

Evaluating Contactless Fingerprint Segmentation for Interoperable Biometric Identification Systems

Avaliação de Segmentação de Impressões Digitais sem Contato para Sistemas de Identificação Biométrica Interoperáveis

Euclides Napoleão Arcoverde Neto^{1,2*}, Thiago Lopes Trugillo da Silveira³

Abstract: Fingerprint recognition is a popular and cost-effective biometric technology. Most existing systems require specialized hardware to capture fingerprints, but contactless fingerprint capture can be performed using smartphone images. This method reduces costs and enhances hygiene and usability. A new image processing workflow is needed to facilitate contactless fingerprint capture, with segmentation a critical step. This study explores the feasibility of using smartphone images for this purpose. It evaluates four deep learning models – SSD MobileNetV2, SSD MobileNetV2 FPN Lite, Mask R-CNN, and U-Net – using the ISFPDv2 dataset. Results show that Mask R-CNN performed best in segmenting fingertip regions, while SSD MobileNetV2 had the highest recognition accuracy against traditional fingerprint databases. These findings demonstrate the potential of deep learning for effective contactless fingerprint recognition using smartphones.

Keywords: Object Detection — Image Segmentation — Fingerprint Recognition

Resumo: O reconhecimento de impressões digitais é uma tecnologia biométrica popular e de baixo custo. Embora a maioria dos sistemas existentes exija hardware especializado para a captura das digitais, a captura sem contato pode ser realizada por meio de imagens obtidas por smartphones, o que reduz custos e melhora a higiene e a usabilidade. Para viabilizar essa abordagem, é necessário um novo fluxo de processamento de imagens, sendo a segmentação uma etapa crítica. Este estudo investiga a viabilidade do uso de imagens de smartphones para essa finalidade, avaliando quatro modelos de *deep learning* – SSD MobileNetV2, SSD MobileNetV2 FPN Lite, Mask R-CNN e U-Net – com o uso da base de dados ISFPDv2. Os resultados indicam que o Mask R-CNN apresentou o melhor desempenho na segmentação das regiões das pontas dos dedos, enquanto o SSD MobileNetV2 obteve a maior acurácia de reconhecimento em comparação com imagens de bancos de dados tradicionais de impressões digitais. Esses resultados demonstram o potencial do uso de *deep learning* para o reconhecimento eficaz de impressões digitais sem contato obtidas por meio de smartphones.

Palavras-Chave: Detecção de Objetos — Segmentação de Imagens — Reconhecimento de Digitais

¹ Serviço Federal de Processamento de Dados (SERPRO), Brasil

² Instituto de Informática, Universidade Federal do Rio Grande do Sul (UFRGS), Brasil

³ Departamento de Computação Aplicada, Universidade Federal de Santa Maria (UFSM), Brasil

*Euclides Napoleão Arcoverde Neto: euneto@gmail.com

*Thiago Lopes Trugillo da Silveira: drtitsilveira@gmail.com

DOI: <http://dx.doi.org/10.22456/2175-2745.147966> • Received: 03/06/2025 • Accepted: 26/12/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

Fingerprint recognition has become a widely adopted and cost-effective biometric authentication method [1]. However, traditional fingerprint capture systems are based on specialized hardware [2], limiting accessibility and affordability. Recent advances in mobile technology have paved the way for contactless fingerprint capture using smartphone cameras. Contactless fingerprint recognition systems eliminate the need for direct finger-to-sensor contact, offering advantages in hygiene,

usability, and user acceptance [3].

Despite these advantages, contactless fingerprint capture introduces new challenges, particularly in uncontrolled environments [3]. Smartphone images often contain complex backgrounds, varying illumination conditions, and possible finger misalignment, making it difficult to accurately segment the fingerprint from the surrounding scene. Effective fingerprint segmentation is a critical step in the fingerprint recognition process, as it involves isolating the fingerprint region from the rest of the image and preparing it for further analysis

and comparison [4]. The quality of segmentation directly impacts the accuracy of subsequent recognition algorithms.

Comparison of fingerprints obtained from contactless sensors with those obtained from contact-based sensors requires specific methods [4]. This interoperability has become a significant challenge. One of the most important limiting factors in the adoption of contactless fingerprints is cross-compatibility with legacy databases of contact-based fingerprints [3].

This work investigates segmentation techniques applied to fingertip isolation in smartphone images. We evaluate the performance of deep learning models designed for generic object detection and segmentation. The aim is to determine how well these models can accurately identify and segment fingertip regions within smartphone images captured in controlled and uncontrolled settings. In addition, we explore the interoperability of segmented fingerprints with traditional contact-based fingerprint databases, evaluating recognition accuracy using a standard fingerprint matcher [5].

2. Related Work

This section overviews existing research on fingerprint segmentation from smartphone images and contact-to-contactless fingerprint matching, highlighting the diverse approaches employed. We focus on methods suitable for mobile applications, emphasizing computational efficiency and performance.

Grosz et al. [3] present an end-to-end system, called C2CL, which matches contactless fingerprints with contact-based fingerprints. They propose a pipeline consisting of segmentation, contrast enhancement, 500 ppi scale normalization, deformation correction, and an adaptation of DeepPrint [6] for contact-contactless matching. The authors used datasets like ISFPDv2 [7, 8]. To segment the fingertips, they trained a U-Net segmentation network [9]. Kunsuk and Areekul [10] proposed two automatic rescaling methods for finger photo image conversion to perform touchless and touch-based fingerprint verification interoperability. The authors aim to rescale an input finger photo image to 500 dpi fingerprint verification systems. Regarding the segmentation stage in the pipeline, they manually segmented the fingertips of all finger photo images in the ISFPDv2 database.

Ruzicka et al. [11] show a processing pipeline for contactless fingerprint recordings that improves sensor interoperability between contactless and contact-based fingerprint sensors. They propose a novel deep learning neural network architecture based on U-Net [9] and MobileNetV2 [12], which is used to segment the fingertip and also to detect a reference point based on the fingerprint core. They developed a novel deep learning network architecture closely related to U-Net to implement pose correction. It was able to solve both sub-tasks for pose correction: Image segmentation and reference point localization and outperformed both Squeeze U-Net [13] as well as EfficientUNet++ [14], which are both contemporary image segmentation networks. Malhotra et al. [8] highlight the challenges of recognizing finger-selfies acquired by a smart-

phone camera. A Deep Scattering Network [15, 16] based finger-selfie representation is proposed, which is matched using a Random Decision Forest classification. Segmentation involves finding a binary mask in a captured finger-selfie that represents the distal phalanges of the finger and combining the region covariance-based saliency with skin color measurements for effective finger-selfie segmentation.

Priesnitz et al. [17] adapt DeepLabv3+ [18] to the requirements of fingertip segmentation. The resulting comparison with the color-based baseline system shows superior segmentation results of the hand area and demonstrates the feasibility of direct fingertip segmentation. Kauba et al. [4] introduce a comprehensive and in-depth analysis of the toolchain stages necessary for a mobile fingerprint recognition solution, namely image capturing, fingerprint pre-processing, fingerprint contactless to contact-based mapping of fingerprints, and feature extraction and comparison. The authors evaluated various techniques for fingerprint segmentation, dividing them into three approaches: traditional, detection, and segmentation. Regarding traditional approaches, they tested skin segmentation, color histogram segmentation, and Gaussian Mixture Model. In terms of detection approaches, SSD MobileNetV2 [19, 12] and a quantized version of SSD MobileNetV2 (which uses integers instead of floating points) were trained and tested. Finally, the authors tested Mask R-CNN [20], Deeplab [21], Segnet [22], and HRNet [23] segmentation models. Kauba et al. demonstrate that the Mask R-CNN and Deeplab models were the best segmentation approaches tested. Concerning detection models, SSD MobileNetV2 Quantized Greyscale and Mask R-CNN RGB performed best.

3. Methodology

This section describes our methodology. Prominent detection and segmentation techniques were considered [12, 24, 9, 20]. The performance of fingerprint segmentation was evaluated in terms of the IoU metric. The quality of segmented fingerprints was assessed using NFIQ values [25, 26], and the accuracy of comparing these fingerprints with instances obtained through traditional sensors was evaluated using the Neurotechnology VeriFinger SDK 13.0 [5].

3.1 Dataset

This work considers the IIITD SmartPhone Finger-selfie version 2 (ISFPDv2) database [7, 8], which consists of over 17,000 fingerprint photos and 2,432 fingerprint images obtained from capture sensors of 304 unique fingers. The fingerprint photos were captured using two smartphones: OnePlus One (OPO) and Micromax Canvas Knight, while the livescan fingerprints (with contact) were acquired using the Secugen Hamster IV at 500 PPI. The images were captured in controlled and uncontrolled indoor environments, and outdoors without flash during the day and with flash at night. All images were obtained with camera autofocus turned on.

Based on the challenges considered, the database is divided into four subsets: (1) white background, (2) natural

background, (3) resolution variation, and (4) cross-domain (contact fingerprints). The database is organized into three directories: BI, RES, and LS. In the BI directory, all images were obtained with 13 megapixels of resolution, varying the environments and lighting conditions (indoor and outdoor, with and without flash). In the RES folder, images are obtained with a white background in indoor environments but with varying image resolution (5, 8, and 13 megapixels). The LS directory contains the livescan images. The ISFPDv2 database has a single finger per image.

We used a subset of the images in the BI directory, representing various environments. It is an important characteristic for performing fingertip segmentation in real-world settings. The NI (Natural Indoor) and NO (Natural Outdoor) images were obtained with a natural background and indoor and outdoor environments, respectively. WI (White Indoor) and WO (White Outdoor) images were obtained with a white background and in indoor and outdoor environments, respectively.

3.2 Data preprocessing

The ISFPDv2 database lacks annotations on fingers or fingertips within images, where the minutiae of fingerprints used by biometric systems for verification and identification are found. Marking the fingertips is the primary focus of this study and can be represented by a polygon or a binary mask that represents the object of interest.

Therefore, the initial step in this study was to generate these masks for a subset of the ISFPDv2 database. To annotate the images and enable the training, validation, and testing of models with a significant amount of data, we implemented an interactive tool to expedite the ground truth generation process. This process uses the Segment Anything Model (SAM) [27] to assist in finger segmentation. Using the mask generated by SAM, the user can define the location of the fingertips.

Figure 1 shows the pipeline for generating masks that correspond to the fingertips. The process begins with the presentation of an image to the user (Figure 1(a)). At this point, the user is asked to select a point in the image that represents the finger (prompt), as in the red marker in Figure 1(a). SAM processes the prompt and delivers three segmentation options, as seen in Figure 1(b). Again, the user needs to interact by choosing the option that best represents the segmentation. In this case, the third option was considered the best segmentation mask. Then, we keep the largest connected component in the mask, discarding the rest. In possession of the mask that represents the entire finger, Figure 1(c), the angle of inclination of the object is calculated with respect to the horizontal axis. The original image is rotated (to compensate for this calculated angle) and presented to the user to choose a cut point to segment only the fingertip. In the example presented in Figure 1, the calculated angle was 13.83° , which means that rotation was necessary to align the finger horizontally. At this point, the user must choose a point in a location close to the illustrative red line shown in Figure 1(d). Presenting the rotated image is necessary for the fingertip to be better

represented since a vertical cut is made on the mask passing through the point selected by the user. Finally, the mask representing the ground truth is generated considering the actual rotation angle, as observed in Figure 1(e).

Masks were generated for 3,023 files randomly selected¹, used in the training and testing of models. All masks related to fingers horizontally pointing to the right.

3.3 Models Training

The 3,023 images and respective annotated masks were divided into training, validation, and testing sets for training purposes. Seventy percent of the samples were assigned for training. The remaining samples were divided equally between the validation and testing sets, each receiving 15% of the samples. It is important to note that this allocation utilized balanced samples concerning the four ways to acquire the database.

Following a comprehensive review of the literature, two detection and two segmentation models were selected for analysis. The models were chosen for their suitability in fingerprint segmentation [4] and their optimization in terms of memory and processing consumption, allowing direct execution on smartphones [4]. We will discuss these models next.

3.3.1 Detection Models

SSD MobileNetV2 implementation used is based on the Object Detection API within the TensorFlow package [28], which provides a unified approach for multiple object detection frameworks capable of real-time inference across various platforms, including smartphones. All training images were resized to 300×300 pixels and augmented, considering the following transformations: horizontal flip with probability 50%, 90-degree rotation with probability 50%, brightness adjustment with $\delta_{max} = 0.2$, contrast adjustment with $\delta_{min} = 0.8$ and $\delta_{max} = 1.25$, and saturation adjustment with $\delta_{min} = 0.8$ and $\delta_{max} = 1.25$. The loss function used has two components: the weighted smoothed L1 [29, 30] and the weighted sigmoid loss [31]. The network was trained using the Root Mean Square Propagation (RMSprop) optimizer with a batch size of 16 for 50,000 iterations. The validation dataset was used to monitor the performance of the model during training. The pre-trained detection model on the COCO 2017 dataset [32], named SSD MobileNet v2 320×320 , was utilized according to the TensorFlow 2 Detection Model Zoo [33].

SSD MobileNetV2 FPN Lite was based on the Object Detection API within the TensorFlow package (TensorFlow 2 Detection Model Zoo), specifically the SSD MobileNet V2 FPN Lite 320×320 model. All parameters used for training were identical to those used in training the SSD MobileNetV2 model described above.

3.3.2 Segmentation Models

U-Net available at Keras library [34] was utilized. For training, all images were resized to 384×384 pixels, normalized

¹Files list available at: SUPPRESSED FOR DOUBLE BLIND REVIEW.

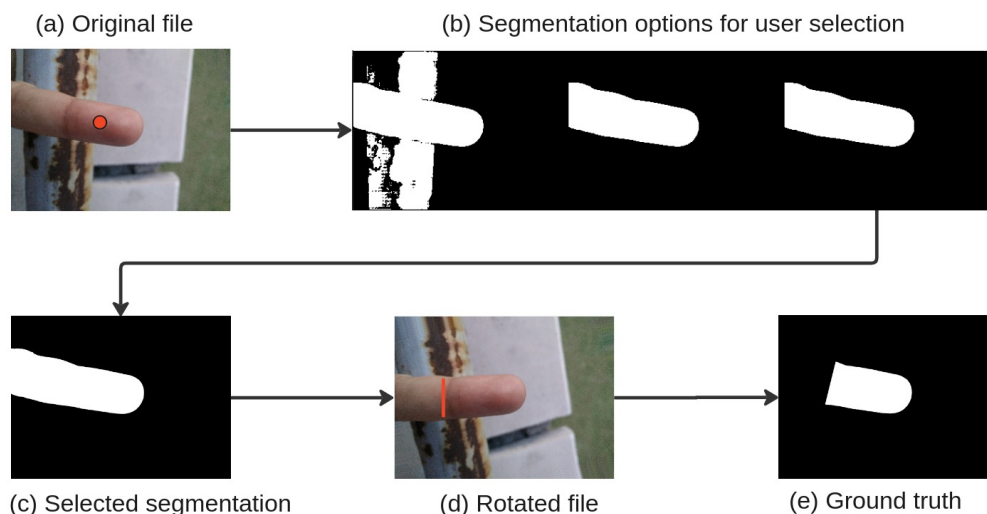


Figure 1. Process for ground truth generation.

and augmented considering the following transformations: horizontal flip with probability 50%, rotations of up to 5° , vertical and horizontal shift of up to 2%, shear range equal to 40, zoom range up to 20%, and constant fill mode. The model has 32 filters, uses batch normalization, dropout of 30%, dropout change per layer turned off, and 4 layers. A checkpoint was also defined to save the best models based on the loss of the validation set, keeping only the best models generated during training. For training, the Adam optimizer was defined with default parameter settings, the binary cross-entropy loss function [35] with IoU and IoU thresholded validation metrics, a batch size of 8, and 50 epochs.

Mask R-CNN was trained using the pre-trained COCO Instance Segmentation Baselines with Mask R-CNN - R101-FPN [36] model, provided by the Detectron2 framework [37]. During training and inference, input images were resized to 1052×780 pixels and augmented considering the following transformations: horizontal flip with probability 50%, random rotation from 0 to 90° , random brightness, contrast and saturation adjustment with minimum intensity of 0.9 and maximum of 1.1. The Stochastic Gradient Descent (SGD) function [35] was used with an initial learning rate of 0.00025, a momentum of 0.9, and a weight decay of 0.0001 to train the model. The network was trained for 2,500 iterations with a batch size of 1 due to memory limitations. Default hyperparameters were used for training.

3.4 NFIQ and Comparison Score Calculation

NFIQ is a crucial tool in the quality assessment of the fingerprint image [26]. NFIQ was conceived to address the need for an objective metric to evaluate the quality of fingerprint images. NFIQ version 1 established a significant standard in this field, introducing metrics and effective software for quantitative evaluation. However, the continuous evolution of technology and the growing demands for precision motivated the development of NFIQ version 2.

NFIQ 1.0 and NFIQ 2.0 [25, 26] values were obtained to evaluate the quality of the detected and segmented fingerprints. Comparisons were performed using instances from the ISFPDv2 database [7, 8], which contains fingerprint images acquired with traditional sensors. The quality assessment was conducted using the Neurotechnology VeriFinger SDK 13.0 [5]. According to the NIST standard, NFIQ 1.0 classifies fingerprint images into five discrete classes, where a value of 1 corresponds to the highest image quality and a value of 5 corresponds to the lowest. In contrast, the VeriFinger SDK represents NFIQ 1.0 scores using a byte-encoded scale with predefined values of 100 (excellent), 80 (very good), 60 (good), 40 (fair), and 20 (poor), which correspond inversely to the NIST NFIQ 1.0 classes from 1 to 5, respectively. For NFIQ 2.0, the SDK provides an integer quality score in the range [0, 100], where higher values indicate better fingerprint image quality. In both cases, the SDK also returns special values (254 and 255) to indicate whether the NFIQ score was calculated or if the computation failed.

A standard preprocessing methodology was defined for segmented fingerprints across the detection and segmentation models to compare NFIQ and score quality. For all models, the results were obtained by maintaining part of the background (simulating a detection model) and without a background (simulating a segmentation model), as can be observed in Figure 2. Figs. 2 (a), (b), (c), and (d) show the results of the detection models, with (a) and (b) referring to the SSD MobileNetV2 model and (c) and (d) referring to the SSD MobileNetV2 FPN Lite model. Figs. 2 (e), (f), (g), and (h) present the results of the segmentation models, with (e) and (f) referring to the U-Net model and (g) and (h) referring to the Mask R-CNN model. We obtain an image with and an image without background from each model. This methodology enabled a comparative analysis of all models using the same technique to determine the object of interest: detection (based on bounding boxes with a background) and segmentation (only the object without a background). For obtaining

images with a background from the detection models, the result of the detection itself was used, which already returns a bounding box. The bounding limits of the mask obtained were used for the segmentation models. To obtain segmented images from the detection models, an elliptical mask was used, whose center is defined from the center of the bounding box and the axes as the width and height of the bounding box minus 20% of the width and height values returned by the models.

Based on Kauba et al. [4], we implement a procedure that aims to enhance the structure of the ridges and valleys of contactless fingerprint images to achieve the best possible compatibility with contact-based fingerprint images. Initially, a 90° counterclockwise rotation was performed so that the fingerprints were in the same orientation as those acquired by sensors in the ISPFdv2 database. Next, the contactless fingerprint image is converted to grayscale, followed by horizontal mirroring, which is necessary due to the differences in the contactless and contact acquisition methods. Contactless samples resemble a horizontally mirrored version of contact samples, which needs to be compensated for. Finally, an adaptive histogram equalization with contrast limiting (CLAHE) [38] is applied to enhance the contrast of the grayscale image, followed by a color inversion. The final result of this process is shown in Figure 2. These images were used to obtain the values of NFIQ 1 and 2 and calculate the score by comparing each of them with the instances captured by contact.

The biometric feature matching algorithm of the Neurotechnology VeriFinger SDK 13.0 [5] provides a similarity score. The higher the score, the higher the probability that the features obtained belong to the same person. The higher the threshold value λ for this score, the more similar the feature sets must be to produce positive results during matching. The matching threshold is related to the matching algorithm's false acceptance rate (FAR). The higher the threshold, the lower the FAR, and the higher the false rejection rate (FRR) and vice versa.

4. Results and Discussion

This section presents the results obtained. Section 4.1 shows the detection and the results of segmentation using the intersection over union (IoU) metric. Section 4.2 shows the results of the fingerprint quality in terms of NFIQ 1 and 2 and the performance of the fingerprint recognition obtained from smartphones that were detected/segmented compared to instances obtained by traditional contact sensors.

4.1 Detection and Segmentation Models

Table 1 shows the results obtained in the test subset for both the detection and segmentation models. Note that while the Mask R-CNN model is used for segmentation in this study, the result obtained in the detection stage before segmentation is also presented. It can be observed that among the detection models evaluated, SSD MobileNetV2 shows the best performance, achieving an average IoU value of 0.83. Thus, based

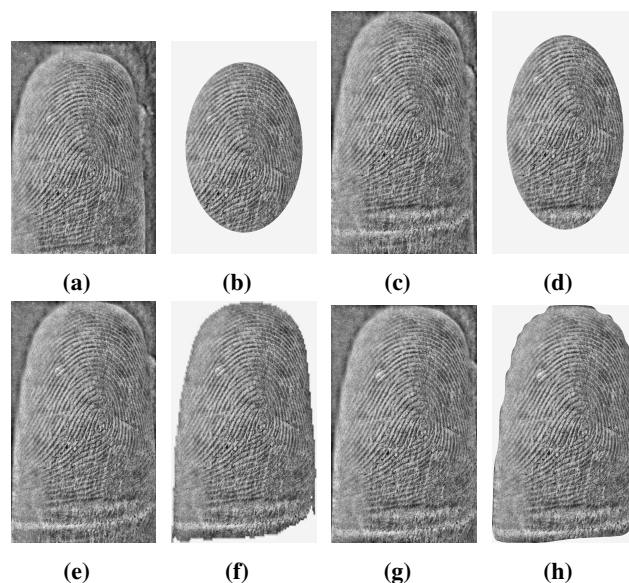


Figure 2. Fingerprint images used for NFIQ and comparison score evaluation. (a), (c), (e), and (g) contain background. (b), (d), (f), and (h) are segmented.

on the results presented, SSD MobileNetV2 emerges as the detection model with the best performance. However, all three models yielded satisfactory results for detecting fingertips, and depending on the application context, any of them can be recommended. If detection is performed directly on smartphones, the SSD MobileNetV2 FPN Lite model might be recommended.

This study does not assess processor and memory usage performance.

The Mask R-CNN model exhibited slightly superior segmentation results to the U-Net model. The Mask R-CNN model achieved an average IoU value of 0.96, while the U-Net model obtained 0.94. Mask R-CNN and U-Net are popular image segmentation methods with significant results on various benchmarks [39, 40]. Mask R-CNN can be used for general-purpose images, including realistic photographs, medical images, and satellite images, and typically requires more training data than U-Net. However, U-Net was initially designed for biomedical images, but can be adapted to other domains. Thus, considering the presented results, both models demonstrated good performance, with Mask R-CNN slightly better for fingertip segmentation.

4.2 Quality and Performance Evaluation for Recognition

Table 2 presents the average values obtained for NFIQ 1 and NFIQ 2 for the 458 test samples. It is important to mention that two images were generated for each detected/segmented fingerprint: one containing background information (Background column) and another removing the background information, attempting to keep only the fingertips as the object of interest (Segmentation column). It can be seen that the

Table 1. Average detection/segmentation performance in terms of IoU values.

Method	Model	Bouding Box	Mask
Detection	SSD MobileNetV2	0.83	
	SSD MobileNetV2 FPN Lite	0.78	
Segmentation	Mask R-CNN	0.79	0.96
	U-Net		0.94

U-Net model, using the segmentation approach, achieved the best performance in terms of the average value of NFIQ 1, reaching a value of 67.69. It is also worth mentioning that the SSD MobileNetV2 FPN Lite model presented the best result for the background approach, obtaining the average NFIQ 1 value of 66.81. For NFIQ 2, the SSD MobileNetV2 FPN Lite model presented the best average of all models with a value of 47.65 for the background approach. Regarding the segmentation method, the best result was obtained using segmented fingerprints using the Mask R-CNN model with a value of 45.41. In the context of a global evaluation of the NFIQ 1 and NFIQ 2 values, the techniques applied in this study for fingerprint detection and segmentation reached the NFIQ 1 values between 60 and 70, which means good to very good quality [25, 5]. In relation to NFIQ 2 values, the methods presented values whose quality is very close to good. Thus, the fingerprints detected and segmented by the models used in this study generally present sufficient quality for recognition purposes. Overall, NFIQ 1 scores tended to be slightly higher for the “segmented” versions compared to the “background” images, whereas NFIQ 2 values generally exhibited slightly inferior results for the “segmented” fingerprints. Although the segmentation achieved a high overlap accuracy (IoU = 0.96), visual inspection alone did not allow a definitive identification of the factors contributing to the observed degradation in NFIQ 2 scores. One plausible explanation is that, despite accurate foreground extraction, the segmentation process may remove peripheral ridge information near the fingertip boundaries, which can negatively affect quality metrics that are more sensitive to ridge continuity and local orientation consistency, as is the case for NFIQ 2. Conversely, the presence of background regions in non-segmented images may inadvertently contribute spurious texture or contrast variations that positively influence the quality assessment, potentially leading to inflated NFIQ 2 scores without reflecting true biometric utility. It is also important to note that no geometric normalization or unwarping models were applied in this study, such as those proposed in [41, 42]. As a result, residual perspective distortion and fingertip curvature effects remain uncorrected and may disproportionately impact the segmented images, where boundary information plays a more critical role. A more in-depth analysis of failed or degraded matches in the segmented set, as well as the incorporation of unwarping strategies, is left as future work.

Table 3 presents the average score values when comparing segmented fingerprints from smartphone images with instances of these fingerprints obtained using traditional touch

sensors. In this context, the model that presented the best result was SSD MobileNetV2, which used the approach with a background, receiving an average score of 75.78. Concerning the segmentation approach, the SSD MobileNetV2 model also presented the best result, reaching an average score of 73.80. In addition to the score, Table 4.3 also presents the number of images with scores higher than the threshold of $\lambda = 48$. Fingerprints whose scores were not equal to or greater than this threshold were considered “non-match” and, therefore, were not used to calculate the average score.

Note that the model that managed to achieve the recognition of the largest number of fingerprints was SSD MobileNetV2 FPN Lite for the segmented approach, with a total of 71 fingerprints. Regarding the background approach, the SSD MobileNetV2 FPN Lite model also achieved the highest number of matches exceeding the threshold. When crossing the number of fingerprints that exceeded the threshold and the score, it is possible to observe that the model that presented the highest average scores was also the one that had the smallest number of matches that exceeded the defined threshold.

We observed that some images in the ISPFdv2 database present fingers with blurred fingerprints. In such cases, the segmentation process works well, but both the NFIQ and the score values are expected to show poor results, reflecting a limitation of the capture process.

5. Conclusion

This study assessed detection and segmentation techniques applied to fingertip isolation in images acquired by smartphones in controlled and uncontrolled environments. The chosen models – namely SSD MobileNetV2 and SSD MobileNetV2 FPN Lite for detection and Mask R-CNN and U-Net for segmentation – were selected due to their suitability for the proposed detection and segmentation tasks. The evaluation used images from ISPFdv2, complemented by the generation of SAM-based ground truth masks for the fingertips.

Among the models evaluated, Mask R-CNN and SSD MobileNetV2 exhibited the best performance for fingertip segmentation and detection, respectively. We also assessed the quality of fingerprints and their use in biometric comparison engines. NFIQ measurements were taken to evaluate fingerprint quality and a score matching was obtained for contact-based fingerprints using the Neurotechnology VeriFin-sdk 13.0. U-Net model yielded the best results in terms of NFIQ 1. SSD MobileNetV2 FPN Lite model performed better in NFIQ 2. SSD MobileNetV2 model presented the

Table 2. NFIQ scores (versions 1 and 2) for segmented fingerprints. NFIQ 1 values are reported using the VeriFinger SDK encoding (20–100), which corresponds inversely to the NIST NFIQ 1 classes (1 = highest quality, 5 = lowest quality). Background and Segmented denote background and segmented images, respectively.

Model	NFIQ 1 ↑		NFIQ 2 ↑	
	Background	Segmentation	Background	Segmentation
SSD MobileNetV2	63.06	60.31	47.31	39.93
SSD MobileNetV2 FPN Lite	66.81	67.16	47.65	41.07
U-Net	63.06	67.69	47.23	44.61
Mask R-CNN	64.98	65.26	46.96	45.41

Table 3. Average Score and Fingerprints Counts ($\lambda = 48$). Background and Segmentation stand for Background and Segmented.

Model	Score		Count	
	Background	Segmentation	Background	Segmentation
SSD MobileNetV2	75.78	73.80	55	56
SSD MobileNetV2 FPN Lite	72.69	70.10	60	71
U-Net	74.98	70.93	58	68
Mask R-CNN	73.36	70.66	57	63

highest average score, but achieved the fewest successful comparisons. Otherwise, fingerprints segmented with the SSD MobileNetV2 FPN Lite model achieved the highest number of matches.

In the future, we will investigate deblurring or enhancing methods in our pipeline to improve fingerprint quality and interoperability with contact-based fingerprint systems. Another exciting avenue we will explore is to train a network for the scoring and matching contact and contactless fingerprint.

Acknowledgements

The authors express their gratitude to the Serviço Federal de Processamento de Dados, Brazil, for partially funding this research.

Author contributions

Author Euclides Arcoverde contributed to the development of the code and writing the article. Author Thiago contributed to the text refinement.

References

- [1] DABOUEI, A. et al. Deep Contactless Fingerprint Unwarping. In: *ICB*. [S.l.: s.n.], 2019.
- [2] CARNEY, L. A. et al. A Multi-Finger Touchless Fingerprinting System: Mobile Fingerphoto and Legacy Database Interoperability. In: *ICBBE*. [S.l.: s.n.], 2017. ISBN 9781450354844.
- [3] GROSZ, S. A. et al. C2CL: Contact to Contactless Fingerprint Matching. *IEEE Trans. on Information Forensics and Security*, 2022. ISSN 15566021.
- [4] KAUBA, C. et al. Towards using police officers' business smartphones for contactless fingerprint acquisition and enabling fingerprint comparison against contact-based datasets. *Sensors*, 2021. ISSN 14248220.
- [5] NEUROTECHNOLOGY. *Neurotechnology VeriFinger SDK v13.0*. 2023. (<https://www.neurotechnology.com/verifinger.html>). Access: Oct. 17, 2023.
- [6] ENGELSMA, J. J.; CAO, K.; JAIN, A. K. Learning a Fixed-Length Fingerprint Representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 43, n. 6, 2021.
- [7] SANKARAN, A. et al. On smartphone camera based fingerphoto authentication. *BTAS*, 2015.
- [8] MALHOTRA, A. et al. On Matching Finger-Selfies Using Deep Scattering Networks. *IEEE Trans. on Biometrics, Behavior, and Identity Science*, v. 2, 2020. ISSN 26376407.
- [9] RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Medical Image Computing and Computer-Assisted Intervention*. [S.l.: s.n.], 2015. ISBN 978-3-319-24574-4.
- [10] KUNSUK, S.; AREEKUL, V. Finger Photo Rescaling for Interoperability of Touchless and Touch-based Fingerprint Verification. In: *SITIS*. [S.l.: s.n.], 2023.
- [11] UHL, A. et al. Improving Sensor Interoperability between Contactless and Contact-Based Fingerprints Using Pose Correction and Unwarping. *IET Biometrics*, 2023. ISSN 2047-4938.
- [12] SANDLER, M. et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: *CVPR*. [S.l.: s.n.], 2018.
- [13] BEHESHTI, N.; JOHNSON, L. Squeeze U-Net: A Memory and Energy Efficient Image Segmentation Network. In: *CVPR*. [S.l.: s.n.], 2020.
- [14] SILVA, J. L. et al. Encoder-Decoder Architectures for Clinically Relevant Coronary Artery Segmentation. 2021.

- [15] BRUNA, J.; MALLAT, S. Classification with scattering operators. In: *CVPR*. [S.l.: s.n.], 2011.
- [16] BRUNA, J.; MALLAT, S. Invariant Scattering Convolution Networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013.
- [17] PRIESNITZ, J. et al. Deep Learning-Based Semantic Segmentation for Touchless Fingerprint Recognition. In: *Pattern Recognition*. [S.l.: s.n.], 2021. ISBN 978-3-030-68793-9.
- [18] CHEN, L.-C. et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: *Computer Vision – ECCV*. [S.l.: s.n.], 2018. ISBN 978-3-030-01233-5.
- [19] LIU, W. et al. SSD: Single Shot MultiBox Detector. In: LEIBE, B. et al. (Ed.). *Computer Vision – ECCV*. [S.l.: s.n.], 2016. ISBN 978-3-319-46448-0.
- [20] HE, K. et al. Mask R-CNN. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2020.
- [21] CHEN, L. C. et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2018. ISSN 01628828.
- [22] BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2017.
- [23] SUN, K. et al. High-Resolution Representations for Labeling Pixels and Regions. *CoRR*, 2019.
- [24] KUMAR, S.; KUMAR, R.; SAAD. Real-Time Detection of Road-Based Objects using SSD MobileNet-v2 FPNlite with a new Benchmark Dataset. In: *iCoMET*. [S.l.: s.n.], 2023.
- [25] TABASSI, E.; WILSON, C.; SCHLENOFF, C. *Fingerprint Image Quality*. [S.l.], 2004. Disponível em: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=905710.
- [26] TABASSI, E. et al. *NIST Fingerprint Image Quality 2*. [S.l.], 2021. Disponível em: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=920087.
- [27] KIRILLOV, A. et al. Segment Anything. *arXiv:2304.02643*, 2023.
- [28] HUANG, J. et al. Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. In: *CVPR*. [S.l.: s.n.], 2017. ISSN 1063-6919.
- [29] HUBER, P. J. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, Institute of Mathematical Statistics, v. 35, n. 1, p. 73–101, 1964.
- [30] REN, S. et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2017. ISSN 01628828.
- [31] LIN, T.-Y. et al. Focal Loss for Dense Object Detection. In: *ICCV*. [S.l.: s.n.], 2017.
- [32] LIN, T. et al. Microsoft COCO: Common Objects in Context. *CoRR*, 2014.
- [33] TENSORFLOW. *TensorFlow 2 Detection Model Zoo*. 2023. https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md. Access: Sept. 25, 2023.
- [34] ZAK, K. *U-Net implementations in Keras*. 2020. <https://github.com/karolzak/keras-unet>. Access: Sept. 25, 2023.
- [35] PRINCE, S. J. D. *Understanding Deep Learning*. MIT Press, 2023. Disponível em: <http://udlbook.com>.
- [36] LIN, T.-Y. et al. Feature Pyramid Networks for Object Detection. In: *CVPR*. [S.l.: s.n.], 2017.
- [37] WU, Y. et al. *Detectron2*. 2019. <https://github.com/facebookresearch/detectron2>.
- [38] ZUIDERVELD, K. J. Contrast Limited Adaptive Histogram Equalization. In: *Graphics Gems IV*. [S.l.: s.n.], 1994. ISBN 0123361559.
- [39] QUOC, T. T. P.; LINH, T. T.; MINH, T. N. T. Comparing U-Net Convolutional Network with Mask R-CNN in Agricultural Area Segmentation on Satellite Images. In: *NICS*. [S.l.: s.n.], 2020.
- [40] JUNG, S. et al. Benchmarking Deep Learning Models for Instance Segmentation. *Applied Sciences*, 2022. ISSN 2076-3417.
- [41] CUI, Z.; FENG, J.; ZHOU, J. *Monocular 3D Fingerprint Reconstruction and Unwarping*. 2022. Disponível em: <https://arxiv.org/abs/2205.00967>.
- [42] SÖLLINGER, D.; UHL, A. Optimizing contactless to contact-based fingerprint comparison using simple parametric warping models. In: *2021 IEEE International Joint Conference on Biometrics (IJCB)*. [S.l.: s.n.], 2021. p. 1–7.