Real-Time Dorsal Hand Recognition based on Smartphone

Mohamed I. Sayed¹, Mohamed Taha¹, and Hala H. Zayed¹,
¹Computer Science Department, Faculty of Computers and Artificial Intelligence, Benha University
Corresponding author: Mohamed I. Sayed (mohamed.sayed@fci.bu.edu.eg).

ABSTRACT The integration of biometric recognition with smartphones becomes necessary for increasing security, especially in financial transactions such as online payment. Vein recognition of dorsal hand is outstanding than other methods like palm, finger, and wrist as it has a wide area to be captured and does not have any wrinkles. Most of the current systems that depend on dorsal hand vein recognition are not work in real-time and have poor results. In this paper, a dorsal hand recognition system working in real-time is proposed to achieve good results with a high frame rate. A contactless device consists of a Universal Serial Bus (USB)-Camera, infrared LEDs, and connected to a smartphone used to collect our dataset. The dataset contains 2200 images collected from both hands of 100 persons. The captured images are processed with light algorithms to improve real-time performance and increase frame rate. The algorithm used for feature detection and extraction is Oriented FAST and Rotated BRIEF (ORB) with K-Nearest Neighbors (K-NN) matching to match features. Another benchmark called Poznan University of Technology (PUT) dataset is used to measure the proposed system efficiency. The result obtained from experimental testing showed that the proposed system had a low Equal Error Rate (EER) with 4.33% and a high frame rate of 29 frames per second (Fps).

INDEX TERMS Biometric system, Dorsal hand, Features extraction, Features matching, Near-Infrared, Veins Recognition, Smartphone.

I. INTRODUCTION
Biometric recognition is one of the most rising areas in the research field in the last years, with an increasing expansion in identification/authentication systems in our daily life: online purchases, bank account access, and smartphone unlocking. Biometric recognition systems are now used worldwide, and they are steadily replacing traditional ways of authentication such as user ID and passwords. These systems provide more prominent security and comfort to the users than conventional methods. Two types of biometrics are physiological (e.g., facial, iris, fingerprint, and vascular) and behavioral (e.g., voice, gesture, keystroke dynamics, and gait). One of the biometric recognition areas that attract researchers nowadays is vein recognition. The advantage of vein recognition systems is only used on alive persons. It depends on capturing veins by illuminating veins with infrared light [1].

In 1543, Andreas Vesalius proposed that the veins in our bodies are changing in their structures and positions. A professor called Arrigo Tamassia, a specialist in forensic medicine, stated that in humans, no vein pattern on the back of the hand is identical in any person, even if they are twins [2]. To perform vein recognition, it needs to build an appropriate device for scanning and capturing a vein's pattern. The main advantage of using veins pattern in authentication is that it needs the person's knowledge as they cannot be seen under normal light. Also, veins are found inside the body, making it immune to changes because it is protected from environmental effects. Research in the veins authentication field has shown that the authentication systems can have higher accuracy [3]. Veins' pattern can be captured from distinct hand parts, like a finger, palm, dorsal, and wrist. The biometric characteristics have received enormous interest in the latest years due to these vein patterns. They are hidden under the body's skin, making them difficult to forge [4].

Finger recognition is the most common method in biometric recognition. Still, it has some disadvantages as there is difficulty in using it because the finger needs to be positioned accurately on the scanning device. Also, the finger has a smaller surface area to be scanned, which means fewer feature
points, leading to difficulty recognizing the veins correctly [5]. On the contrary, dorsal or palm veins occupy a wider area and more feature points to be extracted for recognition. In this case, it will result in better security and higher accuracy [6]. Since we can capture palm vein patterns, and palm prints that are unique to each person, we will have good information for the recognition process [7].

In contrast, the palm has thicker skin tissues and wrinkles that could affect the shape of the vein patterns, whereas the dorsal hand is wrinkle-free. Also, the dorsal hand skin contains melanin, which reflected infrared light used for imaging veins since it penetrates deeper into the skin to deliver good image quality [8].

Researchers are becoming increasingly interested in dorsal hand vein biometrics because of their characteristics. We can summarize these characteristics in the following points: it does not need any contact with the capture device, cannot be forged, does not change over time, and provides high accuracy [9]. Recognition systems of the dorsal hand rely on how the collected images captured from the device are good. Since the dorsal hand veins are concealed within the skin, so they are undetectable under normal light. Near-infrared (NIR) light is used to distinguish veins from the back of the hand that can be done by illuminating the hand with NIC light. Since hemoglobin absorbs infrared light, the veins look darker than the other areas around the veins, which can then be captured as dark pixels by the camera [10].

With the increasing usage of smartphones in all fields of our lives, especially in global health, a biometric solution that can be easily combined with mobile devices is needed. Devices use biometric recognition like fingerprint scanners introduced in notebooks computers and all the current and upcoming smartphones produced nowadays. PalmSecure [11] and Finger Vein Authentication [12] are two available commercial devices depending on veins recognition for authentication systems from Hitachi corporation that need direct contact between the hands and devices.

In 2019, the LG company introduced the first smartphone using palm veins for biometric recognition: LG G8S ThinQ [13]. Although palm vein recognition in this smartphone is slower than traditional authentication methods such as facial or fingerprint recognition, it is an excellent start to relying on veins for biometric recognition.

As a result of the spread of the Coronavirus in the world known as COVID-19, it has become imperative to have biometric systems based on contactless devices. This is what will be introduced in this paper.

The main contributions of our proposed work are stated as follows:

1) A real-time system for dorsal hand vein recognition based on smartphones for collecting and processing data.
2) We used a low-cost contactless hardware device that was introduced in our previous work for veins detection [14] to create a new dataset of dorsal hands.
3) We proposed a novel recognition system that combines the Oriented FAST and Rotated BRIEF (ORB) algorithm with K-Nearest Neighbors (K-NN) for features extraction and matching, which yielded good results.
4) We conducted experiments using our dataset and PUT [15] dataset to evaluate the proposed recognition system efficiency.
5) We used different smartphones in specification to evaluate the proposed system and obtain the computational time efficiency.
6) The results showed the proposed system has a low equal error rate. Also, the proposed system has a low computational time, which makes it suitable for running on smartphones at a high frame rate.

The remaining parts of the paper are stated as follows: Section 2 shows the state-of-art dorsal hand recognition systems. Section 3 describes the proposed system. Discussion of the experimental results and evaluation of the proposed system in section 4. Finally, Section 5 states conclusions and future work.

II. RELATED WORK

In previous works, researchers depended on algorithms with heavy computational time in pre-processing steps or feature extraction and matching processes, whereas they did not provide good results. Also, we notice the lack of biometric systems integrated with mobile phones, whereas we can depend on them in processing, in addition to the capabilities that can be provided, such as ease of portability. Some of the previous work of vein recognition of dorsal hand are discussed here:

The authors introduced a new device to capture dorsal hand veins with good quality [16]. They used 40 infrared LEDs with 940nm of wavelength for illumination. A camera ADMK 22BUC03 was used, which is sensitive to infrared light with 744×480 pixels. For gathering their dataset, the hand must be placed at a fixed position. They captured five images from both hands of 50 persons in two different sessions. A CLAHE technique was used only for pre-processing to enhance image contrast. For feature extraction, they used two local and six global feature extraction techniques. They used cross-correlation to compare features extracted from local techniques and Sparse Representation Classifier (SRC) to compare features extracted from global techniques. After testing the eight techniques, the results stated that the best result among the other techniques was Log-Gabor (LG), with an EER of 7.67%. The main drawback is that Log Gabor’s technique requires a very high computational time and storage space [17].

In [18], the authors presented a recognition system based on a dorsal hand using a smartphone camera. They built a box for infrared illumination and used a special camera sensitive to infrared (mvBlueFox IGC). The database consists of 624 images from 52 persons. Pre-processing started with extracting the Region of Interest (ROI), then two smoothing
filters, median, and Gaussian were used for removing noise, followed by CLAHE for image enhancement. Three techniques: Hessian-Laplace, Harris-Laplace, and Difference of Gaussian (DoG), were used for detecting key points from images. The scale-invariant feature transform (SIFT) was used to compute the descriptor of detected features. To evaluate their system, a smartphone that was not unmodified was used to collect 312 images compared with their database. It was noticed that after the test was done, DoG with SIFT had an EER of 5.6% and false accept rate (FAR) of 61.8%, whereas Hessian-Laplace had an EER of 6.16% FAR of 49.03%, and Harris-Laplace had an EER of 7.73% and FAR of 46.1%. These values are considered to be very high for biometric recognition applications. Also, all techniques used in this work require significant computational time or give a low accuracy rate.

In [20], they presented a near-infrared illumination attached with mobile for hand recognition. A smartphone is called Google Nexus 5 smartphone was used in this study. It is a modified mobile by EigenImaging company specialized in modifying smartphones for near-infrared imaging applications [21]. This modification is a bit costly as it raises the original price of the smartphone to approximately three times. The dataset involved 920 images collected from 31 persons in two sessions. They started their proposed algorithm by extracting ROI with a size of 512×512 pixels, image enhancement with CLAHE, High-Frequency Emphasis Filtering (HEF), and Circular Gabor Filter (CGF). Maximum Curvature (MC), Principal Curvature (PC), Gabor Filter (GF), and SIFT techniques were used to extract features. For the whole dataset, the best result came from MC with an EER of 24.30% over other techniques, which is high. The algorithms proposed in this study have a high complexity time which is not suitable for mobile phones subsequently will affect the frame rates.

In [23], the authors used biometric graph matching (BGM) for the dorsal hand vein recognition system. They used a Complementary Metal Oxide Semiconductor (CMOS) Camera and 6 near-infrared LEDs to collect their dataset. The dataset is called XJTU-A and contains 789 images from 57 persons. A Projection Vein Finder VIVO500S was also used to collect another dataset. The dataset is called XJTU-B 448 images from 56. For pre-processing, they extracted the ROI followed by morphological opening operation. The curvature point algorithm was used for segmentation followed by vein skeleton extraction. To perform matching, they used BGM that consists of three phases: Feature graph registration, Feature graph matching, and Distance measurement. The results showed that the proposed system had an EER of 5.34% for the XJTU-A dataset and 5.46% for the XJTU-B dataset.

Kumar et al. developed a system to identify people through the dorsal hand veins with medical insurance to help them receive treatment in the hospitals affiliated with the insurance without paying [24]. A NIR camera VF620 with a wavelength of 850nm connected to a laptop was used to collect their dataset. They collected 1000 vein patterns from 250 persons. The morphological opening operation was used after extracting ROI to remove dark areas in the image. Then a histogram equalization was done for image enhancement. For the identification process, they used SIFT algorithm. The EER obtained from the testing is about 4.35%. Although the result obtained was good, the SIFT is a patent algorithm and not free also, it is not the best choice for real-time applications as it has large computations and a sluggish running time.

In [26], Garcia and Sanchez proposed a real-time biometric system based on wrist veins for recognition. They used their proposed hardware setup in [27] to collect their dataset called UC3M-CV1 that contains 1200 images collected from 50 persons in two different sessions. They proposed an algorithm called Three-Guideline Software (TGS) to set some instructions for the users to place their hands in specific positions within three lines on the screen for creating the dataset. Pre-processing techniques started with CLAHE then the result was filtered by three smoothing filters: Gaussian, median, and an average of size 11×11. Three algorithms (ORB, Speeded Up Robust Features (SURF), and SIFT) were used for features extraction in identification and verification. Brute Force Matcher (BFM) with simple distance threshold was used for matching features that were extracted from ORB, and Fast Library for Approximate Nearest Neighbor (FLANN) was used with SURF and SIFT. The results stated that the recognition system had an EER of 21.76, 32.29, and 39.94% for the whole dataset using SIFT, SURF, and ORB, respectively. Also, they applied their proposed algorithm to the PUT database [15], which contains 1200 images collected from 50 persons. The results obtained had a high EER value of 34.13% using ORB and a low EER of 15.93% using SIFT.

In [28], the same authors introduced a wrist recognition system based on smartphones. The smartphones used in this study were Pocophone F1© and Mi 8© from Xiaomi. The two smartphones came with built-in infrared LEDs and a near-infrared camera for face recognition. They used the two smartphones to collect 2400 images from 50 persons. They used the same algorithms from their previous work [26]. The results stated that the recognition system had an EER of 18.72, 25.19, and 34.85% for the whole dataset using SIFT, SURF, and ORB, respectively. We can note that the SIFT outperformed the other algorithms but with a high error rate and low frame rate of 3-4 frame per second (Fps) that is not suitable for real-time processing.

In [29], their work was based on Convolution Neural Network (CNN) for dorsal hand recognition system. They have collected 4000 good-quality images from 200 persons, and 2000 low-quality images from 100 persons. Different CNN architectures were used: AlexNet, VGG-16, and VGG-19. The results showed that the system had an EER of 6.502% obtained from low-quality images and 1.006% from good-quality images using VGG Net-19 with fine-tuning. Although
the CNN methods achieved a good result, they take too much time and resources for training and testing processes, affecting the frame rate [30].

In [31], the authors introduced a cross-device to recognize dorsal hand veins using coarse-to-fine matching. They used two different devices to collect 2000 images with a size of 640×480 from 100 persons. Pre-processing steps started with grayscale conversion followed by segmentation. The segmentation threshold value depends on the gradient of the image and the mean value of a window of N×N. The SIFT algorithm is applied to extract features from the collected images and Distinctive Efficient Robust Features (DERF) to compute keypoint descriptors. The system had an EER of 3.20%. Although the result was good, the algorithms used in pre-processing and feature extraction and matching require high computational time [25].

III. PROPOSED WORK

As shown in Fig. 1, the proposed system for the recognition process consists of two main phases: The Database construction and identification or verification phase. The database construction phase starts with capturing images from users, and these images are processed with some pre-processing methods: Grayscale conversion, Removing noise, and Sharpening. Then, the features are extracted from images using the ORB algorithm and sent to the database to be stored for comparison in the identification or verification phase. In identification, it starts with capturing a new frame from the camera. The same pre-processing methods are done on the new frame then the features are extracted. The next step will be matching the feature extracted from the new frame and the whole features stored in the database using K-NN matching to identify the user by his name and which hand is scanned. The same processes in the identification phase are done in the verification phase, except it starts with asking for the user’s id. The extracted features are compared with the user's features stored in the database only. All processes in the identification or verification phase are processed in real-time on an Android smartphone.

A. HARDWARE SETUP

Contactless hardware is used to capture images for the proposed system and was implemented and tested in our previous work [14]. The hardware contains three main components: USB-Camera, Infrared LEDs, and a smartphone mobile. The USB-Camera has a resolution of 640×480 pixels and was modified to be sensitive to infrared light by removing the infrared cut-off filter with another one that only wavelengths above 700nm can pass see Fig. 2(a). For illumination, 20 infrared LEDs are mounted on a test board and connected to a power supply see Fig. 2(b). The dorsal hand is illuminated with infrared LEDs then the camera starts capturing frames and sends them to the smartphone mobile see Fig. 2(d) via On-The-Go cable (OTG) for further processing.
B. DATASET COLLECTION

This hardware configuration is used to collect the dataset. It consists of 2200 dorsal hand images captured from both hands of 100 persons: 47 males and 53 females. The user places his hand in front of the USB-Camera at 10-13 cm away. The dorsal hand is illuminated with infrared LEDs then the camera starts capturing images (11 samples from each hand) and sends them to the smartphone mobile via On-The-Go cable (OTG) for processing. The captured images are resized to 320×240 and stored in .png format to retain the quality of the images. Pre-processing steps are performed to enhance the collected images. The ORB algorithm is applied to detect keypoints and compute descriptors of the images to be stored in the database. The dataset was collected indoors and under ambient light. Fig. 3(a) shows samples of the left dorsal hand, and Fig. 3(b) shows samples of the right dorsal hand of the same users.

C. PRE-PROCESSING STEPS

The main goal of pre-processing steps is to improve and enhance the quality of the image for more processing. The pre-processing methods used in this work have a low computational time to increase processing efficiency in real-time. The dorsal hand is illuminated with infrared radiation and the camera starts to capture frames of the dorsal hand then sends them to the smartphone for processing. The pre-processing steps start with grayscale conversion, as shown in Figure 4(a), shows the conversion result from RGB format to grayscale. CLAHE method is used to enhance image contrast as it outperforms the traditional Global Histogram Equalization (GHE) method that will not give good results because of the difference in illumination in the captured frames, which will affect the next pre-processing steps. CALHE counts on regional contrast and produces a good result especially, in biometric imaging systems [32], [33] see Figure 4(b). A smoothing filter called a median filter is used to smooth images and reduce noises produced from the CLAHE method. The median filter is the best way to eliminate salt and pepper noise and retain the edges of images [34]. Fig. 4(c) shows the result of smoothing the image. For further enhancement, the image is sharpened by the unsharp masking technique. Unsharp masking is a simple filter used for sharpening images. It can be done through two steps: first, blurring the original image, then second, subtracting the blurred image from the original image to detect and present edges using the following Equation (1).

\[ g(x,y) = (k + 1) \times f(x,y) - k \times f_{\text{smooth}}(x,y) \]  

Where:

- \( g(x,y) \) is the output image,
- \( k \) is constant of scaling,
- \( f(x,y) \) is the original image,
- \( f_{\text{smooth}}(x,y) \) is the blurred image.

D. ORB FEATURES EXTRACTION AND MATCHING

An ORB algorithm is used to extract features from the images developed in 2011 [35]. The reasons behind using this algorithm: rotation and scale invariance, impervious to noise, highly effective results, free to use, low memory usage. Also, its computational cost is very low, making it suitable for mobile to process in real-time [36]. ORB mainly consists of two main phases:

- Detecting keypoints using the FAST algorithm.
- Computing descriptors using the BRIEF descriptor.
1) FAST ALGORITHM
The detection process starts using the FAST, a Features from Accelerated Segment Test algorithm for detecting key points [37]. The FAST algorithm depends on finding a pixel with a significant difference in brightness with its neighborhood pixels to be a corner point. Then it can be considered a key point. Harris corner [38] is then applied to calculate the Harris scores on every corner point extracted from FAST based on the intensity difference around that corner point. These scores are sorted to select the top corner points. Because FAST does not have multi-scale features, a multi-scale pyramid of the image adds scale invariance to the detected points. Also, FAST does not have an orientation component so, the intensity weighted centroid [39] is computed to add direction to the detected points. After applying the FAST algorithm, Fig. 5 shows the extracted key points (50 keypoints as an example).

![FIGURE 5. keypoints extracted from the right hand of user18 using ORB]

2) BRIEF DESCRIPTOR
After detecting the key points using the FAST algorithm, the BRIEF descriptor is used for computing and generating descriptors of detected keypoints [40]. BRIEF (Binary Robust Independent Elementary Features) is a binary string vector that contains only 0's and 1's. First, a Gaussian kernel is used for smoothing the image to make BRIEF resistant to noise. The binary descriptor vector is created by comparing random pixels obtained from a $31 \times 31$ patch around the detected keypoint. If the intensity value at point $y$ is greater than at point $x$, it sets the value to 1; otherwise, the value is 0, as shown in Equation (2). Because the BRIFE descriptor lacks rotation invariance, the ORB uses the steer BRIFE to compute the direction of detected keypoints added to the descriptor vector [35].

$$\tau(p; x, y) = \begin{cases} 1, & p(x) < p(y) \\ 0, & \text{otherwise} \end{cases}$$  \hfill (2)

where:
- $\tau$ is the binary test between point $x$ and point $y$.
- $p(x), p(y)$ the intensity value at point $x$ and $y$.

All these steps of detecting key points and computing descriptors of these points to generate the descriptor vector of each point are performed on the dataset. Then, the descriptor vector is saved to the database to be used later in the matching process.

3) FEATURE MATCHING
In the matching process, the primary purpose is to find similarities between two images. After implementing the ORB algorithm to obtain the binary descriptor from the query image captured from the user, the binary descriptors already exist in the database. BFM with K-NN is used to classify key points produced by ORB. It provides good results with less computational time [41]. The distance between descriptors is measured using Hamming distance. K-NN is used to find the two best matches with the smallest distance for each key point in the query image then Lowe's ratio [42] is applied to obtain the good matches. The match point is accepted if the distance ratio between the two nearest neighbors is less than a predefined threshold or rejected if the distance ratio is more significant than the predefined threshold, as shown in Equation (3):

$$match = \begin{cases} \text{good match,} & d_1 < R \\ \text{incorrect match,} & d_2 > R \end{cases}$$  \hfill (3)

Where:
- $match$ is the feature comparison function between two descriptors,
- $d_1$ the distance between a descriptor and first nearest neighbor,
- $d_2$ the distance between a descriptor and second nearest neighbor
- $R$ the Lowe's ratio

After matching the test image with all descriptors in the database and obtaining suitable matches for each descriptor, the max number of suitable matches will specify which person matched with the test image. Fig. 6(a) shows suitable matches obtained from matching two different samples of the right hand of user18 also Fig. 6(b) shows the suitable matches between left and right of the same user.

![FIGURE 6. Results of suitable matches of user18: (a) two samples from the right hand. (b) two samples from left and right hand.]

VOLUME XX, 2017
IV. RESULTS AND DISCUSSION

Some tests were conducted to evaluate the proposed recognition system and its computational-time efficiency. Firstly, our database is used for the evaluation of system recognition performance. Secondly, the PUT database [15] is used to evaluate the proposed system. Finally, we tested the proposed system on different smartphones to obtain time efficiency. In this section, reports and experimental results will be discussed.

A) RECOGNITION PERFORMANCE

False Accept Rate (FAR) and False Reject Rate (FRR) were calculated to determine system performance. FAR where imposter users are accepted wrongly, and it is calculated using Equation (4). FRR where genuine users are rejected wrongly, and it is calculated using Equation (5). The dataset with 2200 images had 2000 (100 persons × 2 dorsal hands × 10 samples) genuine scores and 398000 (100 persons × 2 dorsal hands × 199 patterns × 10 samples) impostor scores.

\[
FAR = \frac{\text{False Accepted}}{\text{total impostor score}}
\]  
\[\text{(4)}\]

\[
FRR = \frac{\text{False Rejected}}{\text{total genuine score}}
\]  
\[\text{(5)}\]

The equal error rate is calculated to measure the performance of the proposed system using FAR and FRR. The EER is where FRR and FAR intersect in the Detection Error Trade-off Curve (DET). Fig. 6 plots the FRR in the orange line and FAR in the blue line according to the threshold values. The EER value obtained according to Fig. 7 was 4.33%.

Fig. 8 presents the Receiver Operating Curve (ROC), it is plotting False Positive Rate (FPR) calculated using Equation (6), and True Positive Rate (TPR) calculated using Equation (7) to measure classification performance. The area under the ROC curve (AUROC) was 0.99. The higher is the AUROC, the better the performance.

\[
FPR = \frac{\text{False Positive}}{\text{True Negative+False Positive}}
\]  
\[\text{(6)}\]

\[
TPR = \frac{\text{True Positive}}{\text{True Positive+False Negative}}
\]  
\[\text{(7)}\]

Table 1 shows a comparison between the proposed system and other previous systems in terms of the methodology used, the number of images in the dataset, and EER values. The systems that work in real-time were [26] and [28], while the other systems work offline.

We can note from the table that all systems almost used the same methods in pre-processing with some differences, while in features extraction, different methods were used. As in [16], eight methods were used to extract the features and a sparse representation classifier was also used for the classification process, but the method that gave better results was LG, with an EER value of 7.67%. In [18] three methods were used to detect keypoints with Sift for extracting keypoints descriptors, and the best result was DoG with SIFT with an EER value of 5.19%. Also, in [20], more than one method was used in feature extraction, but the method with good results was MC with an EER value of 24.30%. In [23], BGM was used for feature extraction and matching. The EER value obtained was 5.34% for the XJTU-A dataset and 5.46% for the XJTU-B dataset. In [24], some morphological operations were used for image enhancement and SIFT for features extractions producing an EER value of 4.35%. In [26] and [28], three methods: SIFT, SURF, and ORB were used for feature extraction while the EER value obtained from using ORB was too high in [26] was 39.94% and in [28] was 34.85%. In [29], VGG-19 was used with fine-tuned while the result obtained for good-quality images was 1.006% of EER, while for low-quality images was 6.502%. In [31], they used SIFT and DERF for feature extraction and description. The value of EER obtained was 3.20%.

Also, we can see from the table that the proposed system has a promising performance with an EER value of 4.33%. The good results are obtained from pre-processing methods that delivered better-enhanced images and the use of ORB for features extraction with K-NN in the matching process.
TABLE 1. Comparative Results between the proposed system and previous systems

| Study                     | Year | Methodology                                      | Dataset images | EER % |
|---------------------------|------|-------------------------------------------------|----------------|-------|
| Raghavendra et al. [16]   | 2015 | CLAHE + Log-Gabor (LG) - SRC                    | 1000           | 7.67  |
| Patil et al. [18]         | 2016 | CLAHE + DoG with SIFT                           | 936            | 5.60  |
| Debiasi et al. [20]       | 2018 | CLAHE + (HFE) + (CGF) + (MC) + (PC) + (GF) +(SIFT) | 920            | 24.3  |
| Zhong et al. [23]         | 2019 | Segmentation + vein skeleton extraction + BGM   | 789            | 5.34  |
| Kumar et al. [24]         | 2019 | morphological operation + SIFT                  | 1000           | 4.35  |
| Garcia and Sanchez [26]   | 2020 | CLAHE + Median + Gaussian + average smooth filter + ORB | 1200           | 39.94 |
| Garcia and Sanchez [28]   | 2020 | CLAHE + Median + Gaussian + average smooth filter + ORB | 2400           | 34.85 |
| Kumar et al. [29]         | 2021 | VGG-19 with fine tuned                          | 2000           | 6.502 |
| Wang et al. [31]          | 2021 | SIFT + DERF + Coarse-to-fine matching           | 2000           | 3.20  |
| Proposed System           | 2021 | CLAHE + median filter + Unsharp Mask + ORB      | 2200           | 4.33  |

B) THE PROPOSED SYSTEM PERFORMANCE
We used a benchmark dataset (PUT) to test the proposed system efficiency in the recognition process. The PUT is a public dataset for wrist veins that contains 1200 images collected from 50 persons: 4 samples for each hand in 3 sessions. The dataset has a genuine score of 1100 (50 persons × 2 wrist hands × 11 samples) and an impostor score of 108900 (50 persons × 2 dorsal hands × 99 patterns × 11 samples). The terms FAR and FRR were calculated and plotted in Fig. 9. The EER obtained from testing the proposed system on PUT was 8.50%.

![PUT database](image)

FIGURE 9. Result of the testing proposed system on PUT

Table 2 illustrates that our proposed system outperforms the proposed system in [59] that was also tested on the PUT dataset, and the value of EER obtained was 34.13%. The result obtained in [59] due to more than one smoothing filter was applied: median, Gaussian, and average filters, which suppressed the image details which affected the feature extraction and matching process. Also, using BFM in feature matching with simple distance threshold whereas we use BFM with K-NN in matching that delivers better results in the matching process.

| System           | Database | Method                                      | EER % |
|------------------|----------|---------------------------------------------|-------|
| Garcia and Sanchez [26] | PUT     | CLAHE + Median + Gaussian + average smooth filter + ORB + BFM | 34.13 |
| Proposed System  | PUT      | CLAHE + median filter + Unsharp Mask + ORB + BFM with K-NN | 8.50  |

C) THE PROPOSED SYSTEM COMPUTATIONAL TIME
This work is based on smartphones due to the rapid growth in their specifications and computing capabilities, making them a good choice for processing many data and the ease of mobility and portability of these devices.

We developed a mobile application to implement the proposed system and measure its efficiency of real-time processing. We implement the proposed system with OpenCV library: a library used for computer vision solutions. Fig. 10 shows screenshots of the mobile application screens. The start screen in Fig. 10(a) contains two options: verification and user identification. The user can choose the verification option for the verification process by pressing the verification icon, and he will move to Fig. 10(b).
FIGURE 10. Mobile Application: a) start screen. b) verification screen. c) screenshot from real-time identification process. d) screenshot from the real-time verification process.

In Fig. 10(b), the user can enter his id and press verify button to move to Fig. 10(d), where the user places his hand in front of the USB-Camera, the hand is illuminated with infrared lights. The camera starts to capture frames, and the pre-processing is done, then features are extracted from the hand, and it matched with the features of the user stored in the database to be verified. A message will appear to the user whether the verification process has been successful or not. For the identification process, the user can choose the identification option from Fig. 10(a) by pressing the identification icon, and he will move to Fig. 10(c). The same processes are done in the identification option. The features extracted from the user in the identification option are compared against all features stored in the database. If the Identification process succeeds, a message will appear to the user with the user’s name and hand.

The mobile application has been tested on more than one smartphone with different specifications to obtain the efficiency of the proposed system. Table 3 illustrates the list of mid-range smartphones: Samsung A20 [43], Realme 5 pro [44], and Poco X3 pro [45] used in the test, with reasonable prices and good specifications. We can conclude from the table that the smartphones differ in terms of CPU, RAM, Storage, and the android version.

The proposed system is evaluated in terms of computational time. We calculate each operation separately on the smartphones mentioned in Table 2. Table 4 shows the results obtained from comparing two samples from our dataset after applying the proposed system. The table shows that the proposed system is characterized by a low computational time while comparing it with work in [28]. They used two smartphones with the exact specifications Xiaomi© Pocophone F1 and Xiaomi© Mi 8 (CPU Octa-Core 2.8 GHz, RAM 6GB, and 64 GB of storage). The total time obtained from applying their proposed system (CLAHE, median, Gaussian, average filter, and ORB) for Pocophone F1 was (52 ms) and for Mi 8 was (46 ms).

The frame rate per second of the proposed system is shown in Table 5. The table shows the proposed system has a high frame rate for the real-time process that outperforms the proposed algorithm in [28]. They achieved 13-15 Fps in the verification phase and 11 Fps in the identification phase using the ORB algorithms on Pocophone F1 and Mi 8 from Xiaomi. We achieved 27 Fps in the verification phase and 12 Fps in the identification phase. The verification process has a high frame rate because the matching is done 1:1. The matching is done between a user against all users stored in the database in the identification process.

| Specifications | Samsung A20 | Realme 5 pro | Poco X3 pro |
|----------------|-------------|--------------|-------------|
| CPU            | Samsung Exynos (octa-cores 1.56 GHz) | Qualcomm Snapdragon 712 (Octa-cores 2.30 GHz) | Qualcomm Snapdragon 860 (Octa-cores 2.96 GHz) |
| GPU            | Mali-G71 MP2 | Adreno 616  | Adreno 618  |
| RAM            | 3 GB        | 8 GB         | 6 GB         |
| Storage        | 32 GB       | 128 GB       | 128 GB       |
| Software       | Android 9   | Android 10   | Android 11   |

| Proposed algorithm | Samsung A20 (ms) | Realme 5 Pro (ms) | Poco X3 Pro (ms) |
|--------------------|------------------|-------------------|------------------|
| Pre-processing     | 21               | 14                | 9                |
| Feature Extraction | 17               | 8                 | 7                |
| Feature matching   | 4                | 3                 | 1                |
| Total              | 42               | 25                | 17               |

The frame rate per second of the proposed system is shown in Table 5. The table shows the proposed system has a high frame rate for the real-time process that outperforms the proposed algorithm in [28]. They achieved 13-15 Fps in the verification phase and 11 Fps in the identification phase using the ORB algorithms on Pocophone F1 and Mi 8 from Xiaomi. We achieved 27 Fps in the verification phase and 12 Fps in the identification phase. The verification process has a high frame rate because the matching is done 1:1. The matching is done between a user against all users stored in the database in the identification process.
TABLE 5. Frame Rate of verification and Identification process (Fps)

|          | Samsung A20 | Realme 5 pro | Poco X3 pro |
|----------|-------------|--------------|-------------|
| Verification | 22          | 27           | 29          |
| Identification | 8           | 12           | 22          |

V. CONCLUSION
In this paper, we proposed a system that works in real-time for biometric recognition of the dorsal hand veins based on a smartphone, which can be used in online payments or smartphone unlocking to increase security. We use contactless hardware that consists of a USB-Camera, infrared LEDs, and a smartphone to collect the dataset. We used this hardware to collect 2200 images of both hands from 100 people. We proposed a system for the biometric recognition process that consists of three main parts: Pre-processing, Feature extraction, and Feature matching. In the pre-processing, a CLAHE, median filter, and the unsharp mask are used to reveal the details of the image, especially the veins. ORB algorithm is used to detect and extract the features from images to store them in the database. BFM with K-NN is used for the matching process. The proposed system was tested and evaluated using the collected dataset to determine the efficiency of the system. The EER value obtained from the results is 4.33%. Also, we test the proposed system on the PUT dataset, and the obtained EER value was 8.50%. On the other hand, the Android application was developed and tested on more than one smartphone to determine the computational time efficiency. The results showed that the proposed system is suitable for working in real-time at a frame rate of 29 Fps. These promising results indicate that we can depend on vein recognition to identify people also, we can integrate it with smartphones to replace traditional methods as vein patterns cannot be forged.

We will increase our dataset and use deep learning techniques to increase recognition efficiency in future work.

REFERENCES

[1] K. Shaheed et al., A Systematic Review on Physiological-Based Biometric Recognition Systems: Current and Future Trends. Springer Netherlands, 2021.
[2] A. Uhl, “State of the Art in Vascular Biometrics,” in Stereotactic and Functional Neurosurgery, vol. 88, no. 4, A. Uhl, C. Busch, S. Marcel, and R. Veldhuis, Eds. Cham: Springer International Publishing, 2020, pp. 3–61.
[3] S. Jose and A. A. Bini, “Towards Building a Better Biometric System Based on Vein Patterns in Human Beings,” Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019, pp. 1–4, 2019.
[4] O. Toygar, F. O. Babalola, and Y. Bitirim, “FYO: A novel multimodal vein database with palmar, dorsal and wrist biometrics,” IEEE Access, vol. 8, pp. 82461–82470, 2020.
[5] M. A. Ahmed, E.-S. M. El-Horbaty, and A.-B. M. Salem, “Intelligent Techniques for Matching Palm Vein Images,” Egyptian Computer Science Journal, vol. 39, no. 1, pp. 1–14, 2015.
[6] D. Zhang, Z. Guo, and Y. Gong, “Multispectral Biometrics Systems,” in Multispectral Biometrics, Cham: Springer International Publishing, 2016, pp. 23–35.
[7] M. E. C. Villarina and N. B. Linsangan, “Palm vein recognition system using directional coding and back-propagation neural network,” Lecture Notes in Engineering and Computer Science, vol. 2420, pp. 796–801, 2015.
[8] R. G. Jakubietz, M. G. Jakubietz, D. Kloss, and J. G. Gruenert, “Defining the Basic Aesthetics of the Hand,” Aesthetic Plastic Surgery, vol. 29, no. 6, pp. 546–551, Dec. 2005.
[9] R. Zhang, D. Huang, Y. Wang, and Y. Wang, “Improving feature based dorsal hand vein recognition through Random Keypoint Generation and fine-grained matching,” Proceedings of 2015 International Conference on Biometrics, ICB 2015, pp. 326–333, 2015.
[10] A. M. Badawi, “Hand vein biometric verification prototype: A testing performance and patterns similarity,” Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPC’06, vol. 1, pp. 3–9, 2006.
[11] M. Watanabe and T. Endoh, “Palm vein authentication technology and its applications,” Proceedings of the ..., pp. 4–5, 2005.
[12] M. Himaga and K. Kou, “Finger vein authentication technology and financial applications,” Advances in Biometrics: Sensors, Algorithms and Systems, pp. 89–105, 2008.
[13] “LG G8S ThinQ | LG UK.” https://www.lg.com/uk/mobile-phones/lg-LMG810EAW (accessed Jun. 18, 2021).
[14] M. Taha, M. Ibrahim, and H. Zayed, “Digital Vein Mapping Using Augmented Reality,” International Journal of Intelligent Engineering and Systems, vol. 13, no. 6, pp. 512–521, Dec. 2020.
[15] R. Kabaciński and M. Kowalski, “Vein pattern database and benchmark results,” Electronics Letters, vol. 47, no. 20, p. 1127, 2011.
[16] R. Raghavendra, J. Subiriyala, and C. Busch, “Hand dorsal vein recognition: Sensor, algorithms and evaluation,” IST 2015 - 2015 IEEE International Conference on Imaging Systems and Techniques, Proceedings, no. 3, pp. 3–8, 2015.
[17] L. Wang, C. Zhang, Z. Liu, S. Bin, and H. Tian, “Image feature detection based on phase congruency by Monogenic filters,” 26th Chinese Control and Decision Conference, CCDC 2014, pp. 2033–2038, 2014.
[18] I. Patil, S. Bhilare, and V. Kanhangad, “Assessing vulnerability of dorsal hand-vein verification system to spoofing attacks using smartphone camera,” ISBA 2016 - IEEE International Conference on Identity, Security and Behavior Analysis, 2016.
[19] S. S. Agaian, M. Mulawka, R. Rajendran, S. P. Rao, K. M. Shreyas Kamath, and S. Rajeev, “A comparative study of image feature detection and matching algorithms for touchless fingerprint systems,” IS and T International Symposium on Electronic Imaging Science and Technology, no. i, pp. 1–9, 2016.
[20] L. DeBiasi, C. Kauba, B. Prommegger, and A. Uhl, “Near-infrared illumination add-on for mobile hand-vein acquisition,” 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems, BTAS 2018, pp. 1–9, 2018.
[21] “Google Nexus 5 NIR Enabled – Eigen Imaging.” https://www.eigenimaging.com/collections/frontpage/products/g oogle-nexus-5-®-nir-enabled (accessed Jun. 18, 2021).
[22] A. Ajay and P. K. Singh, “Novel Digital Image Water Marking Technique Against Geometric Attacks,” International Journal of Modern Education and Computer Science, vol. 7, no. 8, pp. 61–68, 2015.
[23] D. Zhong, H. Shao, and S. Liu, “Towards application of dorsal hand vein recognition under uncontrolled environment based on biometric graph matching,” IET Biometrics, vol. 8, no. 2, pp. 159–167, 2019.
[24] R. Kumar, R. C. Singh, and A. K. Sahoo, “SIFT based dorsal vein recognition system for cashless treatment through medical insurance,” International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10 Special Issue, pp. 444–451, 2019.
[25] X. Xie et al., “A study on fast SIFT image mosaic algorithm based on high-resolution images,” ISNN 2015 - IEEE 12th International Conference on Signal Processing, pp. 991–994, 2015.
on compressed sensing and wavelet transform,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 6, no. 6, pp. 835–843, 2015.

[26] R. Garcia-Martin and R. Sanchez-Reillo, “Wrist vascular biometric recognition using a portable contactless system,” *Sensors (Switzerland)*, vol. 20, no. 5, 2020.

[27] R. Garcia-Martin, R. Sanchez-Reillo, and J. Enrique Suarez-Pascual, “Wrist vascular biometric capture using a portable contactless system,” in *Proceedings - International Carnahan Conference on Security Technology*, Oct. 2019, vol. 2019-October.

[28] R. Garcia-Martin and R. Sanchez-Reillo, “Vein Biometric Recognition on a Smartphone,” *IEEE Access*, vol. 8, pp. 104801-104813, 2020.

[29] R. Kumar, R. C. Singh, and S. Kant, “Dorsal Hand Vein-Biometric Recognition Using Convolution Neural Network,” vol. 1165, no. November, D. Gupta, A. Khanna, S. Bhattacharyya, A. E. Hassamien, S. Anand, and A. Jaiswal, Eds. Singapore: Springer Singapore, 2021, pp. 1087–1107.

[30] M. Jogi, Mohana, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva, “Feature extraction using convolution neural networks (CNN) and deep learning,” 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTETCT 2018 - Proceedings, pp. 2319–2323, 2018.

[31] Y. Wang, X. Cao, and X. Miao, “Cross-device recognition of dorsal hand vein images by two-stage coarse-to-fine matching,” *The Visual Computer*, Jun. 2021.

[32] K. Zuiderveld, *Contrast Limited Adaptive Histogram Equalization*. Academic Press, Inc., 1994.

[33] A. M. Reza, “Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement,” *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol. 38, no. 1, pp. 35–44, 2004.

[34] S. Khan and D. Lee, “An adaptive dynamically weighted median filter for impulse noise removal,” *EURASIP Journal on Advances in Signal Processing*, vol. 2017, no. 1, p. 67, Dec. 2017.

[35] E. Rueblee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” in *2011 International Conference on Computer Vision*, Nov. 2011, pp. 2564–2571.

[36] Z. Wang, Z. Li, L. Cheng, and G. Yan, “An improved ORB feature extraction and matching algorithm based on affine transformation,” in *Proceedings - 2020 Chinese Automation Congress, CAC 2020*, Nov. 2020, pp. 1511–1515.

[37] E. Rosten and T. Drummond, “Machine learning for high-speed corner detection,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 3951 LNCS, 2006, pp. 430–443.

[38] C. Harris and M. Stephens, “A Combined Corner and Edge Detector,” in *Proceedings of the Alvey Vision Conference 1988*, 1988, pp. 23.1-23.6.

[39] P. L. Rosin, “Measuring Corner Properties,” *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 291–307, Feb. 1997.

[40] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, “BRIEF: Binary Robust Independent Elementary Features,” 2010, pp. 778–792.

[41] N. B. Linsangan, R. R. Estacio, and J. L. Torres, “Comparing local invariant algorithms for dorsal hand vein recognition system,” *ACM International Conference Proceeding Series*, pp. 27–31, 2019.

[42] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[43] “Samsung Galaxy A20 - Specs and Features | Samsung LEVANT.” https://www.samsung.com/levant/smartphones/galaxy-a-series/a20/specifications (accessed Oct. 1, 2021).

[44] “realme 5 Pro Specs - realme (United Kingdom).” https://www.realme.com/uk/realme-5-pro/specs (accessed Oct. 1, 2021).

[45] “POCO X3 Pro.” https://www.mi.com/uk/poco-x3-pro/specs (accessed Oct. 1, 2021).