The Effect of Pre-processing on a Convolutional Neural Network Model for Dorsal Hand Vein Recognition

Omar Tarawneh¹, Qotadeh Saber², Ahmed Almaghthawi³, Hamza Abu Owida⁴, Abedalhakeem Issa⁵, Nawaf Alshdaifat⁶, Ghait Jaradat⁷, Suhaila Abuowaida⁸, Mohammad Arabiat⁹

Abstract—There are numerous techniques for identifying users, including cards, passwords, and biometrics. Emerging technologies such as cloud computing, smart gadgets, and home automation have raised users’ awareness of the privacy and security of their data. The current study aimed to utilize the CNN model augmented with various pre-processing filters to create a reliable identification system based on the DHV. In addition, the proposed implementing several pre-processing filters to enhance CNN recognition accuracy. The study used a dataset of 500 hand-vein images extracted from 50 patients, while the dataset training was done using the data augmentation technique. The accuracy of the proposed model in this study in classifying images without using image processing showed that 70% was approved for training. Moreover, the results indicated that using the mean filter to remove the noise gave better results, as the accuracy reached 99% in both training conditions.

Keywords—CNNs; preprocessing; dorsal hand vein; recognition; CNN; authentication

I. INTRODUCTION

There are several methods for users’ authentication identification [1], such as cards, passwords, and biometrics [2]. Emerging technologies such as cloud computing, smart devices, and house automation [3] have created a dramatic consciousness and awareness among users regarding the privacy and security of their data. Hence, the traditional authentication and identification methods were not sufficient and reliable because they could be easily hacked. Moreover, the biometrics system is one of the most reliable and secure methods for personal data authentication and identification. This system uses the unforgeable and unique personal characteristics based on body components or measurements to ensure no copying or stealing of personal data [4]. In addition, biometrics technology is considered an effective and essential solution for authentication identification [5]. Biometric systems are reliable technologies that can recognise individuals’ unique characteristics effectively. These developing automatic technologies have become an appropriate alternative to traditional security methods. In addition, biometrics refers to identifying individuals based on physiological or behavioural characteristics. In addition, physiological biometrics uses the physiological features in the human body to perform facial recognition, fingerprints, iris, finger sweat, and dorsal vein recognition. The behavioural values are recognised based on human behavioural characteristics such as signature, gait, and voice recognition. Therefore, the biometric features are unique, and veins are difficult to construct by hackers [6]. Recent years have witnessed significant attention towards one of the emerging biometrics technologies, namely the biometrics of dorsal hand vein (DHV) [5]. In addition, significant attention has turned to the “dorsal hand vein” pattern because it is contactless, stable, unique, universal, and has the simplicity of detecting liveness [7]. Moreover, the pattern of DHV is a physiological feature through which people can be defined and distinguished from others. The critical step in the biometric system is the image feature extraction [6]. In this paper the researchers argued that vein recognition systems have become the focus of attention for many researchers since it is a new research track in pattern biometrics that uses the physiological features in the human body, while behavioural values are recognised based on human behavioural characteristics such as signature and voice recognition. In addition, the human hand has been considered a promising component for biometric-based identification and authentication systems for many decades. The unique characteristics of the hand vein make it difficult to forge the patterns. This study proposes a hand-side recognition framework based on deep learning and biometric authentication using the hashing method [8].

Vein recognition involves hand dorsal veins [5] and finger veins [7]. The main goal of the “dorsal hand vein” biometric system is to get an E-signature by utilising a well-secured signature device. Besides, [9] argued that CNNs are common in the “ImageNet large scale visual recognition competition (ILSVRC 2012)” due to their ability for identification and computational efficiency. In addition, different studies were conducted to investigate the best extraction of vein images through various methods. Other methods utilised pre-processing to obtain further enhancement for the image before using feature extraction to search for matches and make a comparison. In recent years, CNNs have shown significant competency and performance in image classification. CNNs are used to reduce the image processing early stage and recognise and classify the images of palm veins [9]. The present study suggests utilising...
the CNN model augmented with various pre-processing filters to create a reliable identification system based on the personal DHV. Besides, the study proposes several filters for pre-processing to enhance CNN’s accuracy in recognising different depths. Moreover, different filters might be used to promote the images prior to entering CNN. Part of these filters can sharpen and smooth the image and remove the noise. In addition, the “Vascular Pattern-Based Biometric” deals with the patterns formed by the blood vessels located inside the human body, where the patterns inside the human fingers and hands are the most widely used body parts [10]. This is commonly denoted as finger- and hand-vein recognition. The experiments will be carried out on a well-known dataset of hand-vein images prepared by the researcher Ahmad Badawi of the University of Tennessee, USA [7]. The dataset consisted of 500 images extracted from 50 patients, with 10 images for each. The ten images are divided into five images for the left hand and five for the right hand. The research community is aware that more samples in the training dataset result in a better training model for CNN, which improves recognition accuracy. Therefore, due to the lack of large datasets of dorsal hand-vein images, it is possible to increase the dataset size by increasing the variation of a single image in the training dataset using a technique called data augmentation. In data augmentation, an image can be rotated, scaled, cropped, and mirrored as many times as needed to obtain several variations of the same image, increasing the dataset size [10].

The proposed study aims to stem from the need to determine and identify the most appropriate system, method, and technique to be used in dorsal-hand vein recognition for authentication. The first section involves the introduction. The second section presents the previous works related to the paper topic. Section three will detail the paper idea and discuss the study method and implementation. Section four presents the proposed model results with a discussion. Section five presents the paper’s conclusion and provides recommendations for future works.

II. LITERATURE REVIEW

This section will review the previous studies related to the current research to explore the proposed models for voice-hand recognition. The researchers in [11] aimed to discuss the approaches taken from other research on pre-processing, feature extraction, and classification stages specifically for recognising individual identity. Furthermore, the study aimed to address the strengths and weaknesses of this approach using machine learning to determine the future direction and fill the gap in the previous research. The researchers found that machine learning techniques have a high potential to be the future research direction in this field, and a new method of finger-vein identification should be proposed to overcome the weaknesses of the previous research. The researchers proposed a model for the HVR based on CNN for tackling the tasks of vein recognition, while the original CNN passed through three modifications by the researchers. At first, the researchers imported the regularised “radial basis function (RBF) networks” to the CNN for task recognition. In the second rank, the researcher used the self-growth strategy to train the feature learning layers. Also, to get the final model, the researchers came up with an algorithm for parameter learning and relearning in the new model’s added layer to make the distinguishing features and the best classification results available at the time. The results of the lab database of hand veins achieved a recognition rate of 89.43% in testing and 91.25% in training. In contrast, the comparative experiment with the CNN model and hand feature showed their effectiveness in the proposed model for DHVR [12]. In this paper [13] had proposed the method of dorsal hand vein recognition based on CNN. The study compared the rate of recognition for several CNN depth models and analysed the impact of dataset size on the rate of recognition of dorsal hand veins. In the first rank, the researcher extracted the interest region (ROI) of the images of the dorsal hand vein. Besides, the study utilised the “Contrast Limited Adaptive Histogram Equalisation” (CLAHE) and “Gaussian Smoothing Filter Algorithm” to make a pre-process for the images. Next, the reference “Visual Geometry Group (VGG)” depth CNN and “CaffeNet AlexNet” were trained to extract the features of the image; the researchers then applied a logistic regression for identification. The results of the experiment, which was applied to two datasets different in size, demonstrated that the size of the dataset and the depth of the network influenced the rate of recognition. Still, in different degrees as well, the recognition rate of dorsal hand vein based on “VGG-19” was 99.7%. Recently, the work by Samala et al. (2018) showed that it is possible to use multistage fine-tuned CNN to build a mass classification methodology for digital breast tomosynthesis (DBT). The methodology used multistage transfer learning by using several layers and selecting the best combination. In the beginning, the CNN that was tuned on the ImageNet dataset was implemented on DBT data, and the results were recorded in the multistage CNN that was fine-tuned on the DBT dataset. The CNN classification layers were used with various freeze patterns to extract the optimal combination that gives the highest accuracy. Six different combinations of transfer networks with different freeze patterns for convolutional layers were tested. Compared with single-stage learning, multistage transfer learning improved the results with the fewest variations. The authors [6] study a dorsal hand vein recognition system using a convolutional neural network, which is. This system automatically shows how to extract features from original images without pre-processing, using the pre-trained CNN models (Alex Net) to extract features from the layers. It was found that Alex Net reaches a 100% recognition rate, and using transfer learning gives more accurate rates than when using the pre-trained CNN model for feature extraction. The researchers expected that this work would benefit new methods, paving the way for many benefits in the fields of other biometrics and dorsal hand vein identification. The goal of the Rossan study was to show that biometrics of the dorsal hand veins are what motivate researchers to use different methods for processing the vein pattern, figuring out its features, and matching. The researchers added that processing steps play an important role in a biometric security system, allowing users to access features needed for later stages. Furthermore, they have considered that it is mandatory to investigate pre-processing factors that might affect a biometric system’s performance. The researchers found that different techniques provide different results, with varying impacts on the later stages. A well-defined extracted vein pattern will improve pattern performance, leading to more secure biometric authentication. The researchers [14] wanted to look into palm, hand, and finger vein recognition for automated personal authentication. They also looked at previous work to present an analysis of hand vein pattern.
recognition to make vein pattern authentication more accurate and faster. The researchers discovered that some tools—such as image acquisition, pre-processing, feature extraction, and matching methods—extract and analyse object patterns. They recommended that integrating biometric modalities can solve uni-modal system limitations and achieve higher performance.

In this study conducted to present the method of DHV recognition based on CNN [15]. The researcher compared two trained CNNs from end-to-end to the architectures of deep learning (ResNet, VGG, AlexNet, and SqueezeNet). The researcher implemented the learning transfer and the techniques for fine-tuning to reach identification based on DHV. The conducted experiments were implemented to identify the training behaviour and accuracy of the network architectures. The system was evaluated and trained through the “North China University of Technology (NCUT)” database, which involves images of low contrast and low resolution. Reasonably, there was a need to adopt different steps of pre-processing to find out the impact of a set of methods for image quality enhancement, for instance, “inhomogeneity correction, ordinal image encoding, and Gaussian smoothing. The results of the study indicated that deep learning training based on feature extraction achieved higher performance compared to other DHV identification systems. At the same time, the inhomogeneity image correction, which is one of the pre-processing steps, increased the accuracy by 2–3 percent. [16] aimed to show that personal or identical verification is a fundamental issue for providing authentication or security. The researchers found that biometric template protection is one of the most critical issues in securing today’s biometric system through a hybrid method for finger vein biometric recognition based on a deep learning approach. The study found that each part of the model provides concealable template ability, discrimination, and security. Hence, the proposed model is an enhancement over most existing permutation-based cancellable biometrics and machine learning-based finger vein recognition systems.

In this study [9] proposed the recognition method of palm vein based on CNN, which includes four steps: image matching, feature extraction, pre-processing, and image acquisition. This proposal aims to decrease the steps of recognition processing for the images of palm veins. In addition, the images of the palm vein were extracted through near-infrared light. The study relied on two datasets. The first dataset subjects were 50 individuals, and the researchers collected 20 images per individual, for a total of 1000 images. The captured image size shall be 224*224 and 227*227 based on VGG. Net and AlexNet, respectively, while the captured image size is 640 x 480 pixels. The second dataset’s subjects are 63 individuals, and the researcher captured 1260 images from them. The false rejection rate (FRR) of the first dataset is 0.6%, and the results indicated that AlexNet, VGG-16, and VGG-19 models have proved the deep learning advantages in the image field. The second dataset has a false rejection rate (FRR) of 0.3%. The image contrast was increased, features were emphasised, and the CNN three types were pushed to reach 99% through CHALE pre-processing. Using several graphic cards, the training time would have a significant impact on accuracy. The researchers have trained the VGG depth CNN and AlexNet networks to extract the features of images. Finally, the recognition rate of palm veins using AlexNet, VGG-16, and VGG-19 reached 96%, 97.5%, and 98.5%, respectively. The authors [17] aimed to test the approach of CNN-based recognition for the patterns of DHV. The researchers have tuned VGGNet-16 on four DHV image datasets (low, medium, and good quality) and augmented images (false images and genuine matching). The four datasets involve right- and left-hand DHV images. The researchers compared the results of the proposed model with the results of CNN models such as VGG-19 and VGG Face. The results indicated that the recognition accuracy of the proposed model utilising VGGNet-16 was 99.60% for images of good quality. In comparison, the recognition quality for images of medium quality was 98.46%, and the images of low quality were 97.99%. This paper conducted a study to investigate the impact of pre-processing on the CNN of image segmentation in the medical context [18]. It was the study’s goal to find out how well pre-processing worked on a performance model by testing it consistently across 24 different pre-processing configurations on three different medical datasets (Knee, Liver, and Brain). Prior to training on CNN, different configurations were applied (re-sampling, bias correction, region of interest, and normalisation). Consequently, within the same dataset, the performance between configurations varied by 64%. Therefore, to enhance the performance model, the pre-processing shall be adjusted for particular segmentation applications.

III. Methodology

The latest advancements in digital signals and computing technologies have enabled automated identification of humans based on their behavioural, psychological, and biological features [19]. Moreover, biometric systems refer to systems that allow access to resources based on behavioural, psychological, and physical traits [20]. Besides, the security systems are increasing rapidly, while vein recognition, which is one of the biometric system identification methods, has become an authentic identification method [21]. Convolutional neural networks (CNNs) are considered one of the neural network types utilised for strong correlation data modelling, such as the studies of the earth, time series multivariate, and images [22], [23], [24], [25]. Moreover, the CNNs have achieved significant results in terms of image classification and object detection [26]. Besides, CNN can accomplish the main image’s actual representation and get its visual straight from the picture’s pixels through small reprocessing [6]. In machine learning, data pre-processing is an essential step for enhancing the quality of the data and extracting meaningful insights [27]. This technique involves the raw data organising and cleaning for the models of machine learning training and building [28]. Moreover, the pre-processing involves database acquisition, importing the critical libraries, importing the dataset, handling missing values, identifying them, and encoding the data. Due to the lack of a large DHV image dataset, there is a need to increase the dataset size, and this could be achieved through dataset augmentation to increase the variation of a single image in the dataset training. The data augmentation allows image rotation, scaling, cropping, and mirroring as many times as needed to obtain several variations of the same image, increasing the dataset size. The proposed model in this study utilises the CNN model augmented with various pre-processing filters to create a reliable identification system based on the DHV. The outcomes illustrate the effect of pre-processing techniques on a convolutional neural network model for enhancing dorsal
hand vein recognition. The study problem is identifying and determining the appropriate technique and method for DHV recognition for authentication.

A. The Proposed Dataset

Based on the study problem mentioned previously, it is important to implement data augmentation and data pre-processing to achieve a better quality and authentic DHV image identification system. The proposed experiments will be applied to a previously identified dataset of hand-vein images prepared by the researcher [10]. The dataset consisted of 500 images extracted from 50 patients, with 10 images for each. The 10 images are divided into 5 images for left-hand and 5 images for right-hand. It is known in the research community that the more samples in the training dataset, the better the training model produced by the CNN, hence the better recognition accuracy obtained [7].

B. Procedures and Methodologies

The proposed system CNN was trained during the training process, and the classification was performed during the training process as follows:

- Dataset: A set of images was used in the test, and a dataset of hand-vein images prepared by [29] that contains 500 images was used. The images were taken by 50 people, with 10 images per person. The 10 images were per person, divided into 5 images for left-hand and 5 images for right-hand [10].

- Dataset Preparation: Taking images of the hands by the region of interest, then pre-processing these images to extract features using convolution architecture, helping in the extraction process. Applying filter image processing to extract features from the original image without pre-processing. Applying CNN and tearing eliminates the work of selecting features artificially because CNN can select and express the depth feature of the image and ensure the accuracy of image features. Applying the classification of the DHVR using the pre-trained CNN models (AlexNet), error-correcting output codes, and the K-nearest neighbour algorithm for better classification.

- Research model: Fig. 1 illustrates the proposed model key steps for recognition hand vein using the effect of pre-processing on the Convolutional Neural Network. The proposed model has various phases: The first phase involves the pre-processing operations that are required for the input image processing and includes image size reduction, image conversion to grey level, and finally the removal of the noise. The next phase involves histograms, smoothing, equalisation, and normalisation. These processes are utilised for image colour optimisation and adjusting. Then the image is passed to both the proposed model and the AlexNet model at the same time to identify the features of the image, and then the features are classified and evaluated to determine which one is better.

- Performance Measures of Image Retrieval Time: Precision recalls are the curves the model will be able to plot for each image and has been commonly used method. If precision is at x-axis and recall is at y-axis, then top right corner area will show the best performance of the algorithm under study. However, there are other methods such as ROC curves and f-measure, and more interestingly you can use statistical parametric or non-parametric tests such as ANOVA, McNemar’s test, Friedman Test or Quad Test. The required time for retrieving the image equals the required time for model building based on the process data calculation and analysis to be calculated before modifying the model. The positive predictive value is precision. It shapes a critical point from the instances related between retrieved instances from the results of the process as shown in the following.

\[
\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of irrelevant and relevant images}}
\]  

Therefore, this study considered evaluating the performance of the proposed model using precision, recall, f-measure, ROC, and ANOVA for demonstrating its efficiency and accuracy in recognizing dorsal hand vein images.

IV. RESULTS AND DISCUSSION

In this chapter, the previous experiments were conducted to find the best way to detect the features of the image with high accuracy in a short time. Pre-processing was used before the image entered CNN, and the pre-processing was not used. The impact of the processing on the accuracy of the results was measured, and then the comparison between the AlexNet and the proposed models was made. The results are based on
a comparison of the CNN model (AlexNet, Proposed) and the presence or absence of image pre-processing.

A. The Effect of Image Processing on the Results

The effect of image processing on the results Before entering the images into the proposed CNN model, the images are entered in several stages to extract the vein pattern from the images to increase the accuracy of the model. Table I shows the accuracy of the model in classifying images without using image processing. 70% was approved for training, 30% for testing, and 80% to 20% were also approved. The pre-

TABLE I. THE MODEL ACCURACY IN IMAGE CLASSIFICATION

<table>
<thead>
<tr>
<th>Classification</th>
<th>30% for the test</th>
<th>20% for the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>69%</td>
<td>76%</td>
</tr>
</tbody>
</table>

processing was conducted before entering the images into the CNN, and in the second case, the pre-processing was not used. The processing effect on the accuracy of the results, then a comparison between AlexNet and VGGNet). A set of images was used in the test, and the rest of the images were used in the training, but the overall image set consisted of 500 images. As shown in Chapter Four, the image processing in the proposed model was in stages, where the noise was removed in the first stage, then normalisation of the values and application of histogram equalisation on the image before entering the proposed CNN model. The second table shows the use of mean and medium filters to remove noise. A total of 30% of the images were used, and the results were 150 images in the classification, 80% of the images in the training, and an accuracy of 76%. With Mean Filter In the first case, 20% of the images were used for classifications, and 80% were used for training with a mean filter, and the accuracy was 96% in the second case, we used 30% of the images used for classification, and 70% of the images were used with a mean filter, so the accuracy was 99%. With Medium Filter In the first, 20% of the images were used for classification and 80% were used for training with a medium filter, so the accuracy was 96% in the second case, 30% of the images were used for classifications, and 70% were used for training with a medium filter, so the accuracy was 98%.

TABLE II. RESULTS OF MEAN AND MEDIUM FILTERS

<table>
<thead>
<tr>
<th></th>
<th>20% for the test</th>
<th>30% for the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>Precision with</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>98%</td>
<td>96%</td>
</tr>
</tbody>
</table>

Table II shows that using the mean filter to remove the noise gave better results, as the accuracy reached 99% in both training conditions.

B. Comparison with Other Approaches

When applying the proposed CNN model to the processed images and several well-known networks (AlexNet and VGGNet), the processed images were adopted according to the proposed steps in this research, and the image size was modified to suit each network. Table III shows the accuracy of the proposed model in classifying images compared to other networks. Table II: Comparing the proposed CNN model with several known tools. In the proposed CNN model, 20% of the images were extracted for classification, and 80% of the images were used for training, and we have entered the proposed model. Hence, the accuracy of the classification was 99%. In this case, 30% of the images were extracted for classification, and 70% of the images were used for training. We entered the proposed model, so the accuracy was 99% in the Alexnet model CNN, 20% of the images were used for classification, and 8% of the images were used for training; we entered the AlexNet, so the accuracy result was 98%. In this case, 30% of the images were taken for classification, and 70% were used for training; we entered AlexNet, so the accuracy was 96%. In the VGG net model CNN, 20% of the images were taken for classification and 80% were used for training. We entered the VGGNet, and the accuracy was 96%. In this case, 30% of the images were used for classification and 70% were used for training, and we entered the VGGNet, and the classification accuracy was 95.

C. Overall Results

Table I shows the model’s accuracy in classifying images without using image processing; the results showed that 70% was approved for training, 30% for testing, and 80% to 20% were also approved. Secondly, Table II illustrates the use of mean and medium filters to remove the noise in the proposed model; the results indicated that using the mean filter to remove the noise gave better results, as the accuracy reached 99% in both training conditions. Finally, by applying the CNN proposed model and several networks (AlexNet and VGGNet), the processed images were adopted according to the proposed steps in this research, and the image size was modified to suit each network. The third table shows the accuracy of the proposed model in classifying images compared to other networks.

TABLE III. COMPARING THE PROPOSED CNN MODEL WITH SEVERAL KNOWN TOOLS

<table>
<thead>
<tr>
<th></th>
<th>20% for the test</th>
<th>30% for the test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed CNN</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>ALEXNET</td>
<td>98%</td>
<td>96%</td>
</tr>
<tr>
<td>VGGNET</td>
<td>96%</td>
<td>95%</td>
</tr>
</tbody>
</table>

V. Conclusion

In conclusion, the study found that implementing the proposed model increased the accuracy of classifying images without using image processing. In addition, the results indicated that using the mean filter to remove the noise gave better results, as the accuracy reached 99% in both training conditions. Finally, by applying the CNN proposed model and several networks (AlexNet and VGGNet), the processed images were adopted according to the proposed steps in this research, and the image size was modified to suit each network. Therefore, it is recommended that pre-processing be implemented on a convolutional neural network to enhance dorsal hand vein recognition. Moreover, future work is advised
to conduct further studies to enhance the accuracy of dorsal hand vein recognition using different techniques.

REFERENCES


