Article

Enhanced Multimodal Biometric Recognition Based upon Intrinsic Hand Biometrics

Syed Aqeel Haider [1],*, Yawar Rehman [2] and **S. M. Usman Ali [2]**

[1] Department of Computer & Information Systems Engineering, Faculty of Electrical & Computer Engineering, NEDUET, Karachi 75270, Pakistan
[2] Department of Electronic Engineering, Faculty of Electrical & Computer Engineering, NEDUET, Karachi 75270, Pakistan; yawar@neduet.edu.pk (Y.R.); uashah68@neduet.edu.pk (S.M.U.A.)
* Correspondence: hsaqeel@neduet.edu.pk

check for
updates

Abstract: In the proposed study, we examined a multimodal biometric system having the utmost capability against spoof attacks. An enhanced anti-spoof capability is successfully demonstrated by choosing hand-related intrinsic modalities. In the proposed system, pulse response, hand geometry, and finger–vein biometrics are the three modalities of focus. The three modalities are combined using a fuzzy rule-based system that provides an accuracy of 92% on near-infrared (NIR) images. Besides that, we propose a new NIR hand images dataset containing a total of 111,000 images. In this research, hand geometry is treated as an intrinsic biometric modality by employing near-infrared imaging for human hands to locate the interphalangeal joints of human fingers. The L2 norm is calculated using the centroid of four pixel clusters obtained from the finger joint locations. This method produced an accuracy of 86% on the new NIR image dataset. We also propose finger–vein biometric identification using convolutional neural networks (CNNs). The CNN provided 90% accuracy on the new NIR image dataset. Moreover, we propose a robust system known as the pulse response biometric against spoof attacks involving fake or artificial human hands. The pulse response system identifies a live human body by applying a specific frequency pulse on the human hand. About 99% of the frequency response samples obtained from the human and non-human subjects were correctly classified by the pulse response biometric. Finally, we propose to combine all three modalities using the fuzzy inference system on the confidence score level, yielding 92% accuracy on the new near-infrared hand images dataset.

Keywords: convolutional neural network; intrinsic modalities; multimodal biometric system; the fuzzy inference system

1. Introduction

Automated verification and authentication is a widely addressed issue across the globe nowadays. In this regard, human physical and behavioral characteristics are under research. Fingerprints, irises, faces, and finger veins are some physical biometric traits. On the other side, human gait, keystrokes, and handwriting may be treated as behavioral biometric traits. In the proposed research, we have focused on physical biometric characteristics (e.g., hand geometry, finger–vein, and pulse response). Hand geometry has been used for identification in automated biometric systems. It is classified as a medium-level, dependable biometric modality. This means that it moderately satisfies all the characteristics which are required to qualify as a biometric modality [1]. Certain characteristics must be satisfied to some extent to be qualified as a biometric modality [2–4], which are as follows:

- The first characteristic is universality, meaning that the characteristic chosen as the biometric should be globally present in humans. Except for disabled people, features related to chosen modalities are present in all living human beings with no limitation of place.
- Next is collectability, meaning the biometric trait must have quantified, measurable value so that it may be captured from a human without any difficulty. The three modalities of focus fulfill this condition.
- Then there is permanence, meaning a biometric modality must remain invariable within a specified period so the systems based on that modality may be considered reliable for a specified time limit. This quality is also present in the considered modalities.
- Acceptability is another characteristic that must be satisfied. Acceptability means users should feel comfortable with the biometric capturing setup. It is not an easy task to assure that each user is satisfied, but the setup should be made easier and hassle-free. Our proposed biometric capturing setup is simple in that users have to place their hands on a pad for image and pulse response capturing.
- The fifth trait is liveliness, which means that the human character cannot be duplicated by using some non-living thing. In other words, the modality must only be gatherable from a living human body. Finger–vein and pulse response biometrics fulfill this quality perfectly. In the case of the hand geometry biometric, it is possible to form a hand from non-living material with the same features as a living human. This deficiency of the hand geometry biometric is dealt with by capturing images using a mounted camera with a near-infrared filter. In the captured hand image, the phalangeal joints will have brighter regions while the rest of the image will be darker. This way of treating the hand geometry biometric may be called the phalangeal biometric. Therefore, for the phalangeal biometric, it is not an easy task to falsely produce the same NIR images of the human hand by using non-living materials.
- Finally, it must be hard to counterfeit, which is another essential and desirable characteristic. This means that a modality must be less and less vulnerable to spoofing attacks. The finger–vein biometric has proven capable against spoof attacks. On the other hand, the hand geometry biometric is easier to break for a spoof in comparison with a phalangeal biometric-based system. The main reason behind this fact is that, for the phalangeal biometric, we are going to locate bone joints within the human hand by using NIR images, while in a simple hand geometry biometric, normal camera images are processed to extract external features related to the human hand. It also clarifies that the phalangeal joint biometric is intrinsic, while hand geometry is an extrinsic biometric modality.

2. Related Works

Biometric traits play an important role in the identification and classification of humans. For us humans, recognition and classification is a very simple job with the help of evolved and dedicated biological neural networks. However, due to the infancy of the machine learning age, these tasks are quite difficult for a computer to achieve flawlessly. Hence, throughout the computer vision literature, we find researchers trying to achieve better classification accuracy using different biometric traits. The face is one of the most popular biometric traits used for human classification. In [5], Q. Feng et al. used various classifiers, and their accuracies were compared using face biometric traits. Facial marks were also used by C. Zeinstra, R. Veldhuis, and L. Spreeuwers as a biometric trait in [6]. The authors used the Bayesian classifier, and better accuracy was reported at around twenty-five facial marks per face. This paper concluded that the accuracy of face classification is dependent on the number of facial marks. Hence, it implies that biometric traits play an important role in the identification of humans.

Gait recognition [7,8] was also used by S. M. Darwish, X. Wang, and S. Feng as a biometric trait for human identification and recognition. Fingerprint recognition [9] with the Support Vector Machine (SVM) classifier was used by M. Komeili, N. Armanfard, and D. Hatzinakos as another biometric trait. Surprisingly, in [10], E. Maiorana and P. Campisi processed electroencephalogram (EEG) signals

and used those as a biometric trait. Iris recognition was researched by J. Peng et al., N. Ahmadi, and G. Akbarizadeh [11,12] and considered a powerful trait in the classification of humans.

In [13], M. Chaa, Z. Akhtar, and A. Attia introduced a three-dimensional palmprint and used it for human classification. The authors of [14], S. Veluchamy, and L. R. Karlmarx, used finger knuckles and finger veins as the novel biometric traits. The authors reported that they achieved 96% classification accuracy using these traits.

The researchers in [15] worked on remote photoplethysmography for detecting spoof attacks using fake finger–vein data. In [16], a multimodal biometric system was proposed based on finger vein and finger shape biometric modalities using NIR images. An efficient technique for the enhancement of NIR finger–vein images was proposed in [17]. The authors of [18] presented a robust technique for region of interest localization in the finger–vein biometric system. One research article [19] discussed the implementation of deep learning techniques for the proposed multimodal biometric system combining the iris, face, and finger–vein biometrics. In [20], the researchers proposed feature level fusion of the finger–vein and fingerprint biometrics. They used an NIR imaging device for capturing images.

The pulse response biometric is a recently researched biometric. Rasmussen et al. [2] introduced this biometric. According to the authors, the pulse response biometric may be effectively used as it satisfies the aforementioned characteristics of biometric modality.

In [21], the researchers introduced a new methodology for finger vein authentication using a convolutional neural network and supervised discrete hashing. They claimed to investigate its performance using well-known CNN architectures in other domains (e.g., light CNN, VGG16, Siamese, and a CNN with Bayesian inference-based matching). They performed a comparative analysis between the proposed and existing methods and claimed unmatched performance by the proposed system. For comparative analysis, the researchers used a publicly available two-session finger vein database.

The authors of [22] presented a novel convolutional neural network-based finger–vein identification system. They tested the proposed CNN by using four publicly available databases of finger vein images. They claimed that their proposed system was almost independent of the quality of the images captured and analyzed. They also claimed to achieve an accuracy of identification beyond 95% for all of the four considered databases of images.

The researchers in [23] introduced a biometric recognition system, based on the dorsal hand vein, using a convolutional neural network. They tried CNNs of different depths to compare the recognition rate. They also analyzed the effect of the dataset size on the recognition rate. They reported extraction of the region of interest (ROI) from the image. Then, these ROI images were preprocessed by using contrast limited adaptive histogram equalization (CLAHE) and a Gaussian smoothing filter algorithm. Features were extracted by using Reference-CaffeNet, AlexNet, and VGG-depth CNN. In the last stage, they applied logistic regression for identification. They performed experiments on datasets of two different sizes and reported achieving different degree effects on the recognition rate. They claimed to observe a 99.7% recognition rate by VGG19 for the dorsal hand vein. They also reported a declined recognition rate of 99.52% for SqueezeNet.

The authors of [24] proposed a robust finger–vein recognition method. They employed different databases and considered environmental changes, based on the convolutional neural network. They claimed to maintain new finger vein databases during this research. They performed the experiments by using these databases and the openly available SDUMLA-HMT finger vein database. They claimed to observe better performance in comparison with the conventional systems.

In [25], the authors proposed a multimodal biometric system based upon three modalities. They used fingerprint, finger–vein, and face biometrics to develop an accurate and efficient identification system. They tested the performance by using the publicly available SDUMLA-HMT dataset. The researchers preprocessed images for fingerprints and finger veins in the first step using their respective algorithms. In the second step, the convolutional neural networks were used to extract features for all three modalities of focus separately. After that, the softmax classifier was used for fingerprints and faces while the random forest classifier was used for the finger–vein

biometric. Then, matching scores were generated for each of the three modalities. They declared the recognized subject after performing score-level fusion and comparing the overall matching score with a predefined threshold.

3. Proposed System

In our proposed work, we used the pulse response, hand geometry (in the form of the interphalangeal joint biometric), and finger–vein biometric modalities for the detection of a live human body to improve performance against counterfeit attempts. The block diagram of the proposed system is shown in Figure 1. Our paper consists of the following key contributions:

- A new near-infrared (NIR) hand images dataset is proposed, with a dataset containing a total of 111,000 images acquired from 185 humans. The dataset will be uploaded online and made freely available (under approval of the ethical committee).
- A novel hand geometry biometric system is implemented. The proposed system uses image processing, pixel clusters, and centroid distances for hand geometry recognition. The proposed system achieved 86% accuracy on the proposed NIR hand dataset.
- Finally, the fuzzy rule-based system combines all three biometrics for human recognition using hand images. The proposed method combines the pulse response, hand geometry, and finger–vein biometrics using fuzzy rule-based inference and gives a cumulative confidence score. The proposed method achieved 92% accuracy on the proposed dataset.
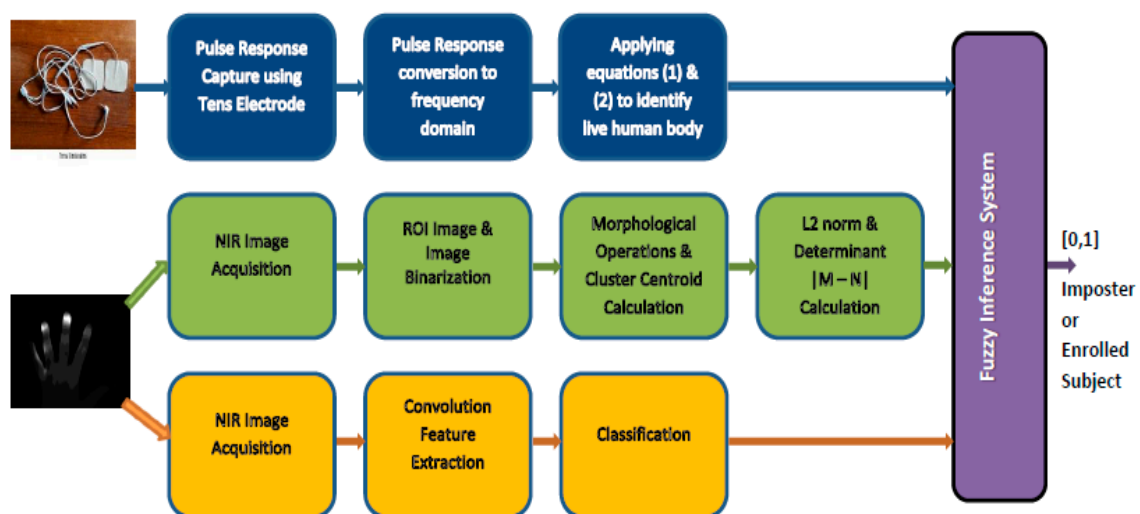
Figure 1. Block diagram of the proposed method. Blue = pulse response; green = hand geometry; and yellow = finger–vein.

The steps involved for each one of the selected biometric modalities are described in the following discussion.

3.1. Pulse Response Biometric

In our proposed work, we developed a device for applying a train of pulses to the human palm and capturing the response from the other hand. The authors in [2] used brass electrodes for applying and capturing signals, while we have used more sophisticated and reliable Tens electrodes for applying and capturing signals. We also tried brass electrodes, but those were unable to capture desired signals better than the Tens electrodes.

The biometric signal capturing device was made with the help of INA128P (High Precision, Low Power Instrumentation Amplifier), TL071CN (Low Noise, JFET Input Operational Amplifier), and

7660S (Super Voltage Converter). The square wave pulses of a 200Hz frequency were applied through the Tens electrode on the palm of the first hand, and the response was captured for 5 ms from the other hand. A sampling rate of 10 kHz was used. The device and the Tens electrode pairs are shown in Figure 1.

The captured signal was fed into a computer using the microphone audio interface. The signal was read and processed further by using MATLAB. MATLAB recorded the input signal at a sampling rate of 10 kHz for 5 ms and stored it in a 50×1 matrix. The frequency of the applied pulse was 200 Hz because the response was captured for 5 ms, such that for a single iteration of the applied pulse, the response was captured.

In the next step, a fast Fourier transform (FFT) was applied to the acquired signal. The FFT transformed the time domain signal into the frequency domain. The frequency domain response was stored in a database maintained in Microsoft Excel. For an applied pulse, we got a column of 50 cell entries. A total of 50 iterations were performed for each subject to record his or her response signals, which resulted in a 50×50 matrix.

When we analyzed a response from any human, it could easily be observed that the 26th cell entry of the column was zero for almost all of the columns. In addition to that, we observed that 25th and 27th cell values were the same, as well as the 24th, 23rd, 22nd, and so on until the 2nd cell value were the same as the 28th, 29th, 30th, and so on until the 50th cell value, respectively. However, when we captured the pulse response from any non-living thing, the responses were all zero response patterns. We tried and validated our pulse response biometric state by capturing the response of wood, metal, and plastic surfaces.

Therefore, for identifying that the response was captured from a living human, we performed cell-to-cell subtraction as shown in Equation (1) below:

$$\sum_{j=2}^{25} (\text{CELL}(j) - \text{CELL}(52 - j)) \ = \ 0 \tag{1}$$

In addition to this, it was also required to check that all the entries in all 50 columns were not zeros, as shown in Equation (2) below:

$$\sum_{i=1}^{50} \text{CELL}(i) \ \neq \ 0 \tag{2}$$

These equations were used to differentiate between a living human response and an artifice body response.

In Equations (1) and (2), CELL denotes the Excel sheet cell entry for the captured response.

3.2. Hand Geometry as an Intrinsic Biometric

Our approach to deal with hand geometry was to further enhance the anti-spoofing capability of the proposed system. We treated hand geometry, which is an extrinsic modality, as an intrinsic modality.

Hand geometry is a proven biometric modality, and many researchers have worked on it in unimodal as well as multimodal biometric systems. In [15–18], J. Svoboda, O. Klubal, and M. Drahansky [26], T. A. Budi Wirayuda et al. [27], J. Svoboda, M. M. Bronstein, and M. Drahansky [28], and R. Srikantaswamy [29] treated hand geometry as an extrinsic modality. That is why all the disadvantages related to extrinsic physical biometric modalities apply to their work. The main disadvantage is the vulnerability to spoof attacks. Some false user may replicate another user's hand geometry by using some fake hand made of artificial material.

In this research work, we considered the index, middle, and ring fingers for gathering raw data, using a near-infrared (NIR) camera, in the form of images. Once an image was captured using an NIR camera, it was pre-processed and went through a few math functions for the generation of identification and authentication results.

Our proposed algorithm comprises the stages as shown in Figure 2. Each step is explained in the following discussion.
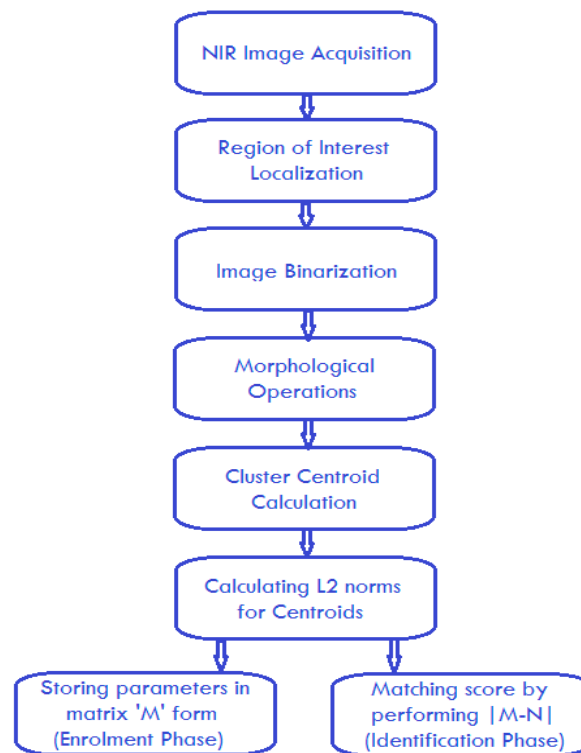
Figure 2. Hand geometry algorithm flow diagram.

3.2.1. Image Acquisition

Allied Vision's Manta G145B—GigE—Camera was used for capturing NIR images. It is a monochrome camera, having a 1.4 megapixel resolution and giving an image of 1360×1024 pixels. A near-infrared (NIR) filter was attached to this camera for capturing NIR images. The camera was mounted on a specially designed metal stand, facing downward. A lighting source of a constant intensity was placed at the bottom of the stand. We designed a lighting pad with high-intensity LEDs placed beneath only the human finger joints. The focus of the camera and the distance between the camera and the lighting source was adjusted at the start of the image acquisition phase and remained constant for all the images taken from volunteers. Under these specially designed lighting conditions, NIR images for the right hand were captured and stored in the database for further processing. Figures 3a and 4a show the captured images.

3.2.2. Region of Interest Localization

As reported in Section 3.2.1., we used a metal stand to mount the camera. The design of the image capturing setup was such that there was no need to crop or align the captured image. Therefore, in this step, a predefined pixel region of the captured image was extracted. Our pixel region of interest (ROI) was from row 230 to 700 and from column 300 to 990. In this way, our ROI image covered the index, middle, and ring fingers of the volunteer. The example ROI images are shown in Figures 3b and 4b.

3.2.3. Image Binarization

In this step, the output image of the previous step was converted into a black and white image. After iterative preprocessing of the images, for getting the best results, an intensity threshold value of 100 was selected for the applied lighting conditions for image binarization. If some different intensity lighting source were to be used, then the threshold would also need to be adjusted accordingly.

Figures 3c and 4c show the threshold output image of this step. It is visible in the images that, at the location of the interphalangeal joint, there are brighter regions, whereas other portions are darker.
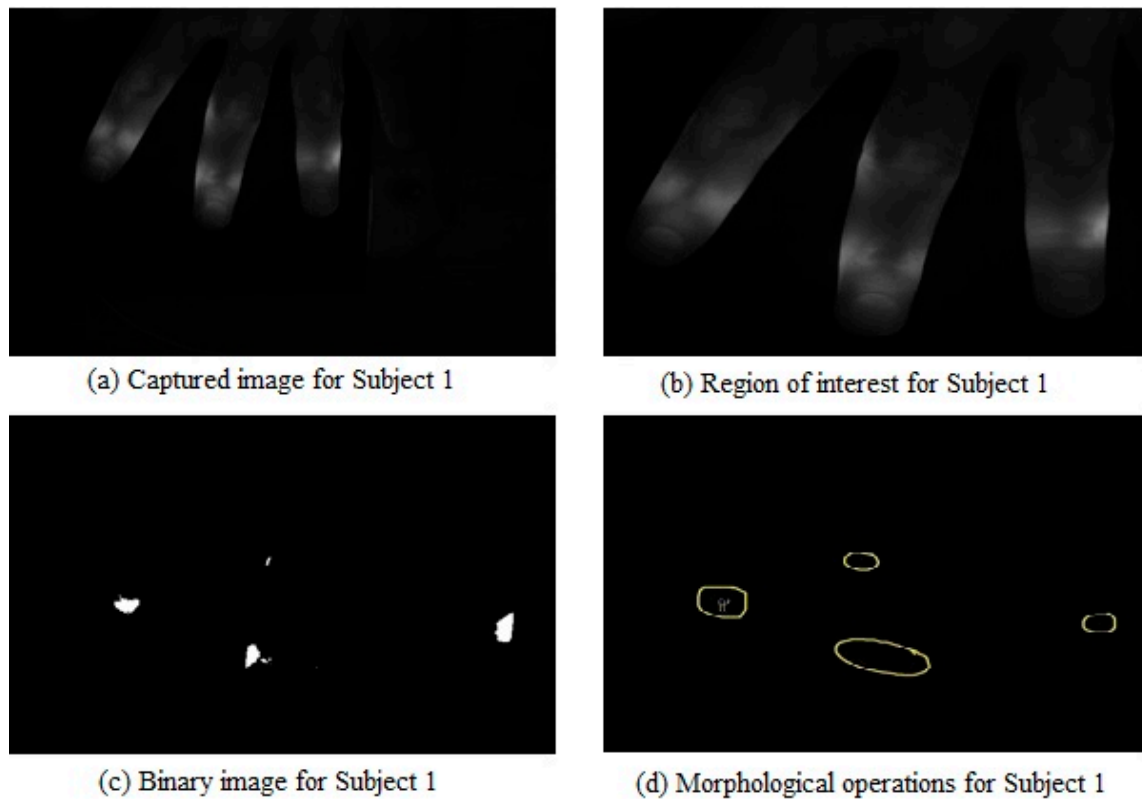

(a) Captured image for Subject 1

(b) Region of interest for Subject 1

(c) Binary image for Subject 1

(d) Morphological operations for Subject 1

Figure 3. Images for Subject 1.

(a) Captured image for Subject 2

(b) Region of interest for Subject 2

(c) Binary image for Subject 2

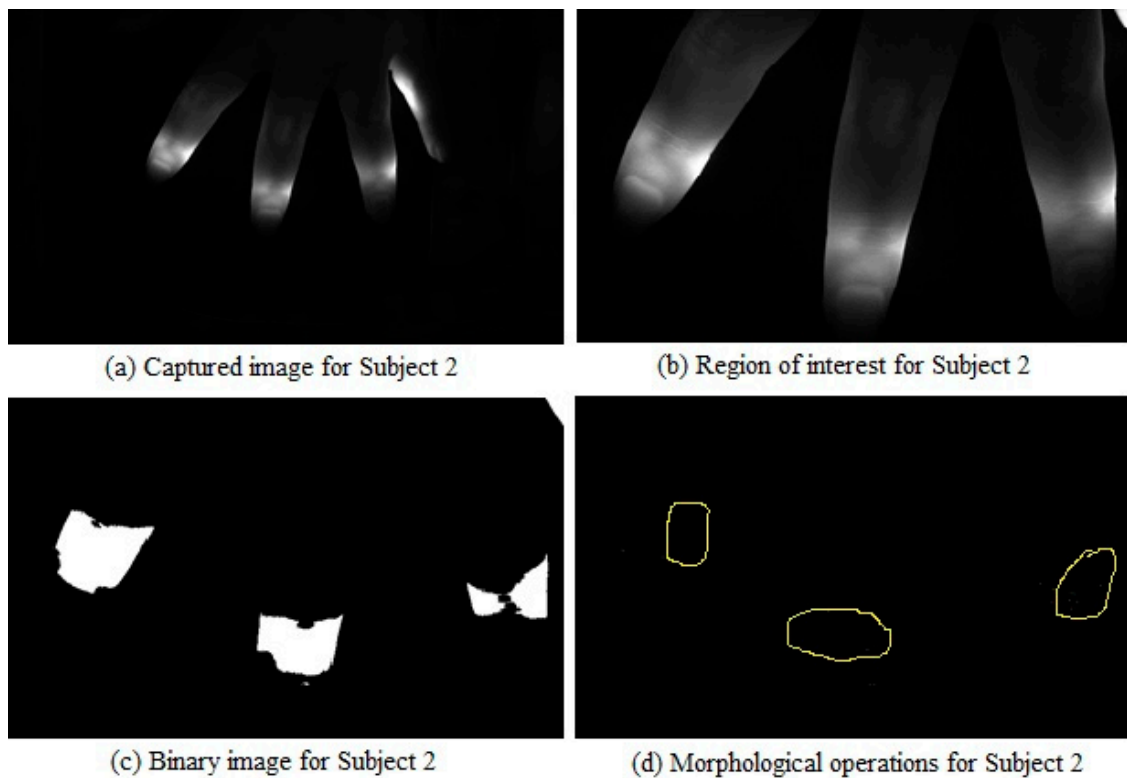(d) Morphological operations for Subject 2

Figure 4. Images for Subject 2.

3.2.4. Morphological Operations

By using morphological operations, the brighter regions were converted into a bright spot. For this, two morphological operations, called fill and shrink, were applied.

The morphological fill operation removed any darker spot surrounded by a brighter region in the threshold image. This fill operation limited the number of scattered brighter regions. As a result, the unnecessary smaller, brighter regions were merged to form a bigger, brighter region.

The morphological shrink operation converted the brighter region into a bright spot. Therefore, by applying the two selected morphological operations, we got an image with bright spots representing each brighter region. The output images for this step are shown in Figures 3d and 4d.

3.2.5. Cluster's Centroids Calculation

It may be seen clearly that, after performing the morphological operations, there was a cluster of white dots for each brighter region. Now, we have to calculate the distances between the brighter regions. Therefore, we needed to locate the centroid for each brighter region. To achieve this task, we grouped the neighboring white dots and located the single point to get a single white dot instead of a cluster.

For the formation of the clusters, we started with locating the white dot (pixel position) having the lowest row and column location. Then, we located the white dots situated within +X rows and +Y columns to form a group of pixels or clusters. We worked with values of X and Y of 140, 100, 70, and 40. The centroid of the located cluster was then calculated by adding the row pixel positions together and then dividing it by the total number of pixels within the cluster. This gave the pixel row position for the centroid of the cluster. Besides this, the column pixel positions were added together and then divided by the total number of pixels within the cluster. This gave the pixel column position for the centroid of the cluster.

The very next white dot located outside the cluster was considered as the starting point of the next cluster. In this way, we located four clusters and calculated their centroids.

3.2.6. Calculating Parameters

In this final step, the Euclidean distances between the located four centroids of bright regions were calculated. Matrix M, shown in Equation (3), was formed by arranging our data in such a way that it may be stored in an organized database. Later on, this matrix is used for generating matching score results by performing subtraction with Matrix N (see Section 4.3), followed by taking the determinant for the resultant Matrix S. Matrix M had six entries for inter-centroid distances, as listed below:

$$M = \begin{bmatrix} d1 & d2 & d3 \\ d4 & d5 & d6 \\ 2 & 2 & 2 \end{bmatrix} \tag{3}$$

d1 = Distance between Centroids 1 and 2
d2 = Distance between Centroids 1 and 3
d3 = Distance between Centroids 1 and 4
d4 = Distance between Centroids 2 and 3
d5 = Distance between Centroids 2 and 4
d6 = Distance between Centroids 3 and 4

3.2.7. Storing and Matching Parameters

The matrix containing the calculated parameters (i.e., Matrix M) was stored in the database for each user during the enrollment phase. In the matching and authentication phase, the matrix calculated on runtime (i.e., Matrix N) was matched with the matrices stored in the database by performing element-to-element subtraction. The output matrix (i.e., Matrix S) was converted into a

determinant value. The subject with the lowest determinant value was picked as the most probable authenticated person.

3.3. Finger–Vein Biometric

We proposed to learn the finger–vein biometric that was introduced in [14]. However, the authors in [14] used finger vein information after filtering those images by several handcrafted filters. Eventually, they learned the features using the K-SVM classifier. We, on the other hand, proposed to learn the finger vein information directly from near-infrared (NIR) hand images using a convolutional neural network (CNN). During the experiments, we found out that the CNN also had learned hand phalangeal joint biometric (PJB) information. This biometric trait captures information based on the distance between the distal phalange joint of the middle finger and the proximal phalangeal joints of the index and little fingers. These distances can be translated into a relationship, which may provide distinct characteristics or features about humans. In Section 3.2, we proposed a handcrafted technique (i.e., hand geometry as an intrinsic biometric). It finds the distance between finger joints by using NIR images to calculate cluster distance relationships.

3.3.1. Methodology

In this section, we will discuss the data set collection, augmentation, and CNN architecture used for NIR image training.

3.3.2. Dataset

The dataset was collected from a total of 185 subjects. Two hundred images were collected from each subject. Once the images were collected, the data was augmented using random values of translation, angular rotation, size, and horizontal flipping of the images. These augmentations provided us with four hundred more images per subject, in addition to the two hundred collected images per subject. Hence, the total number of images per subject reached six hundred. Samples from the dataset are shown in Figure 5. Our proposed NIR dataset consisted of 185 subjects, 600 images per subject, and a total of 111,000 images.
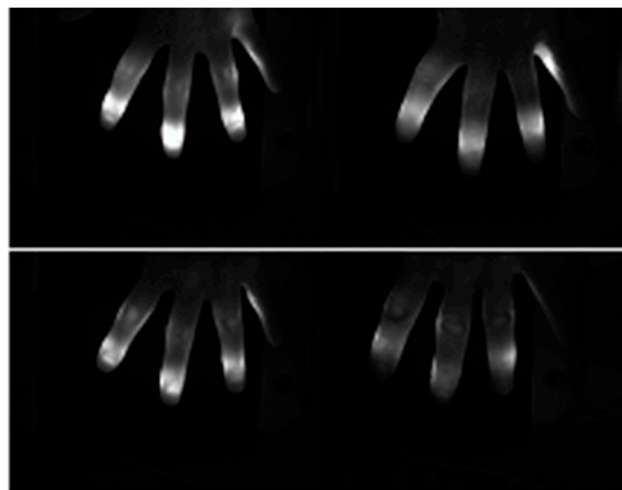

Figure 5. Four sample subjects from the near-infrared (NIR) dataset.

3.3.3. Convolutional Neural Network Architecture

We used the framework of Alexnet [30], which consisted of eight layers. The first five were the convolutional layers. In these layers, various filters were convolved with NIR images to extract distinct features from all the classes. These filters extracted rich information from the images, such as edges, curves, and colors. In NIR images, vein information is represented by the darker regions

inside the fingers, whereas the non-vein region is represented by the brighter regions inside the fingers. Once extraction was completed, those discriminative features were forwarded to the remaining three fully connected layers, where weights were converged using a loss function in several iterations.

By default, the architecture of the convolutional neural network (CNN) used two graphics processing units (GPUs) to speed up the processing. In our experiments, we used a single GPU; hence, all the operations were performed inside the sole GPU. We randomly split the data into 80% training data and 20% validation data. It took about 10 h on an INTEL core i5 with 8GB RAM and an NVIDIA 1050 GPU to train 185 classes containing 80% of the total images. Each image was of a $224 \times 224 \times 3$ pixel size, as per the CNN architecture requirement. The training was completed using four epochs, having a batch size of sixty-four images. We used a constant learning rate of 0.001 throughout the training. The classification accuracy was evaluated on the validation set containing 20% of the total data.

3.4. Fuzzy Logic System

We designed a fuzzy logic system to combine the outputs of all the biometric modalities that were used in our proposed system (i.e., pulse response, hand geometry, and finger–vein biometrics). The fuzzy system will render a confidence value on a scale of 0–1. The values near 0 represent that the person is an imposter, and values near 1 represent that the person belongs to the predicted class by the hand geometry and finger–vein systems. This combination of all the biometric traits is shown in Figure 6 with the help of a block diagram. Figure 6 shows that the fuzzy inference system takes input values from all three modalities and gives a confidence value in the range of 0–1.
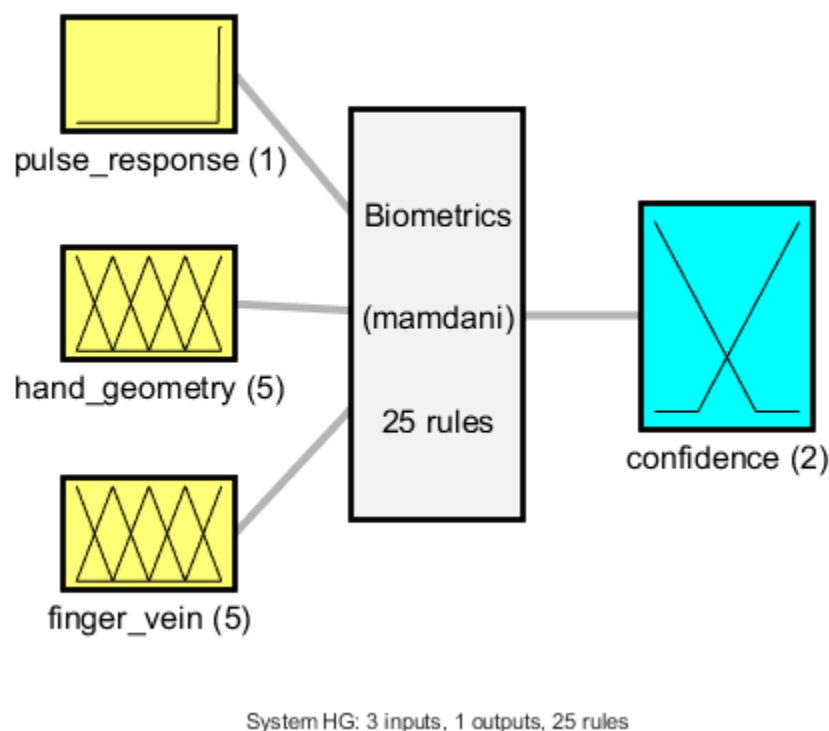

System HG: 3 inputs, 1 outputs, 25 rules

Figure 6. Fuzzy biometric fusion system.

There were three steps to making the fuzzy logic system. First, we converted real values into fuzzy linguistic variables (e.g., 20 km/h to slow speed, 100 km/h to high speed, and 200 km/h to very high speed). Next, we designed a rule set that conformed to our needs. Finally, the answer calculated by fuzzy logic had to be de-fuzzified to make it understandable in the real world. The following section will elaborate on the procedure of loading all the modalities into the fuzzy logic system to derive a cumulative answer.

3.4.1. Fuzzification of Pulse Response

As we mentioned earlier, the pulse response biometric only identifies whether the hand is of a living person or not. For this purpose, Equations (1) and (2) need to be true for a living person and false for a non-living person. Figure 7a shows the fuzzification of the pulse response trait. Although pulse response will either be true or false, we assigned all true values to be above 0.98. Once it was certain that a person was living, the system would go further to evaluate other modalities.
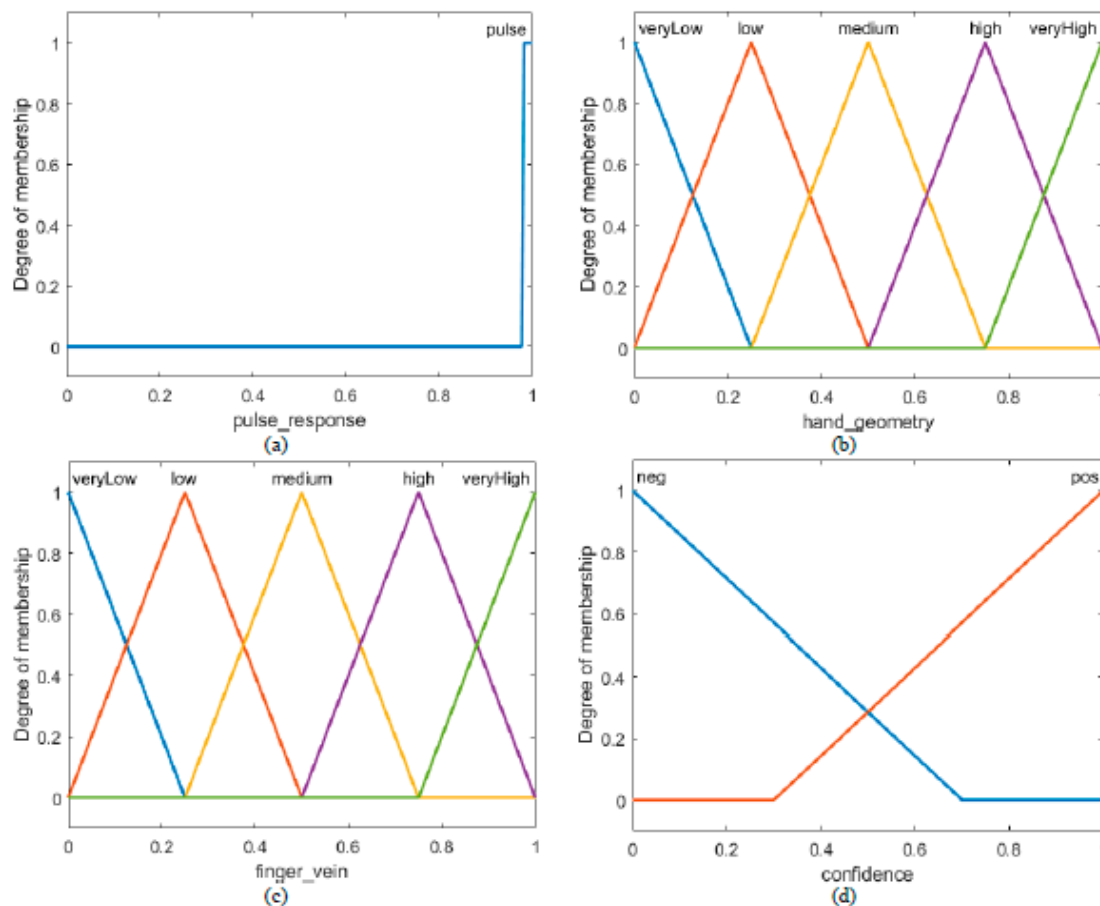
Figure 7. Fuzzification of inputs and outputs: (**a**) Fuzzification of Pulse Response System; (**b**) Fuzzification of Hand Geometry System; (**c**) Fuzzification of Finger-Vein System; (**d**) Fuzzification of the Output.

3.4.2. Fuzzification of Hand Geometry

As mentioned in Section 3.2, cluster distances were calculated and stored in Matrix M, represented by Equation (3) in the enrollment phase. Then, during the identification phase (see Section 4.2), Matrix N was arranged for the candidate subject, and hence the determinant of Matrix S (see Equation (5)) was calculated. The value of the determinant gave us information about the NIR image belonging to one of the various classes. We fuzzified the hand geometry output by normalizing the determinant values in the range of 0–1. A total of five linguistic variables were assigned in the said range (i.e., very low, low, medium, high, and very high). Those variables were all assigned triangular membership functions, as shown in Figure 7b.

3.4.3. Fuzzification of Finger Vein

As mentioned in Section 3.3, the convolutional neural network (CNN) provided an output in the form of a confidence score for each subject class. The score was in the range of 0–1. Similar to hand

geometry fuzzification, a total of five linguistic variables were assigned in the said range (i.e., very low, low, medium, high, and very high). All of those variables were assigned triangular membership functions, as shown in Figure 7c.

3.4.4. Fuzzification of Output (Confidence Value)

The output of the fuzzy system would be a confidence value in the range 0–1. We defined two triangular membership functions (i.e., pos and neg), as shown in Figure 7d. It should be noted that the fuzzy membership value (on the Υ-axis) of neg decreased when the confidence increased from 0 to 1 and vice versa.

3.4.5. Designing a Fuzzy Inference System

Once all inputs (biometric modalities) were fuzzified, fuzzy rule-based inference was designed. A total of twenty-five rules were incorporated in the fuzzy inference system. Rules were designed and tweaked to give an advantage to the output of the finger–vein CNN-based system. For example, if the hand geometry system identified an input NIR image with the membership function of medium to a different class other than the finger–vein CNN-based system with the membership function of medium, the decision of the latter would be considered. In another example, if the hand geometry system identified an input NIR image with the membership function of very high to a different class other than the finger–vein CNN-based system with the membership function of very low, only then would the decision of the former be considered.

3.4.6. De-Fuzzification

The centroid method was used to de-fuzzify the answer obtained from the fuzzy rule-based inference system. The de-fuzzified value would be in the range of 0–1.

4. Experimental Setup

Our proposed dataset was formed by capturing NIR hand images and pulse response information from 185 volunteers. We captured these in two sessions, separated by a four-week interval. In the first session, one hundred NIR hand images and twenty-five pulse response instances per subject were captured. In the second session, one hundred more NIR hand images and twenty-five pulse response instances from the same volunteers were captured. Those two hundred images were augmented using translation, angular rotation, size, and horizontal flipping of the images. This yielded a total of six hundred NIR hand images per subject and a total of 111,000 images. The dataset was split into 80% training and 20% testing sets.

4.1. Pulse Response Biometric Setup

In the pulse response setup, we captured the pulse response biometric from each subject. Those volunteers were called, and their pulse response biometrics were measured and stored. A sample of a captured pulse response is shown in Figure 8. The response shown was captured at a sampling rate of 5000 samples per second. Hence, there are 25 entries per column in the response shown in Figure 8, and only nine columns are shown. It is easily visible that the second entry within each column is the same as 25th and 3rd, 4th, 5th, and so on, until the 13th entries, which are the same as the 24th, 23rd, 22nd, and so on until the 14th entries. The response for 10,000 samples per second is not shown here because of the limitation of space. We found that every living body had the property mentioned in Section 3.1. It was also verified that non-living bodies all had zero responses when the same type of pulse was applied.

3.835938	-3.65625	5.101563	4.0625	-4.03125	-3.53125	4.515625	-1.30469	0.46875
-0.23735	0.241795	-0.40916	-0.28911	0.296245	0.232269	-0.43321	0.219969	-0.18449
0.04328	-0.02174	-0.13245	-0.17939	0.221732	0.04444	-0.04662	0.158052	-0.12855
0.048243	0.009012	-0.07539	-0.08753	0.133849	0.030467	-0.10013	0.117028	-0.10148
0.012862	-0.02619	-0.08585	-0.07873	0.078352	0.03729	-0.06104	0.085359	-0.10471
0.039415	-0.02138	-0.08835	-0.09226	0.084445	-0.00483	-0.05135	0.082031	-0.10099
0.036465	-0.02209	-0.06474	-0.08296	0.087343	-0.0051	-0.07638	0.100344	-0.07955
0.016613	-0.02478	-0.07266	-0.08997	0.091133	0.007012	-0.06824	0.092377	-0.06518
0.00778	-0.006	-0.05155	-0.10064	0.088438	-0.00075	-0.06463	0.079927	-0.07063
0.008878	-0.00885	-0.06561	-0.08635	0.097153	0.008177	-0.0532	0.089626	-0.07883
-0.00426	0.013563	-0.07962	-0.08352	0.075711	0.012641	-0.07756	0.082031	-0.07479
-0.02254	0.012416	-0.08413	-0.09245	0.097509	0.010024	-0.07821	0.063647	-0.06484
-0.01189	0.022219	-0.07173	-0.08473	0.077778	0.026799	-0.07303	0.067889	-0.05922
-0.01189	0.022219	-0.07173	-0.08473	0.077778	0.026799	-0.07303	0.067889	-0.05922
-0.02254	0.012416	-0.08413	-0.09245	0.097509	0.010024	-0.07821	0.063647	-0.06484
-0.00426	0.013563	-0.07962	-0.08352	0.075711	0.012641	-0.07756	0.082031	-0.07479
0.008878	-0.00885	-0.06561	-0.08635	0.097153	0.008177	-0.0532	0.089626	-0.07883
0.00778	-0.006	-0.05155	-0.10064	0.088438	-0.00075	-0.06463	0.079927	-0.07063
0.016613	-0.02478	-0.07266	-0.08997	0.091133	0.007012	-0.06824	0.092377	-0.06518
0.036465	-0.02209	-0.06474	-0.08296	0.087343	-0.0051	-0.07638	0.100344	-0.07955
0.039415	-0.02138	-0.08835	-0.09226	0.084445	-0.00483	-0.05135	0.082031	-0.10099
0.012862	-0.02619	-0.08585	-0.07873	0.078352	0.03729	-0.06104	0.085359	-0.10471
0.048243	0.009012	-0.07539	-0.08753	0.133849	0.030467	-0.10013	0.117028	-0.10148
0.04328	-0.02174	-0.13245	-0.17939	0.221732	0.04444	-0.04662	0.158052	-0.12855
-0.23735	0.241795	-0.40916	-0.28911	0.296245	0.232269	-0.43321	0.219969	-0.18449

Figure 8. Pulse response captured at 5000 samples per second.

4.2. Hand Geometry Biometric Setup

In the hand geometry setup, the same dataset split of the test set was used. NIR hand images in the test set were used to validate the respective subject class. In the identification phase, we calculated the six distances, as discussed in Section 3.2, for every test set candidate and substituted those six values in the following equation:

$$N = \begin{bmatrix} d1 & d2 & d3 \\ d4 & d5 & d6 \\ 1 & 1 & 1 \end{bmatrix} \tag{4}$$

This matrix N is used to calculate matrix S in the following equation:

$$S = M - N \tag{5}$$

Then, the absolute determinant was calculated and stored in the Excel file, which was declared as the final matching score. The enrolled user (or subject) with the lowest value of the final matching score with the corresponding class was regarded as the identified person or subject. It should be noted that M was calculated and saved for 165 subjects out of 185 in the enrollment phase. Images for all 185 subjects—165 enrolled and 20 non-enrolled subjects—were used for the matching phase to observe the performance of the proposed algorithm. After iterative experiments, we selected a minimum threshold of 200 for the final matching score to differentiate between the enrolled users (or subject) and the imposters. If matching scores exceeded the threshold, the candidate was identified as an imposter. Finally, matching scores and class labels obtained from all the test images were stored.

4.3. Finger–Vein Biometric Setup

For the finger–vein setup, firstly, the CNN was trained using the training set. Then, the trained model was used to validate the test set. The confidence scores and class labels from all test sets were stored. After this, the pulse response output in the form of 0 and 1, hand geometry normalized output in the range of 0–1, and finger–vein output scores in the range of 0–1 were given to the fuzzy system.

5. Experimental Results

5.1. Pulse Response Biometric

According to the pulse responses captured from a live human body and non-living materials like wood, plastic and metal, it was observed that 99% accuracy was demonstrated in identifying the live human body in comparison with a non-living material body.

5.2. Hand Geometry Biometric

According to the proposed algorithm, it was observed that an accuracy of 86% was achieved in identifying the correct subject when hand geometry was used in unimodal settings. A false acceptance ratio (FAR) of 0.18 and a false rejection ratio (FRR) of 0.17 were observed when experiments were performed for the whole image database collected from the 185 subjects (165 enrolled subjects and 20 non-enrolled subjects).

5.3. Finger–Vein Biometric

We evaluated our proposed method in the NIR hand images dataset. In addition to the CNN described in Section 3.3, VGG16 [31] and VGG19 [32] were also tested on the NIR hand images dataset. Furthermore, we used precision–recall metrics for reporting the results. More discussion about the evaluation metrics is in the following section.

5.3.1. Evaluation Metrics

We used precision–recall as the evaluation metrics and reported the results as shown in Figure 9. Each of the results was reported using the all-versus-one strategy (i.e., all samples of one class in the testing images were considered as positives, and all the remaining samples were considered as negatives). An algorithm was run on the test images. Values of true positive (TP), false positive (FP), and false negative (FN) were saved for all the classes. This process was repeated until all 185 subject classes were done. Precision and recall values were calculated using the values of TP, FP, and FN, saved for one run. After that, the classifier threshold was changed, and the whole process was repeated a few times.

After obtaining a few values of precision–recall, we plotted those values and calculated the area under the curve using trapezoid calculations. This area under the curve (AUC) reflects the classifier accuracy, as shown in Figure 9.

5.3.2. Experiments with NIR Dataset

First of all, we evaluated our proposed method of hand geometry discussed in Section 3.2. For different values of pixels in the X-Y direction, corresponding accuracies were obtained. It would be noted that for the cluster size of 140 pixels, the proposed algorithm gave an AUC of around 86%, as shown in Figure 9a. In our opinion, the accuracy of the proposed hand geometry algorithm was quite impressive, keeping in mind the accuracies of CNNs.

As mentioned earlier, we compared the accuracies of three different CNNs for the finger–vein biometric. We trained and tested Alexnet, VGG16, and VGG19. We found out that the CNNs achieved higher accuracies than their hand geometry counterpart, as shown in Figure 9b.
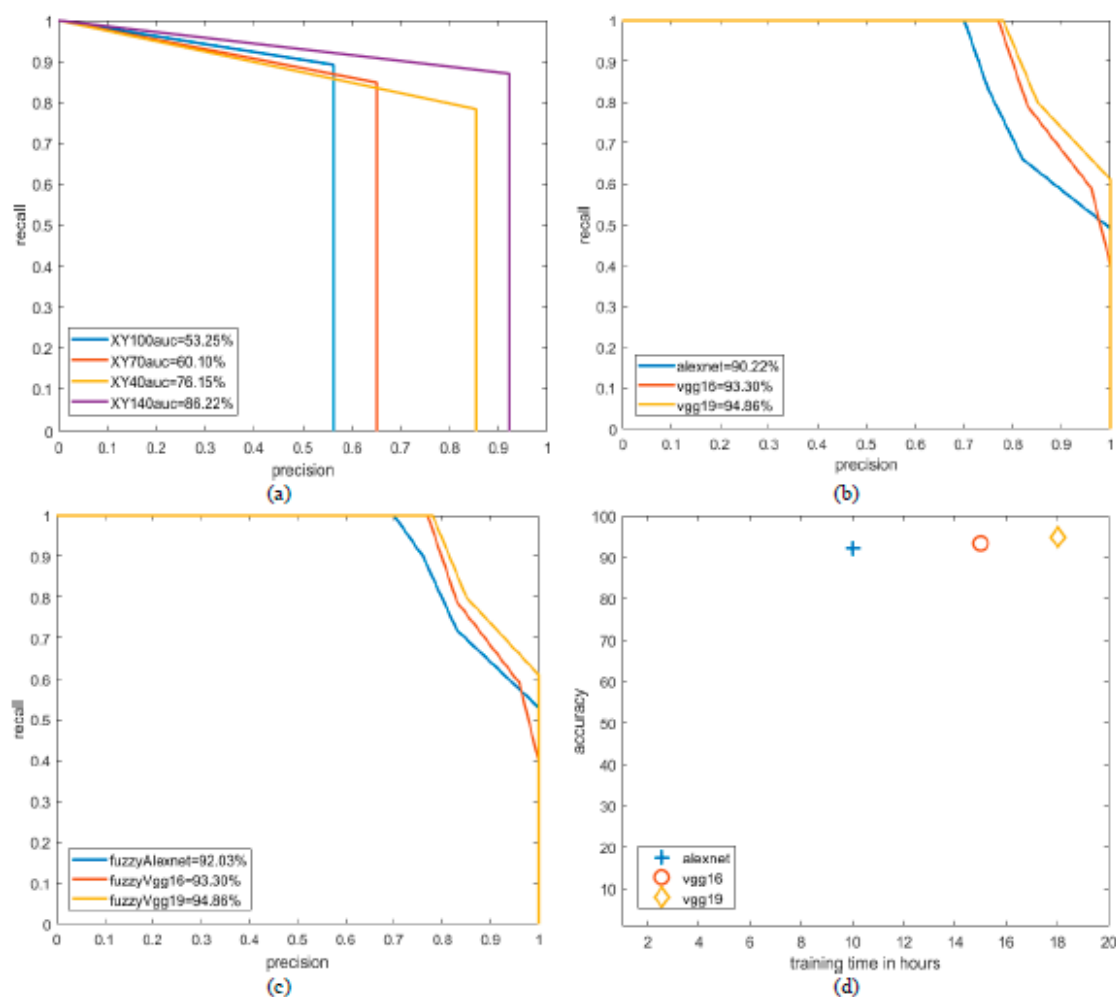
Figure 9. ROC curves and training time vs. accuracy: (**a**) PR Curves of the Proposed Hand Geometry System; (**b**) PR Curves of Convolutional Neural Networks tested on NIR Dataset; (**c**) PR Curves of Convolutional Neural Networks fused with the Fuzzy System; (**d**) Training Time vs Accuracy of the Convolutional Neural Networks on NIR Dataset.

Table 1 lists the performance parameters, accuracy, and training time for the final combined holistic fuzzy system by employing all three CNNs. For the hand geometry biometric, 140-pixel clustering was kept constant. When Alexnet was selected for the finger–vein biometric, the overall output accuracy of the system increased to 92%, which was a 2% increase from the sole finger–vein biometric using Alexnet, as shown in Figure 9c.

Table 1. Performance comparison.

Name of Method	Accuracy	Training Time
Fuzzy Alexnet	92.03%	≈10 h
Fuzzy VGG16	93.30%	≈15 h
Fuzzy VGG19	94.86%	≈18 h

We noted that the fuzzy inference rule helped complement the hand geometry and finger–vein biometric systems positively. We also observed that VGG16 and VGG19 had no influence on the fuzzy inference rules, and the accuracy remained the same. We noted that this was due to the strong decision score assigned by the aforementioned algorithms.

Figure 9d shows the accuracy vs. training time for all three CNNs. Alexnet took about 10 h for training on 80% of the total dataset, whereas VGG16 and VGG19 took 16 h and 18 h on the same training set, respectively. During the training phase, we noted that Alexnet dealt with 60 million parameters, VGG16 with 138 million parameters, and VGG19 with 144 million parameters. Our GPU had limited memory, and all of the dataset could not be loaded into the GPU's memory in one shot. Due to this, all the CNNs were spending the majority of the time transfering parameter inter-memory. This implies that, with a larger GPU memory size, training can be made faster for the aforementioned algorithms.

5.3.3. Biometric Performance Parameters

Figure 10a,b show the performance evaluation of the proposed fuzzy Alexnet in terms of the false acceptance rate (FAR) and false rejection rate (FRR) vs. the threshold of the classifier and the genuine acceptance rate (GAR) vs. FAR biometric performance metric graphs, respectively.
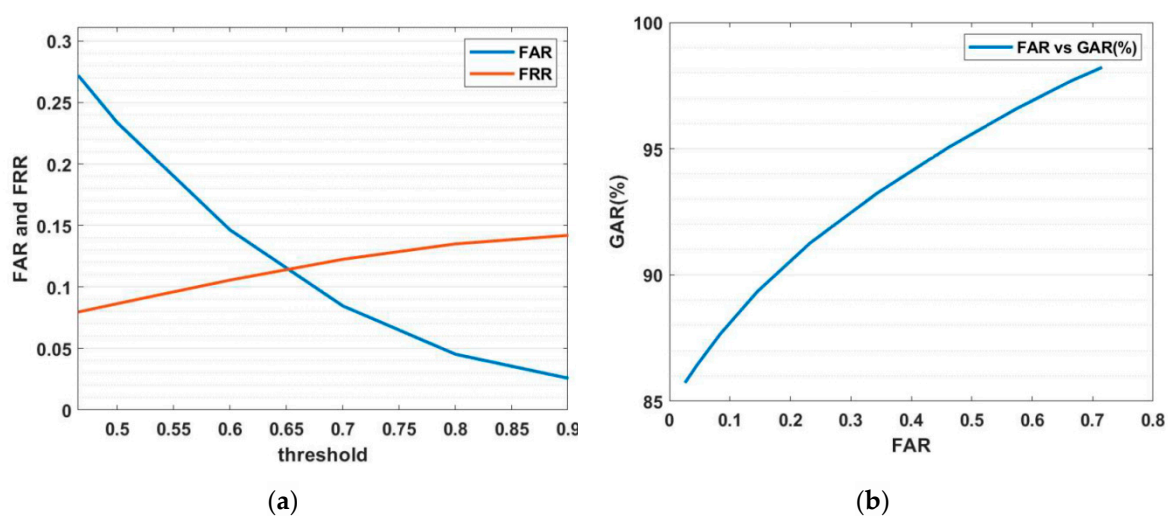

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 10. (**a**) Equal error rate; (**b**) GAR vs. FAR.

In Figure 10a, the FAR, as well as the FRR, was plotted against different values of the sensitivity threshold of the classifier. This plot helped in finding out the desirable values of the FAR and FRR from the graph. At the point of intersection, we have FAR = FRR = 0.113. Hence, we achieved an equal error rate (EER) equal to 1 where the FAR and FRR are at same value, which is the desired performance parameter for any biometric system.

The graph shown in Figure 10b explains the performance of our proposed multimodal biometric system by plotting the genuine acceptance rate against the false acceptance rate. It describes the effect of increasing the accuracy on the FAR of the system by adjusting the sensitivity of the classifier.

6. Conclusions

In this paper, we proposed a robust anti-spoofing system using biometrics modalities. The pulse response biometric filtered the non-living material very efficiently, as its demonstrated accuracy was 99%. A new near-infrared (NIR) hand images dataset, containing a total of 111,000 NIR hand images collected from 185 human subjects, was also proposed. Besides that, we formulated a handcrafted technique for hand geometry recognition that achieved 86% accuracy on the NIR hand dataset. We also implemented the finger–vein biometric system using convolutional neural networks. Finally, a novel fuzzy rules-based biometric system was proposed, which achieved an accuracy of 92% on the proposed NIR hand images dataset. During the experiments, we found out that convolutional neural networks like VGG16 and VGG19 solely achieved accuracies near the proposed fuzzy rules-based biometric system at the trade-off of training time. For future work, we plan to make a stronger fuzzy system that can correct more classification errors with the help of its wider rule base canvas.

7. Patents

In this research, we maintained a dataset of near-infrared images for the human hand. This dataset was collected from 185 subjects.

8. On-Request Dataset

The dataset maintained may be provided if requested by submitting through email a scanned copy of the signed form attached in Appendix A. The form must be signed by the requesting research personnel, as well as the legal officer of the researcher's institution.

Author Contributions: There are three authors for this article. The first author is S.A.H., the second author is Y.R., and S.M.U.A. is the third one. The concept of the research was generated by first author and discussed with third author for necessary modifications; the methodology was selected by the first and second authors; coding was performed by the first and second authors; validation of the results was carried out by all the three authors; formal analysis and investigation were performed by the first and second authors; resources for capturing near-infrared images and pulse responses were fetched by the first author under the guidance and facilitation of third author; biometric data collection was performed by the first and second authors with the support of the third author; all the authors arranged the volunteers for dataset collection; the original draft was prepared by the first and second authors; review and editing was performed by the first and second authors under the guidance and feedback of third author; visualization was performed by the first and second authors; the third author supervised the whole span of research; project administration was performed by the first author with the help of the third author; funding was applied for and acquired by the first author under the suggestion of the third author. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Appendix A

Application and Database License Agreement for

Near-Infrared Hand Images Database Access

Emerging Technologies Research Group

Department of Electronic Engineering—NED University of Engineering & Technology

The database of near-infrared human hand images was collected during the Ph.D. research project under the approval of the Research Ethics committee of NED University of Engineering & Technology. The informed consent of all the participating volunteers was fetched. The participants' ages ranged from 19–55 years.

In order to facilitate researchers working in the field of finger–vein and hand geometry biometrics, this dataset is available through proper channel requests. All requests to acquire this database must be submitted by the researcher's institution on behalf of an individual researcher or the research laboratory or center. In order to receive a copy of this database, a legal officer of the researcher's parent institution must sign this document and agree to observe the restrictions stated below. In addition to other possible counteractions, failure to comply with these restrictions may result in the withdrawal of permission to use this database, as well as denial of access to additional facilities provided by our university and research group. The submission of a completed license agreement does not automatically guarantee access to the database. The submitted request is to be approved by the statutory body of the NED University of Engineering & Technology before providing time-limited access to the database. The distribution of this database may be further restricted to comply with local legal requirements.

Consent

The researcher and the countersigning official agree to the following terms and conditions:

1. Redistribution and modification: The database, in whole or in part, will not be further distributed, published, copied, or disseminated in any way or circumstances whatsoever, whether for academic, research, or commercial use, without prior approval from the Principal Investigator and statutory body of the NED University of Engineering & Technology. Researchers will not forge the originality of the images for any type of violation of copyright for the database.

2. Publication Requirement: The researchers intending to include more than 10 images from the provided database in reports, papers, and other documents to be published or released must first obtain approval in writing from the Principal Investigator.

3. Citation: The researchers getting help from the provided database in their experimental or literature review are required to include the citation of the published paper, whose details will be provided with the approval of the database request.

4. Publications to NEDUET: The researchers requesting the database must submit a copy of all the reports, and articles that are for public or general release that use this database must be forwarded immediately upon release or publication to the Department of Electronic Engineering, NED University of Engineering & Technology.

Authorized Signature of the Researcher

Authorized Signature and Stamp of the Institution's Legal Officer

---

Organization Name, Complete Address, and Contact E-mail Address:

Please submit the scanned copy of the signed license agreement to the following official e-mail: hsaqeel@neduet.edu.pk; Syed Aqeel Haider, Research Scholar, Department of Electronic Engineering, NED University of Engineering & Technology, Karachi—Pakistan.

References

1. Derawi, M.O. Smartphones and Biometrics—Gait and Activity Recognition. Ph.D. Thesis, Gjovik University College, Gjovik, Norway, 2012.

2. Rasmussen, K.B.; Roeschlin, M.; Martinovic, I.; Tsudik, G. *Authentication Using Pulse- Response Biometrics*; University of Oxford: Oxford, UK, 2014.

3. Yang, L.; Yang, G.; Yin, Y.; Xi, X. Exploring soft biometric trait with finger vein recognition. *Neurocomputing* **2014**, *135*, 218–228. [CrossRef]

4. Yang, J.; Shi, Y. Finger–vein ROI localization and vein ridge enhancement. *Pattern Recognit. Lett.* **2012**, *33*, 1569–1579. [CrossRef]

5. Feng, Q.; Yuan, C.; Pan, J.-S.; Yang, J.-F.; Chou, Y.-T.; Zhou, Y.; Li, W. Superimposed sparse parameter classifiers for face recognition. *IEEE Trans. Cybern.* **2016**, *47*, 378–390. [CrossRef]

6. Zeinstra, C.; Veldhuis, R.; Spreeuwers, L. Grid-based likelihood ratio classifiers for the comparison of facial marks. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 253–264. [CrossRef]

7. Darwish, S.M. Design of adaptive biometric gait recognition algorithm with free walking directions. *IET Biom.* **2016**, *6*, 53–60. [CrossRef]

8. Wang, X.; Feng, S. Multi-perspective gait recognition based on classifier fusion. *IET Image Process.* **2019**, *13*, 1885–1891. [CrossRef]

9. Komeili, M.; Armanfard, N.; Hatzinakos, D. Liveness detection and automatic template updating using fusion of ECG and fingerprint. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1810–1822. [CrossRef]

10. Maiorana, E.; Campisi, P. Longitudinal evaluation of EEG-based biometric recognition. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 1123–1138. [CrossRef]

11. Peng, J.; Aved, A.J.; Seetharaman, G.; Palaniappan, K. Multiview boosting with information propagation for classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 657–669. [CrossRef]

12. Ahmadi, N.; Akbarizadeh, G. Hybrid robust iris recognition approach using iris image pre-processing, two-dimensional gabor features and multi-layer perceptron neural network/PSO. *IET Biom.* **2017**, *7*, 153–162. [CrossRef]

13. Chaa, M.; Akhtar, Z.; Attia, A. 3D palmprint recognition using unsupervised convolutional deep learning network and SVM classifier. *IET Image Process.* **2019**, *13*, 736–745. [CrossRef]

14. Veluchamy, S.; Karlmarx, L.R. System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier. *IET Biom.* **2016**, *6*, 232–242. [CrossRef]

15. Bok, J.Y.; Suh, K.H.; Lee, E.C. Detecting Fake Finger-Vein Data Using Remote Photoplethysmography. *Electronics* **2019**, *8*, 1016. [CrossRef]

16. Kim, W.; Song, J.M.; Park, K.R. Multimodal Biometric Recognition Based on Convolutional Neural Network by the Fusion of Finger-Vein and Finger Shape Using Near-Infrared (NIR) Camera Sensor. *Sensors* **2018**, *18*, 2296. [CrossRef]

17. Bernacki, K.; Moroń, T.; Popowicz, A. Modified Distance Transformation for Image Enhancement in NIR Imaging of Finger Vein System. *Sensors* **2020**, *20*, 1644. [CrossRef]

18. Yao, Q.; Song, D.; Xu, X. Robust Finger-vein ROI Localization Based on the 3σ Criterion Dynamic Threshold Strategy. *Sensors* **2020**, *20*, 3997. [CrossRef]

19. Alay, N.; Al-Baity, H.H. Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Face, and Finger Vein Traits. *Sensors* **2020**, *20*, 5523. [CrossRef]

20. Lv, G.-L.; Shen, L.; Yao, Y.-D.; Wang, H.-X.; Zhao, G.-D. Feature-Level Fusion of Finger Vein and Fingerprint Based on a Single Finger Image: The Use of Incompletely Closed Near-Infrared Equipment. *Symmetry* **2020**, *12*, 709. [CrossRef]

21. Xie, C.; Kumar, A. Finger Vein Identification Using Convolutional Neural Network and Supervised Discrete Hashing. In *Deep Learning for Biometrics*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 109–132.

22. Das, R.; Piciucco, E.; Maiorana, E.; Campisi, P. Convolutional neural network for finger-vein-based biometric identification. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 360–373. [CrossRef]

23. Wan, H.; Chen, L.; Song, H.; Yang, J. Dorsal hand vein recognition based on convolutional neural networks. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; pp. 1215–1221.

24. Hong, H.G.; Lee, M.B.; Park, K.R. Convolutional neural network-based finger-vein recognition using NIR image sensors. *Sensors* **2017**, *17*, 1297. [CrossRef]
25. Mehdi Cherrat, E.; Alaoui, R.; Bouzahir, H. Convolutional neural networks approach for multimodal biometric identification system using the fusion of fingerprint, finger-vein and face images. *PeerJ Comput. Sci.* **2020**, *6*, e248. [CrossRef]
26. Svoboda, J.; Klubal, O.; Drahansky, M. Biometric recognition of people by 3D hand geometry. In Proceedings of the The International Conference on Digital Technologies 2013, Zilina, Slovakia, 29–31 May 2013; pp. 137–141.
27. Budi Wirayuda, T.A.; Kuswanto, D.H.; Adhi, H.A.; Dayawati, R.N. Implementation of feature extraction based hand geometry in biometric identification system. In Proceedings of the 2013 International Conference of Information and Communication Technology (ICoICT), Bandung, Indonesia, 20–22 March 2013; pp. 259–263.
28. Svoboda, J.; Bronstein, M.M.; Drahansky, M. Contactless biometric hand geometry recognition using a low-cost 3D camera. In Proceedings of the 2015 International Conference on Biometrics (ICB), Phuket, Thailand, 19–22 May 2015; pp. 452–457.
29. Srikantaswamy, R. Fusion of fingerprint, palmprint and hand geometry for an efficient multimodal person authentication system. In Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Bangalore, India, 21–23 July 2016; pp. 565–570.
30. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in neural information processing systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
31. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
32. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]