*applied sciences*

MDPI

*Article*

# A Video-Based Mobile Palmprint Dataset and an Illumination-Robust Deep Learning Architecture for Unconstrained Environments

**Betül Koşmaz Sünnetci [1,*], Özkan Bingöl [2], Eyüp Gedikli [3], Murat Ekinci [4], Ramazan Özgür Doğan [5], Salih Türk [1] and Nihan Güngör [6]**

[1] Department of Software Engineering, Gümüşhane University, Gümüşhane 29100, Türkiye; sturk@gumushane.edu.tr
[2] Department of Software Development, Karadeniz Technical University, Trabzon 61100, Türkiye; bingolo@ktu.edu.tr
[3] Department of Computer Engineering, Trabzon University, Trabzon 61335, Türkiye; eyupgedikli@trabzon.edu.tr
[4] Department of Computer Engineering, Karadeniz Technical University, Trabzon 61100, Türkiye; ekinci@ktu.edu.tr
[5] Department of Artificial Intelligence Engineering, Trabzon University, Trabzon 61335, Türkiye; dogan@trabzon.edu.tr
[6] Basic Sciences Department, Samsun University, Samsun 55060, Türkiye; nihan.gungor@samsun.edu.tr
* Correspondence: betul.kosmaz@gumushane.edu.tr; Tel.: +90-506-296-60-62

## Featured Application

The proposed framework facilitates secure and user-friendly biometric authentication on mobile platforms, particularly in mobile payment and access control scenarios, under unconstrained conditions.

## Abstract

The widespread adoption of mobile devices has made secure and user-friendly biometric authentication critical. However, widely used modalities such as fingerprint and facial recognition show limited robustness under uncontrolled illumination and on heterogeneous devices. In contrast, palmprint recognition offers strong potential because of its rich textural patterns and high discriminative power. This study addresses the limitations of laboratory-based datasets that fail to capture real-world challenges. We introduce MPW-180, a novel dataset comprising videos of 180 participants recorded on their own smartphones in everyday environments. By systematically incorporating diverse illumination conditions (with and without flash) and natural free-hand movements, MPW-180 is the first dataset to adopt a bring-your-own-device paradigm, providing a realistic benchmark for evaluating generalization in mobile biometric models. In addition, we propose PalmWildNet, an SE-block-enhanced deep learning architecture trained with Triplet Loss and a cross-illumination sampling strategy. The experimental results show that conventional methods suffer over 50% performance degradation under cross-illumination conditions. In contrast, our method reduces the Equal Error Rate to 1–2% while maintaining an accuracy above 97%. These findings demonstrate that the proposed framework not only tolerates illumination variability but also learns robust illumination-invariant representations, making it well-suited for mobile biometric authentication.

**Keywords:** mobile palmprint recognition; in-the-wild biometric dataset; illumination-invariant deep learning

check for updates

Academic Editor: Rui Araújo

Received: 16 September 2025
Revised: 9 October 2025
Accepted: 20 October 2025
Published: 23 October 2025

**Citation:** Koşmaz Sünnetci, B.; Bingöl, Ö.; Gedikli, E.; Ekinci, M.; Doğan, R.Ö.; Türk, S.; Güngör, N. A Video-Based Mobile Palmprint Dataset and an Illumination-Robust Deep Learning Architecture for Unconstrained Environments. *Appl. Sci.* **2025**, *15*, 11368. https://doi.org/10.3390/app152111368

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

The increasing reliance on mobile devices in everyday life has made secure and convenient verification of digital identities more critical than ever. Biometric authentication, which is rapidly replacing traditional methods such as passwords or PINs, offers a stronger layer of security by exploiting unique physiological and behavioral traits [1,2]. Despite their widespread deployment on mobile platforms, dominant modalities such as fingerprint and face recognition still face notable limitations in terms of spoof resistance, reliability, and usability [3–5]. For example, facial recognition systems can be bypassed using covertly captured images, whereas fingerprint sensors are vulnerable to latent prints lifted from surfaces such as glass [3]. These vulnerabilities, combined with factors such as sensor degradation, contamination, and low user acceptance, particularly in touch-based systems, highlight the urgent need for novel and more robust biometric solutions.

Palmprint recognition is emerging as a robust and promising biometric modality owing to its rich textural structure—comprising principal lines, wrinkles, and ridges—along with its high inter-individual discriminative power and relatively large surface area [1,6,7]. Furthermore, palmprint-based systems are attractive for practical applications because they can achieve high accuracy even when implemented using low-cost devices [8,9].

Palmprint recognition has gained remarkable momentum in recent years, largely owing to advances in deep neural network architectures [7,10–12]. Nevertheless, much of this progress has relied on datasets collected under highly constrained conditions, including fixed scanners, controlled illumination, and cooperative behavior. Such idealized laboratory settings do not adequately capture the challenges inherent in real-world scenarios [11,13]. Key obstacles include device-induced variability in image quality [14–16], sudden changes in illumination [17–20], variations in free-hand positioning [3,21–23], and motion blur caused by natural hand movement [10].

To address this gap, recent studies have increasingly focused on datasets that incorporate real-world challenges by acquiring images using mobile devices [13,24,25]. Capturing palmprints using consumer-grade smartphones provides notable advantages in terms of convenience and hygiene, thereby supporting their adoption in everyday applications such as mobile payment systems. This trend highlights the significant practical potential of palmprint recognition technology and indicates its growing acceptance in real-world applications.

Nevertheless, as emphasized in several studies, even these next-generation datasets suffer from important shortcomings, such as limited device diversity [14–16], incomplete coverage of illumination variations [20,26], and insufficient representation of the dynamics associated with free-hand movements [3,22,27]. These limitations indicate that, particularly in contactless systems, adverse factors, most notably illumination variability, still exert a substantial influence on system performance.

To address these critical gaps, this study introduces a comprehensive framework that facilitates the transition from controlled experiments to real-world applications. Our first contribution is a data collection methodology that enables participants to record videos in everyday environments using their own smartphones across diverse makes and models. This approach yields a video-based, in-the-wild dataset that systematically captures real-world challenges, including device diversity, uncontrolled illumination and unconstrained hand movement. To the best of our knowledge, the resulting dataset is the first of its kind, comprising complete video sequences rather than single frames and exhibiting a high degree of device heterogeneity. This contribution provides a realistic benchmark for advancing the robustness and generalization of palmprint recognition models under unconstrained mobile conditions.

The second contribution of this study is the development of an end-to-end deep learning pipeline, PalmWildNet, which combines an SE-block-enhanced backbone with

a triplet loss-based metric learning strategy. The pipeline is explicitly optimized for cross-illumination conditions using a novel positive sampling mechanism that pairs instances acquired under different lighting scenarios. This design enables the model to learn illumination-invariant representations, thereby improving its robustness and generalization capability across diverse real-world environments.

The remainder of this paper is organized as follows: Section 2 reviews the related work. Section 3 introduces the MPW-180 dataset and the data collection methodology. Section 4 presents the proposed PalmWildNet architecture and training strategies. Section 5 describes the experimental setup and reports the results, followed by Section 6, which presents discussion, conclusion and summarizes potential directions for future research.

## 2. Related Works

### 2.1. The Evolution of Palmprint Recognition: From Conventional Methods to Deep Learning

Palmprint recognition has a long and well-established history in the field of biometrics. Early studies primarily focused on handcrafted feature extractors using manual feature engineering. Notable examples include methods designed to capture micro-texture patterns, such as Local Binary Patterns (LBP) [28–30] and techniques that encode the orientation of principal lines and wrinkles (e.g., PalmCode [31], FusionCode [32], CompCode [33], DoGCode [34], OrdinalCode [35,36], RLOC [37], and BOCV [38]). In addition, methods that measure specific textural responses, including Gabor filters [39–41] and principal line extraction [42–46], have been widely adopted in the literature.

To address the high dimensionality of palmprint data, subspace learning methods have been extensively employed. These include Principal Component Analysis (PCA) [47–50], Linear Discriminant Analysis (LDA) [51,52], Independent Component Analysis (ICA) [49,50,53], and Locality Preserving Projections (LPP) [54,55], along with their numerous variants such as 2DPCA [56], 2DLDA [57,58], Kernel PCA [51], Kernel LDA [22,55], and 2DLPP [59,60]. Although these conventional approaches have achieved considerable success in controlled laboratory settings, they remain sensitive to environmental perturbations, such as illumination, orientation, scale, and partial occlusion, and therefore exhibit limited generalization capability.

Over the past decade, the advent of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized palmprint recognition. CNNs have largely overcome the limitations of handcrafted feature engineering through their ability to learn hierarchical features autonomously from raw data. Pretrained architectures such as ResNet [61], VGGNet [62], and DenseNet [63] have been successfully applied to palmprint recognition using transfer learning [64–67]. This strategy substantially improves the generalization capability of models, particularly when dealing with limited datasets.

In addition to the general success of deep learning, numerous CNN-based models have been specifically designed or adapted for biometric tasks, such as palmprint recognition. These models were customized to capture the distinctive features, patterns, and textures of palm images more effectively. Such customization typically involves tailoring the network architecture, loss functions, or preprocessing steps to meet the unique requirements of a task. Prominent models in this field, which have demonstrated strong performance across multiple experiments, include PalmNet [68], which integrates Gabor filters and PCA into the CNN framework; PalmRCNN [64], built upon ResNet and InceptionNet architectures; CompNet [69], known for using learnable Gabor kernels and having a small number of model parameters; DHN [70], which transfers summary codes learned from the teacher model to the student model using knowledge distillation; RegPalm [71], which is designed to reduce pattern variance for large-scale palmprint recognition.

However, despite these architectural advances, a fundamental problem remains: the training and evaluation of such models primarily rely on controlled and homogeneous datasets. Typically acquired with fixed scanners, consistent illumination, and cooperative participants, these datasets fail to capture the complexity and variability inherent in real-world applications.

### 2.2. Mobile Platforms and In-the-Wild Palmprint Recognition

The limitations of existing palmprint datasets in reflecting real-world performance have shifted research attention toward mobile platforms and less constrained scenarios in recent years. Leading this transition, datasets such as XJTU-UP [70] and BJTU_PalmV2 [72] were collected using five different smartphone models, providing a crucial foundation for studying the cross-sensor variability. Similarly, datasets such as the MPD [73] and NTU-CP-v1 [65] were acquired using two distinct devices. Together, these efforts mark the initial steps toward moving beyond single-device acquisition protocols and examining the influence of heterogeneous hardware and software on recognition performance.

However, despite incorporating device diversity, these seminal studies faced a key methodological limitation: data acquisition was restricted to a small, pre-selected set of device models (e.g., two or five). This constraint introduces the risk of models developing device-specific biases based on the unique characteristics of this limited set. Moreover, it offers only a partial view of how such systems would generalize to a truly unconstrained ecosystem encompassing hundreds of diverse, unseen device models encountered in the wild.

In the literature, the most unconstrained example of palmprint data collection is the NTU-PI-v1 [74] database, which was compiled from online images. Containing samples from 1093 individuals, this dataset reflects a truly in-the-wild setting because it was collected without user cooperation or adherence to any acquisition protocol. However, this extreme lack of control renders the dataset unsuitable for biometric-recognition tasks. The limited number of samples per identity, inability to systematically regulate illumination and pose conditions, and uncertainties in label reliability restrict its utility primarily to tasks such as palmprint detection. This underscores the persistent challenge of balancing realism and methodological rigor when constructing palmprint recognition datasets.

### 2.3. The Challenge of Illumination Variation

Illumination variability remains one of the most fundamental challenges in mobile and unconstrained environments such as smart homes. The literature reports a wide spectrum of approaches to address this issue, ranging from classical image processing techniques to advanced deep learning models. Classical methods, such as Contrast Limited Adaptive Histogram Equalization (CLAHE) [75] and Retinex theory [76], often prove inadequate under complex, non-uniform illumination conditions. In contrast, deep learning has introduced more robust solutions, including

- Data Augmentation, which enhances training sets through random modifications in brightness and contrast.
- Metric Learning, which brings images of the same subject under different lighting conditions closer in the feature space.
- Image Fusion combines images captured with and without flash.

Although these strategies have been widely explored in the face recognition literature, they have not been systematically applied to palmprint recognition without contact. In particular, the performance implications of harsh flash illumination versus ambient light remain insufficiently addressed, representing a significant gap in this field.

This study was designed to address this critical research gap. We present a comprehensive investigation that systematically examines mobile palmprint recognition under variable and uncontrolled illumination. To this end, we introduce, to the best of our knowledge, the first video-based mobile palmprint dataset captured in the wild. In parallel, we propose a deep learning model specifically developed to address the challenges posed by this dataset, demonstrating strong robustness to illumination variability. Through these contributions, we seek to advance palmprint biometrics beyond constrained laboratory settings and into the realistic and challenging conditions of everyday life.

## 3. Dataset: MPW-180 (Mobile Palmprint in the Wild—180)

To mitigate the transient distortions inherent in in-the-wild scenarios, such as motion blur and focus shifts, a video-based acquisition strategy was adopted instead of static image capture. This approach produces hundreds of potential Region of Interest (ROI) frames for each subject from a single video. It not only provides the rich data pool required to train deep learning architectures but also enables the selective curation of the sharpest frames, thereby improving the overall quality of the dataset used for training the model. Based on this strategy, a structured data collection protocol was designed to simulate unconstrained mobile usage. Each participant recorded approximately 30 s videos under four distinct scenarios, systematically covering the key aspects of palmprint biometrics.

These scenarios combined two primary illumination conditions (LED flash enabled vs. ambient light) with two biometric modalities (right vs. left hand). This design yielded four sub-datasets: HR_FT (right hand with flash), HR_FF (right hand without flash), HL_FT (left hand with a flash), and HL_FF (left hand without flash). The two illumination conditions defined—with flash and ambient light—were a deliberate methodological strategy to isolate the fundamental domain shift problem. The "ambient light" (without flash) scenario was designed to capture a broad spectrum of uncontrolled, real-world conditions. Recordings took place in participants' natural environments (e.g., home and office) without restrictions on ambient lighting properties. Consequently, this condition inherently includes complex illumination factors, such as mixed interior lighting, variable shadow casting, and diverse color temperatures.

Informed consent was obtained from all participants after they were briefed on the study's aims, procedures, and the anonymous use of their data for scientific research. All data were fully anonymized by removing any Personally Identifiable Information (PII) and assigning each participant a unique, non-traceable ID.

Before data collection, the participants were instructed to position their palms within the camera's field of view and move their hands slowly during recording to introduce natural variations in pose, scale, and focus. The instructed movements included lateral (side-to-side), vertical (up-and-down), and depth (back-and-forth) motions, as well as finger movements toward and away from the camera. No further guidance or interventions were provided during the recordings. Data acquisition took place in the participants' natural environments (e.g., home and office), without restrictions on background complexity or ambient lighting properties (intensity, direction, and color temperature). This unconstrained setting ensured that the dataset authentically reflected the challenges of real-world conditions.

This study introduces a novel, large-scale, video-based dataset named Mobile Palmprint in-the-Wild (MPW-180), collected from 180 participants. The MPW-180 dataset constitutes the first phase of a larger data collection effort involving more than 500 volunteers, for whom all data processing and annotation pipelines have been carefully completed. To ensure consistency and reproducibility, this study focused exclusively on the 180-subject subset, for which all verification and annotation procedures were finalized. The MPW-180

dataset will be made publicly available upon acceptance of this paper to facilitate reproducibility and advance research on mobile palmprint recognition. The dataset contains four video recordings per participant, and its key statistics are summarized in Table 1.

**Table 1.** Statistical Summary of the MPW-180 Dataset.

| Parameter | HR_FT | HR_FF | HL_FT | HL_FF | Total |
|---|---|---|---|---|---|
| Number of Participants | 180 | 180 | 180 | 180 | 720 |
| Number of Videos | 180 | 180 | 180 | 180 | 720 |
| Average Video Duration (s) | 31.98 | 32.53 | 31.76 | 32.35 | - |
| Minimum Video Duration (s) | 12.05 | 15.03 | 12.10 | 16.08 | - |
| Maximum Video Duration (s) | 67.10 | 59.03 | 58.30 | 62.23 | - |
| Total Number of Frames | 193,386 | 192,804 | 193,089 | 189,635 | 768,914 |
| Average Number of Frames | 1074 | 1071 | 1072 | 1053 | - |
| Total Selected ROIs | 165,466 | 165,676 | 163,011 | 160,873 | 655,026 |
| Average Selected ROIs | 919 | 920 | 905 | 893 | - |
| Video Resolution | from 480 × 864 to 2160 × 3840 | | | | |
| ROI Resolution | from 128 × 128 to 1585 × 1585 | | | | |
| Videos per Participant | 4 | | | | |
| Selected ROIs per Participant | 3639 | | | | |

The MPW-180 dataset was collected using 180 distinct smartphone devices from 11 different brands. Beyond brand diversity, the distribution of specific device models was also broad. The most frequently represented models were the iPhone 11 (n = 29; Apple Inc., Cupertino, CA, USA), Redmi Note 8 (n = 19; Xiaomi Corp., Beijing, China), Redmi Note 8 Pro (n = 7), and Galaxy A32 (n = 5; Samsung Electronics Co., Ltd., Suwon, Republic of Korea). The dataset also encompasses a wide range of less common models, including the iPhone 13/14 series, Galaxy A54/A34, Huawei P30/P40 (Huawei Technologies Co., Ltd., Shenzhen, China), and devices from Poco (Xiaomi Corp., Beijing, China), GM (Istanbul, Türkiye), Infinix (Infinix Mobility, Hong Kong, China), Oppo (Guangdong Oppo Mobile Telecommunications Corp., Dongguan, China), and Reeder (Reeder Technology, Samsun, Türkiye). This heterogeneity is essential for ensuring that recognition systems are robust against device-centric variations arising from differences in camera hardware and internal image processing pipelines. The distribution of device brands represented in the dataset is shown in Figure 1.

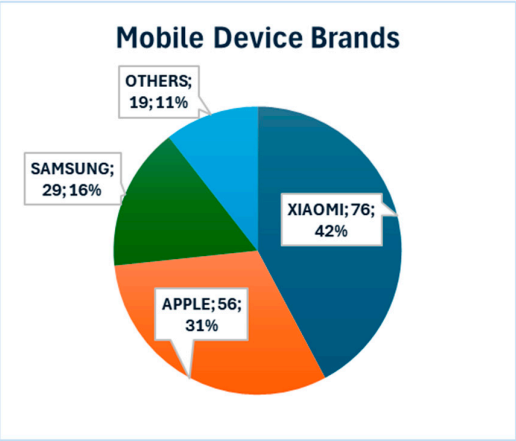



**Figure 1.** Distribution of smartphone brands represented in the MPW-180 dataset, highlighting the heterogeneity of the devices used during data acquisition.

To underscore the novel contribution of the MPW-180 to the literature, Table 2 presents a systematic comparison with other prominent mobile palmprint datasets. As the table indicates, MPW-180 is unique in being the only large-scale, video-based dataset acquired from

participants using their own personal smartphones in a fully unconstrained bring-your-own-device (BYOD) setting. Unlike earlier datasets that were collected under controlled conditions with limited device diversity and static image capture, MPW-180 captures the true variability of real-world mobile scenarios. The scale, diversity of devices, and use of unconstrained video acquisition collectively establish the MPW-180 as a critical benchmark for advancing research on robust, contactless palmprint recognition in realistic environments.

**Table 2.** Comparison of prominent mobile-based and contactless palmprint datasets.

| Dataset | Acquisition Device | Subjects | Hand (R/L) | Illumination Variation | Video/ Image | Samples | Resolution | Background |
|---|---|---|---|---|---|---|---|---|
| COEP [77] | Dijital Camera (1 model) | 168 | Right | N/A | Image | 1344 | 1600 × 1200 | Controlled |
| Stereo Palmprint [22] | Webcam (2 identical cameras) | 138 | Right | White LED | Image | 2760 | 640 × 480 | Uncontrolled |
| KTU Contactless [78] | CCD Camera (1 model) | 145 | Right | White LED | Image | 1752 | 768 × 576 | Uncontrolled |
| IITD Touchless [79] | CCD Camera (1 model) | 230 | Both | Circular Fluores-cent | Image | 2601 | 800 × 600 | Controlled |
| CASIA Palmprint [35] | CMOS Camera (1 model) | 312 | Both | LED | Image | 5502 | 640 × 480 | Controlled |
| PolyU_3D2D [80] | 3D Digitizer (1 model) | 177 | Right | Structured light | Image & 3D Data | 3540 | 640 × 480 | Controlled |
| PolyU-IITD Contactless [81] | Digital Camera (N/A) | 600 | Both | Without flash | Image | 6000 | from 1280 × 960 to 4352 × 3264 | Controlled |
| XJTU-UP [70] | Smartphone (5 models: iPhone 6S, Huawei Mate8, LG G4, Samsung Galaxy Note5, MI8) * | 100 | Both | With & without flash | Image | 20,000 | from 3264 × 2448 to 5312 × 2988 | Uncontrolled |
| MPD [73] | Smartphone (2 Models: Huawei ve Xiaomi) | 200 | Both | With & without flash | Image | 16,000 | 224 × 224 (ROI) | Uncontrolled |
| BJTU_PalmV2 [72] | Smartphone (5 models: iPhone 6, Nexus 6p, Huawei Mate8, Nubia Z9, Xiaomi Redmi 1S) * | 148 | Both | N/A | Image | 2663 | Normalized: 3264 × 2448 | Uncontrolled |
| NUIG_Palm1 [82] | Smartphone (5 models: N/A) | 81 | Both | With & without flash | Image | 1616 | from 8 Mega-pixel to 16 Mega-pixel | Uncontrolled |
| PRADD [83] | Smartphone (2 models) and 1 digital camera | 100 | Right | Without flash | Image | 12,000 | from 2048 × 1536 to 3264 × 2448 | Controlled |
| Mobile PV1 [84] | Smartphone (Multiple types of mobile phones: N/A) | 210 | Both | With & without flash | Image | 13,440 | N/A | Uncontrolled |
| SYEnet Palmprint [85] | Smartphone (1 model) | 200 | Both | With & without flash | Image | 6400 | 1280 × 720 | Uncontrolled |
| Birjand MPD [24] | Smartphone (1 model) | 41 | Both | With flash | Image | 1640 | 3264 × 2448 | Controlled |
| Sapienza MPD [24] | Smartphone (1 model) | 110 | Both | With flash | Image | 1800 | 3264 × 2448 | Controlled |
| Tongji Contactless Palmprints [77] | CCD Camera (1 model) | 300 | Both | White LED | Image | 12,000 | 800 × 600 | Controlled |
| NTU-CP-v1 [65] | Different Cameras (2 models) | 328 | Both | Without flash | Image | 2478 | from 420 × 420 to 1977 × 1977 | Uncontrolled |
| HIT-NIST-V1 [72] | Smartphone (5 models: Meizu 16S, Huawei Nova5Pro, Huawei Mate8, iPhone 6 and Xiaomi Redmi 1S) * | 162 | Both | Without flash | Image | 3020 | from 960 × 2448 to 1280 × 4208 | Uncontrolled |

**Table 2.** *Cont.*

| Dataset | Acquisition Device | Subjects | Hand (R/L) | Illumination Variation | Video/ Image | Samples | Resolution | Background |
|---|---|---|---|---|---|---|---|---|
| NTU-PI-v1 [74] | Internet Images (N/A) | 1.093 | Both | N/A | Image | 7781 | from 30 × 30 to 1415 × 1415 | Uncontrolled |
| MPW-180 (Our Dataset) | Smartphone (180 distinct devices: Each participant used their own device) | 180 | Both | With & without flash | Video | 655,026 | from 480 × 864 to 2160 × 3840 | Uncontrolled |

* The manufacturers and origins of the devices listed in this table are provided as follows—iPhone 6 and iPhone 6S (Apple Inc., Cupertino, CA, USA); Huawei Mate8, Huawei Nova5Pro (Huawei Technologies Co., Ltd., Shenzhen, China); LG G4 (LG Electronics Inc., Seoul, Republic of Korea); Samsung Galaxy Note5 (Samsung Electronics Co., Ltd., Suwon, Republic of Korea); MI8 and Xiaomi Redmi 1S (Xiaomi Corporation, Beijing, China); Nexus 6P (Google LLC, Mountain View, CA, USA); Nubia Z9 (ZTE Corporation, Shenzhen, China); Meizu 16S (Meizu Technology Co., Ltd., Zhuhai, China).

Representative hand images of multiple subjects in the dataset are shown in Figure 2. As illustrated, the extracted video frames were fully unconstrained and captured without strict subject cooperation. This results in substantial variations in hand postures; for example, fingers may appear closely grouped in some frames, while in others, they are noticeably spread apart. In addition, the samples demonstrated pronounced differences in ambient illumination, ranging from strong overexposure to dim lighting. These variations highlight the authenticity of the dataset and the challenges it poses in developing robust palmprint recognition systems.



**Figure 2.** Representative video frames from participants in the MPW-180 dataset, illustrating variations in hand posture and illumination conditions.

## 4. Proposed Method

### *4.1. Overall Pipeline*

The end-to-end workflow of the proposed system is depicted schematically in Figure 3. The pipeline begins with video frames captured by a mobile device camera, which serves as the input to the system. In the first stage, the hand and palm regions are automatically detected, and a ROI is extracted for recognition. The ROI was then resized to a fixed resolution, and its pixel values were normalized to reduce geometric and photometric distortions. After preprocessing, the ROI was passed to a deep convolutional neural network, PalmWildNet, which served as the backbone of the system. Designed specifically for mobile palmprint recognition, PalmWildNet transforms the input into a discriminative feature representation by hierarchically learning the rich textural and linear patterns that are characteristic of the palm.

The final stage of the system, output generation, was designed to operate in two distinct modes to accommodate the experimental scenarios examined in this study. The first mode, Classification-Based Recognition, employs an N-class Softmax layer, where features extracted by PalmWildNet are used to produce probabilistic predictions of the subject's identity. This configuration served as a baseline for performance evaluation and for assessing the challenges introduced by cross-illumination conditions. The second mode, which constitutes the core of our framework, is the Metric Learning-Based Embedding. In this

setting, the Softmax layer was replaced with an embedding layer that projected palmprint images into a D-dimensional Euclidean space. The primary objective of this approach is to learn an identity-discriminative and generalizable feature representation that remains robust under varying light conditions. Together, these two output mechanisms form the foundation for the comparative analyses presented in the next sections.

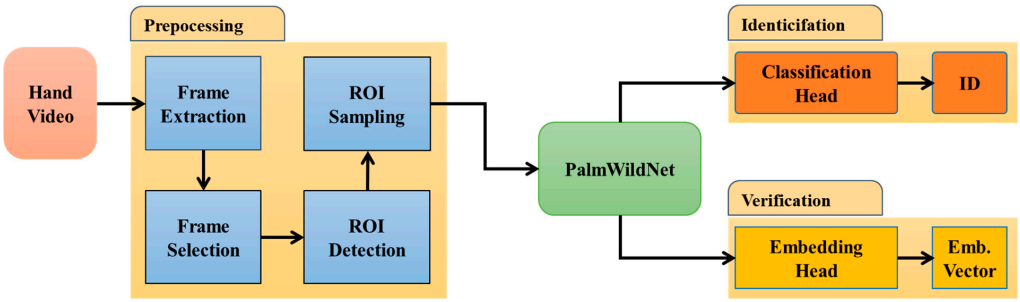



**Figure 3.** Schematic representation of the proposed mobile palmprint recognition pipeline.

### *4.2. Data Pre-Processing*

Transforming raw video data into a standardized format suitable for deep learning is a fundamental prerequisite for the success of the recognition system. The stages of this preprocessing workflow are shown in Figure 4. The primary objective of this offline phase is to generate high-quality, geometrically consistent palm ROI images from raw video streams while correcting for pose distortions caused by natural hand postures. This workflow consists of four key steps that explicitly account for the physiological structure of the hand: (1) temporal sampling and frame extraction, (2) frame selection, (3) ROI detection, and (4) ROI sampling.

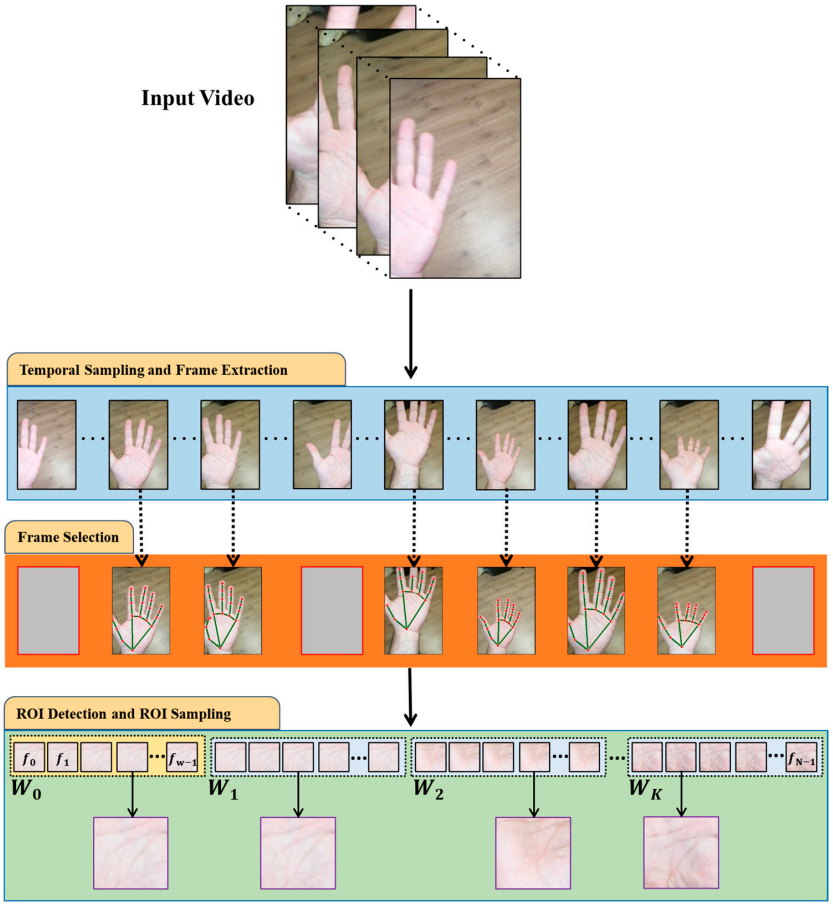



**Figure 4.** Preprocessing pipeline for generating palm ROI images from raw video data, including temporal sampling, frame selection, ROI detection, and ROI sampling.

#### 4.2.1. Temporal Sampling and Frame Extraction

Each video recording was decomposed into its constituent frames. To ensure temporally consistent sampling, a robust strategy was adopted that accounted for the variable frame rates (FPS) often observed in mobile devices. Given a target FPS, the system leverages the video timestamps to perform resampling. This method is more accurate than conventional fixed-interval frame skipping and preserves a consistent temporal density even during slow or fast hand movements. When reliable timestamp information is unavailable because of codec limitations, the system defaults to an index-based sampling approach that uses the video's reported FPS.

#### 4.2.2. Frame Selection

After frame extraction, Google's MediaPipe library [86,87] was used to identify frames that were suitable for further processing and to localize the palm region within each frame. MediaPipe is designed to detect 21 hand landmarks with high accuracy for both hands, as shown in Figure 5.

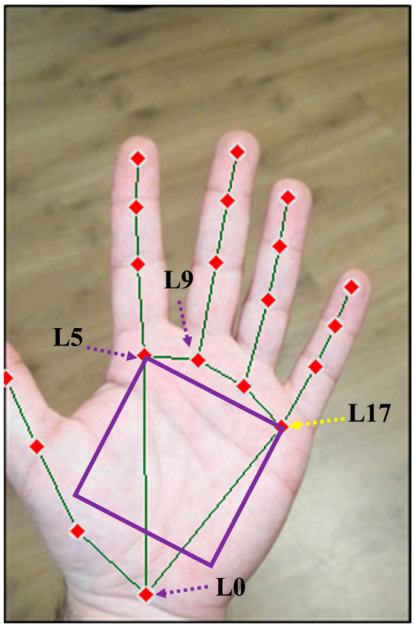



**Figure 5.** Hand landmark detection using the MediaPipe framework. The 21 anatomical landmarks, highlighted in red, correspond to physiologically meaningful joint positions of the hand. The palm ROI, delineated in purple, is geometrically derived from these landmarks to ensure consistent and orientation-invariant localization across different poses and illumination conditions.

The selection of the MediaPipe Hands framework for keypoint detection is a critical design choice, offering distinct advantages over conventional segmentation or valley-point-based ROI localization methods. The success of palmprint ROI extraction traditionally depends on accurately detecting the valleys between fingers, which are highly susceptible to pose variations, nonideal finger spacing, and motion blur, often resulting in erroneous valley points [88–91]. In contrast, MediaPipe, which has been validated in complex hand-based academic studies, such as sign language recognition, demonstrates superior landmark detection stability [92–95]. By employing a highly optimized, real-time convolutional neural network architecture, it robustly extracts 3D skeletons and 2D keypoints despite the pose and scale variations inherent in unconstrained video capture. This proven robustness to device heterogeneity and illumination changes is vital for the in-the-wild characteristics of the MPW-180 dataset, thereby supporting the reliability and reproducibility of the

corpus [96,97]. Furthermore, the success of MediaPipe in detecting palmprint ROI has been demonstrated in several studies [98–100].

Nevertheless, it is worth noting that MediaPipe may experience certain limitations, such as failing to detect landmarks when the hand is too close to the camera or when fingers are not visible. Likewise, right–left hand confusion can occur in frames where the thumb is occluded. However, these issues can be readily mitigated because the hand side is known for each video, allowing such cases to be automatically identified during preprocessing. Frames in which the MediaPipe model failed to reliably detect a hand, such as those affected by excessive motion blur, partial hand visibility, or poor illumination, were immediately discarded because no reliable ROI could be extracted. The frame selection process is further refined by excluding frames for which the subsequent ROI detection step (described in the following section) is unsuccessful.

#### 4.2.3. Geometrically Consistent ROI Detection

While most palmprint studies determine the ROI using the valley points between the index and middle fingers and between the middle and ring fingers, these points are highly sensitive to hand posture. Instead, we leveraged the intrinsic anatomical stability of the metacarpophalangeal (MCP) joints by utilizing the $L_5$ (INDEX_FINGER_MCP) and $L_{17}$ (PINKY_MCP) keypoints to define the ROI quadrilateral. This method takes advantage of the inherent anatomical stability of the MCP joints to establish a baseline that is less affected by natural finger spacing and wrist movements. This geometric consistency is further reinforced by using the wrist ($L_0$) and middle finger joint ($L_9$) as reference points to orient the normal vector ($\vec{n}$) toward the palm (Equations (1) and (2)). By relying on the consistent keypoint coordinates provided by the deep learning-based MediaPipe model, we ensured that the generated ROI consistently captured the feature-rich central region of the palm. Consequently, a stable ROI can be produced even under hand rotation or camera perspective distortion.

Our methodology defines a pose- and orientation-invariant ROI quadrilateral by using the $L_5$ and $L_{17}$ landmarks. Specifically, the top edge of the ROI was defined as the line segment connecting $L_5$ and $L_{17}$. The Euclidean distance between these two points determines the side length, $s = \|L_{17} - L_5\|_2$, of the square to be extracted, where $L_i$ represents the vector of pixel coordinates for the i-th landmark. As the target ROI is a square, the other two vertices are located by adding a vector of length s that is perpendicular to the top edge and directed towards the palm. To determine the correct palm orientation ($n$) of the two possible normal vectors ($n_1, n_2$) perpendicular to the $L_5\vec{L}_{17}$ vector, the one closest to a reference point ($L_{ref}$) is chosen, which is the midpoint between the wrist ($L_0$) and the middle finger joint ($L_9$). This ensured that the ROI remained within the palm, regardless of the direction in which the hand was facing.

$$L_{\text{ref}} = \frac{L_0 + L_9}{2} \tag{1}$$

$$\vec{n} = \underset{\vec{n'} \in \{\vec{n_1}, \vec{n_2}\}}{argmin} \left\| \left( \frac{L_5 + L_{17}}{2} + s.\vec{n'} \right) - L_{\text{ref}} \right\|_2 \tag{2}$$

As a result of these steps, a source rectangle, $Q_{src} = \{L_5, L_{17}, L_{17} + s.n, L_5 + s.n\}$, is defined, which is not usually a square in the image plane owing to perspective, but corresponds to a square region of the palm. Then, using the $Q_{src}$ rectangle, the palm ROI image was both cropped and oriented by warping. This operation is accomplished by defining a square target rectangle, $Q_{dst} = \{(0,0),\ (S-1,0),\ (S-1,S-1),\ (0,S-1)\}$, with a side length $S$ of pixels. Here, $S$ is determined by the roi_size parameter or dynamically adjusted

according to the top side length s. Then, a $3 \times 3$ perspective transformation matrix $M$ is calculated that maps the vertices of $Q_{src}$ to the vertices of $Q_{dst}$. Finally, the matrix $M$ is applied to the original square. This process always transforms the palm region $Q_{src}$ into a square image, regardless of the hand position. Bicubic interpolation was used during this transformation to maximize the image quality. Furthermore, frames in which the ROI region critical for recognition could not be generated were excluded from the list. The frames to be used can be selected by simultaneously identifying all the desired ROI corner points. Therefore, this process was performed by considering the presence and location of hand landmarks.

#### 4.2.4. ROI Sampling

Consecutive frames extracted from raw video data are highly correlated in the temporal domain. Directly using these near-duplicate frames for deep learning model training introduces two critical issues [88] that must be addressed. First, it substantially increases the risk of overfitting, thereby reducing the effective diversity of the dataset. Second, variations in the video duration, recording quality, and device settings across participants resulted in significant imbalances in the number of valid ROIs obtained per video. To address these challenges and ensure the creation of a balanced, high-quality dataset, we developed a two-stage quality-based subsampling strategy.

The first step of this strategy is to divide the time series $F = \{f_0, f_1, \ldots, f_{N-1}\}$, consisting of $N$ ROIs obtained from each video, into $K$ consecutive, non-intersecting windows ($W$), from which a single ROI is selected. Here, $K$ is a parametric value representing the total number of samples targeted for that video (for training, validation, and testing). The window width was calculated as $w = \lfloor N/K \rfloor$. Thus, the i-th window is defined as $W_i = \left\{ f_{(i-1)w+1}, \ldots, f_{iw} \right\}$. This approach ensures temporal diversity by ensuring that the samples are spread over the entire video duration.

The second and most critical step of the subsampling strategy involves selecting the most informative frame within each window $W_i$. i.e., the frame with the highest feature potential for model training. The Laplacian variance was employed as a focus metric. This measure quantifies image sharpness, which is directly associated with the prominence of edges and preservation of high-frequency textural details, both of which are essential for discriminative palmprint feature extraction. The Laplacian operator, a second-order derivative operator, is highly sensitive to high-frequency regions such as edges and fine textural details. For a two-dimensional image $I(x, y)$, the Laplacian $\nabla^2 I$ is defined as:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{3}$$

The sharpness score of an image, $S(I) = Var\left(\nabla^2 I(x, y)\right)$, is calculated as the variance of the Laplacian-applied image. A higher value of $S(I)$ indicates that the image has more edge and texture details and is therefore more sharply focused. The proposed selection procedure calculates the score $S\left(f_j\right)$ for all frames within each window $W_i$ and selects a single frame, $f_i^*$, with the highest score:

$$f_i^* = \underset{f_j \in W_i}{\operatorname{argmax}} S\left(f_j\right) \tag{4}$$

The proposed windowing and max-sharpness selection strategy offers three principal advantages. First, it exposes the model to a broader range of data by substantially reducing the redundancy of consecutive, highly correlated frames. Second, it ensured dataset balance across participants and recording scenarios by uniformly selecting a fixed number ($K$) of samples from each video. Third, it actively filters out frames affected by motion blur or

defocus artifacts, thereby improving the training quality and maximizing the likelihood of working with ROIs that possess the highest feature extraction potential. Figure 6 illustrates representative examples of the ROI acquisition process for a single participant.

To prevent potential information leakage during temporal windowing, care was taken to ensure that no frame from the same video appeared in more than one partition. However, owing to the sequential nature of video data, the last frame of one window may be temporally adjacent to the first frame of the subsequent window. This minimal overlap was intentionally permitted to maintain temporal continuity and enhance diversity among the samples. Furthermore, this effect is considered negligible, as any pair of windows is separated by a sequence of frames at least one full window length apart.



(**a**) (**b**) (**c**) (**d**)

**Figure 6.** Representative video frames and corresponding ROIs from a participant in the MPW-180 dataset: (**a**) right hand with flash, (**b**) right hand without flash, (**c**) left hand with flash, and (**d**) left hand without flash.

#### 4.2.5. Quality Control and Data Curation

Following automated ROI extraction and quality-based subsampling, a manual verification and curation stage was conducted to further enhance the quality and consistency of the dataset. During this rigorous review process, ROI images that were judged by an expert to lack sufficient discriminative information for biometric recognition were systematically removed.

The elimination criteria were defined by objective rules and targeted three major categories of defects:

- Severe motion blur or illumination artifacts, including overexposed frames in which texture details are lost and underexposed frames in which noise dominates.
- Incorrect ROI localization, where the algorithm mistakenly selected non-palm regions (e.g., fingers, wrist, or back of the hand) owing to non-standard hand positions or extreme perspective distortions.
- Partial or complete occlusion of the palm, where key outlines or textural areas were obscured by external objects.

All ROI images that passed this quality control stage were preserved in their original square resolution and were stored as independent files. For reproducibility, the curated dataset was organized into a hierarchical folder structure indexed by user identity. Figure 7 shows representative examples of manually curated images.



**Figure 7.** Examples of ROI excluded during the manual quality control process, illustrating common defects such as severe motion blur, illumination artifacts, incorrect ROI localization, and palm occlusion.

### *4.3. On-the-Fly Data Normalization and Augmentation*

To enhance the generalization capability of the model and improve its robustness against variations such as illumination, data augmentation and normalization operations were applied on-the-fly by the data loader during training. This dynamic approach allows the generation of an effectively infinite number of data variations without increasing the storage requirements on the disk. Two distinct transformation pipelines were defined for the training and evaluation stages of the study.

Training Transformations: During the training phase of the model, the following transformations were randomly applied to each ROI image:

- All ROI images were rescaled to 224 $\times$ 224 pixels to match the input size of the PalmWildNet architecture.
- To increase the model's robustness to small changes in hand position and orientation, slight geometric distortions were applied to the images, such as random rotations within $\pm$8 degrees and random affine transforms of $\pm$2%.
- To simulate slight variations in ambient illumination, the brightness and contrast values of the images were randomly adjusted within $\pm$15%.
- A sharpening filter was applied with a 50% probability to emphasize the fine crease and texture details on the palm and encourage the model to focus on these distinctive features.
- Finally, the images were converted to a tensor in a single-channel grayscale format, and the pixel values were normalized to the range [$-$1, 1] using a mean of 0.5 and standard deviation of 0.5.

Validation and Test Transformations: For validation and testing, a deterministic preprocessing pipeline was employed to ensure a consistent and reproducible performance evaluation. This pipeline included only resizing (224 $\times$ 224), grayscale conversion, tensor conversion, and normalization. No augmentation techniques were applied in these phases.

### *4.4. PalmWildNet Deep Network Architecture*

The PalmWildNet architecture, developed for primary performance analyses in this study, was designed to capture rich and discriminative palmprint patterns while maintaining a balance between model depth and computational efficiency. The design is inspired by prior work [101] and follows the philosophy of ResNet [61]. Central to this architecture is the integration of Squeeze-and-Excitation (SE) blocks [102] into each residual module, introducing a channel-wise attention mechanism that adaptively recalibrates feature responses. This hybrid design enables the model to learn deep spatial hierarchies and dynamically emphasize the most informative feature channels, thereby enhancing its

representational power. The overall structure of PalmWildNet is illustrated in Figure 8, and its key components are summarized in Table 3.

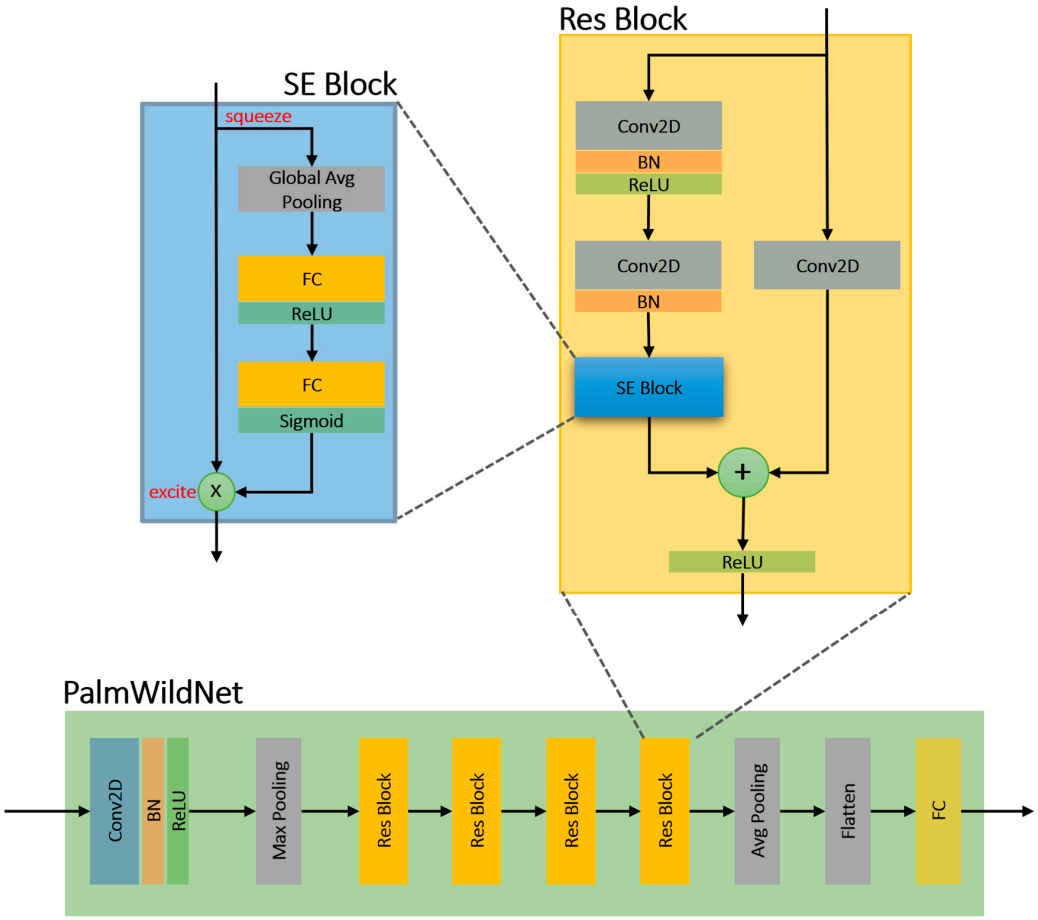



**Figure 8.** PalmWildNet Deep Network Architecture.

**Table 3.** PalmWildNet Layers.

| Layers | Block Type (×Number) | Output Size | Stride | SE Block | Parameters (M) |
|---|---|---|---|---|---|
| Stem | Conv7 × 7 → BN → ReLU → MaxPool | 64 × 56 × 56 | 2then 2 | – | 0.003 |
| Layer1 | ResBlock(64) ×2 | 64 × 56 × 56 | 1 | ✓ | 0.149 |
| Layer2 | ResBlock (128) ×2 | 128 × 28 × 28 | 2 → 1 | ✓ | 0.530 |
| Layer3 | ResBlock (256) ×2 | 256 × 14 × 14 | 2 → 1 | ✓ | 2.116 |
| Layer4 | ResBlock (512) ×2 | 512 × 7 × 7 | 2 → 1 | ✓ | 8.459 |
| Head | Avg.Pool. → Flatten → Dropout → FC | 512 × 1 × 1 | – | – | 0.092 |
| | | | | Total | 11.350 |

The PalmWildNet architecture comprises four main stages built upon SE-enhanced residual blocks. The network accepts an input of size 224 × 224 × 1 and progressively reduces the spatial resolution while increasing the number of channels (64 → 128 → 256 → 512) across these stages. The final feature map was vectorized via a Global Average Pooling (GAP) layer and passed through a fully connected layer with N = 180 output classes.

This design allows the model to dynamically adapt to each of the inputs. For instance, in cases where an ROI contains strong specular reflections that cause certain channels to overreact, the SE block suppresses their influence by reducing the corresponding scaling score. Conversely, channels that capture discriminative biometric information, such as principal palm lines or distinctive textural patterns, are assigned higher weights. By doing

so, the model learns to emphasize meaningful signal features while attenuating noise, thereby improving recognition robustness under unconstrained conditions.

#### 4.4.1. Squeeze-and-Excitation (SE) Block

Palmprint ROIs obtained in in-the-wild environments contain unwanted information, such as illumination artifacts (e.g., highlights and shadows) and sensor noise, in addition to the biometric signature. The SE block offers an elegant solution to this problem. For a feature map, the SE mechanism recalibrates the features by generating a channel-by-channel importance map, thus highlighting the information that is important for recognition. It consists of two main stages: Squeeze and Excitation. In the Squeeze stage, the $F_{sq}$ operator reduces the $H \times W$ value in each channel to a single scalar value and extracts a channel-by-channel global statistic ($z_c$). This information summarizes the overall activation state of the channel [102]:

$$z_c = F_{sq}(X_c) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{i,j,c} \quad (5)$$

In the excitation phase, the descriptor vector is passed $z_c$ through a gate mechanism consisting of two fully connected layers. This mechanism learns nonlinear inter-channel dependencies and produces an importance score $s_c$ in the range (0, 1) for each channel [102].

$$s_c = F_{ex}(z, W) = \sigma(W_2.\delta(W_1 z)) \quad (6)$$

The sigmoid function $\sigma$ here allows these scores to function as a gate. The original feature map $X_c$ is rescaled by multiplying these scores by channel $(s_c.X_c)$.

#### 4.4.2. Residual Block

The most important elements in determining palmprint identity are high-frequency texture details such as fine wrinkles and pores. In very deep networks, this sensitive information tends to be lost or "softened" as it propagates through layers. Residual connectivity is used to convey this low-level texture information to the deepest layers of the network by carrying the input x directly to the output. Furthermore, the complex network of lines in palmprints can create gradient surfaces that are difficult to learn. Residual connectivity prevents vanishing gradients, allowing even deeper and more complex models to be trained stably. The fundamental mathematical expression for residual blocks is that the output $y$ of a block is the sum of a nonlinear transformation applied to its input $x$ and the input itself: $y = F(x, \{W_i\}) + x$.

#### 4.4.3. Classification Head

The feature map produced by the final residual layer was transformed into a one-dimensional representation using a GAP layer. The GAP offers two main advantages: it substantially reduces the number of trainable parameters compared to fully connected flattening operations, and it enforces a stronger correspondence between the learned feature maps and class categories, thereby improving generalization. To further mitigate the risk of overfitting, a dropout layer with a rate of 20% was applied, randomly deactivating neurons during training to encourage redundancy reduction and feature robustness. The resulting feature vector was then passed through a fully connected layer consisting of N = 180 neurons, where each neuron corresponded to one subject in the MPW-180 dataset. Finally, this output is provided to the softmax activation function, which converts the raw scores into normalized class probabilities. This end-to-end configuration not only enables reliable classification of palmprint identities but also provides a strong baseline for evaluating the robustness of the proposed PalmWildNet architecture under variable illumination and device conditions.

4.4.4. Embedding Architecture for Metric Learning

The main hypothesis of our study is that a biometric representation robust to non-identifying variations, such as illumination, can be learned more effectively within a metric learning framework than through traditional classification approaches. To this end, the PalmWildNet architecture was transformed from a classifier to an embedding function that maps each palmprint image to a unique vector in a D-dimensional Euclidean space that densely encodes the identity information. This transformation is achieved by a targeted modification of the output layer only, while preserving the PalmWildNet backbone and, consequently, its learned rich feature extraction capabilities. Specifically, the classification header is defined in Section 4.4.3. was removed from the architecture, which assigns probabilities to 180 classes. Instead, a single embedding layer is added that linearly maps the 512-dimensional feature vector from the backbone to a D-dimensional embedding vector (e) (set as D = 128 in this study).

The output of this layer, the embedding vector e, is subjected to an L2 normalization process to ensure that all representations in the feature space lie on a unit hypersphere. This process sets the Euclidean norm of each embedding vector to 1, as shown in Equation (7):

$$\hat{e} = \frac{e}{\|e\|_2} \tag{7}$$

This process projects all embedding vectors onto a unit hypersphere. This normalization for ROIs eliminates variations in the vector magnitude that might be caused by the image quality (e.g., slight blur) or contrast. This allows the model to focus on encoding all identity information in the direction of the vector rather than its magnitude. This makes the representation more robust to small variations in the ROI quality.

The real power of this embedding architecture emerges when it is combined with the Triplet Loss [103] function. The goal of the training process is to structure the feature space according to two fundamental principles: (1) minimizing the intra-class variance and (2) maximizing the inter-class distance. Triplet Loss generates a robust learning signal based on relative distances to achieve this goal.

Training is performed using a triplet of data samples at each step: an Anchor $(x_a)$, a Positive $(x_p)$, and a Negative $(x_n)$ sample. The most critical component of our methodology is the selection strategy for the positive samples. For each anchor sample, a positive sample is deliberately selected from the same identity but under the opposite illumination condition (e.g., the anchor is with flash, and the positive sample is without flash). This cross-illumination positive-sampling strategy allows the model to actively minimize illumination variance. A negative sample is selected from any image with an identity different from that of the anchor.

The model generates embedding vectors $(e_a, e_p, e_n)$ from these three inputs. The objective of the Triplet Loss function is to ensure that the squared Euclidean distance between the anchor and positive is smaller than the squared Euclidean distance between the anchor and negative by at least one margin ($\alpha$). This is mathematically expressed as follows:

$$d(e_a, e_p)^2 + \alpha < d(e_a, e_n)^2 \tag{8}$$

The loss function that satisfies this condition is expressed as follows:

$$L_{triplet}(e_a, e_p, e_n) = \max\left(0, \underbrace{\|e_a - e_p\|_2^2}_{\text{intra-class var.}} - \underbrace{\|e_a - e_n\|_2^2}_{\text{inter-class dist.}} + \alpha\right) \tag{9}$$

This training strategy directly addresses the fundamental problem of palmprint ROIs in the wild environment. If the model is trained using a standard approach, it tends to learn two separate subclusters in the feature space for with and without flash images of the same person. Our choice of "cross-illumination positive example" gives the model a clear instruction: ignore visual differences due to illumination conditions and close the gap between these two subclusters, merging them into a single, dense, identity-specific cluster.

Mathematically, minimizing the range $\|e_a - e_p\|_2^2$ in the loss function directly translates to minimizing the within-class variance resulting from illumination conditions. Simultaneously, maximizing the term $\|e_a - e_n\|_2^2$ allows for a clear separation of different identities, independent of illumination. In this way, PalmWildNet becomes not only a feature extractor but also a generator that produces an illumination-disentangled biometric signature from ROIs collected under challenging and variable conditions. The training process was optimized to minimize the Equal Error Rate (EER) metric, which measures the practical success of this approach.

#### 4.4.5. Training Procedure and Hyperparameters

The experiments were conducted on an NVIDIA RTX 4090 GPU using PyTorch 2.0.0+cu118 with CUDA 11.8 and cuDNN 8.7, implemented in Python 3.9.16 [104] on a 64-bit Windows environment. The optimization target was the Cross-Entropy Loss function, which is a standard approach for multi-class classification problems. The model parameters were updated using the Adam optimization algorithm [105], which offers adaptive learning rates and is widely used in contemporary deep learning applications. The optimization process was initiated using a relatively low initial learning rate of $1 \times 10^{-4}$.

Training was performed using the TripletMarginLoss function. As formulated in Equation (9), this function aims to ensure that the distance between cross-illuminated positive examples of the same identity and negative examples of different identities exceeds the predefined margin ($\alpha$). In our study, this margin value was set to $\alpha = 0.5$. The Adam algorithm was selected for optimization. The most fundamental difference from the classification model is the manner in which the generalization performance is monitored during the training process.

To ensure more stable convergence in the later stages of training and to refine the optimization process in both the classification and metric learning models, the ReduceLROnPlateau learning rate scheduler was integrated into the training process. This strategy monitors the validation loss and reduces the current learning rate by a factor of 10 if no improvement is detected in this value for five epochs. This dynamic adjustment allows the model to take more precise steps at narrow local minima on the loss surface.

Two primary regularization techniques were used together to prevent overfitting and increase the generalization capacity of the model. First, a dropout layer with a $p = 0.2$ ratio was applied to the classification head to prevent inter-neuron co-adaptation. Second, and more importantly, an Early Stopping mechanism was implemented to prevent overtraining of the model and to find the model with the best generalization point. This mechanism monitored the validation loss and automatically terminated the training process when no improvement was observed over 10 consecutive epochs, thereby saving the best-performing model weights. All training processes were run with a batch size of 64 and for a maximum of epochs until the Early Stopping criterion was met.

## 5. Experiments and Results

This section presents the results of the extensive experimental studies conducted to evaluate the effectiveness and robustness of the methodology presented in Section 4. The experiments were designed to analyze the performance of the designed PalmWildNet

architecture under both ideal and challenging conditions, to reveal the severity of the cross-illumination problem, and to prove the solution that our metric learning-based approach provides to this problem.

### 5.1. Dataset Splitting Strategy

To consistently and reproducibly evaluate the performance of our model, an image-independent approach was adopted; that is, an image could not appear in more than one set (training, validation, or test) simultaneously. Each user had a different number of ROIs. However, to create a balanced dataset, we attempted to partition the data equally across all the sets. The selection of ROIs for a user is detailed in Section 4.2. In this process, a total of 80 ROI images were used equally for each user. The images were randomly shuffled and divided into 50 training, 15 validation, and 15 test images. The determined ROI images were consistently used across all experiments.

### 5.2. Evaluation Metrics

The final performance of the models in our study was evaluated using two main protocols that are widely accepted in the biometric system literature. For the closed-set identification (1:N) scenario, we report the Rank-1 and Rank-5 accuracies, which represent the percentage of times the model correctly matches the true identity with its most likely prediction for an image in the test set. For the more practical 1:1 verification scenario, the system performance was measured using the Equal Error Rate (EER) and True Positive Rate (TPR) at a given False Positive Rate (FPR).

In addition to the final performance metrics, a set of diagnostic metrics was used to monitor the training process. These metrics are Training and Validation Loss and Training and Validation Accuracy, which are calculated at the end of each epoch. These diagnostic metrics were used to detect potential problems, such as overfitting, and to guide regularization mechanisms, such as Early Stopping, and were utilized in all experiments.

### 5.3. Experiment 1: Performance Analysis of PalmWildNet Architecture on Various Datasets

The first experiment was designed to evaluate the effectiveness of the proposed PalmWildNet architecture in addressing the general palmprint recognition problem and contextualize its contribution within the broader literature. To this end, the model's generalization capability and recognition performance were systematically assessed across five publicly available benchmark datasets with diverse acquisition characteristics: Tongji Contactless Palmprints [73], COEP [77], Birjand MPD [24], BJTU_PalmV2 [72], and IITD Touchless [79].

Across all benchmark datasets, PalmWildNet was trained from scratch on the designated training split and evaluated on a held-out test split. To accommodate heterogeneity in subject counts, image availability, and acquisition conditions, we adopted dataset-specific partitioning atop a general 70/15/15 (training/validation/test) random split. For datasets with very few images per subject, a subject-wise leave-two-out protocol was used; for example, in the IITD Touchless dataset (~6 images per subject), one image per subject was assigned to validation, one to testing, and the remainder to training. In the COEP dataset, one image per participant was reserved for validation, two for testing, and the remaining five or six images were used for training. To enhance generalization, especially under data scarcity, we applied a consistent on-the-fly augmentation pipeline across all datasets and increased the number of training epochs for the smallest datasets.

For datasets comprising multiple acquisition sessions (e.g., Tongji, BJTU, and BMPD), all sessions were merged into a single pool to avoid session-specific biases and ensure consistent evaluation. Furthermore, to eliminate variability due to handedness, only right-hand images were used in all datasets. This protocol provides a rigorous and fair basis for

comparing PalmWildNet with existing palmprint datasets, and the performance results are reported in Table 4.

**Table 4.** Performance of PalmWildNet Model on Different Datasets.

| Dataset | Participants | Rank-1 Acc. (%) | Rank-5 Acc. (%) | EER (%) | SOTA (*) |
|---|---|---|---|---|---|
| Tongji [73] | 300 | 100.00 | 100.00 | 3.3 | 99.72 (RR **) [106] |
| COEP [77] | 167 | 98.77 | 100.00 | 0.4 | 0.612 (EER) [107] |
| Birjand MPD [24] | 41 | 99.04 | 100.00 | 5.1 | - |
| BJTU_PalmV2 [72] | 148 | 96.93 | 98.63 | 3.7 | 5.26 (EER-BLDC) [12]<br>91.88 (RR-LLDP) [108] |
| IITD Touchless [79] | 230 | 96.17 | 99.57 | 1.1 | 99.71 (RR) [84] |

* SOTA: State-of-The-Art, ** RR: Recognition Rate.

The results clearly demonstrate that the PalmWildNet architecture achieves highly competitive state-of-the-art (SOTA) performance on widely adopted benchmark datasets. Notably, a perfect recognition accuracy of 100.00% was obtained on the Tongji Contactless dataset, whereas extremely low Equal Error Rates (EERs) of 0.40% and 1.10% were achieved on the COEP and IITD Touchless datasets, respectively. These outcomes underscore the strong discriminative capacity of the SE-residual block design of PalmWildNet.

Despite the competitive performance achieved by PalmWildNet, the model did not surpass the best published results on two specific datasets: IITD (96.17% vs. 99.71%) and BJTU_PalmV2 (96.93% vs. 98.63%). This discrepancy warrants a deeper discussion. The lower result on the IITD Touchless dataset is highly susceptible to estimation variance because of its peculiar structure, which provides only a single test image per subject. In contrast, the top-performing methods on this dataset are often highly specialized, manually tuned, or simple feature-based approaches (e.g., LBP variants) that are over-optimized for a controlled environment and specific acquisition protocol, achieving minimal variation. Similarly, the slightly lower Rank-1 accuracy on BJTU_PalmV2 compared to its SOTA (96.93% vs. 98.63%) is likely due to the specific, fine-grained feature emphasis of the original SOTA method (e.g., Bifurcation Line Direction Coding or Local Line Directional Pattern). These feature-engineered methods were precisely tuned to the BJTU acquisition protocol.

However, a generalized deep learning architecture, such as PalmWildNet, designed for high illumination and device variability, naturally exhibits a slightly broader generalization margin when tested using such restrictive single-shot protocols. Nevertheless, the ~ 97% accuracy achieved is still competitive and confirms the model's fundamental ability to extract distinguishing palm features, even when trained solely on standard laboratory data. To further illustrate the recognition performance of PalmWildNet, ROC curves for the evaluated datasets are presented in Figure 9.

The key conclusion from this experiment is that the proposed PalmWildNet architecture is not a niche solution tailored to a single dataset but rather a general-purpose, robust, and high-performance framework for palmprint recognition. By establishing the fundamental strength of the architecture across diverse benchmarks, we demonstrated its broad applicability and strong generalization capability. Building on this foundation, the next stage of our study shifts focus to one of the most challenging open problems in the literature: palmprint recognition under in-the-wild conditions. Accordingly, the second experiment evaluated the performance of PalmWildNet when confronted with heterogeneous real-world acquisition scenarios represented in our newly introduced MPW-180 dataset.

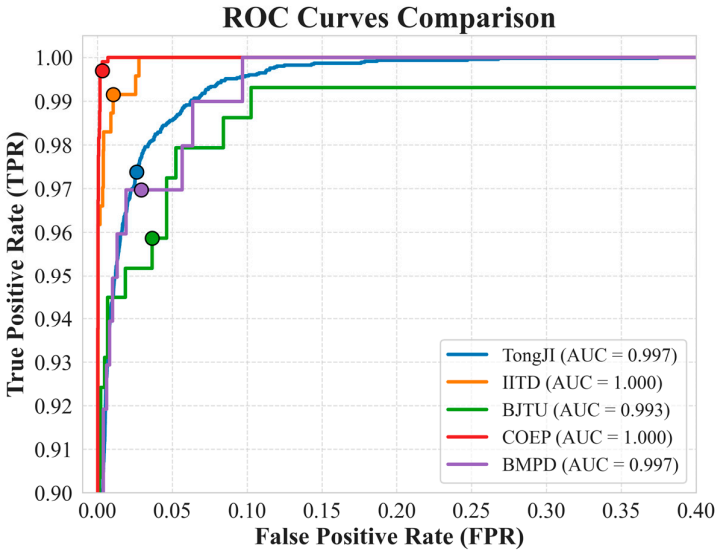



**Figure 9.** ROC curves generated by PalmWildNet on five publicly available palmprint datasets (Tongji, COEP, Birjand MPD, BJTU_PalmV2, and IITD).

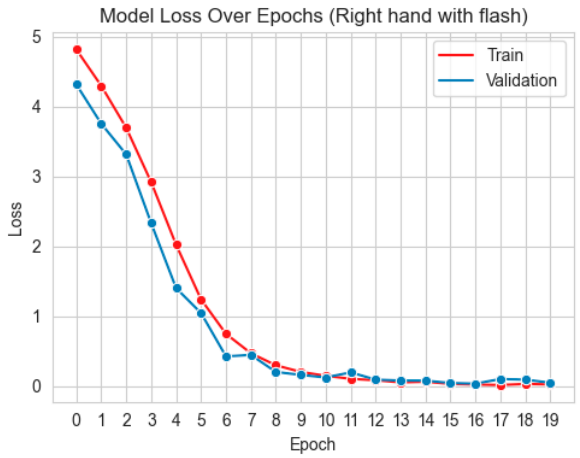
### *5.4. Experiment 2: Performance Analysis Under Matched-Illumination Conditions*

The primary objective of the first experiment was to evaluate the fundamental learning capacity and discriminative power of the PalmWildNet architecture under idealized matched-illumination conditions. This setting simulates scenarios in which the model encounters images captured under identical illumination conditions during both training and testing, thereby establishing a baseline for subsequent and more challenging cross-illumination experiments.

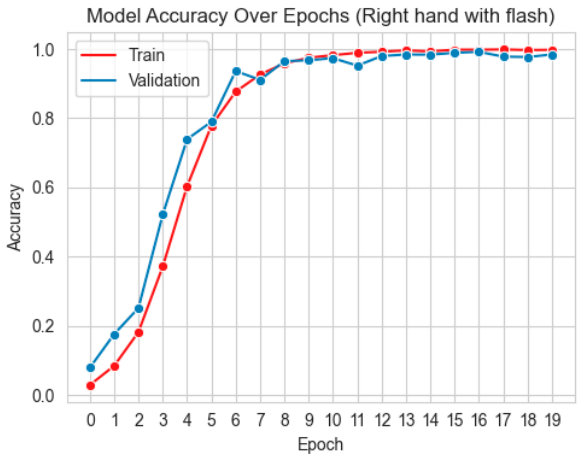
To this end, four independent training and testing sessions were conducted using the four principal subsets of the MPW-180 dataset (HR_FT, HR_FF, HL_FT, and HL_FF, as defined in Section 3). In each case, PalmWildNet was trained exclusively on data from a single illumination condition and evaluated using previously unseen test samples from the same condition. Training employed Cross-Entropy Loss with the Adam optimizer, along with the hyperparameters specified in Section 4.

The loss and accuracy curves obtained during training are shown in Figure 10. In each row, the first plot depicts the epoch-wise loss values, whereas the second plot shows the corresponding accuracy values. The training and validation loss curves demonstrated consistent learning dynamics without evidence of overfitting. All models were trained for 20 epochs, which was sufficient for the convergence. Notably, from approximately the 13th epoch onwards, the performance metrics stabilized, indicating that the model had reached its optimal operating point.

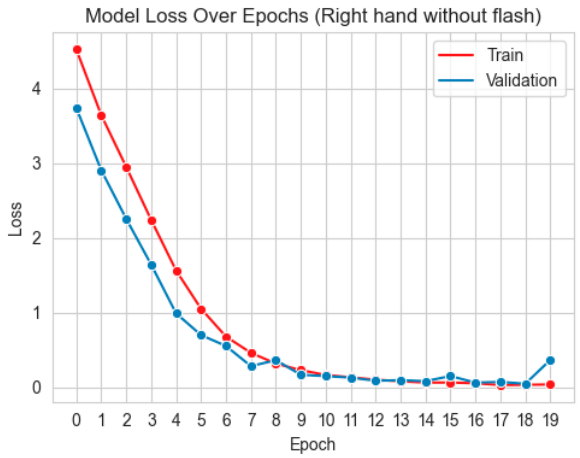
Upon completion of the training, the results of the comprehensive evaluations conducted on the test sets using the best-performing model weights are reported in Table 5. For completeness, the table also includes outcomes under cross-illumination conditions; however, these results are reserved for a detailed discussion in the following section.



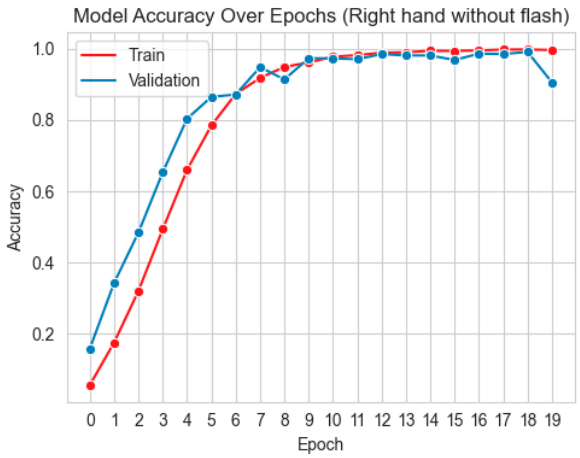

(**a**)



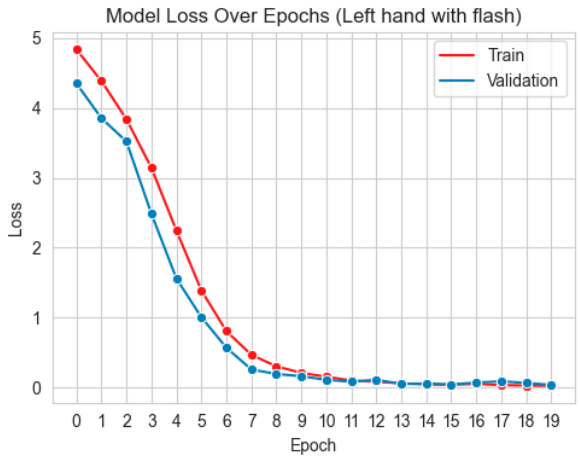

(**b**)



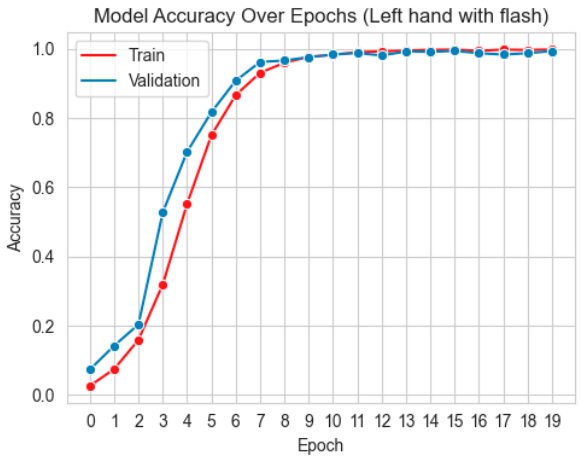

(**c**)




(**d**)

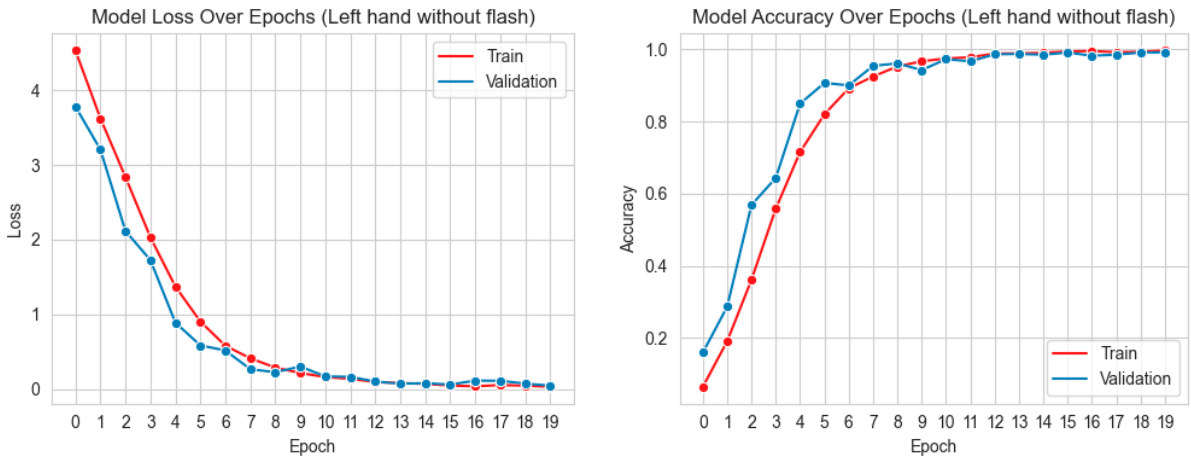



(**e**)




(**f**)

**Figure 10.** *Cont.*




(**g**) (**h**)

**Figure 10.** Training and validation performance of PalmWildNet under matched-illumination conditions using four MPW-180 sub-datasets. The plots depict (**a**) HR_FT loss, (**b**) HR_FT accuracy, (**c**) HR_FF loss, (**d**) HR_FF accuracy, (**e**) HL_FT loss, (**f**) HL_FT accuracy, (**g**) HL_FF loss, and (**h**) HL_FF accuracy, respectively.

**Table 5.** Comprehensive Recognition Performance of PalmWildNet Under Matched-Illumination Conditions.

| Sub Dataset | Rank-1 Acc (%) | Rank-5 Acc (%) |
|---|---|---|
| HR_FT | 99.39 | 99.89 |
| HR_FF | 99.28 | 99.67 |
| HL_FT | 99.06 | 99.61 |
| HL_FF | 98.83 | 99.89 |
| Mean | 99.14 | 99.77 |

The results presented in Table 5 clearly demonstrate that the PalmWildNet architecture achieves exceptionally high recognition performance under conditions in which the illumination of the training and test data is consistent. Rank-1 accuracies exceeding 99.5% and Rank-5 accuracies approaching 100% across all experimental scenarios confirm that the architecture effectively captures palmprint patterns, primarily owing to the synergy of residual connectivity and SE-based channel attention. Notably, the Rank-5 accuracy surpassing 99.9% implies that even in rare instances where the top prediction is incorrect, the correct identity remains almost always within the top five candidates. This highlights a very high level of system reliability, which is particularly important in mission-critical biometric applications.

Beyond recognition accuracy, the system also demonstrated operational efficiency, with an average inference time of approximately 5 ms per image on a modern GPU. This finding indicates that PalmWildNet is computationally lightweight enough to be deployed in practical and real-time applications without sacrificing recognition quality. Additionally, the marginally higher accuracy observed in datasets acquired under flash illumination further supports the hypothesis that standardized lighting conditions and increased visibility of fine palmprint textures facilitate more effective model training.

However, the most critical insight from this experiment is that outstanding performance under idealized matched-illumination conditions does not guarantee comparable success in real-world biometric scenarios. In practice, it is common for users to enroll under one illumination setting (e.g., with flash) and later attempt to authenticate under a different setting (e.g., ambient light only). This raises the central challenge of cross-illumination robustness, which can significantly impact the reliability of any mobile biometric system.

To address this issue, the following experiment investigated the generalization capacity of PalmWildNet under more demanding and realistic cross-illumination conditions.

### 5.5. Experiment 3: Performance of Cross-Illumination Scenarios

The primary objective of this experiment was to rigorously evaluate the robustness and generalization capacity of PalmWildNet under mismatched illumination conditions, a challenge that is unavoidable in real-world biometric applications. This scenario reflects a practical case where a user may register (enroll) with the system under one illumination setting (e.g., with flash) but subsequently attempt authentication under a different setting (e.g., without flash). To investigate this, the models trained independently on each of the four subsets in Experiment 2 were tested using the cross-illumination protocol. Specifically, a model trained exclusively on the HR_FT subset was evaluated using the HR_FF test set, and the same cross-condition evaluations were performed for all subset combinations.

The results obtained under these mismatched illumination conditions contrast sharply with the nearly perfect recognition rates achieved in Experiment 2. The observed performance degradation highlights the significant impact of illumination variability on contactless palmprint recognition and emphasizes the importance of developing illumination-invariant learning strategies. The detailed results of this evaluation are presented in Table 6.

**Table 6.** Cross-illumination recognition performance of PalmWildNet on MPW-180 dataset.

| Training Dataset | Test Dataset | Rank-1 Acc (%) | Rank-5 Acc (%) |
|---|---|---|---|
| HR_FT | HR_FF | 43.56 | 62.56 |
| HR_FF | HR_FT | 35.67 | 53.33 |
| HL_FT | HL_FF | 50.22 | 67.22 |
| HL_FF | HL_FT | 34.28 | 51.00 |
| Mean | | 40.93 | 58.53 |

The results in Table 6 provide compelling evidence of the inadequacy of standard training approaches for deployment in real-world biometric systems. Although Experiment 2 achieved near-perfect performance with an average Rank-1 accuracy exceeding 99%, this value dropped precipitously to 40.93% under cross-illumination conditions. Even the Rank-5 accuracy, which typically indicates the reliability of a model, remained at only 58.53%. This means that in nearly half of the authentication attempts, the correct identity did not appear among the top five predictions, which is an unacceptable outcome for practical security applications.

The underlying reason for this dramatic decline is the domain-shift problem. When trained exclusively on flash-illuminated images, the model primarily captures high-frequency textural details, strong contrast, and specular reflections. These features become unreliable in without-flash environments, where images are softer, noisier, and lack discriminative cues. Conversely, models trained on without-flash data learn to accommodate variable shadows and low-contrast patterns but fail when exposed to the crisp, high-contrast structures present in flash-acquired images.

An additional observation is the asymmetry of the performance degradation. Models trained on with-flash data and tested on with-flash data achieved slightly better performance (~47% average Rank-1 accuracy) than the reverse scenario (~35%). This indicates that the richer, higher-quality information contained in flash images fosters more generalizable representations than those derived from noisier ambient-light data.

Overall, this experiment highlights a fundamental vulnerability in current palmprint recognition systems: their heavy dependence on the consistency of illumination. These findings underscore the urgent need for illumination-invariant learning strategies, such as

metric learning, domain adaptation, and data fusion techniques, to ensure robust performance in unconstrained mobile environments.

### 5.6. Experiment 4: Evaluation of the Proposed Method for Illumination Durability

In the previous experiment (Section 5.4), it was quantitatively demonstrated that the baseline PalmWildNet model failed to generalize under cross-illumination conditions, leading to a dramatic decline in performance. As discussed in Section 4, the incorporation of Residual and SE blocks partially mitigates illumination sensitivity, and the on-the-fly augmentation pipeline includes random brightness adjustments. These mechanisms explain why the recognition rates in Table 6 are approximately 50%. However, these results remain far below the accuracy thresholds required for reliable biometric authentication.

To overcome this limitation, we propose a dual-level strategy that addresses both the data and methodological dimensions of the problem.

1. Data-level strategy: The HR_FT and HR_FF datasets were merged into a unified training pool to expose the model to both illumination domains and provide sufficient visual diversity for robust feature learning. This pooling was applied across the training, validation, and testing partitions for both the right and left hands, thereby ensuring that each identity was represented under both flash and no-flash conditions during learning. This strategy is designed to force the model to learn illumination-invariant representations by explicitly integrating the variability into the training data.

2. Methodology-Level Strategy: Conventional Cross-Entropy Loss does not enforce geometric consistency within the learned feature space. As a result, the model may generate metrically distinct sub-clusters for with and without-flash samples of the same identity while still maintaining high classification accuracy within the decision boundary. Such a configuration yields tolerant but not truly robust representation. To address this limitation, we replaced the final classification layer with an embedding layer and trained the network using a Triplet Loss function. Strategic cross-illumination sampling (anchor-positive pairs from different illumination domains and negatives from other identities) was employed to explicitly penalize illumination-induced variance in the embedding space. This modification enforces discriminative consistency across domains, thereby embedding illumination invariance directly into learned feature representations.

The final performance of this enhanced framework is reported in Table 7, evaluated using both identification metrics (Rank-1 and Rank-5) and verification metrics (EER) to provide a comprehensive assessment of the illumination robustness of the system.

**Table 7.** Recognition performance of PalmWildNet trained under combined illumination conditions and optimized using Triplet Loss.

| Training Dataset | Test Dataset | Rank-1 Acc (%) | Rank-5 Acc (%) | EER (%) |
|---|---|---|---|---|
| HR_Fusion (HR_FT and HR_FF) | HR_FT | 96.67 | 98.56 | 1.36 |
| | HR_FF | 97.00 | 98.33 | 2.00 |
| | HR_Fusion | 97.00 | 98.44 | 1.72 |
| HL_Fusion (HL_FT and HL_FF) | HL_FT | 97.22 | 98.56 | 1.73 |
| | HL_FF | 95.78 | 97.78 | 2.38 |
| | HL_Fusion | 97.00 | 98.17 | 2.14 |

The consistently high Rank-1 accuracy values—exceeding 95% in all scenarios—confirm that the learned embedding space achieves a clear separation between identity clusters. This is further evidenced by the effectiveness of the 1-NN classifier, which benefits from the large inter-class distances in the feature space. From the perspective of authentication performance, the Equal Error Rate (EER) values reported in Table 7 range between 1.36% and 2.38%, which is remarkable given the unconstrained and heterogeneous nature of

the MPW-180 dataset. Such low EER values provide compelling quantitative evidence that Triplet Loss not only maximizes inter-class separation but also minimizes intra-class variance, effectively consolidating with and without flash samples of the same identity into a single dense cluster.

To further substantiate these findings, the Receiver Operating Characteristic (ROC) and Detection Error Trade-off (DET) curves for both hands are presented in Figure 11. The ROC curves (Figure 11a,b) consistently approach the upper-left corner of the plot, indicating that the system achieves high true positive rates (TPR) even at very low false positive rates (FPR)—a hallmark of safe and reliable biometric performance. Similarly, the DET curves (Figure 11c,d), plotted on a logarithmic scale, provide a more granular view of the error behavior. The proximity of these curves to the origin confirms that the system maintains low error rates across all operating points, further validating the robustness of the proposed illumination-invariant embedding space.

The consistency of the performance across modalities is particularly noteworthy. The close similarity between the results obtained from the right-hand (HR) and left-hand (HL) models demonstrates that the proposed methodology is not biased toward a specific hand modality but rather constitutes a generalizable and modality-independent approach. Moreover, the absence of significant fluctuations in performance across the with-flash (_FT), without-flash (_FF), and fused (_Fusion) test sets further confirms that the model is effectively invariant to illumination differences, exhibiting a stable recognition performance across all experimental conditions.

Holistically, these findings indicate that the severe "performance crash" observed in Experiment 3, where recognition accuracy declined by more than 50%, has been completely resolved. This outcome validates the central hypothesis of our study: by leveraging strategically sampled cross-illumination positives within the Triplet Loss framework, the model is explicitly forced to treat illumination variation as irrelevant noise within the feature space. Consequently, PalmWildNet no longer considers whether an image is acquired with or without flash but instead concentrates exclusively on the intrinsic, identity-defining textural patterns of the palm. This provides compelling evidence that the proposed framework achieves not only illumination tolerance but also true illumination robustness, thereby representing a significant step toward the practical, real-world deployment of mobile palmprint recognition systems.

To obtain a deeper understanding of the quantitative performance of the proposed model and further analyze the structure of its learned feature space, we conducted additional visual examinations of the test sets. These analyses include the match score distributions presented in Figure 12 and Cumulative Matching Characteristic (CMC) curves presented in Figure 13. The verification capability of the system is primarily determined by the distinction between "Genuine" pairs (samples belonging to the same identity) and "Impostor" pairs (samples belonging to different identities).

Figure 12 illustrates the probability density distributions of the L2 distance scores across three test scenarios (with flash, without flash, and combined) for both the right- and left-hand datasets. As can be clearly observed, the distributions of genuine pairs (green) are highly concentrated within a very narrow region close to zero, whereas the distributions of impostor pairs (red) are spread across higher distance values. The minimal overlap between these two distributions provides strong visual evidence of the model's ability to achieve a clear identity separation. This visual separation directly explains the extremely low Equal Error Rate (EER) values reported in Table 7. The vertical dashed line in each graph marks the optimal decision threshold corresponding to the EER, further highlighting the effectiveness of the model in balancing the false acceptance and false rejection rates.

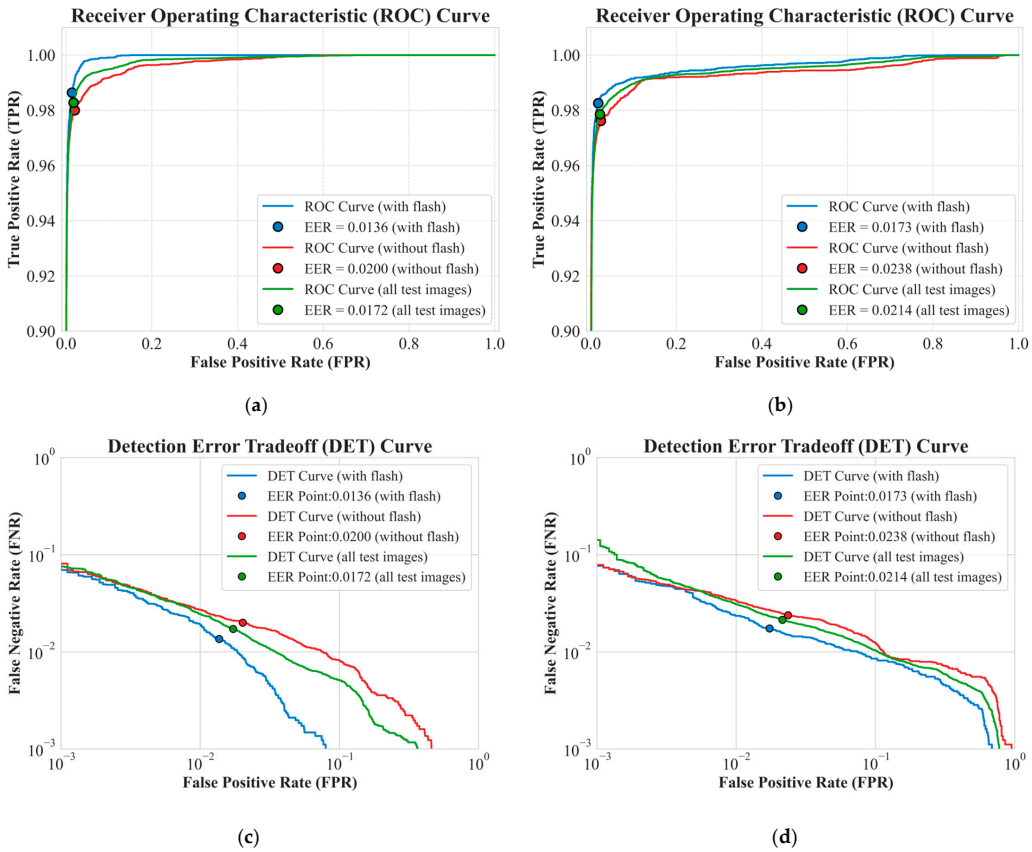



**Figure 11.** Receiver Operating Characteristic (ROC) and Detection Error Trade-off (DET) curves of PalmWildNet under different illumination conditions (with flash, without flash, and fused). (**a**) ROC curve for right-hand images; (**b**) ROC curve for left-hand images; (**c**) DET curve for right-hand images; (**d**) DET curve for left-hand images.

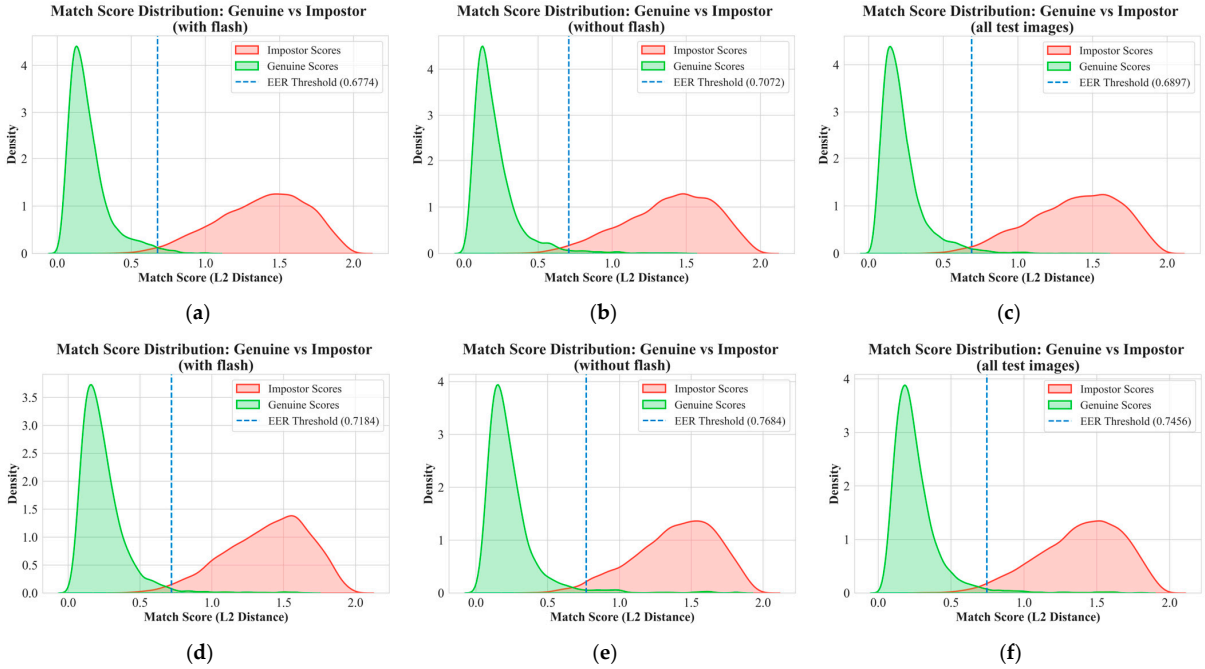



**Figure 12.** L2 distance score distributions of Genuine and Impostor pairs across different test scenarios. (**a**) Right hand with flash, (**b**) right hand without flash, (**c**) right hand fused, (**d**) left hand with flash, (**e**) left hand without flash, and (**f**) left hand fused. The dashed vertical lines denote the Equal Error Rate (EER) thresholds.

The closed-set identification capability of the proposed model was further assessed using Cumulative Matching Characteristic (CMC) curves, as illustrated in Figure 13. These curves quantify the probability of retrieving the correct identity within the top-k ranked candidates, thereby providing a practical measure of recognition reliability for the model. For both right- and left-hand experiments, the curves exhibited an exceptionally high starting point, with Rank-1 accuracy reaching approximately 97%. The performance rapidly saturates beyond this point, surpassing 99% for Rank-2 and Rank-3, and remains consistently high thereafter.

This behavior demonstrates that in rare cases where the model does not assign the correct identity at the top rank, the true identity is almost always contained within the top few candidates. Such a steeply rising CMC profile is a hallmark of robust recognition systems, confirming that the proposed framework achieves high accuracy and practical reliability in candidate retrieval, which is an essential requirement for large-scale biometric deployments.

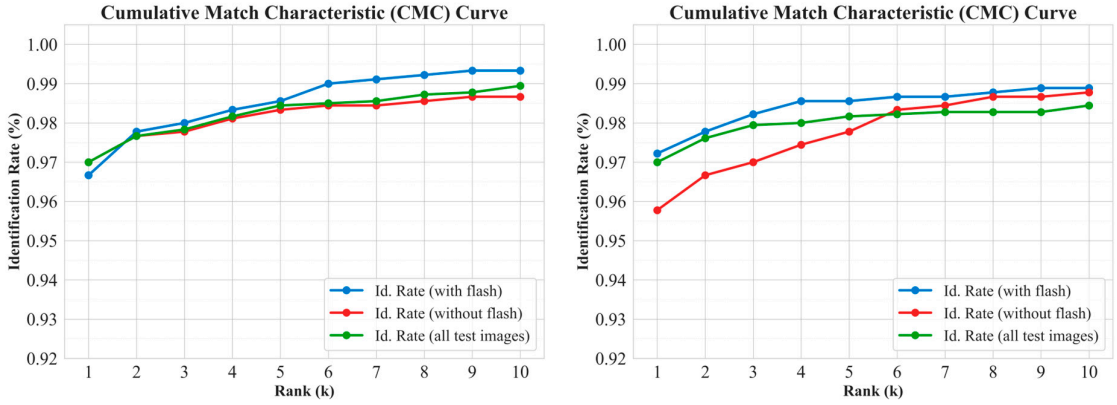



**Figure 13.** CMC curves of the proposed model on combined test sets for (**a**) right-hand and (**b**) left-hand images.

### *5.7. Computational Efficiency Analysis*

To evaluate the computational efficiency of the proposed PalmWildNet, its inference performance was measured in both the GPU and CPU environments. All experiments were conducted on an NVIDIA RTX 4090 GPU (24 GB VRAM) (NVIDIA, Santa Clara, CA, USA) and an Intel Core i9-13900K CPU (Intel, Santa Clara, CA, USA) with 64 GB RAM. The average inference time was calculated to be 5.1 ms per ROI on the GPU and 87.3 ms per ROI on the CPU (single-threaded). These results confirm that PalmWildNet can achieve real-time performance on modern desktop computers. Furthermore, because the proposed method is designed with lightweight convolutional and SE modules, it can be further optimized for mobile and embedded environments through quantization, pruning, or mixed-precision inference. This indicates that the model architecture is highly suitable for real-time mobile biometric applications, which is the primary focus of this study.

## 6. Discussion and Conclusions

This study addresses two of the most pressing challenges in mobile palmprint recognition: uncontrolled illumination and diverse devices. Our experiments reveal that conventional CNN-based approaches collapse under illumination mismatches, with accuracy dropping by more than 50% when training and test conditions differ. This provides a clear quantitative demonstration of the susceptibility of traditional models to domain shifts. In contrast, the proposed cross-illumination sampling strategy combined with a Triplet Loss-based metric learning framework effectively mitigates this vulnerability, yielding Equal Error Rates (EER) in the 1–2% range and maintaining recognition accuracies above

97%. These results confirm that the proposed approach does not merely tolerate illumination variability but learns genuinely invariant representations, successfully disentangling identity-specific features from the illumination artifacts.

This study makes two contributions. From a data perspective, the MPW-180 dataset, collected in a fully unconstrained, BYOD setting, represents the most diverse mobile palmprint benchmark to date, offering cross-sensor variability that closely mirrors real-world conditions. It provides a challenging yet necessary platform for evaluating the algorithmic generalization in unconstrained scenarios. From a methodological perspective, the PalmWildNet architecture, enhanced with SE blocks and coupled with Triplet Loss optimized via cross-illumination sampling, delivers a robust framework for in-the-wild biometric recognition. Additionally, the observed inference time of ~5 ms per image demonstrates that the system achieves high accuracy and the computational efficiency required for practical deployment, including real-time applications.

Nevertheless, this study has several limitations. Although MPW-180 incorporates wide device and illumination diversity, other challenges, such as correcting for perspective distortions, integrating multimodal palmprint cues (e.g., line- and texture-based features), and developing advanced strategies for optimal frame selection from video sequences, were not directly addressed in this study. Moreover, the video-based nature of the dataset opens new opportunities for future work, particularly in integrating liveness detection mechanisms to enhance security and reliability.

Although MediaPipe provides high keypoint accuracy in frames in which hand visibility is sufficient, we observed two primary failure scenarios. The first is the left-/right-hand ambiguity in images where the thumb is not clearly visible; however, we discarded these frames entirely because our acquisition protocol allowed us to know which hand the video belonged to. The second and more significant limitation arises in frames where the subjects move their hands too close to the camera, causing the fingers to be not fully visible in the frame. Despite these frames often containing rich, high-resolution palmprint ROI information, MediaPipe fails to detect keypoints, leading to the automatic exclusion of these potentially valuable samples during frame selection. Given the high volume of collected video frames (over 768,000 in total), we considered the data loss acceptable within the scope of the current study. Nevertheless, we acknowledge that this is a significant shortcoming in maximizing ROI extraction. In our future work, we plan to address this specific challenge by developing a novel approach for accurate palm localization in zoomed-in or partially occluded hand images.

Although palmprint characteristics remain constant throughout life, demographic factors such as age, sex, and ethnicity can indirectly influence recognition performance by influencing skin texture and image quality. Despite the diversity of our dataset, these effects were not analyzed in this study. To the best of our knowledge, systematic studies on the effect of age on palmprint biometrics are very limited in the literature.

Although the current dataset design permits minimal temporal adjacency between consecutive windows to preserve natural motion continuity, future studies using MPW-180 will incorporate stricter non-overlapping rules. This adjustment ensures stricter inter-window independence and supports future analyses of the impact of temporal sampling strategies on biometric video datasets.

In conclusion, this study demonstrates that palmprint recognition can be successfully transitioned from controlled laboratory conditions to dynamic and unpredictable real-world environments. These findings underscore the necessity of moving beyond conventional classification paradigms toward approaches that explicitly model and compensate for unwanted variability. The MPW-180 dataset and PalmWildNet framework together establish both a compelling benchmark and a practical methodological reference

for the community, representing a significant contribution to the development of reliable, real-world-ready mobile biometric systems.

**Author Contributions:** Conceptualization, Ö.B. and M.E.; methodology, Ö.B. and E.G.; software, Ö.B. and B.K.S.; validation, E.G. and R.Ö.D.; formal analysis, N.G.; investigation, S.T.; resources, S.T. and R.Ö.D.; data curation, B.K.S. and S.T.; writing—original draft preparation, Ö.B. and B.K.S.; writing—review and editing, Ö.B. and B.K.S.; visualization, E.G. and R.Ö.D.; supervision, Ö.B. and M.E. and E.G.; project administration, Ö.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the TÜBİTAK-1001 Program (Project No: 122E402). However, all responsibility for this paper belongs to the author. The financial support received from TÜBİTAK does not imply that the content of the paper is scientifically approved by TÜBİTAK. We would like to thank TÜBİTAK for their support of this project.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Scientific Research and Publication Ethics Committee of Gümüşhane University (protocol code: 2022/5, date of approval: 28 July 2022, official document number: E-95674917-108.99-123997). Informed consent was obtained from all the participants. The participants were explicitly informed that their palm videos would be used solely for academic research purposes. To ensure privacy and ethical compliance, the data were anonymized at the point of collection, and personally identifiable information (PII), such as names, device IDs, and timestamps, was permanently removed. The data are stored on the institution's secure servers and are accessible only to authorized researchers.

**Informed Consent Statement:** Written informed consent was obtained from all participants included in the study.

**Data Availability Statement:** The source code for the PalmWildNet model and the experimental pipeline are openly available in the GitHub repository at: https://github.com/bingolo/PalmWildNet (accessed on 19 October 2025), v0.5.0-alpha. The code is licensed under Apache License 2.0. This repository also includes sample images and direct access links to the datasets. The MPW-180 dataset used in this study can be accessed through the TÜBİTAK Open Archive (Aperta). The access link to the Aperta system is as follows: https://aperta.ulakbim.gov.tr (accessed on 19 October 2025). The dataset is licensed under CC BY-NC 4.0 (Creative Commons Attribution-NonCommercial 4.0 International), restricting its use to academic and non-commercial purposes only.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study, collection, analyses, or interpretation of data, writing of the manuscript, or decision to publish the results.

## References

1. Guerrache, F.; Haddadou, H. A Quality-Based Context-Switching Framework for Palmprint Recognition. *Multimed. Tools Appl.* **2025**, *84*, 21991–22008. [CrossRef]
2. Rane, M.E.; Bhadade, U.S. Multimodal Score Level Fusion for Recognition Using Face and Palmprint. *Int. J. Electr. Eng. Educ.* **2025**, *62*, 37–55. [CrossRef]
3. Liang, X.; Zhang, D.; Lu, G.; Guo, Z.; Luo, N. A Novel Multicamera System for High-Speed Touchless Palm Recognition. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *51*, 1534–1548. [CrossRef]
4. Kaur, S.; Chaudhary, G.; Dinesh Kumar, J. PCANet Based Biometric System with Fusion of Palmprint and Dorsal Hand Vein. *J. Intell. Fuzzy Syst.* **2022**, *42*, 841–849. [CrossRef]
5. Yu, T.; Teoh, A.P.; Liao, J.; Wang, C. Show Your Palm to Pay: Are Customers Ready for Palm Print Recognition Technology in Retail Stores in China? *Int. J. Hum. Comput. Interact.* **2025**, *84*, 21991–22008. [CrossRef]
6. Ma, S.; Hu, Q.; Zhao, S.; Wu, W.; Wu, J. Multiscale Multidirection Binary Pattern Learning for Discriminant Palmprint Identification. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2506412. [CrossRef]
7. Stoimchev, M.; Ivanovska, M.; Štruc, V. Learning to Combine Local and Global Image Information for Contactless Palmprint Recognition. *Sensors* **2022**, *22*, 73. [CrossRef]

8. Türk, Ö.; Çalışkan, A.; Acar, E.; Ergen, B. Palmprint Recognition System Based on Deep Region of Interest Features with the Aid of Hybrid Approach. *Signal Image Video Process.* **2023**, *17*, 3837–3845. [CrossRef]
9. Hardalac, F.; Yaşar, H.; Akyel, A.; Kutbay, U. A Novel Comparative Study Using Multi-Resolution Transforms and Convolutional Neural Network (CNN) for Contactless Palm Print Verification and Identification. *Multimed. Tools Appl.* **2020**, *79*, 22929–22963. [CrossRef]
10. Chai, T.; Prasad, S.; Yan, J.; Zhang, Z. Contactless Palmprint Biometrics Using DeepNet with Dedicated Assistant Layers. *Vis. Comput.* **2023**, *39*, 4029–4047. [CrossRef]
11. Jia, W.; Ren, Q.; Zhao, Y.; Li, S.; Min, H.; Chen, Y. EEPNet: An Efficient and Effective Convolutional Neural Network for Palmprint Recognition. *Pattern Recognit. Lett.* **2022**, *159*, 140–149. [CrossRef]
12. Macan, H.F.; Bingöl, Ö.; Türk, S.; Ekinci, M. Super-Resolution Based ROI Enhancement Approaches for Palmprint Recognition. In Proceedings of the 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Türkiye, 21–22 September 2024; pp. 1–5.
13. Zhao, S.; Fei, L.; Wen, J.; Zhang, B.; Zhao, P.; Li, S. Structure Suture Learning-Based Robust Multiview Palmprint Recognition. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 8401–8413. [CrossRef] [PubMed]
14. Zhu, Q.; Xin, G.; Fei, L.; Liang, D.; Zhang, Z.; Zhang, D.; Zhang, D. Contactless Palmprint Image Recognition Across Smartphones with Self-Paced CycleGAN. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 4944–4954. [CrossRef]
15. Shen, L.; Zhang, Y.; Zhao, K.; Zhang, R.; Shen, W. Distribution Alignment for Cross-Device Palmprint Recognition. *Pattern Recognit.* **2022**, *132*, 108942. [CrossRef]
16. Liu, C.; Shao, H.; Zhong, D. Learning Discriminative Palmprint Anti-Spoofing Features via High-Frequency Spoofing Regions Adaptation. *IET Image Process.* **2025**, *19*, e70029. [CrossRef]
17. Chaa, M.; Akhtar, Z. 3D Palmprint Recognition Using Tan and Triggs Normalization Technique and GIST Descriptors. *Multimed. Tools Appl.* **2021**, *80*, 2263–2277. [CrossRef]
18. Jing, K.; Zhang, X.; Song, G. Correntropy-Induced Discriminative Nonnegative Sparse Coding for Robust Palmprint Recognition. *Sensors* **2020**, *20*, 4250. [CrossRef]
19. Gonzalez-Soler, L.J.; Zyla, K.M.; Rathgeb, C.; Fischer, D. Contactless Hand Biometrics for Forensics: Review and Performance Benchmark. *EURASIP J. Image Video Process.* **2024**, *2024*, 26. [CrossRef]
20. Cho, S.; Oh, B.S.; Toh, K.A.; Lin, Z. Extraction and Cross-Matching of Palm-Vein and Palmprint from the RGB and the NIR Spectrums for Identity Verification. *IEEE Access* **2020**, *8*, 4005–4021. [CrossRef]
21. Sünnetci Koşmaz, B.; Bingöl, Ö.; Gedikli, E.; Ekinci, M. Depth Estimation Based Automatic Pose Correction in Hand Images. In Proceedings of the ICHORA 2025-2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Ankara, Türkiye, 23–24 May 2025; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2025.
22. Bingöl, Ö.; Ekinci, M. Stereo-Based Palmprint Recognition in Various 3D Postures. *Expert Syst. Appl.* **2017**, *78*, 74–88. [CrossRef]
23. Zhang, D.; Lu, G.; Li, W.; Zhang, L.; Luo, N. Palmprint Recognition Using 3-D Information. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2009**, *39*, 505–519. [CrossRef]
24. Izadpanahkakhk, M.; Razavi, S.M.; Taghipour-Gorjikolaie, M.; Zahiri, S.H.; Uncini, A. Novel Mobile Palmprint Databases for Biometric Authentication. *Int. J. Grid Util. Comput.* **2019**, *10*, 465–474. [CrossRef]
25. Amrouni, N.; Benzaoui, A.; Zeroual, A. Palmprint Recognition: Extensive Exploration of Databases, Methodologies, Comparative Assessment, and Future Directions. *Appl. Sci.* **2024**, *14*, 153. [CrossRef]
26. Zhu, Q.; Xu, N.; Zhang, Z.; Guan, D.; Wang, R.; Zhang, D. Cross-Spectral Palmprint Recognition with Low-Rank Canonical Correlation Analysis. *Multimed. Tools Appl.* **2020**, *79*, 33771–33792. [CrossRef]
27. Liao, F.; Wu, T.; Gao, F.; Leng, L. Translational Calibration in Region-of-Interest Localization for Palmprint Recognition. *Vis. Comput.* **2024**, *40*, 7293–7305. [CrossRef]
28. Wang, X.; Gong, H.; Zhang, H.; Li, B.; Zhuang, Z. Palmprint Identification Using Boosting Local Binary Pattern. In Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006.
29. Ong Michael, G.K.; Connie, T.; Jin Teoh, A.B. Touch-Less Palm Print Biometrics: Novel Design and Implementation. *Image Vis. Comput.* **2008**, *26*, 1551–1560. [CrossRef]
30. Ojala, T.; Pietika, M.; Ma, T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [CrossRef]
31. Zhang, D.; Kong, W.-K.; You, J.; Wong, M. Online Palmprint Identification. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 1041–1050. [CrossRef]
32. Kong, A.; Zhang, D.; Kamel, M. Palmprint Identification Using Feature-Level Fusion. *Pattern Recognit.* **2006**, *39*, 478–487. [CrossRef]
33. Kong, A.W.-K.; Zhang, D. Competitive Coding Scheme for Palmprint Verification. In Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK, 26 August 2004.

34. Wu, X.; Wang, K.; Zhang, D. Palmprint Texture Analysis Using Derivative of Gaussian Filters. In Proceedings of the 2006 International Conference on Computational Intelligence and Security, Guangzhou, China, 3–6 November 2006.
35. Sun, Z.; Tan, T.; Wang, Y.; Li, S.Z. Ordinal Palmprint Represention for Personal Identification. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005.
36. Sun, Z.; Wang, L.; Tan, T. Ordinal Feature Selection for Iris and Palmprint Recognition. *IEEE Trans. Image Process.* **2014**, *23*, 3922–3934. [CrossRef]
37. Jia, W.; Huang, D.S.; Zhang, D. Palmprint Verification Based on Robust Line Orientation Code. *Pattern Recognit.* **2008**, *41*, 1504–1513. [CrossRef]
38. Guo, Z.; Zhang, D.; Zhang, L.; Zuo, W. Palmprint Verification Using Binary Orientation Co-Occurrence Vector. *Pattern Recognit. Lett.* **2009**, *30*, 1219–1227. [CrossRef]
39. Pan, X.; Ruan, Q.Q. Palmprint Recognition Using Gabor-Based Local Invariant Features. *Neurocomputing* **2009**, *72*, 2040–2045. [CrossRef]
40. Tabejamaat, M.; Mousavi, A. Generalized Gabor Filters for Palmprint Recognition. *Pattern Anal. Appl.* **2018**, *21*, 261–275. [CrossRef]
41. Kong, W.K.; Zhang, D.; Li, W. Palmprint Feature Extraction Using 2-D Gabor Filters. *Pattern Recognit.* **2003**, *36*, 2339–2347. [CrossRef]
42. Li, M.; Wang, H.; Liu, H.; Meng, Q. Palmprint Recognition Based on the Line Feature Local Tri-Directional Patterns. *IET Biom.* **2022**, *11*, 570–580. [CrossRef]
43. Fei, L.; Xu, Y.; Zhang, B.; Fang, X.; Wen, J. Low-Rank Representation Integrated with Principal Line Distance for Contactless Palmprint Recognition. *Neurocomputing* **2016**, *218*, 264–275. [CrossRef]
44. Luo, Y.T.; Zhao, L.Y.; Zhang, B.; Jia, W.; Xue, F.; Lu, J.T.; Zhu, Y.H.; Xu, B.Q. Local Line Directional Pattern for Palmprint Recognition. *Pattern Recognit.* **2016**, *50*, 26–44. [CrossRef]
45. Huang, D.S.; Jia, W.; Zhang, D. Palmprint Verification Based on Principal Lines. *Pattern Recognit.* **2008**, *41*, 1316–1328. [CrossRef]
46. Li, W.; Zhang, B.; Zhang, L.; Yan, J. Principal Line-Based Alignment Refinement for Palmprint Recognition. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2012**, *42*, 1491–1499. [CrossRef]
47. Bai, X.; Gao, N.; Zhang, Z.; Zhang, D. 3D Palmprint Identification Combining Blocked ST and PCA. *Pattern Recognit. Lett.* **2017**, *100*, 89–95. [CrossRef]
48. Ekinci, M.; Aykut, M. Palmprint Recognition by Applying Wavelet-Based Kernel PCA. *J. Comput. Sci. Technol.* **2008**, *23*, 851–861. [CrossRef]
49. Connie, T.; Jin, A.T.B.; Ong, M.G.K.; Ling, D.N.C. An Automated Palmprint Recognition System. *Image Vis. Comput.* **2005**, *23*, 501–515. [CrossRef]
50. Connie, T.; Teoh, A.; Goh, M.; Ngo, D. Palmprint Recognition with PCA and ICA. *Proc. Image Vis. Comput.* **2003**, *163*, 227–232.
51. Song, L.; Huang, H. Revisiting Competitive Coding Approach for Palmprint Recognition: A Linear Discriminant Analysis Perspective. *arXiv* **2022**, arXiv:2206.15349. [CrossRef]
52. Wang, Y.; Ruan, Q. Kernel Fisher Discriminant Analysis for Palmprint Recognition. In Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, China, 20–24 August 2006.
53. Shang, L.; Huang, D.S.; Du, J.X.; Zheng, C.H. Palmprint Recognition Using FastICA Algorithm and Radial Basis Probabilistic Neural Network. *Neurocomputing* **2006**, *69*, 1782–1786. [CrossRef]
54. Gui, J.; Jia, W.; Zhu, L.; Wang, S.L.; Huang, D.S. Locality Preserving Discriminant Projections for Face and Palmprint Recognition. *Neurocomputing* **2010**, *73*, 2696–2707. [CrossRef]
55. Lu, J.; Tan, Y.P. Improved Discriminant Locality Preserving Projections for Face and Palmprint Recognition. *Neurocomputing* **2011**, *74*, 3760–3767. [CrossRef]
56. Zhao, Z.Q.; Huang, D.S.; Jia, W. Palmprint Recognition with 2DPCA+PCA Based on Modular Neural Networks. *Neurocomputing* **2007**, *71*, 448–454. [CrossRef]
57. Li, M.; Yuan, B. 2D-LDA: A Statistical Linear Discriminant Analysis for Image Matrix. *Pattern Recognit. Lett.* **2005**, *26*, 527–532. [CrossRef]
58. Xu, S.; Suo, J.; Ding, J. Improved Linear Discriminant Analysis Based on Twodimensional Gabor for Palmprint Recognition. In Proceedings of the 2011 International Conference of Soft Computing and Pattern Recognition (SoCPaR), Dalian, China, 14–16 October 2011; IEEE: New York, NY, USA, 2011.
59. Hu, D.; Feng, G.; Zhou, Z. Two-Dimensional Locality Preserving Projections (2DLPP) with Its Application to Palmprint Recognition. *Pattern Recognit.* **2007**, *40*, 339–342. [CrossRef]
60. Pan, X.; Ruan, Q.Q. Palmprint Recognition with Improved Two-Dimensional Locality Preserving Projections. *Image Vis. Comput.* **2008**, *26*, 1261–1268. [CrossRef]
61. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385. [CrossRef]

62. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556. [CrossRef]
63. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
64. Zhang, L.; Cheng, Z.; Shen, Y.; Wang, D. Palmprint and Palmvein Recognition Based on DCNN and a New Large-Scale Contactless Palmvein Dataset. *Symmetry* **2018**, *10*, 78. [CrossRef]
65. Matkowski, W.M.; Chai, T.; Kong, A.W.K. Palmprint Recognition in Uncontrolled and Uncooperative Environment. *IEEE Trans. Inf. Forensics Secur.* **2019**, *15*, 1601–1615. [CrossRef]
66. Liu, Y.; Kumar, A. Contactless Palmprint Identification Using Deeply Learned Residual Features. *IEEE Trans. Biom. Behav. Identity Sci.* **2020**, *2*, 172–181. [CrossRef]
67. Shao, H.; Zhong, D.; Du, X. Efficient Deep Palmprint Recognition via Distilled Hashing Coding. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 16–17 June 2019.
68. Genovese, A.; Piuri, V.; Plataniotis, K.N.; Scotti, F. PalmNet: Gabor-PCA Convolutional Networks for Touchless Palmprint Recognition. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 3160–3174. [CrossRef]
69. Yang, Z.; Huangfu, H.; Leng, L.; Zhang, B.; Teoh, A.B.J.; Zhang, Y. Comprehensive Competition Mechanism in Palmprint Recognition. *IEEE Trans. Inf. Forensics Secur.* **2023**, *18*, 5160–5170. [CrossRef]
70. Shao, H.; Zhong, D.; Du, X. Deep Distillation Hashing for Unconstrained Palmprint Recognition. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2505613. [CrossRef]
71. Zhong, Y.; Chai, W.; Wang, L.; Zheng, D.; Fu, H.; Ma, H. RegPalm: Towards Large-Scale Open-Set Palmprint Recognition by Reducing Pattern Variance. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 8541–8554. [CrossRef]
72. Chai, T.; Prasad, S.; Wang, S. Boosting Palmprint Identification with Gender Information Using DeepNet. *Future Gener. Comput. Syst.* **2019**, *99*, 41–53. [CrossRef]
73. Zhang, L.; Li, L.; Yang, A.; Shen, Y.; Yang, M. Towards Contactless Palmprint Recognition: A Novel Device, a New Benchmark, and a Collaborative Representation Based Identification Approach. *Pattern Recognit.* **2017**, *69*, 199–212. [CrossRef]
74. Jia, W.; Hu, R.X.; Gui, J.; Zhao, Y.; Ren, X.M. Palmprint Recognition Across Different Devices. *Sensors* **2012**, *12*, 7938–7964. [CrossRef]
75. Pizer, S.M.; Amburn, E.P.; Austin, J.D.; Cromartie, R.; Geselowitz, A.; Greer, T.; ter Haar Romeny, B.; Zimmerman, J.B.; Zuiderveld, K. Adaptive Histogram Equalization and Its Variations. *Comput. Vis. Graph. Image Process.* **1987**, *39*, 355–368. [CrossRef]
76. Land, E.H.; McCann, J.J. Lightness and Retinex Theory. *J. Opt. Soc. Am.* **1971**, *61*, 1–11. [CrossRef]
77. COEP Palmprint Database. Available online: https://www.coep.org.in/resources/coeppalmprintdatabase (accessed on 9 February 2021).
78. Aykut, M.; Ekinci, M. Developing a Contactless Palmprint Authentication System by Introducing a Novel ROI Extraction Method. *Image Vis. Comput.* **2015**, *40*, 65–74. [CrossRef]
79. Kumar, A. Incorporating Cohort Information for Reliable Palmprint Authentication. In Proceedings of the 6th Indian Conference on Computer Vision, Graphics and Image Processing, ICVGIP 2008, Bhubaneswar, India, 16–19 December 2008; pp. 583–590.
80. Kanhangad, V.; Kumar, A.; Zhang, D. A Unified Framework for Contactless Hand Verification. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1014–1027. [CrossRef]
81. Kumar, A. Towards More Accurate Matching of Contactless Palmprint Images under Less Constrained Environments. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 34–47. [CrossRef]
82. Ungureanu, A.-S.; Thavalengal, S.; Cognard, T.E.; Costache, C.; Corcoran, P. Unconstrained Palmprint as a Smartphone Biometric. *IEEE Trans. Consum. Electron.* **2017**, *63*, 334–342. [CrossRef]
83. Shao, H.; Zhong, D.; Du, W. Towards Efficient Unconstrained Palmprint Recognition via Deep Distillation Hashing. *arXiv* **2020**, arXiv:2004.03303. [CrossRef]
84. Zhao, S.; Fei, L.; Zhang, B.; Wen, J.; Cui, J. Toward Mobile Palmprint Recognition via Multi-View Hierarchical Graph Learning. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 101–113. [CrossRef]
85. Ma, S.; Hu, Q.; Zhao, S.; Chen, S.; Jiang, L. SYEnet: Simple yet Effective Network for Palmprint Recognition. *Inf. Sci.* **2024**, *669*, 120518. [CrossRef]
86. Lugaresi, C.; Tang, J.; Nash, H.; McClanahan, C.; Uboweja, E.; Hays, M.; Zhang, F.; Chang, C.-L.; Yong, M.G.; Lee, J. Mediapipe: A Framework for Building Perception Pipelines. *arXiv* **2019**, arXiv:1906.08172. [CrossRef]
87. Zhang, F.; Bazarevsky, V.; Vakunov, A.; Tkachenka, A.; Sung, G.; Chang, C.-L.; Grundmann, M. MediaPipe Hands: On-Device Real-Time Hand Tracking. *arXiv* **2020**, arXiv:2006.10214.
88. Lin, C.; Chen, Y.; Zou, X.; Deng, X.; Dai, F.; You, J.; Xiao, J. An Unconstrained Palmprint Region of Interest Extraction Method Based on Lightweight Networks. *PLoS ONE* **2024**, *19*, e0307822. [CrossRef] [PubMed]
89. Zhu, J.; Zhong, D.; Luo, K. Boosting Unconstrained Palmprint Recognition with Adversarial Metric Learning. *IEEE Trans. Biom. Behav. Identity Sci.* **2020**, *2*, 388–398. [CrossRef]

90. Chai, T.; Wang, X.; Li, R.; Jia, W.; Wu, X. Joint Finger Valley Points-Free ROI Detection and Recurrent Layer Aggregation for Palmprint Recognition in Open Environment. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 421–435. [CrossRef]
91. Liang, X.; Li, Z.; Fan, D.; Li, J.; Jia, W.; Zhang, D. Touchless Palmprint Recognition Based on 3D Gabor Template and Block Feature Refinement. *Knowl. Based Syst.* **2022**, *249*, 108855. [CrossRef]
92. Wali, A.; Shariq, R.; Shoaib, S.; Amir, S.; Farhan, A.A. Recent Progress in Sign Language Recognition: A Review. *Mach. Vis. Appl.* **2023**, *34*, 127. [CrossRef]
93. Aggarwal, D.; Ahirwar, S.; Srivastava, S.; Verma, S.; Goel, Y. Sign Language Prediction Using Machine Learning Techniques: A Review. In Proceedings of the 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Coimbatore, India, 2–4 March 2023; IEEE: New York, NY, USA, 2023; pp. 1296–1300.
94. Sundar, B.; Bagyammal, T. American Sign Language Recognition for Alphabets Using MediaPipe and LSTM. *Procedia Comput. Sci.* **2022**, *215*, 642–651. [CrossRef]
95. Srivastava, S.; Singh, S.; Pooja; Prakash, S. Continuous Sign Language Recognition System Using Deep Learning with MediaPipe Holistic. *Wirel. Pers. Commun.* **2024**, *137*, 1455–1468. [CrossRef]
96. Dong, Y.; Payandeh, S. Hand Kinematic Model Construction Based on Tracking Landmarks. *Appl. Sci.* **2025**, *15*, 8921. [CrossRef]
97. Pu, M.; Chong, C.Y.; Lim, M.K. Robustness Evaluation in Hand Pose Estimation Models Using Metamorphic Testing. In Proceedings of the 2023 IEEE/ACM 8th International Workshop on Metamorphic Testing (MET), Melbourne, Australia, 14 May 2023; IEEE: New York, NY, USA, 2023; pp. 31–38.
98. Turk, S.; Bingol, O.; Dogan, R.O. Selecting the Best Frames from Videos for Palmprint Biometric Verification. In Proceedings of the 8th International Artificial Intelligence and Data Processing Symposium, IDAP 2024, Online, 21–22 September 2024; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2024.
99. Kocakulak, M.; Acir, N. Dynamic ROI Extraction for Palmprints Using Mediapipe Hands. In Proceedings of the 2022 30th Signal Processing and Communications Applications Conference (SIU), Safranbolu, Turkey, 15–18 May 2022; IEEE: New York, NY, USA, 2022; pp. 1–4.
100. Kocakulak, M.; Acır, N. A Contactless Palmprint Imaging System Design Using Mediapipe Hands. *J. Innov. Sci. Eng.* **2023**, *7*, 29–39. [CrossRef]
101. Ettehadi, N.; Kashyap, P.; Zhang, X.; Wang, Y.; Semanek, D.; Desai, K.; Guo, J.; Posner, J.; Laine, A.F. Automated Multiclass Artifact Detection in Diffusion MRI Volumes via 3D Residual Squeeze-and-Excitation Convolutional Neural Networks. *Front. Hum. Neurosci.* **2022**, *16*, 877326. [CrossRef]
102. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
103. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A Unified Embedding for Face Recognition and Clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.
104. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. Pytorch: An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 8026–8037.
105. Kingma, D.P. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
106. Nalamothu, A.; Rayachoti, E. An Effective Feature Selection and Classification Technique for Palmprint Biometric Identification Systems. *Knowl. Inf. Syst.* **2025**, *67*, 8115–8142. [CrossRef]
107. Zhang, C.; Liang, X.; Fan, D.; Chen, J.; Zhang, B.; Wu, B.; Zhang, D. HGAIQA: A Novel Hand-Geometry-Aware Image Quality Assessment Framework for Contactless Palmprint Recognition. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 5039713. [CrossRef]
108. Wang, H.; Cao, X. Palmprint Recognition Using Bifurcation Line Direction Coding. *IEEE Access* **2025**, *13*, 70366–70377. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.