Research Article GU J Sci 37(2): 654-675 (2024) DOI: 10.35378/gujs.1246486

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Shallow Convolutional Neural Network for Gender Classification Based on Hand

Md KHALILUZZAMAN *

Dept. of Computer Science and Engineering, International Islamic University Chittagong, Chattogram - 4318, Bangladesh

Highlights

- This paper recommended a shallow CNN model for classifying the gender from the hand image.
- Analyze the comparative analysis among the proposed model and present state-of-the-art.
- The classification accuracy was improved to a high level of precision.

Article Info

Received: 02 Feb 2023
Accepted: 25 Sep 2023

Keywords

Gender classification
CNN
Dorsal
Palm
11K hand dataset

Abstract

Gender classification based on the hand image is used in computer vision for human-computer communication, hand-based authentication, and identification systems. Beside this, gender classification may be applied for criminal investigations, visual surveillance, and other legal purposes. The traditional manual methods require a lot of time and are susceptible to variable fluctuations. However, for low amounts of data, the deep-learning models are going to be overfitted. In this regard, this work proposes a shallow convolutional neural network (CNN) with a regularization method. Here, different gender classification models are built to detect the gender individually from dorsal and palmar hand images. For that, the 11K hand dataset is divided into four labels, i.e., men dorsal side, women dorsal side, men palm side, and women palm side. These data have been pre-processed by resizing and scaling. Furthermore, a model is developed for classifying gender from the real time data. According to the experimental results, the model developed for the dorsal hand images outperforms the other proposed models and the current state-of-the-art.

1. INTRODUCTION

Gender classification by hand plays a vital role in computer vision. It has many significant applications in computer vision, like human-computer interaction, visual surveillance, forensics applications, and security control [1, 2]. Gender classification in soft biometrics such as the face, hand geometry, hand palm, gait analysis, and others has been used in recent decades [3, 4]. The research community uses these features particularly for pattern detection, human-computer interaction, and machine learning.

Automatic gender classification has increased in various software and hardware due to the growth of online and social networking applications. When dismembered human remains are encountered, archaeological studies [5], i.e., a kind of study that involves the human past, and gender identification from various body parts, are frequently used in criminal investigations. As we know, compared to facial appearance, the variety of hand appearance is lower, and distinctive features possessed by the human hands can reveal gender information [6]. Therefore, gender classification can be done by hand, and will be less complex. Also, in order to increase the reliability and accuracy of hand-based authentication and identification systems, gender information from hand images could be very significant. The gender classification is a significant aspect of the evidence from the examination. Hence, this gender classification technique can be utilized in the examination, as it is a quick procedure.

There are numerous existing challenges to classify the gender from hand images, such as, balance dataset [7], ambiguity between men's and women's hands, deformed hands as data, and the limited research on

*e-mail: khalil@iiuc.ac.bd

gender classification using hand images. As it is also a task of forensics [8], some research uses the ratio between fingers to classify the gender [9].

Traditional gender classification systems work according to a generic pipeline [10]. Various kinds of hand-crafted techniques are computed from images for classifying of different objects with preprocessing [11]. Although these techniques have some advantages, they also have some drawbacks, such as the need for manual supervision during feature extraction and selection. At the present state of the art, many approaches are used that do not deliver great performance for a bigger dataset. However, convolutional neural networks (CNNs) are used in current research to improve performance for the larger dataset with less manual supervision. The development of CNNs, which can outperform traditional approaches, has received a lot of research attention. Salient feature extraction from multiple-stacked deep layers is employed for a wide variety of computer vision and pattern recognition tasks in an end-to-end pipeline as the basis of CNN efficacy.

In this paper, we focus on classifying gender individually with the help of a hand image. For that, deep learning techniques are used to propose a personalized shallow convolutional neural network (CNN). The contribution of this paper is as follows:

- A shallow CNN-based model with deep learning techniques are recommended to classify the gender from the hands data i.e., 11k Hand and own created dataset.
- Analyze the results among the proposed models and other models to observe why our's or another's model worked better.
- The model is applied to real-world data to see the performance of the proposed model on real-world data.

The paper is divided according to the following sections: In section II, the literature review is discussed. In section III, the proposed method is explained. The experimental findings are summarized in section IV. This paper is concluded in section V.

2. RELATED WORKS

In the last few decades, many researchers have developed state-of-the-art methods for classifying gender from hand images. The region and boundary features of the hand, which are based on Zernike moments, Fourier descriptors, and the fusion method, were used earlier to construct gender identification systems [11]. For gender classification, the geometrical characteristics of the palm images are also evaluated [10]. Several techniques are currently being developed to solve this problem using CNNs, such as, a method proposed in [12] for gender recognition and biometric identification using a large dataset, i.e., 11K hands. In this paper, the authors propose a large dataset, which is called 11k hands. The dataset included both the dorsal and palmar sides of the human hand. They trained a CNN model for gender recognition tasks and designed a two-stream CNN to handle gender recognition issues. For biometric identification, they used a set of support vector machines (SVM) as a feature extractor. The dataset contains 11,076 hand images from 190 different people, ranging in age from 18 to 75 years. The paper had an accuracy of 87.4 percent for palm, 91 percent for dorsal, and 94.2 percent for palm and 97.3 percent for dorsal, respectively, for gender recognition and biometric identification.

In [13], five neural network structures are used to classify the gender from palm and dorsal side hand images. For training and testing the models, the authors created their own U-HD dataset. The proposed model shows a better result for the Inception-V3 structure compared to the other structures, with an accuracy of 90.49%. Hand shape based gender classification is presented in [11]. Here, the hand silhouette is divided into six parts, including the fingers and palm. From these parts, efficient features are extracted using the Fourier descriptor and Zernike moments. Finally, the gender is classified using the Eigen spaces. In [14], a deep learning based model for person identification based on the hand image is presented, where both the local and global deep-features are utilized to extract the discriminative features from the convolutional layers. For local feature learning, the uniform partition is performed on the convolutional layers both in the vertical and horizontal directions. In [15], a machine-learning model based on the geometric profile of the finger is proposed for person identification. Here, a preprocessing method is introduced to extract the finger level shape. From that shape, a forward-backward greedy algorithm is used for feature extraction. The features are classified using random forests and k-nearest neighbors. Later, in [16], a modified forward-backward algorithm is presented to classify the person. For person recognition, the authors also used k-nearest neighbor and random forest. This paper achieves 2% higher accuracy with respect to present state-of-the-art.

From hand images, an attribute recognition framework is proposed in [17]. Here, the authors introduce the multi-convolutional neural network for recognizing different attributes such as gender and age. For training and testing the model, the authors used the 11K hand dataset. However, the proposed method achieves lower accuracy for gender than other attributes. The authors of [18] propose five DCNN for gender and ethnicity classification in uncontrolled environments: AlexNet, DenseNet, ResNet-50, SqueezeNet, and VGG-16. The models that are used are fine-tuned to achieve better performance. The evaluation of the models is explained based on the scenarios of segmented, full hand, and palm-print images. From the experimental results, it is revealed that the full hand images are more suitable for gender recognition than the palm-print and segmented images. In another paper [19], three deep learning models, i.e., ResNet-50, VGG-16, and EfficientNet-B3, are used to classify the gender from the hand image fingerprint. Among the models, EfficentNet-B3 achieves the highest accuracy for gender recognition.

A deep learning based system [20] is proposed for recognizing individuals from the hand images. Here, firstly, the image is preprocessed through the CLAHE method. After that, the depth features are estimated using the deep convolutional network. The estimated features are recognized by using the SVM classifier. In [21], authors proposed a method for gender classification based on hand shape. In this paper, the authors investigate a method for resolving the problem of gender classification based on hand shape. In compared to the appearance of the face, the hands are less variable. For hand feature extraction, the hand has been segmented into different regions, with each region represented by a set of features. They used two MPG-7 shape descriptors to represent the geometry of fingers and the palm. ZMs and FDs are two kinds of shape descriptors used in this work to extract information. Here, three classifiers were used to classify the input images. These are minimum distance, K-nearest neighbor, and linear discriminant analysis (LDA). In this paper, they had accuracy rates near 89% and 87%, respectively.

A gender classification method is presented in [22] based on the geometry features of the palm image. In this paper, the authors used support vector machines (SVM) for gender classification purposes. Here, a flatbed scanner was used to acquire palm images. The scanner used in this work is BenQ, which is mainly based on CIS (contact image sensor). After that, an extraction algorithm is used to extract the features. As a learning algorithm, the PSSVM algorithm was used. The paper had an accuracy rate of 85.42% for these experiments.

In [23], authors describe a method for biometric palm recognition that combines hand geometry and palm print information. In this paper, three steps were used for recognition. In step one, preprocessing is used to transform the raw data into a specific target area, which is the ROI. In step two, feature extraction is used to extract two features, i.e., the palm print and hand geometry. In step three, to locate the matching test features, the extracted features are compared to the database. The Standard image processing processes were used in the pre-processing, including gray scaling and thresholding. The accuracy of this paper for the palm print biometric is 69%, for hand geometry, it is 51%. The combination of palm print and hand geometry achieves an accuracy of 89%.

The authors of [24] describe a way to figure out a person's gender based on the size of their hands. To establish the identity of individuals, gender determination is an important criterion. For forensic identification and criminal investigation, it can also be used. In this paper, authors took hand measurements with the help of a standard anthropometric instrument, i.e., a sliding caliper, for collecting each subject’s anthropometric measurements. For the experiment, they used the hand length, handbreadth, and index of each subject. The hand index was calculated by multiplying (hand breadth/hand length) by 100. The hand index was analyzed based on the standard range described by Martin and Saller [9]. The accuracy in hand length was 80% for females and 77% for males. In hand recognition, 82% of females and 81% of males. In

[25] presents a Novel approach for Gender identification from hand dorsal images using computer vision technique. In this paper, the authors used a three-layer convolutional neural network to extract features from hands. They employed the 11K hand dataset. Their model training and validation accuracy are 100% and 100%, respectively.

The researchers are also investigating various modalities in this direction, including face, palm print, ear, gait, etc. [26-29]. Motivated by this advancement, in this work, we have investigated a straightforward shallow deep model for gender classification utilizing a four-category hand dataset.

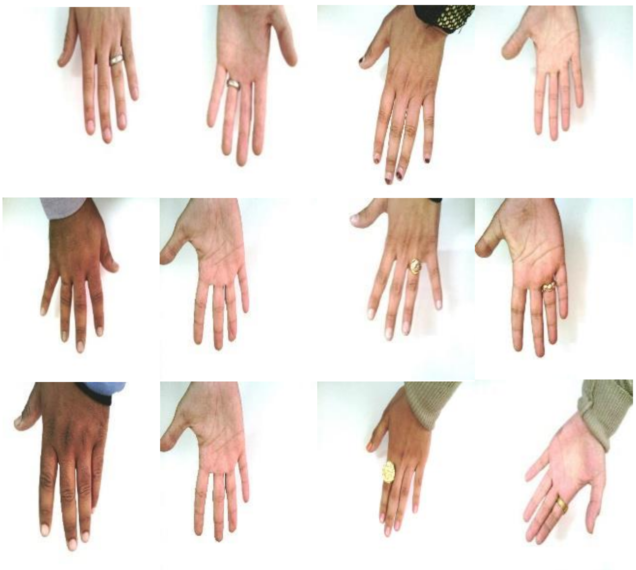
3. THE PROPOSED FRAMEWORK

The main objective of the paper is to classify the gender from hand images. Therefore, it is an image classification problem [30]. In an image classification problem, the input is the image data, and the output shows us what was in the image by predicting with the help of some algorithms. Moreover, the best algorithm for image classification is CNN [26, 31]. In our work, we create a framework for gender classification using male and female hand images as model input. Our framework is based on ConvNet, which is a convolutional neural network. We used our CNN model as the feature extractor of hands to classify the gender. We applied the ridge regularization to reduce the model's overfitting problem. We also used validation while training the model to improve its ability to classify unknown data.

3.1. Dataset

For our work, we used the 11K Hand dataset, which is now a benchmark dataset. In that dataset, there are a total of 11,076 hand images. These images are 1600 by 1200 pixels. The researchers collected hand images from nearly 190 people. From each subject, they collected 58 images, where the dorsal side has 30 images and the Palmer side has 28 images on average. They asked each subject to close and open their fingers so that they could collect different types of hand pictures. With a uniform background, every hand was photographed. The dataset is available at the following link [12]. https://sites.google.com/view/11khands/home

We partition this dataset into four categories: 'men dorsal side', 'women dorsal side', 'men palm side', and 'women palm side'. As we wanted to find out which hand side is better for predicting gender, i.e., dorsal side or palm side, from the 11,076 images, we used 5026 images for the dorsal side of men and women and 5019 images for the palm side of men and women. Some samples from the dataset are shown in Figure 1.



(a) (b) (c (d)

Figure 1. *Sample of 11K dataset: a) Men dorsal side, b) Men palm side, c) Women dorsal side, and d) Women palm side*

3.2. Data Preprocessing

In order to extract important information from the image, or to get a modified or enhanced image, requires performing a number of processes on the data. This process is known as image processing. It used when we have to perform some operation on an image to extract some information from it. Image processing steps and data preprocessing aid is used for reducing noise in data and increasing accuracy. In a dataset, there are images of different shapes. Using a dataset that has different shapes for all images is very difficult to work with. Data preprocessing makes this work easy for us by reshaping all images into the same shape.

Data Resizing

Our original image data is 1200 by 1600 pixels. We resize our data from 1200 by 1600 pixels to 70 by 70. This is because inputs with defined dimensions are often needed for deep learning models. All images are changed to have the same shape by data resizing, maintaining compatibility with the input layer of the model. For that, the images of the dorsal and palm-side hands images were resized to 70 by 70 pixels. Reducing the size of the data eases the computational load during inference and training. Because there are more pixels in large images, there are more processing operations needed.

Pixel Scaling

We scaled every pixel value of our images between 0 and 1. We divided the pixels of our images by 255, to scale them from 0 to 1, which is called the normalization. Pixel normalization is useful for improving feature extraction, stabilizing machine learning training, and improving robustness against noise in image processing. It also improves the comparability of images taken under various lighting conditions and prevents dominant pixel values [32]. Images are excellent for a variety of processing approaches because to its ability to avoid overfitting, facilitate effective convergence in optimization algorithms, and maintain relationships between pixel values [33].

3.3. Real Data Preprocessing

We collected around 300 real images of men's and women's hands to test our model. From one subject, we collected around 24 images. We used a Redmi 9s, a Redmi 7 Pro, and two more different devices to collect our data. To preprocess the real data, we followed the image preprocessing steps. There were some challenges while collecting real data for our work. Due to COVID-19, we were not able to go out and collect our data on our own. As a result, we had to gather this information online with the help of our friends and family. However, these 300 data points are not enough, as our work on deep learning has shown. This is because in deep learning, when there is more data, it helps the model learn more accurately. As the data is not enough, it causes an overfitting problem. To reduce the model's overfitting problem, there are some techniques, such as, data augmentation, ridge regularization, lasso regularization, and dropout.

Now, for our real data, we used the same preprocessing procedure. Before giving an image as an input to a model, it is necessary to preprocess those images to improve the prediction rate and model accuracy. Some steps must be taken to preprocess the signals in order to reduce noise. One of the important steps is to label the data. The real data for men and women are presented in Figures 2 and 3, respectively.

(a) *(b)*

Figure 2. *Sample of real dataset: a) Men dorsal side, and b) Men palm side*

(a) *(b)*

Figure 3. *Sample of real dataset: a) women dorsal side, and b) women palm side*

3.4. Proposed Model

In this section, we will explore the proposed shallow convolutional neural network. The model is designed with three convolutional layers and two dense layers. We designed this network, so that we can extract the features from images efficiently. Our base paper [17] used a 2-steam CNN model to extract the features of the hands. Our model is simpler and performs slightly better compared to the base model.

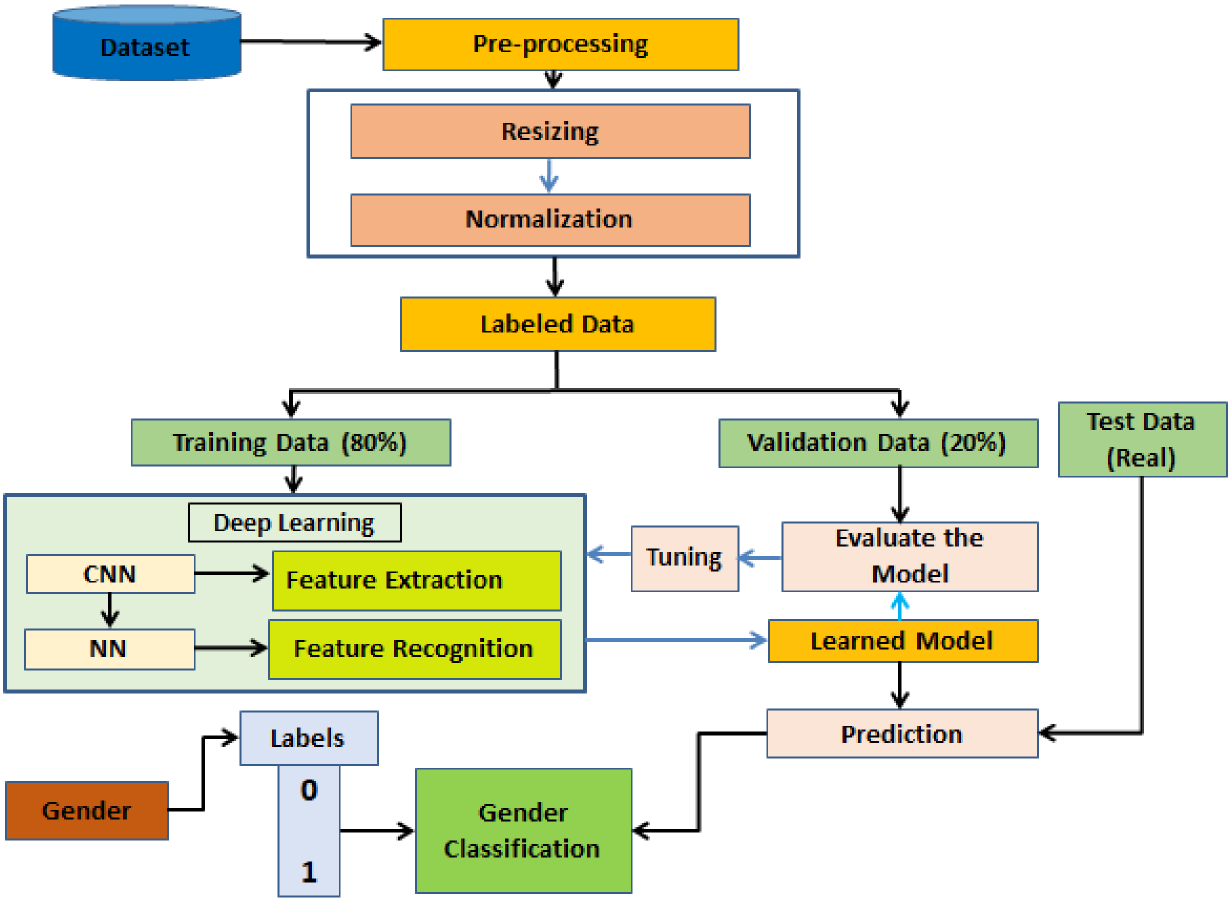


Figure 4(a). *The workflow diagram of the proposed gender classification model*

In our proposed model, we used convolution, max-pooling, ridge regularization, dense layering, and dropout. We used the 11K hands dataset in our proposed model to classify the gender. We also justified which hand side is better for predicting gender, i.e., dorsal or palm. We also test our model on real-world hand images, which are our own collected data. The workflow diagram of the proposed gender classification model is presented in Figure 4(a). A visual representation of our proposed model is shown, in Figure 4(b).

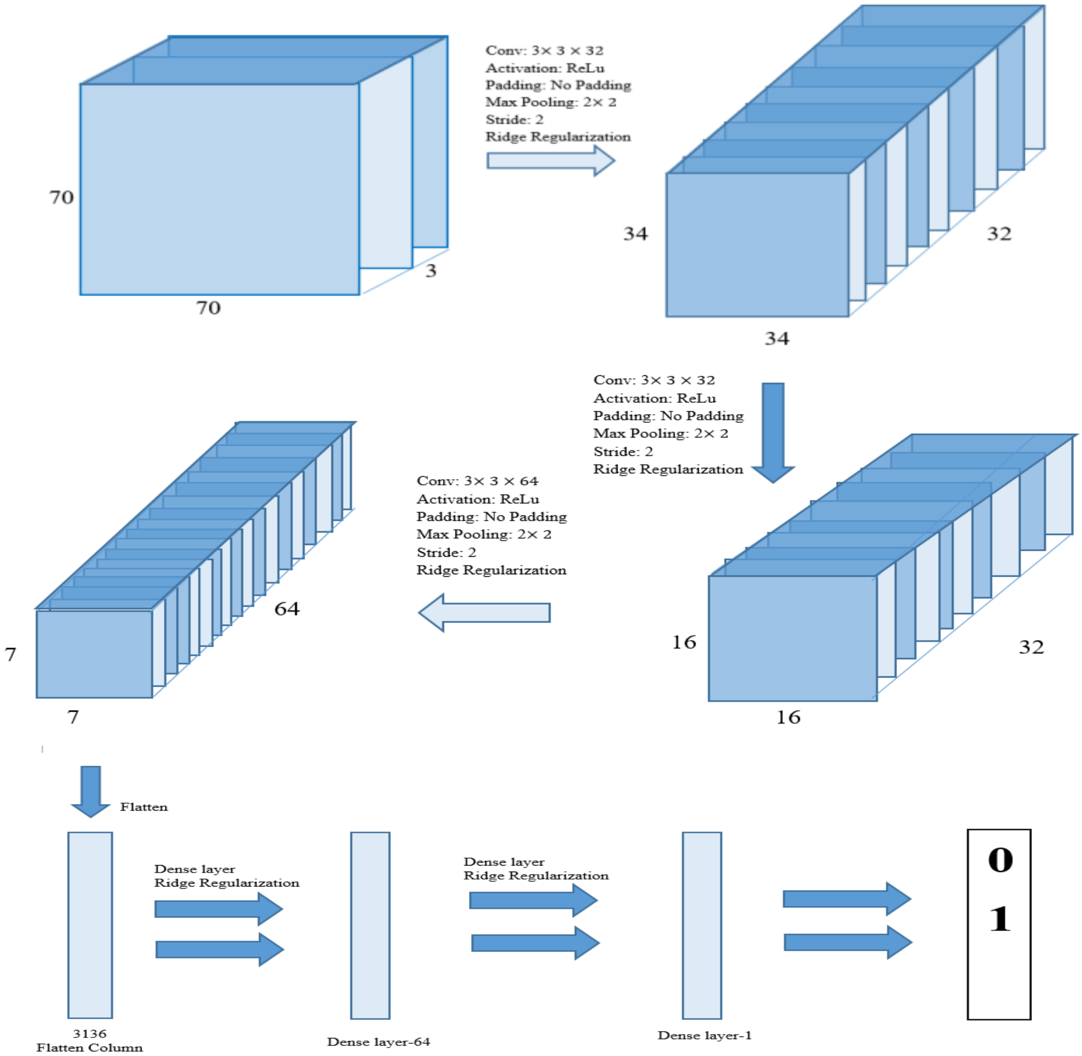


Figure 4(b). *The proposed model for the gender classification from image*

Our first convolutional layer consists of 32 filters with a kernel size of 3. We also employed max pooling with a kernel size of 2. The 32 filters combine 70 by 70 input image sizes. After convolving the input data, it became 68 by 68. Here, 32 is the number of filters we used. To help the network learn the complex pattern of the data, we used the activation function. In our work, we used Rectified Linear Unit (ReLU) activation, as it is the most frequently used activation function in the present state of the art. Any value that is negative is turned into zero by this activation function.

We used the max pooling to reduce the dimensionality of the feature map. The max pooling filter size is 2. This max pooling was used with stride size 2. After applying the max pooling to the image data, we get an image with a size of 34; again, 32 is the number of filters we used. To reduce overfitting, we used L2 regularization with a learning rate of 0.001. It helps reduce the overfitting problem by adding the square value of weights in the cost function.

In the second layer of the convolutional neural network, we again applied 32 filters, just like the first layer. These 32 filters combine the images produced by our first convolutional neural network layer. Therefore, this layer took the 34 size image data as input data. We obtained an image size of 32 after convolving this image. As we need to use an activation function, here we also applied the ReLU activation function to our feature map. In this layer, we again used a max pooling with filter size 2 and stride size 2 to reduce the dimensionality of our feature map. After applying the max pooling we got an image data that size of 16. Again in this layer of a convolutional neural network, L2 regularization is applied with a learning rate of 0.001 to reduce the overfitting problem.

In the third convolutional neural network layer, we used 64 filters, unlike our 1st and 2nd convolutional neural networks. These 64 filters convolve the image obtained from the second convolutional neural network layer, which are 16 layers deep. After convolving this image with three kernels, we got an image with a size of 14. Here, 64 is the number of the filters we use in this layer. As an activation function, here we also applied the ReLU activation function to our feature map. In this layer, we again used max pooling with filter size 2 and stride size 2 to reduce the dimensionality of our feature map. After max pooling, the image size of 14 is converted to 7. To reduce overfitting, we used L2 regularization with a learning rate of 0.001. After three convolutional neural network layers, we applied dropout. We used this to reduce our overfitting problem. We used dropout with the value of 0.4.

After three convolutional neural network layers, we applied flattening to our model. It turned our model from 3D to 1D, and it helps to further computation. After that, we used a dense layer with a size of 64 after flattening our model into one column. It is a fully connected layer [16], where every neuron in one layer is connected to the neurons of the next layer. To reduce overfitting, we used L2 regularization with a learning rate of 0.001. Finally, dense layer-1 is applied, which is the final output layer of the proposed model. For recognizing the predicted output, the softmax classifier is used. The output layer provides a value of 0 for women or 1 for men as a binary classification. The Hyper-parameters used in the proposed model are presented in Table 1.

Table 1. *Hyper-parameters used in the proposed method*

<table>
<tr><th>Layer</th><th>Regularization</th><th>Activation function</th><th>Learning Rate</th><th>Filter Size</th><th>Pooling</th><th>Batch Size</th><th>Optimizer and Loss Function</th><th>Dropout</th></tr>
<tr><td>First CNN Layer</td><td>L2 regularization</td><td>ReLU</td><td rowspan="5">0.001</td><td>32 filter kernel size of 3</td><td>Max</td><td rowspan="5">32</td><td rowspan="5">Adam and Binary Cross Entropy</td><td>No</td></tr>
<tr><td>Second CNN Layer</td><td>L2 regularization</td><td>ReLU</td><td>32 kernel size of 3</td><td>Max</td><td>No</td></tr>
<tr><td>Third CNN Layer</td><td>L2 regularization</td><td>ReLU</td><td>64 kernel size of 3</td><td>Max</td><td>Yes (0.4)</td></tr>
<tr><td>FC-1</td><td>L2 regularization</td><td>ReLU</td><td>64 Neurons</td><td>-</td><td>No</td></tr>
<tr><td>FC-2</td><td>-</td><td>Softmax</td><td>1 Output Neuron</td><td>-</td><td>No</td></tr>
</table>

4. EXPERIMENTAL RESULTS

We applied our proposed model to a benchmark dataset named 11K Hands. Here, total dorsal side hand images are 5026. The dataset is split by percentage, where 80% of the data is used as training data and 20% of the data is used as validation data. Therefore, for dorsal side hand images, we had 4020 images for training and 1006 images for validation. Again, for the palm side, we worked with a total of 5019 images. After splitting the palm side dataset, we got a total of 4016 images for training and 1003 images for validation. For testing we have used the real data. The samples of real data both for dorsal and palm side are shown in Figure 2 and 3 for men and women respectively. The data pre-processing steps are done by OpenCV, which is Python's computer vision library. Our model is implemented using the Python Keras library. To implement this model, we used a device with an Intel® Core™ i3-5005u processor, a CPU @ at 2 GHz, and 4 GB of RAM.

In the experiment, we set the epoch and batch sizes to 30 and 32, respectively. As a loss function and optimization algorithm, in this work, we use binary cross entropy [34] and the Adam optimizer [35]. The loss function is presented in (1). Without regularization, we noticed that our model was overfitted. So, we used ridge regularization, where the learning rate is 0.001 both for dorsal side hand images and palm side hand images. After using ridge regularization, it helped us reduce the overfitting of our model.

$$L = \frac{1}{n} \sum_{j=1}^{n} \quad - [y_j \,. log(p_j) \; + \; (1 - y_j). log(1 - \; p_j)] \qquad (1)$$

where, p_j is the probability of women class, and $(1 - p_j)$ is the probability of men class, n is the number of data samples.

The performance of our model was analyzed with the help of several machine learning metrics such as model accuracy, recall, precision, F1-score, and confusion matrix. The model accuracy is the accuracy gained on the dataset by the proposed model, as shown in (2). Here, total numbers are the sum of the true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The recall is the ratio between correctly identified positives and the total number of positives. The formulation is presented in (3). The precision is defined as the ratio of correctly predicted positives to total predicted positives. The formula is shown in (4). The F1-score is the harmonic relationship between recall and precision. The F1-score is shown in (5). The confusion matrix gives us a visual representation of correctly predicted data and incorrectly predicted data

$$Accuracy = \frac{Total\ number\ of\ TP\ and\ TN}{Total\ numbers} \tag{2}$$

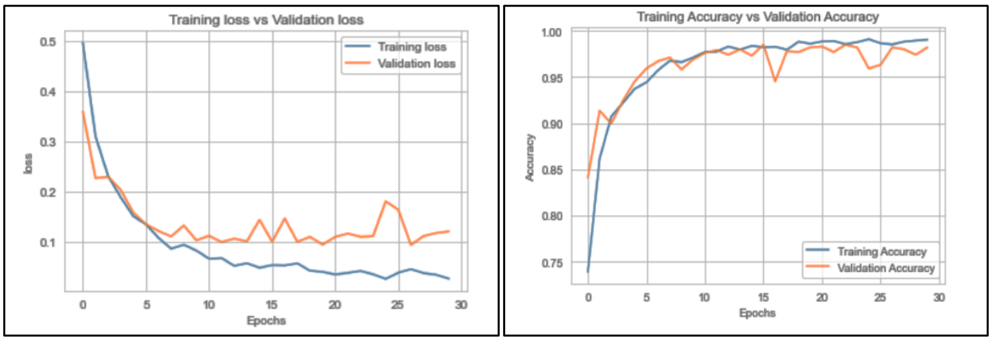
$$Recall = \frac{Total\ number\ of\ TP}{Total\ number\ of\ TP\ and\ FN} \tag{3}$$

$$Precision = \frac{Total\ number\ of\ TP}{Total\ number\ of\ TP\ and\ FP} \tag{4}$$

$$F1 - Score = 2 * \frac{recall * precision}{recall + precision}. \tag{5}$$

4.1. Gender Classification for Dorsal Side Hand Images without Regularization

In this section, we have shown the experimental results of the proposed model for the dorsal side hand image, where regularization is not used. Figures 5(a) and 5(b) show the training and validation loss and accuracy curves for this issue, respectively.



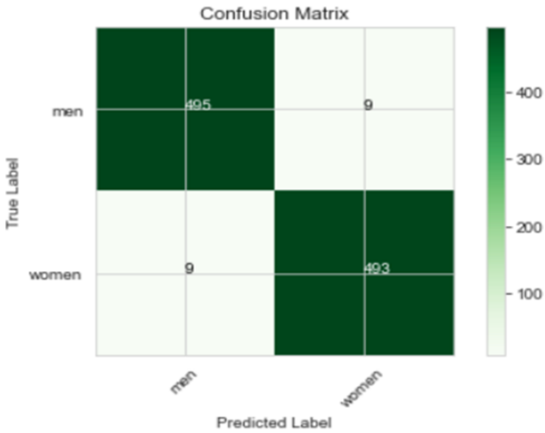
(a) *(b)*

Figure 5. *Loss and accuracy result of training and validation of dorsal side hand images without regularization: a) loss curve, and b) accuracy curve*

The classification report and confusion matrix for the model used for the dorsal side hand images without regularization are shown in Figures 6(a) and 6(b), respectively.

	precision	recall	f1-score	support
men (Class 0)	0.98	0.98	0.98	504
women (Class 1)	0.98	0.98	0.98	502
accuracy			0.98	1006
macro avg	0.98	0.98	0.98	1006
weighted avg	0.98	0.98	0.98	1006

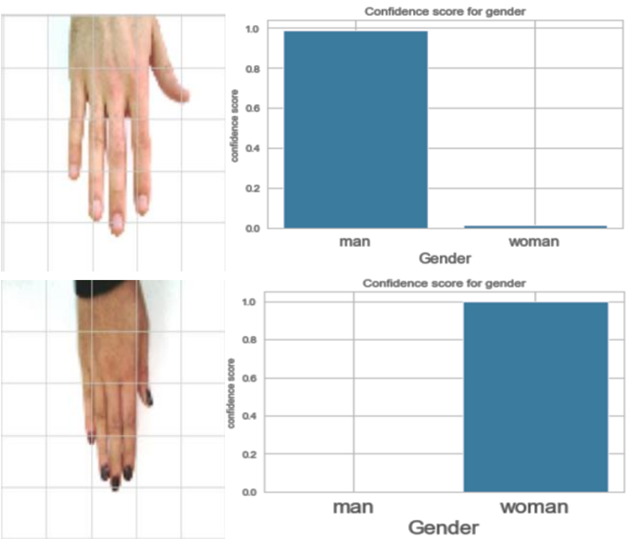
(a)



(b)

Figure 6. *Classification report and confusion matrix of dorsal side hand images without regularization: a) classification report, and b) confusion report*

The confusion matrix gives us a visual representation of correctly predicted data and incorrectly predicted data. For dorsal-side hand images, we can see that our non-regularized model correctly predicted 495 men's hands and 493 women's hands. The processed predicted image with the proposed model's prediction score in a bar chart is shown in Figure 7.

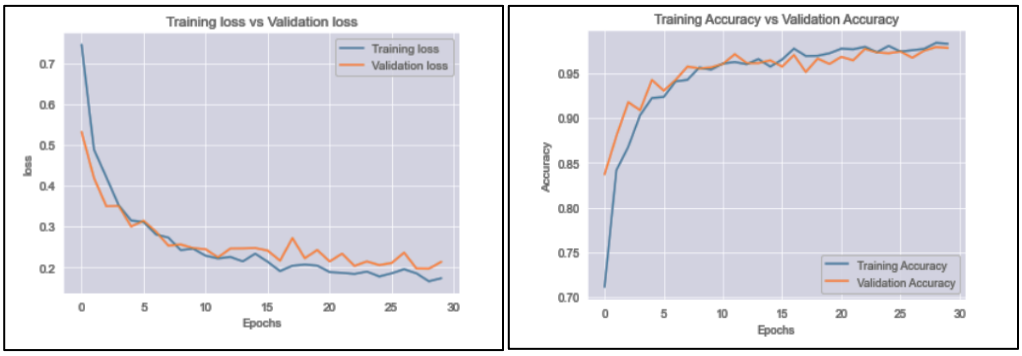


(a) (b)

Figure 7. *Processing example of dorsal side hand images without regularization: a) input image, b) predicted model score in bar chart*

4.2. Gender Classification for Dorsal Side Hand Images with Regularization

In this section, we have shown the processing example and experimental results of the proposed model for the dorsal side hand image, where regularization is used. Figures 8(a) and 8(b) show the training and validation loss and accuracy curves, respectively.



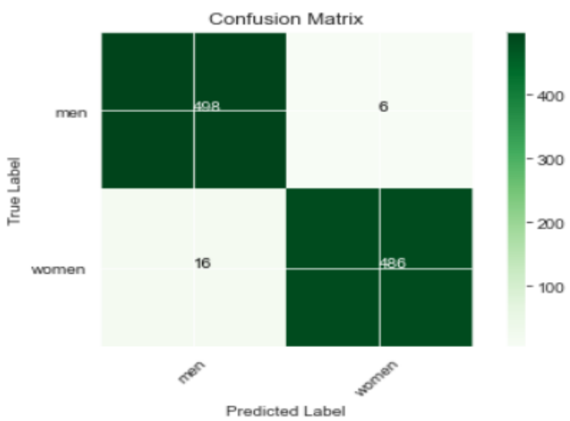
(a) *(b)*

Figure 8. *Loss and accuracy result of training and validation of dorsal side hand images with regularization: a) loss curve, and b) accuracy curve*

The classification report and confusion matrix for the model used for the dorsal side hand images without regularization are shown in Figures 9(a) and 9(b), respectively.

	precision	recall	f1-score	support
men (Class 0)	0.97	0.99	0.98	504
women (Class 1)	0.99	0.97	0.98	502
accuracy			0.98	1006
macro avg	0.98	0.98	0.98	1006
weighted avg	0.98	0.98	0.98	1006

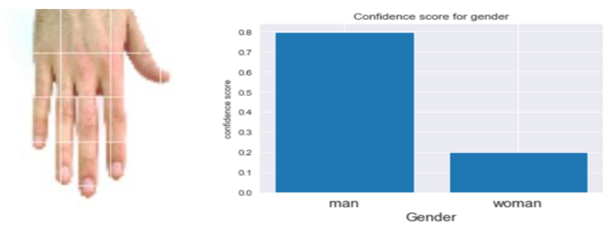
(a)



(b)

Figure 9. *Classification report and confusion matrix of dorsal side hand images with regularization: a) classification report, and b) confusion report*

For dorsal-side hand images, we can see that our regularized model correctly predicted 498 men's hands and 486 women's hands. The processed predicted image with the proposed model's prediction score in a bar chart is shown in Figure 10.



(a) *(b)*

Figure 10. *Processing example of dorsal side hand images with regularization: a) input image, b) predicted model score in bar chart*

4.3. Gender Classification for Palm Side Hand Images without Regularization

In this section, we have shown the processing example and experimental results of the proposed model for the palm side hand image, where regularization is not used. In the experiment, we set the epoch size to 30 and the batch size to 32. Figures 11(a) and 11(b) show the training and validation loss and accuracy curves, respectively.

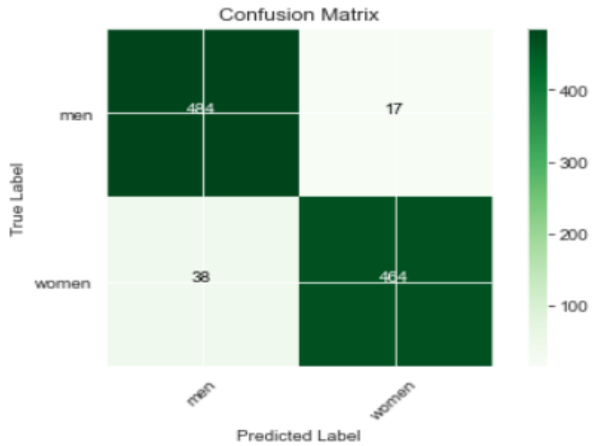
(a) *(b)*

Figure 11. *Loss and accuracy result of training and validation of the palm side hand images without regularization: a) loss curve, and b) accuracy curve*

The classification report and confusion matrix for the model used for the palm side hand images without regularization is shown in Figures 12(a) and 12(b), respectively.

	precision	recall	f1-score	support
men (Class 0)	0.93	0.97	0.95	501
women (Class 1)	0.96	0.92	0.94	502
accuracy			0.95	1003
macro avg	0.95	0.95	0.95	1003
weighted avg	0.95	0.95	0.95	1003

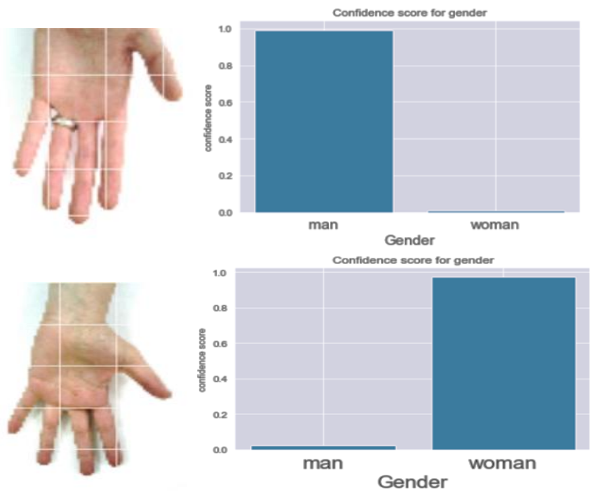
(a)



(b)

Figure 12. *Classification report and confusion matrix of palm side hand images without regularization: a) classification report, and b) confusion report*

In this section, we have shown the processing example and experimental results of the proposed model for the palm side hand image, where regularization is not used. In the experiment, we set the epoch size to 30 and the batch size to 32. The training and validation loss and accuracy curves are depicted in Figures 13(a) and 13(b), respectively.



(a) *(b)*

Figure 13. *Processing example of palm side hand images without regularization: a) input image, b) predicted model score in bar chart*

4.4. Gender Classification for Palm Side Hand Images with Regularization

In this section, we have shown the processing example and experimental results of the proposed model for the palm side hand image, where regularization is used. The loss and accuracy curve of the training and validation is presented in Figures 14(a) and 14(b), respectively.

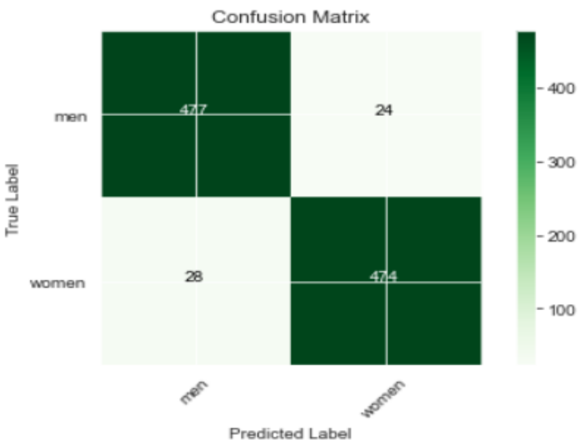
(a) *(b)*

Figure 14. *Loss and accuracy result of training and validation of palm side hand images with regularization: a) loss curve, and b) accuracy curve*

The classification report and confusion matrix for the model used for the palm side hand images with regularization is shown in Figures 15(a) and 15(b), respectively.

	precision	recall	f1-score	support
men (Class 0)	0.94	0.95	0.95	501
women (Class 1)	0.95	0.94	0.95	502
accuracy			0.95	1003
macro avg	0.95	0.95	0.95	1003
weighted avg	0.95	0.95	0.95	1003

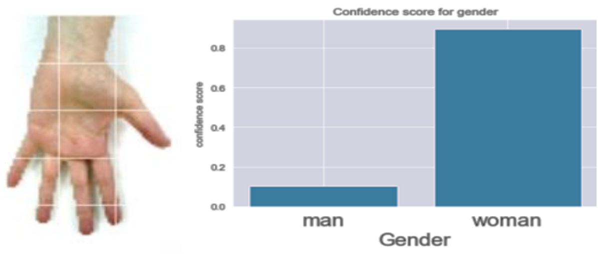
(a)



(b)

Figure 15. *Classification report and confusion matrix of palm side hand images with regularization: a) classification report, and b) confusion report*

For dorsal side hand images, we can see that our model, with regularization, has correctly predicted 477 men's hands and 474 women's hands. The processed predicted image with the proposed model's prediction score in a bar chart is shown in Figure 16.



(a) *(b)*

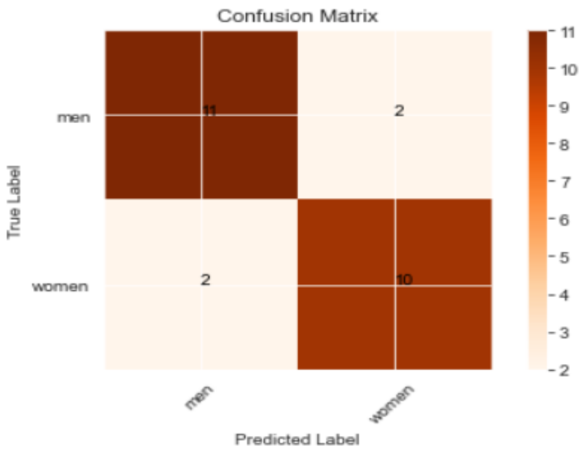
Figure 16. *Processing example of palm side hand images with regularization: a) input image, b) predicted model score in bar chart*

4.5. Gender Classification for Real Data of Palm and Dorsal Side Hand Images

In this section, we have shown the experimental results of the proposed model for the dorsal and palm side images of the hand. Here, we have collected 135 dorsal side hand images and 125 palm side hand images for men and women. The total men's and women's dorsal side hand images are 67 and 68, respectively. Furthermore, the total men's and women's palm side hand images are 65 and 60, respectively. For validation, we took 20% of total hand images, and the remaining 80% were used for training. The model is tested using the other real data that are not used in the training and validation. The total size of the test images are 40. The evaluation of the sample test images are shown in section 4.6. In the experiment, we set the epoch size to 30 and the batch size to 32. The classification report and confusion matrix for the model used for the palm and dorsal side hand images are shown in Figures 17 and 18, respectively.

	precision	recall	f1-score	support
men (Class 0)	0.85	0.85	0.85	13
women (Class 1)	0.83	0.83	0.83	12
accuracy			0.84	25
macro avg	0.84	0.84	0.84	25
weighted avg	0.84	0.84	0.84	25

(a)

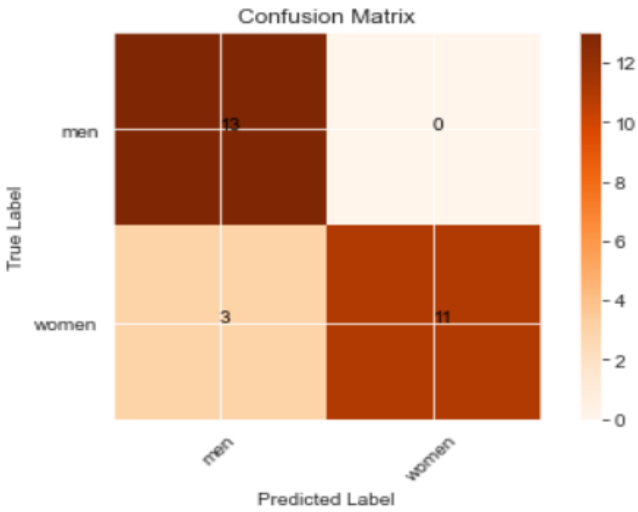


(b)

Figure 17. *Classification report and confusion matrix of palm side hand images for real time data: a) classification report, and b) confusion report*

	precision	recall	f1-score	support
men (Class 0)	0.81	1.00	0.90	13
women (Class 1)	1.00	0.79	0.88	14
accuracy			0.89	27
macro avg	0.91	0.89	0.89	27
weighted avg	0.91	0.89	0.89	27

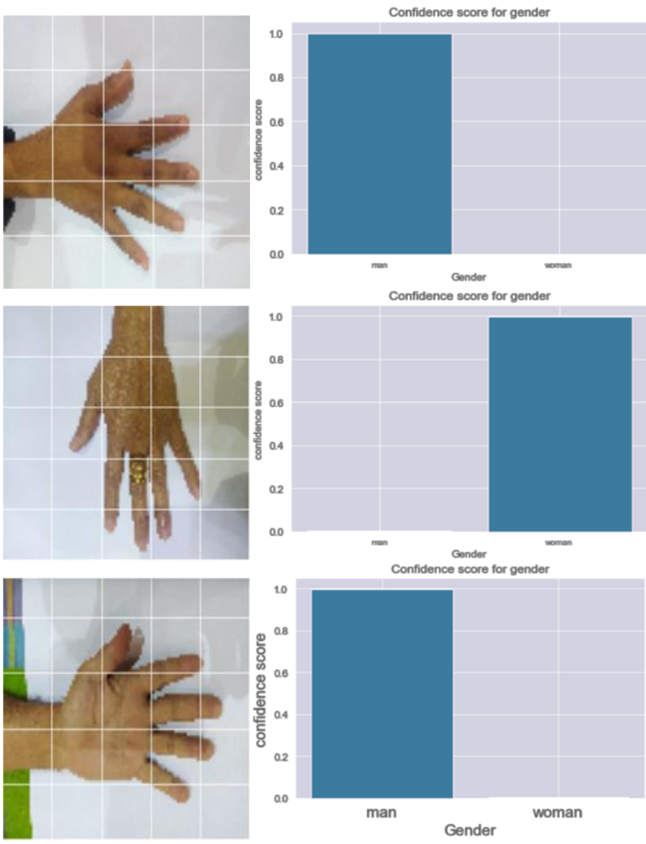
(a)

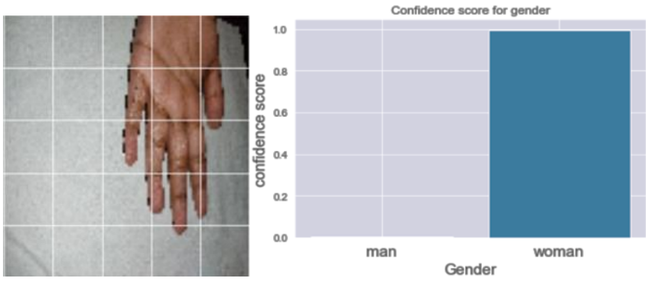


(b)

Figure 18. *Classification report and confusion matrix of dorsal side hand images for real time data: a) classification report, and b) confusion report*

The processed predicted image with the proposed model's prediction score in a bar chart is shown in Figure 19.



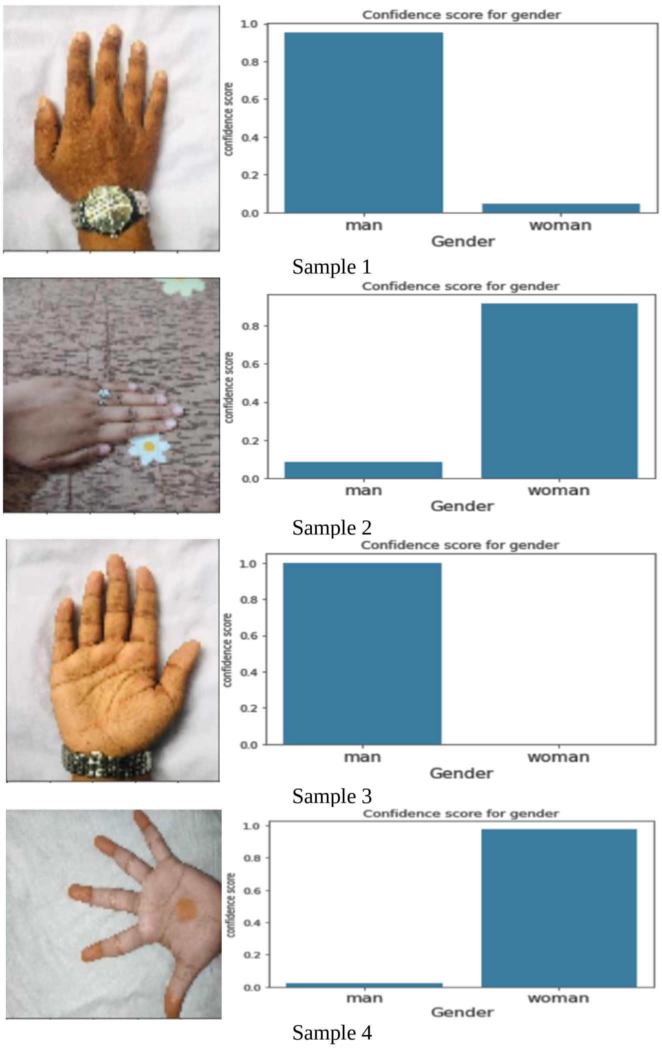


(a) *(b)*

Figure 19. *Processing example of dorsal and palm side hand images: a) input image, b) predicted model score in bar chart*

4.6. Classifying Gender by Hands that are Not in Dataset in Different Environments

In this section, some processing examples are presented to show the efficiency of the proposed model. The samples are taken from different environmental conditions and backgrounds. The processing examples are shown in Figure 20.



(a) *(b)*

Figure 20. *Processing sample examples of dorsal and palm side hand images with different environmental and background conditions: a) input image, b) predicted model score in bar chart*

From Figure 20, it is revealed that different samples are captured in real time under different environmental conditions. Sample 1 is captured from a dorsal side men's hand image where a watch is worn at the hand.

Sample 2 is taken from an environment where the background is not uniform and a ring is worn on the finger. Sample 3 and 4 are taken from palm site hand images, where the male hand image has a watch and the female hand image's fingers are a different color.

4.7. Implementation of ROC Curve to Compare Among Models

A graph is known as a receiver operating characteristic (ROC) curve when it shows us the performance of a model at different thresholds with two parameters: the false positive rate (FPR) and the true positive rate (TPR). The ROC curve is shown in Figure 21. In this figure, the comparison among the proposed models is shown. From the ROC curve, the AUC (area under the ROC curve) score clearly shows that the model of dorsal side hand images gives a better AUC score than the models of palm side hand images.

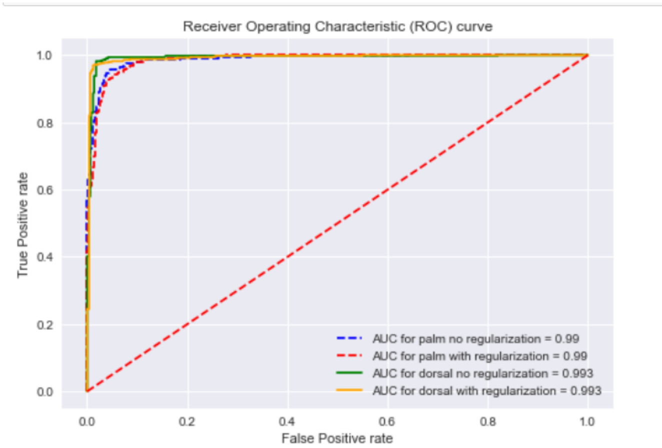


Figure 21. *ROC curve with AUC score of the proposed models*

4.8. Discussions

When we applied our model on the 11K hands dataset, we got some pretty good results. For the dorsal side, with and without regularization, both models give us 98% accuracy. For palm side hand images, proposed models with and without regularization give 95% accuracy. In comparison to the model presented in [13], the authors applied five different models to the dataset to perform the comparative analysis. From the experimental results, it is observed that among the five models, i.e., VGG-16, VGG-19, ResNet-50, Xception, and Inception-V3, Xception-V3 shows 87.35% and 92.62% accuracy for the palm and for dorsal side hand images, respectively, which are the best results among the models. The model used in [17] was also trained and tested with an 11K hand dataset. The model shows 94.31% accuracy both for the palm and dorsal side hand images. If we compare it with Afifi's model [12], our models worked better with respect to this model. Afifi's 2D-CNN model gave accuracy of 87.4% for the palm side hand images, and for dorsal side hand images, it gave 91% accuracy. Again, for real data of dorsal side hand images, the proposed model gives 89% accuracy. For palm-side real hand images, our proposed model gives 84% accuracy. The comparison tables of proposed and present state of the art methods are shown in Table 2.

From the table it is clearly revealed that the proposed model for 11K hands dataset performed better than any other model and was more precise. For palm side classification, it attained an impressive accuracy of 95.00%, and for dorsal side classification, it achieved an even greater accuracy of 98.00%. This demonstrates how well the proposed approach performs in precisely identifying hand motions and patterns. However, the proposed model's accuracy slightly declined when put to the test using real data. It attained a classification accuracy of 89.00% for the palm side and 84.00% for the dorsal side. Although the performance of the proposed model remained excellent, it is clear that real-world data presented different difficulties than the 11K hands dataset. The reason behind this is that the real time images are collected from different environmental conditions with various backgrounds. Another issue is that the model is trained with real time data, which is small in size and captured from different distances.

Table 2. *Comparison among the models*

Model	Palm side accuracy	Dorsal side accuracy
Mukherjee [13]	87.35%	92.62%
Lin [17]	94.31%	94.31%
Afifi [12]	87.40%	91.00%
Proposed model for 11K hands	**95.00%**	**98.00%**
Proposed model for real data	89.00%	84.00%

5. CONCLUSION

In this research work, we propose a three-layer shallow convolutional neural network for classifying gender from an image. In this work, different models are developed for testing the dorsal and palm sides of the hands. A model is developed for classifying gender from the real time data. However, all models are not performing equally well. According to the experimental results, the model that is being developed for the dorsal hand images is showing better results with respect to the other proposed models. If we compare accuracy with the present state of the art, our model shows better accuracy. Our model also shows a pretty good result for real-world hand images. In the future, this work can be extended to recognizing gender with biometric identification. Here is more scope to create a dataset with different environmental conditions and backgrounds.

CONFLICTS OF INTEREST

No conflict of interest was declared by the author.

REFERENCES

[1] Dantcheva, A., Elia, P., Ross, A., “What else does your biometric data reveal? a survey on soft biometrics”, IEEE Transactions on Information Forensics and Security, 11(3): 441–467, (2015).

[2] Dantcheva, A., Velardo, C., D’angelo, A., Dugelay, J.L., “Bag of soft biometrics for person identification”, Multimedia Tools and Applications, 51(2): 739–777, (2011).

[3] Reid, D.A., Nixon, M.S., Stevenage, S.V., “Soft biometrics; human identification using comparative descriptions”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(6): 1216–1228, (2013).

[4] Tome, P., Fierrez, J., Vera-Rodriguez, R., Nixon, M.S., “Soft biometrics and their application in person recognition at a distance”, IEEE Transactions on Information Forensics and Security, 9(3): 464–475, (2014).

[5] Sørensen, M.L.S., “Gender archaeology”, John Wiley & Sons, (2013).

[6] Menache, A., “Understanding motion capture for computer animation and video games”, Morgan Kaufmann, (2000).

[7] Dey, S., and Kapoor, A.K., “Sex determination from hand dimensions for forensic identification”, International Journal of Research in Medical Sciences, 3(6): 1466-1472, (2015).

[8] Aboul-Hagag, K.E., Mohamed, S.A., Hilal, M.A. and Mohamed, E.A., “Determination of sex from hand dimensions and index/ring finger length ratio in Upper Egyptians”, Egyptian Journal of Forensic Sciences, 1(2): 80-86, (2011).

[9] Martin, C., Werner, U., and Gross, H.M., “A real-time facial expression recognition system based on active appearance models using gray images and edge images”, In 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, 1-6, IEEE, (2008).

[10] Wu, M., Yuan, Y., "Gender classification based on geometry features of palm image", The Scientific World Journal, (2014).

[11] Amayeh, G., Bebis, G., Nicolescu, M., "Gender classification from hand shape", 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 1–7, IEEE, (2008).

[12] Afifi, M., "11K Hands: Gender recognition and biometric identification using a large dataset of hand images", Multimedia Tools and Applications, 78(15): 20835-20854, (2019).

[13] Mukherjee, R., Bera, A., Bhattacharjee, D., and Nasipuri, M., "Human gender classification based on hand images using deep learning", Artificial Intelligence: First International Symposium (ISAI 2022), Haldia, (2022).

[14] Baisa, N.L., Jiang, Z., Vyas, R., Williams, B., Rahmani, H., Angelov, P., Black, S., "Handbased person identification using global and part-aware deep feature representation learning", arXiv preprint arXiv:2101.05260, (2021).

[15] Bera, A., Bhattacharjee, D., "Human identification using selected features from finger geometric profiles", IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50(3): 747– 761, (2017).

[16] Bera, A., Bhattacharjee, D., Shum, H.P., "Two-stage human verification using handcaptcha and anti-spoofed finger biometrics with feature selection", Expert Systems with Applications, 171: 114583, (2021).

[17] Lin, Y.C., Suzuki, Y., Kawai, H., Ito, K., Chen, H.T., Aoki, T., "Attribute estimation using multi-cnns from hand images", 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 241–244, IEEE, (2019).

[18] Matkowski, W.M., Kong, A.W.K., "Gender and ethnicity classification based on palmprint and palmar hand images from uncontrolled environment", 2020 IEEE International Joint Conference on Biometrics (IJCB), 1–7, IEEE, (2020).

[19] Rim, B., Kim, J., Hong, M., "Gender classification from fingerprint-images using deep learning approach", International Conference on Research in Adaptive and Convergent Systems, 7–12, (2020).

[20] Yuan, Y., Tang, C., Xia, S., Chen, Z., Qi, T., "Handnet: Identification based on hand images using deep learning methods", 2020 4th International Conference on Vision, Image and Signal Processing, 1–6, (2020).

[21] Amayeh, G., Bebis, G., and Nicolescu, M., "Gender classification from hand shape", 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 1-7, IEEE, (2008).

[22] Wu, M., and Yuan, Y., "Gender classification based on geometry features of palm image", The Scientific World, (2014).

[23] Liliana, D.Y., and Utaminingsih, E.T., "The combination of palm print and hand geometry for biometrics palm recognition", International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS, 12(01), (2012).

[24] Dey, S., and Kapoor, A.K., "Sex determination from hand dimensions for forensic identification", International Journal of Research in Medical Sciences, 3(6): 1466-72, (2015).

[25] Chakrabarty, N., “A novel strategy for gender identification from hand dorsal images using computer vision”, 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 108-113, IEEE, (2019).

[26] Chai, T., Prasad, S., Wang, S., “Boosting palmprint identification with gender information using deepnet”, Future Generation Computer Systems, 99: 41–53, (2019).

[27] Jain, A., Kanhangad, V., “Gender classification in smartphones using gait information”, Expert Systems with Applications, 93: 257–266, (2018).

[28] Yaman, D., Eyiokur, F.I., Sezgin, N., Ekenel, H.K., “Age and gender classification from ear images”, 2018 International Workshop on Biometrics and Forensics (IWBF), 1–7, IEEE, (2018).

[29] Greco, A., Saggese, A., Vento, M., Vigilante, V., “A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff”, IEEE Access, 8: 130771–130781, (2020).

[30] Rawat, W., and Wang, Z., “Deep convolutional neural networks for image classification: A comprehensive review”, Neural Computation, 29(9): 2352-2449, (2017).

[31] Tareef, A., Al-Dmour, H., and Al-Sarayreh, A., “An Automated Deep Learning Framework for Human Identity and Gender Detection”, Journal of Advances in Information Technology, 14(1), (2023).

[32] Brownlee, J., "How to Use Data Scaling to Improve Deep Learning Model Stability and Performance", Machine Learning Mastery, (2021).

[33] Goodfellow, I., Bengio, Y., and Courville, A., “Deep Learning", MIT Press, (2016).

[34] Ahmad, I., and Khan, M., “Hand recognition using palm and hand geometry features”, LAP LAMsBERT Academic Publishing, (2017).

[35] Tani, T.B., Afroz, T., and Khaliluzzaman, M., “Deep Learning Based Model for COVID-19 Pneumonia Prediction with Pulmonary CT Images”, Computational Intelligence in Machine Learning: Select Proceedings of ICCIML 2021, 365-379, Singapore: Springer Nature Singapore, (2022).