. Park. “Highly Sensitive Mutual-Capacitive Fingerprint Sensor With Reference Electrode”. In: *IEEE Electron Device Letters* 43.11 (Nov. 2022). Conference Name: IEEE Electron Device Letters, pp. 1973–1976. ISSN: 1558-0563. DOI: 10.1109/LED.2022.3208560. URL: <https://ieeexplore.ieee.org/document/9899393> (visited on 03/03/2025).
- [274] W. Yang, S. Wang, K. Yu, J. J. Kang, and M. N. Johnstone. “Secure Fingerprint Authentication with Homomorphic Encryption”. In: *2020 Digital Image Computing: Techniques and Applications (DICTA)*. 2020 Digital Image Computing: Techniques and Applications (DICTA). Melbourne, Australia: IEEE, Nov. 29, 2020, pp. 1–6. ISBN: 978-1-7281-9108-9. DOI: 10.1109/DICTA51227.2020.9363426. URL: <https://ieeexplore.ieee.org/document/9363426/> (visited on 05/21/2025).
- [275] S. Yoon, J. Feng, and A. K. Jain. “Altered Fingerprints: Analysis and Detection”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.3 (Mar. 2012). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 451–464. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2011.161. URL: <https://ieeexplore.ieee.org/document/6136517> (visited on 03/03/2025).
- [276] S. Yoon and A. K. Jain. “Longitudinal study of fingerprint recognition”. In: *Proceedings of the National Academy of Sciences* 112.28 (July 14, 2015), pp. 8555–8560. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1410272112. URL: <https://pnas.org/doi/full/10.1073/pnas.1410272112> (visited on 06/18/2024).
- [277] Y. Yu, Q. Niu, X. Li, J. Xue, W. Liu, and D. Lin. “A Review of Fingerprint Sensors: Mechanism, Characteristics, and Applications”. In: *Micromachines* 14.6 (June 14, 2023), p. 1253. ISSN: 2072-666X. DOI: 10.3390/mi14061253. URL: <https://www.mdpi.com/2072-666X/14/6/1253> (visited on 03/06/2024).

- [278] A. Zaafouri, M. Sayadi, and F. Fnaiech. “A developed unsharp masking method for images contrast enhancement”. In: *Eighth International Multi-Conference on Systems, Signals & Devices*. Eighth International Multi-Conference on Systems, Signals & Devices. Mar. 2011, pp. 1–6. DOI: 10.1109/SSD.2011.5767378. URL: <https://ieeexplore.ieee.org/document/5767378> (visited on 03/07/2025).
- [279] I. M. Zaid, M. Halwani, A. Ayyad, A. Imam, F. Almaskari, H. Hassanin, and Y. Zweiri. “Elastomer-Based Visuotactile Sensor for Normality of Robotic Manufacturing Systems”. In: *Polymers* 14.23 (Jan. 2022). Number: 23 Publisher: Multidisciplinary Digital Publishing Institute, p. 5097. ISSN: 2073-4360. DOI: 10.3390/polym14235097. URL: <https://www.mdpi.com/2073-4360/14/23/5097> (visited on 01/08/2024).
- [280] M. D. Zeiler and R. Fergus. “Visualizing and Understanding Convolutional Networks”. In: *Computer Vision – ECCV 2014*. Ed. by D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014, pp. 818–833. ISBN: 978-3-319-10590-1. DOI: 10.1007/978-3-319-10590-1_53.
- [281] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha, M. Li, and A. Smola. *ResNeSt: Split-Attention Networks*. Dec. 30, 2020. DOI: 10.48550/arXiv.2004.08955. arXiv: 2004.08955[cs]. URL: <http://arxiv.org/abs/2004.08955> (visited on 06/10/2024).
- [282] K. Zhang, Z. Liu, J. Hu, and S. Wang. “An Auto-Encoder Based Method for Camera Fingerprint Compression”. In: *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). ISSN: 2379-190X. June 2023, pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10094951. URL: <https://ieeexplore.ieee.org/document/10094951> (visited on 03/06/2025).
- [283] L. Zhang, Z. Xu, C. Barnes, Y. Zhou, Q. Liu, H. Zhang, S. Amirghodsi, Z. Lin, E. Shechtman, and J. Shi. “Perceptual Artifacts Localization for Image Synthesis Tasks”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 7579–7590. URL: https://openaccess.thecvf.com/content/ICCV2023/html/Zhang_Perceptual_Artifacts_Localization_for_Image_Synthesis_Tasks_ICCV_2023_paper.html (visited on 05/21/2025).
- [284] R. Zhang, Y. Yin, W. Deng, C. Li, and J. Zhang. *Deep Learning for Finger Vein Recognition: A Brief Survey of Recent Trend*. July 5, 2022. DOI: 10.48550/arXiv.2207.02148. arXiv: 2207.02148[cs]. URL: <http://arxiv.org/abs/2207.02148> (visited on 03/06/2025).
- [285] W.-P. Zhang, Q.-R. Wang, and Y. Tang. “A wavelet-based method for fingerprint image enhancement”. In: *Proceedings. International Conference on Machine Learning and Cybernetics*. Proceedings. International Conference on Machine Learning and Cybernetics. Vol. 4. Nov. 2002, 1973–1977 vol.4. DOI: 10.1109/ICMLC.2002.1175382. URL: <https://ieeexplore.ieee.org/document/1175382?arnumber=1175382> (visited on 03/03/2025).
- [286] Y. Zhang and Q. Xiao. “An Optimized Approach for Fingerprint Binarization”. In: *The 2006 IEEE International Joint Conference on Neural Network Proceedings*. The 2006 IEEE International Joint Conference on Neural Network Proceedings. ISSN: 2161-4407. July 2006, pp. 391–395. DOI: 10.1109/IJCNN.2006.246708. URL: <https://ieeexplore.ieee.org/document/1716119> (visited on 03/06/2025).

- [287] Q. Zhao, L. Zhang, D. Zhang, N. Luo, and J. Bao. “Adaptive pore model for fingerprint pore extraction”. In: *2008 19th International Conference on Pattern Recognition*. 2008 19th International Conference on Pattern Recognition. ISSN: 1051-4651. Dec. 2008, pp. 1–4. DOI: 10.1109/ICPR.2008.4761458. URL: <https://ieeexplore.ieee.org/document/4761458> (visited on 03/06/2025).
- [288] X. Zheng, Y. Wang, and X. Zhao. “Fingerprint Image Segmentation Using Active Contour Model”. In: *Fourth International Conference on Image and Graphics (ICIG 2007)*. Fourth International Conference on Image and Graphics (ICIG 2007). Aug. 2007, pp. 437–441. DOI: 10.1109/ICIG.2007.63. URL: <https://ieeexplore.ieee.org/document/4297126> (visited on 06/18/2024).
- [289] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang. *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*. July 18, 2018. arXiv: 1807.10165[cs, eess, stat]. URL: <http://arxiv.org/abs/1807.10165> (visited on 06/10/2024).
- [290] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye. “Object Detection in 20 Years: A Survey”. In: *Proceedings of the IEEE* 111.3 (Mar. 2023). Conference Name: Proceedings of the IEEE, pp. 257–276. ISSN: 1558-2256. DOI: 10.1109/JPROC.2023.3238524. URL: <https://ieeexplore.ieee.org/document/10028728> (visited on 03/07/2025).

Dipl. Ing. LAURENZ RUZICKA

27.01.1997

EDUCATION

- TU Wien** 2022 - Ongoing
Doctoral Program Technical Physics
Focus on 3D contactless fingerprint images and the application of machine learning in the context of computer vision.
- TU Wien** 2019-2022
Master Program Technical Physics - Graduated With Honors
Diploma Thesis: "Spatio-Temporal Pattern Matching for 3D Data with an application for automate diagnosis of iRBD"
- TU Wien** 2016-2019
Bachelor Program Technical Physics
Bachelor thesis: "Non-standard node split criteria for Boosted Decision Trees"
Two semesters of *Software & Information Engineering* parallel to physics
- BRG4 Waltergasse** 2007 - 2015
High School Diploma - Graduated with Honors
Focus on biology, chemistry and physics

WORK EXPERIENCE

- Austrian Institute of Technology** June 2022 - Ongoing
PhD Candidate & Junior Scientist
Research regarding contactless fingerprint technology. Work activities include:
- Design and implementation of machine learning model architectures and classical algorithms, as well as project based work with clients.
 - Conceptualization and writing of academic publications.
- Austrian Institute of Technology** Mai 2021 - March 2022
Research Assistant
Employed for writing the diploma thesis. Work activities included:
- Implementation of algorithms for processing 3D data of a time-of-flight sensor, as well as spatio-temporal pattern matching and analysis
 - Selection, implementation and application of state-of-the-art deep learning models and natural language processing tools

TU Wien
Student Assistant

October 2019 - July 2021

Supported and graded students at the courses:
Experimental exercise in physics (I, II and III), as well as computing for technical physics I

Acodemy
Trainer

March 2018 - October 2019

Taught children of ages 6 - 16 on how to program

Karl-Schubert-School
Assistant

September 2015 - Mai 2016

Civil service at a school for mentally disabled children, assisted in teaching

INTERNATIONAL EXPERIENCE

University of Ghent July 2019
International Workshop - How to Run on Water

Karlsruhe Institute of Technology June 2019
International Workshop - Hope to CO₂ - Empowering a Sustainable Future

CERTIFICATES & LANGUAGES

Fundamentals of Accelerated Data Science	August 2025
Webinar Trustless Biometric Sensors	June 2025
Resilience and Well-Being in Academia	April 2025
Writing of competitive applications with focus on Horizon Europe Calls	March 2025
AIT-specific basics of project work	October 2024
Science communication	June 2024
Presentation technique	June 2024
IPRs- protection and utilization	March 2024
BWL compact	April 2024
Data protection and information security	February 2024
NVIDIA DLI: Fundamentals of Deep Learning	October 2024
Python for HPC	December 2022
The art of scientific writing	November 2022
PhD Seminar Hirscheegg	October 2022

German Mother tongue
English Business fluent

TALKS

IFPC '25: International Face and Fingerprint Performance Conference 2nd April 2024
Invited Speaker

Contactless Fingerprint Quality Assessment

PUBLICATIONS

- Sensors** published March 2025
First author
TipSegNet: Fingertip Segmentation in Contactless Fingerprint Imaging
- WPES '24: Proceedings of the 23rd Workshop on Privacy in the Electronic Society** published November 2024
Coauthor
Towards Real-Time Privacy-Preserving Minutiae Matching
- Sensors** published April 2024
First author
Toward Synthetic Physical Fingerprint Targets
- IEEE Transactions on Biometrics Behavior and Identity Science** published March 2024
Coauthor
MCLFIQ: Mobile Contactless Fingerprint Image Quality
- Proceedings of the 17th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 1: BIOSIGNALS** published February 2024
First author
Centrality of the Fingerprint Core Location
- IET Biometrics** published December 2023
First author
Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints Using Pose Correction and Unwarping
- European Journal of Neurology** published August 2023
Second author
Improved automatic identification of isolated rapid eye movement sleep behavior disorder with a 3D time-of-flight camera
- IEEE Engineering in Medicine and Biology Society** published July 2022
Shared first author
TeaSpam: A Novel Method of TEmporal and SPAtial Movement Encoding During Sleep
- Physical Review B** published August 2022
Coauthor
Photoexcitations in the Hubbard model – generalized Loschmidt amplitude analysis of impact ionization in small clusters