International Journal of Advances in Intelligent Informatics
Vol. 8, No. 3, November 2022, pp. 324-336

Determination of gender from fingerprints using dynamic horizontal voting ensemble deep learning approach

Olorunsola Stephen Olufunso [a,1,*], Abraham Eseoghene Evwiekpaefe [a,2], Martins Ekata Irhebhude [a,3]

[a] Department of Computer Science, Nigeria Defence Academy, Kaduna, Nigeria
[1] stevenolorunsola@yahoo.com; [2] contact_abrahama@yahoo.com; [3] mirhebhude@nda.edu.ng
* corresponding author

ARTICLE INFO

Article history
Received October 1, 2022
Revised October 23, 2022
Accepted November 26, 2022
Available online November 30, 2022

Keywords
Fingerprints pattern
Deep learning
Classification
Gender
Soft biometric

ABSTRACT

Despite tremendous advancements in gender equality, there are still persistent gender disparities, especially in important human activities. Consequently, gender inequality and related concerns are a serious problem in our global society. Major players in the global economy have identified the gender identity system as a crucial stepping stone for bridging the enormous gap in gender-based problems. Extensive research conducted by forensic scientists has uncovered a unique pattern hidden in the fingerprint and these distinguishing characteristics of fingerprints can be utilized to determine the gender of individuals. Numerous research has revealed various fingerprint-based approaches to gender recognition. The purpose of this research is to present a novel dynamic horizontal voting ensemble model with hybrid Convolutional Neural Network and Long Short Term Memory (CNN-LSTM) deep learning algorithm as the base learner to automatically determine human gender attribute based on fingerprint pattern. More than four thousand Live fingerprint images were acquired and subjected to training, testing and classification using the proposed model. Result of this study indicated over 99% accuracy in predicting person's gender. The proposed model also performed better than other state-of-the-art model such as ResNet-34, VGG-19, ResNet-50 and EfficientNet-B3 model when implemented on SOCOFing public dataset.

1. Introduction

Despite significant progress in gender equality, there are still key gender gaps particularly in education, health, right to job opportunities and other basic needs of livelihood. Case instance is in Afghanistan under the control of the Taliban, there have been news of violation against women and girls' rights. There are evidences that the Taliban will implement policies that will restrain women and girl from accessing education, confinement to their home or even denying them access to most jobs [1]. Hence, gender inequality and other associated issues are major problem in our world. International donor agencies, such as World Bank, European Union, African Development Bank (AfDB) and many others, identifies gender identity system as a vital stepping-stone for the female, particularly, as a means of empowering and given them access to peculiar services and other privileges as a citizen [2]. An effective gender identification system has been discovered to be an important enabler for attaining a number of key development results toward eradicating gender inequality, poverty and financial exclusion.

The good thing is that there are several methods to verify the identity of people. However, biometric system offers a better approach for personal identification with numerous gains over other methods [3].

Biometric system as the science concerning the numerical analysis of human characteristics [4]. Biometric systems are usually built to recognize a particular human trait such as hand geometry, voice, fingerprint, face, iris, DNA, gait, palm print and keystroke dynamics [5]. Some of these traits have been used in gender based classification researches. In ealier study, Machine learning method was used in identifying gender through facial images [6], [7]. Fingerprints, however, have been acclaimed to be the most known and used biometric solutions to authenticate people in biometric systems [8]. Fingerprint system is a keenly researched area in biometric technologies [9]. It is one of the most well-known biometrics across the globe for personal recognition and identification of computerized systems. Fingerprint identification is popular because of its uniqueness and inherent ease in acquisition [4]. It has been successfully applied in law enforcement and forensics to identify suspects. Fingerprints are composed of many ridges and furrows, and they are a unique marker of a person as no two identical fingerprints have ever been found, that's why we have the statement that every individual has a unique impression of fingerprints and that makes fingerprint based technology widely accepted and much desirable system [10]. From the fingerprint pattern it has been discovered that other ancillary information (e.g., gender, age, race) can be determined [11]. This ancillary information is called soft biometrics. They are soft because they are not sufficient enough to uniquely identify individuals, but can be used to complement the identity information offered by the primary biometric identifiers.

Results from an extensive research study by forensic scientist have discovered unique pattern embedded in the fingerprint and have concluded that a closer look at the minutia of the fingerprint can provide a clue to a person's gender and other rich information about an individual [12]. These unique features of fingerprints can be used in differentiating between individuals by their gender. The earliest mentioning of soft biometrics such as gender was aimed at using it to filter large biometric databases, to limit the number of searches. Detecting particular fingerprint in a large database during fingerprint identification process usually demand high computational complexity with respect to time and hardware resources. However, the search can be aided by knowing, for example, the gender of the individual involved in the search. The field of soft biometrics is increasingly gaining attention over traditional biometrics and recent researches are tilting toward the field as potential replacement for traditional biometric [13]. Reasons behind this development is due to the non-intrusiveness and seamless approach for recognition of soft biometrics features.

Many studies have demonstrated different approaches to human gender recognition based on fingerprint pattern. Recently, ResNet-34 [10], VGG-19, ResNet-50 and EfficientNet-B3 model [14] have been used for automatic fingerprint recognition. The deep learning has greatly improved both precision and detection efficiency when used in image classification applications [15], [16], [17],[18]. Early approach of Neural Network to gender classification based on fingerprint images was proposed by [19]. The researchers analyzed a dataset of 10 fingerprint images for 2200 people of different gender and age using feature extraction techniques. Neural networks, Fuzzy C-means, and linear discriminant analysis were used, and classification results of 88.5%, 80.39%, and 86.5% were obtained, respectively. In [20], an alternative study that aimed at determining the gender based on finger ridge count within a distinct region of Southern India was presented. A sample of 550 rolled-finger fingerprints was taken, in which 275 were male and female each, all of them in the age limits of 18 - 65 years. The application of Bayes' theory in this study revealed that a fingerprint possessing ridge density less than 13 ridges per 25 mm2 is most likely to be of male. Likewise, a fingerprint having a ridge count greater than 14 ridges per 25 mm2 is most likely to be of female. The outcomes of the research indicated that women of South Indian origin have significantly higher ridge count (mean = 14.14) when compared to men (mean = 12.57). The females have an ominously greater ridge count than their male counterparts.

Similar study presented the different techniques for gender identification using fingerprints [21]. The study proposed a novel method to evaluate gender by examining fingerprint using Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT) and Power Spectral Density (PSD). The study employed a dataset of 220 persons within different age and gender and identifies the frequency domain approaches as the most efficient method, because of its flexibility and less computation time required when compared with the spatial domain approaches. Result shows 90% matches for female samples and

79% matches for male samples. Research based on wavelet transforms and Singular Value Decomposition (SVD) methods for fingerprint based gender classification was conducted by [22]. A sample of 3570 fingerprints was taken, in which 1980 fingerprints for males and 1590 fingerprints for females. Authors proposed a new method for classifying gender based on fingerprint images, by using DWT and SVD. The techniques considered the spatial features of the singular value decomposition, which includes the internal structure of the fingerprint images and the frequency features of the wavelet domain which improved performance in gender classification. The result indicated that the female fingerprints show significantly higher classification rates than those of males. Gender classification of 91.67% and 84.69% for male and females, respectively.

In the region of Uttarakhand, fingerprint was used to determine gender [23]. The fingerprints were obtained from125 males and 125 females between the ages of 18-60 years. The mean value of the ridges of the fingerprints collected was calculated. The results indicated that there are significant differences in epidermal ridge density between males and females and also supports the construct that the ridge density of women's fingerprints is statistically significantly greater than men. The outcome of the experiment shows that fingerprint ridge of <12 ridges/25 mm2 are more probable of male while fingerprint ridge of >14 ridges/25 mm2 are more probable of female. A novel method to classify gender from fingerprint using two combined methods of wavelet transformation to extract fingerprint features and pass the output to back-propagation neural network algorithms for the final gender classification was presented by [24]. The experimentation was performed using a fingerprint database of 275 male and 275 female fingerprints and obtained a classification accuracy of 91.45%. In investigating the relationship between the fingerprint ridge densities and the gender of a person, [25] conducted an experiment on the population of persons who were born and lived in northern Malaysia. The sample of the experiment consists of 25 males and 25 females, totaling 50 participants selected from the age group of 18-60 year. The results of the study further supported the claim that that fingerprint ridges of <12 ridges/25mm2 is probable for male and fingerprint ridges >14 ridges/25mm2 is likely to be for female. The outcome of the experiment concluded that in Malaysia also, women tend to have a greater ridges density compared to men.

Further research to understudy and analyzed the fingerprints and gender prediction among medical students of Nepal medical college and teaching hospital [26]. The sample size for the study was 200 medical students (100 males and 100 females) with their ages ranging from 18-25 years. Using the Henry Classification System, the system that allocate individual finger number according to the position in the hand starting with the right thumb and ending with the left little finger as number ten. Fingerprint patterns are classified into Loop (60-70%); Whorl (25- 35%); Arch (5%-7%) and Composite (2- 3%). The results of the study indicated that the loops fingerprint pattern occurred most while the arches fingerprint patterns were the least common. The result further indicated that the males have a higher frequency of loops and females have a higher frequency of whorls. Application of deep learning approach for gender classification based on fingerprint was carried out by [10]. The authors used a pre-trained deep Convolutional Neural Networks (CNNs) for classification. Transfer learning based on ResNets-34 model was employed to train the network (CNN). The experiment was carried out on the publicly available SOCOFing dataset. The CNN attained an accuracy of 75.2% for the classification of gender.

Related study on deep learning approach for gender classification was conducted and 62% performance accuracy on the level of a single minutia was obtained [27]. Also, a study on fingerprint pattern distribution in all ten fingers among Malaysian Tamils for gender determination for use in crime scene investigation [28], was carried out. The study used 280 adult Malaysian Tamils living mostly in Peninsular Malaysia. Out of 280 study subjects, 140 are males and 140 are females with age range of 18 to 55 years. Using inking plate, the fingerprints were examined under a magnifier and the patterns and it indicated that the males had a dominant arch pattern in both hands and left hand comparatively shows higher frequency than the right while the females possessed a higher frequency of whorl pattern compared to males. The result showed 56%, 30.8% and 13% for loop pattern, whorl and least arch pattern respectively. Latest work by [14] employed the use of deep learning methods for gender classification from fingerprint image. The network trained end-to-end learning of 8,000 images. The network was validated on 1,520 images with 360 images for the testing. State-of-the-art models particularly VGG-19, ResNet-50 and EfficientNet-B3 model were used to train from scratch. Among the three models, EfficientNet-B3 model attained the best accuracy of 97.89% for training, 69.86% for validating and 63.05% for testing. Similarly, [29] conducted a study on human gender classification based on each of the five finger type and carried out the performances evaluation of the various trained deep learning models. The study also revealed performance improvement from fusion amongst the finger types. The overall gender classification accuracies of 91.3% was reported for the best proposed fusion scheme.

Summarily, most of the deep learning approaches adopted for classification are complex and sophisticated models, some of which were pre-trained on huge dataset. The complex structure translates into high training cost time and lots of computational resources. There is the need for improvement in model performance accuracy with moderate dataset with simpler architecture and with less computational time. This paper aims to propose a novel dynamic horizontal voting ensemble system with hybrid Convolutional Neural Network and Long Short Term Memory (CNN-LSTM) algorithm as the base learner to automatically recognize human gender trait based on fingerprint pattern. The advantage of LSTMs is in the ability to selectively recollect patterns for an extensive period of time while CNNs are excellent in feature extraction. This makes the hybrid model better than other state-of-the-art models such as CNN, RCNN, Fast RCNN, Faster RCNN and LSTM. Hence the motivation to use Deep CNN-LSTM network in this study. The dataset used for this paper was acquired through a live scan and subjected to preprocessing in order to reduce the noise. The remainder of this article is organized as follows: Section II defined the dataset used and methodology adopted for the study; section III discussed the results obtained from the experiment. Finally, the conclusions and future works are described in section V.

2. Method

Gender classification based on fingerprints is a complex technique; therefore, a robust machine learning approach must be designed for this task. This section elaborates on the design methodology and model analysis. Fig. 1 shows the framework of the proposed Dynamic Horizontal Voting Ensemble (DHVE) approach. Four phases are involved in the implementation; first data collection and preprocessing. Second phase, model development and training. The final two phases are the dynamic ensemble selection and the prediction phase.

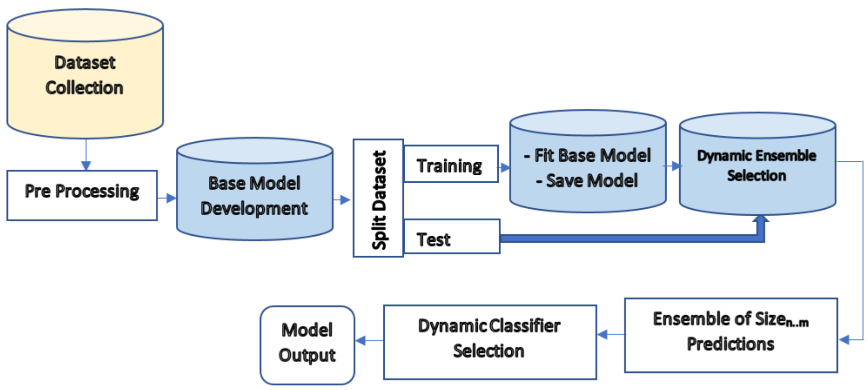

Fig. 1. System Architecture of the Proposed DHVE model for Soft Biometric Classification

2.1. Data Collection

In this study, a total of 450 Nigerian participants had their fingerprints scanned. The ten (10) fingerprints of each individual were acquired, resulting in a total sample size of 4,500 images out of which 2470 are for male subjects while 2030 belong to female subjects. To provide for dataset balance and fairness, only 2030 subset images from male subjects were utilized for the gender classification experiment show in Table 1. The fingerprint images were labeled with features such as image ID, gender, age group, ethnicity, and finger type labels (i.e. the thumb, index, middle, ring, and little finger labels), as depicted in Fig. 2.

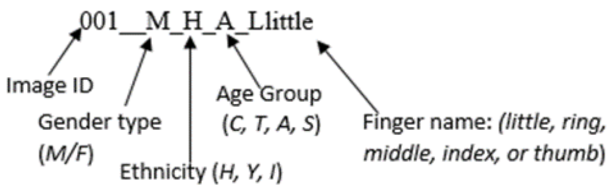

Fig. 2. Fingerprint image attribute

The classification of gender is a two-class (binary) problem: male and female

Table 1. Dataset distribution according to gender

Biometric Class	Training Set	Test Set	Total
Male	1,624	406	2,030
Female	1,624	406	2,030
Total	3,248	812	4,060

2.2. Data Preprocessing

The dataset was first preprocessed using the histogram equalization technique. This was accomplished by evaluating all of the grey levels steadily over the image's histogram [30], [31]. Next, a bilateral filter was applied to further improve the quality of the fingerprint images. Typically, this filter is utilized for image smoothing and denoising while preserving edges. Bilateral filter is an improvement on the Gaussian filter, which frequently obscures vital edge information because it blurs everything, regardless of whether it is noise or an edge. The formula for Gaussian blurring can be expressed as follows:

$$GB[I]_p = \sum_{q\in S} G_\sigma(\parallel p - q \parallel)\ I_q \qquad (1)$$

Where Gσ(x) denote the 2D Gaussian kernel. Gaussian filtering works by calculating the pixel intensity-weighted average of positions next to each other in a way that decreases weight based on the distance to the midpoint p. Pixel q is defined by the Gaussian Gσ(||p − q||), where σ is the element describing the neighborhood size. The bilateral filter, denoted by BF[I], is defined by:

$$BF[I]_p = \frac{1}{W_P} \sum_{q\in S} G_{\sigma_s}(\parallel p - q)\ G_{\sigma_r}\left(I_p - I_q\right) I_q \qquad (2)$$

where Wp is the normalization parameter to ensure pixel weights sum to 1.

$$W_p = \sum_{q\in S} G_{\sigma_s}(\parallel p - q)\ G_{\sigma_r}\left(I_p - I_q\right) I_q \qquad (3)$$

Parameters σs and σr will define the degree of filtering for the image I in equation (2).

2.3. Base Model Architecture

The hybrid Deep Convolutional Neural Network-Long Short Term Memory (Deep CNN-LSTM) model was used as the base model for this experiment show in Fig.3. The structure of the CNN model is made up of two convolutional layers, two maxpooling layers, and two fully-connected layers. The CNN and LSTM models work together by changing the shape of the CNN's output into (batch size, H, W*channel), where H and W are the image's height and width, respectively. This will lead to 3D data that the LSTM layer will use. The reshape sub-routine is activated by the lambda function. The LSTM model is connected to the dense and softmax activation function layers that are used for the final prediction. Each of the used LSTM layers were made with 16 and 96 units respectively. For final classification, the output of the LSTM layer is fed into the fully connected (FC) output layer with a Softmax activation

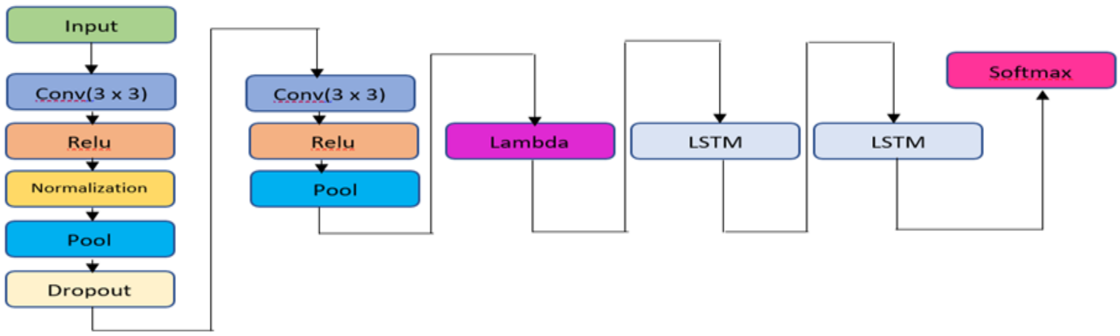

Fig. 3. Architecture of CNN-LSTM

2.4. Proposed Dynamic Selection Scheme for Horizontal Voting Ensemble

In the current horizontal voting ensemble method, ensemble members are chosen on purpose from a predetermined point during a single run. In contrast, the proposed dynamic ensemble selection approach builds a horizontal voting ensemble for prediction by dynamically selecting proficient models based on validation accuracy measure during base learner training on the training set. This method lets the best-performing models be part of the prediction ensemble. In this study, the base model was optimized for 150 epochs with a mini-batch size of 128 and a validation split of 0.2. In Algorithm 1, the proposed method for dynamically selecting the ensemble member for the horizontal voting ensemble is shown. Each training epoch is saved if the model's accuracy is equal or greater than the predetermined threshold set. Then, the best subset of the saved models is chosen based on the required ensemble size needed to make prediction. Algorithm 1 show in Fig. 4.

Input
Data: $D_{set} = D_{Trn} \cup D_{Test}$, $D_{Trn} \cap D_{Test} = \emptyset$
Parameters: N, n
Initialize K_{member}, E_n = [ ]
Initialize *threshold*
Procedure
    for all i in range N:
        Use D_{Trn} for one epoch training
        if $Model_{Accuracy} \geq threshold$:
          model.save(filename(i))
    end for
K_{member} = load_all_model saved
Sort K_{member} by $Model_{Accuracy}$
for i in range (1, n):
    $E_n = E_n \cup K_{member}$ (i)
end if
Output: E_n, n

Fig. 4. Algorithm Dynamic ensemble selection phrase

Algorithm 2 demonstrates the process of dynamically choosing a classifier for the final prediction. An ensemble member formed by algorithm1 was delivered as input to algorithm2. A maximum epoch of 150 was utilized during the experiment, with an ensemble size ranging from 1 to 50. As a result, algorithm 2 provides the general algorithm for the dynamic selection approach to horizontal voting ensemble. The dynamic approach was used at two key points in the development of the model. First, when choosing which model to form part of the ensemble, and second, in determining the best prediction score from either a single classifier or the ensemble score. Algorithm 2 show in Fig. 5.

Input: Trained n members models for voting ensemble E_n= {e_1,..e_n} and evaluate on
test datasets D_{test} = {D_{t1},..,D_{t5}}
Ensemble score En_{score},
Initialization: Preds, $Model_{scores}$ = [ ]
Initialization: $Model_{max}$, Subset, Single_Score
Procedure
for all i in range n:
Subset = En [:i]
if iteration <= i then
Put D_{test} into Subset(iteration), get softmax output vector $pred_i$
Add $pred_i$ to Preds
Iteration = iteration + 1
end if
Pred = $\sum_{predi}^{predi \in Preds} predi$
En_{score} = argmax(Pred)
for j in range (1, i+1):
Single_Score = E_n[j-1]. predict(D_{test})
$Model_{scores}$ = $Model_{scores}$ $\bigcup$ Single_Score
end for
Iteration = 1
end for
$Model_{max}$ = max($Model_{scores}$)
If $Model_{max}$ outperform En_{score} then En_{score} = $Model_{max}$
Output: En_{score}

Fig. 5. Dynamic horizontal voting ensemble

3. Results and Discussion

The classification performance of the model for gender classes Male and Female is shown in Table 2 respectively.

Table 2. Classification performance DHVE model

Fingerprints	Precision	Recall	F1-Score	Support
Female	0.99	0.98	0.99	406
Male	0.98	0.99	0.99	406

Both classes Female and Male have Precision, Recall and F1-score that is close to 1. This result shows that most positive samples and non-positive samples are classified correctly (for Precision values), while most True Positives were identified and classified correctly (for Recall Scores). The F1 score of > 0.5 imply that the model responds well to imbalance data. The proposed model gives an overall accuracy of 99% as shown in Table 3.

Table 3. Overall Classifications Performance of Proposed Model

Classification Parameter	Precision	Recall	f1-score	support
Accuracy			0.99	812
macro avg	0.99	0.99	0.99	812
weighted avg	0.99	0.99	0.99	812

Fig. 6 highlights the confusion matrix of the proposed model. Correct prediction of 402 and 399 were made for the male and female gender respectively. The model however, wrongly classified 7 of the Female gender as Male while 4 of Male fingerprint were wrongly classified as Female. The confusion matrices show the classification performance for each classes.

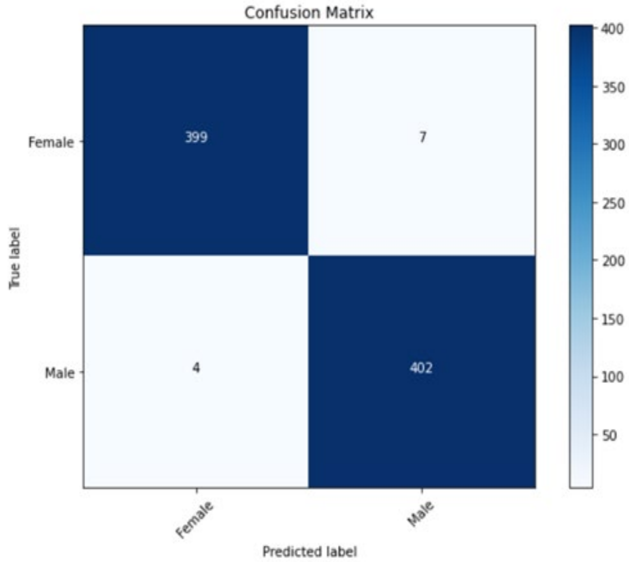

Fig. 6. Confusion Matrix of the Proposed Model

Fig. 7 shows the Receiver Operator Characteristics (ROC) Curve of the proposed DHVE Model. It shows the Probabilistic Curve that plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various thresholds values. The orange line indicates the classifier's ROC curve, and the blue line the point at which TPR = FPR. The Area Under the Curve (AUC) value was near perfect for gender classes under consideration with the ROC curve closer to the coordinate (0,1) at value of 0.99 accuracy.

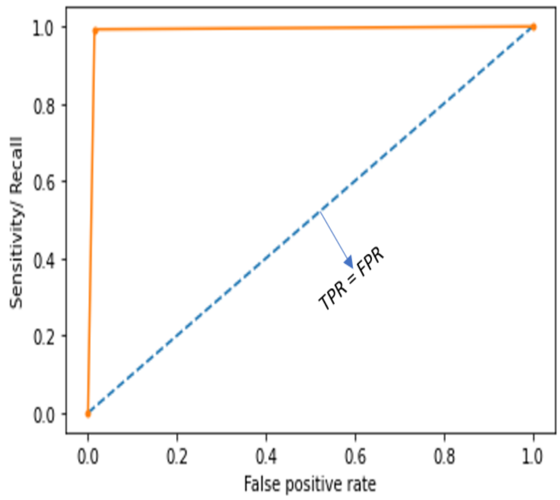

Fig. 7. The Receiver Operator Characteristics (ROC) Curve of the proposed DHVE Model.

Fig. 8 shows the performance comparison of the proposed DHVE model with CNN-LSTM and HVE model. In Fig. 8, the Line Plot summarizing the performance of the base model (CNN-LSTM) in blue colour, horizontal voting ensemble in green colour and the proposed dynamic horizontal voting ensemble model in orange colour. With an ensemble size of 1-50 members, the average accuracy for CNN-LSTM, HVE and DHVE models are 0.977, 0.982 and 0.984. The proposed method result (i.e. Orange line) has 0.984 accuracy score and the flattened from ensemble size 20 shows the model stability in terms of model's sensitivity to fluctuations in the data.

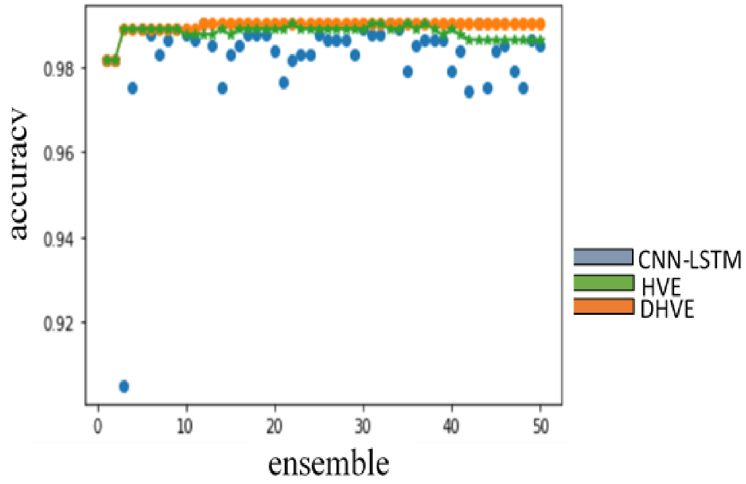

Fig. 8. Line Plot showing DHVE and other Model Performance for Gender Classification

3.1. Comparison with classical networks

To further validate the generalization and classification performance of the proposed model, it was trained on a publicly available dataset, the SOCOFing dataset, and the results were compared to those of other models trained on the same dataset. Previous experiment on the same public dataset includes; ResNet-34 [10], Systematic Pixel Counting (SPC) [32], Deep Convolutional Neural Network (DCNN) [29], Dense Dilated Convolution (DDC)-ResNet [33]. The result of existing models in comparison with the proposed model (DHVE), after training is shown in Table 4.

The 6000 images in the publicly available Sokoto Coventry Fingerprint Dataset include 4770 images of male participants and 1230 images of female participants [34]. To give a fair comparison, we only use a subset of 1230 images of male participants and all available images of females. Moreover, this prevents the model from favoring any particular class and reduces the likelihood of overfitting. The dataset was splited 80, 20 for training and testing respectively. Fig. 9 shows the confusion matrix for the proposed model applied to the SOCOFing dataset. The correct predictions for males and females were 121 and 120, respectively. However, the model incorrectly classified three female fingerprints as male and two male fingerprints as female.

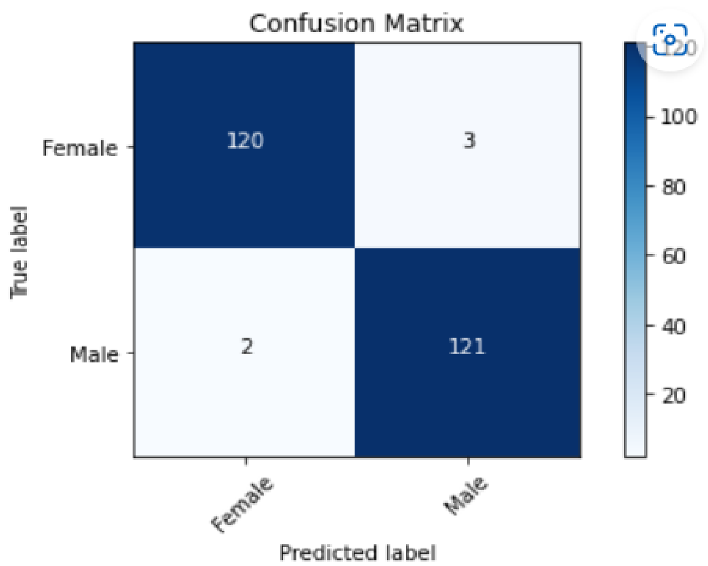

Fig. 9. Confusion Matrix for DHVE model on SOCOFing dataset

Table 4. Classification Performance of DHVE model on SOCOFing Dataset

Classification Parameter	Classes	Precision	recall	f1-score	Support	Classes
	Female	0.98	0.98	0.98	123	
	Male	0.98	0.98	0.98	123	
Accuracy				0.98	246	Accuracy
macro avg		0.98	0.98	0.98	246	macro avg
weighted avg		0.98	0.98	0.98	246	weighted avg

The Comparison with other existing models on the same dataset is shown in Table 5.

Table 5. Comparison of the proposed with existing results

Author(s)	Approach	Model Performance Accuracy (%)
[10]	ResNET-34	75.2
[32]	Systematic Pixel Counting	93.3
[29]	DCNN – based	91.3
[33]	DDC-ResNet	96.5
Proposed Model	**DHVE**	**98.0**

The main results of this study can be summed up as follows: first, the proposed model has performed better than the benchmark study on the SOCOFing dataset [10]. Most existing methods used single precitive methods. However, ensembles are utilized to produce greater predictive performance than a single predictive model on a predictive modeling challenge. This is accomplished by adding bias to the model, which reduces the variance component of the prediction error (i.e. in the context of the bias-variance trade-off). Hence, a major reason our proposed method obtained substantial improvement over existing methods. Performance accuracy has improved from 75.2% to 98% on the same dataset. Secondly, it is an improvement on other existing model trained on same dataset. Thirdly, the proposed model architecture is simpler compare to other complex architecture model such as RESNET-34 used on same dataset. A close examination of Table. 5 reveals that the finding achieved by the proposed model is comparable to those of the existing state-of-the-art model. Despite the complexity of transfer learning-based methods, the proposed model outperforms them with its high level of accuracy and simplicity. Clearly, ResNet and DCNN-based models have a significantly higher computational complexity than the proposed method. Table. 6 provides a comparison of the number of trainable parameters of the most typical deep architectures employed in the study. The cost and time-effectiveness of the proposed model is clearly indicated in Table. 6.

Table 6. Number of trainable parameters of the deep architectures used in the studies

Model	Trainable Parameters
ResNET-34	21.5M
DCNN	7.2M
DDC-ResNet	12.8M
Proposed DHVE	1.7M

4. Conclusion

This study has presented a model on the subject of human gender classification based on fingerprint. We highlighted the importance and the need for more simplified but effective deep learning architecture compare to complex pre-trained architectures. In this paper, we have presented a simpler deep learning ensemble model that is able to perform effectively with other state-of-the-art models. The architecture of the model includes image enhancement techniques by applying histogram equalization and bilateral filter before feeding the fingerprint images into the deep learning model for feature extraction and classification. This paper has improved on the benchemark work on SOCOFing dataset. The proposed deep learning model achieved more than 21% higher accuracy when compared with the pre-trained method used in the benchmark study on SOCOFing dataset. Our future work will include performance analysis of each of the various finger types in predicting human gender classification.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] H. Barr, "List of Taliban Policies Violating Women's Rights in Afghanistan | Human Rights Watch." (accessed Jan. 02, 2020). Available: https://www.hrw.org/news/2021/09/29/list-taliban-policies-violating-womens-rights-afghanistan.

[2] Mahmood, F. G. Forero, J. Jensson, S. K. Davidar, D. Abundo, M. L. Brixi, H. Kucey, and Andrea "Partnering for Gender Equality : Umbrella Facility for Gender Equality - Annual Report 2021." (accessed oct. 21, 2022). Available: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/227871639542714919/partnering-for-gender-equality-umbrella-facility-for-gender-equality-annual-report-2021.

[3] M. A. Alsmirat, F. Al-Alem, M. Al-Ayyoub, Y. Jararweh, and B. Gupta, "Impact of digital fingerprint image quality on the fingerprint recognition accuracy," *Multimed. Tools Appl.*, vol. 78, no. 3, pp. 3649–3688, Feb. 2019, doi: 10.1007/S11042-017-5537-5/METRICS.

[4] X. Yin, Y. Zhu, and J. Hu, "Contactless Fingerprint Recognition Based on Global Minutia Topology and Loose Genetic Algorithm," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, no. 1, pp. 28–41, 2020, doi: 10.1109/TIFS.2019.2918083.

[5] Y. Xu, G. Lu, Y. Lu, and D. Zhang, "High resolution fingerprint recognition using pore and edge descriptors," *Pattern Recognit. Lett.*, vol. 125, pp. 773–779, Jul. 2019, doi: 10.1016/j.patrec.2019.08.006.

[6] R. Golwalkar and N. Mehendale, "Masked Face Recognition Using Deep Metric Learning and FaceMaskNet-21," *SSRN Electron. J.*, pp. 1–8, Nov. 2020, doi: 10.2139/ssrn.3731223.

[7] M. E. Irhebhude, A. O. Kolawole, and H. K. Goma, "A Gender Recognition System Using Facial Images with High Dimensional Data," *Malaysian J. Appl. Sci.*, vol. 6, no. 1, pp. 27–45, Apr. 2021, doi: 10.37231/myjas.2021.6.1.275.

[8] H. B. Kekre and V. A. Bharadi, "Finger-Knuckle-Print verification using Kekre's wavelet transform," in *Proceedings of the International Conference & Workshop on Emerging Trends in Technology - ICWET '11*, 2011, p. 32, doi: 10.1145/1980022.1980030.

[9] Y. Faridah, H. Nasir, A. K. Kushsairy, S. I. Safie, S. Khan, and T. S. Gunawan, "Fingerprint Biometric Systems," *Trends Bioinforma.*, vol. 9, no. 2, pp. 52–58, Sep. 2016, doi: 10.3923/tb.2016.52.58.

[10] Y. I. Shehu, A. Ruiz-Garcia, V. Palade, and A. James, "Detailed Identification of Fingerprints Using Convolutional Neural Networks," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2018, pp. 1161–1165, doi: 10.1109/ICMLA.2018.00187.

[11] Z. Ezzati Khatab, A. Hajihoseini Gazestani, S. A. Ghorashi, and M. Ghavami, "A fingerprint technique for indoor localization using autoencoder based semi-supervised deep extreme learning machine," *Signal Processing*, vol. 181, p. 107915, Apr. 2021, doi: 10.1016/j.sigpro.2020.107915.

[12] S. Kloppenburg and I. van der Ploeg, "Securing Identities: Biometric Technologies and the Enactment of

Human Bodily Differences," *Sci. Cult. (Lond).*, vol. 29, no. 1, pp. 57–76, Jan. 2020, doi: 10.1080/09505431.2018.1519534.

[13] B. Hassan, E. Izquierdo, and T. Piatrik, "Soft biometrics: a survey," *Multimed. Tools Appl.*, pp. 1–44, Mar. 2021, doi: 10.1007/s11042-021-10622-8.

[14] B. Rim, J. Kim, and M. Hong, "Gender Classification from Fingerprint-images using Deep Learning Approach," in *Proceedings of the International Conference on Research in Adaptive and Convergent Systems*, Oct. 2020, pp. 7–12, doi: 10.1145/3400286.3418237.

[15] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/S13244-018-0639-9/FIGURES/15.

[16] L. Zhang *et al.*, "Evaluation and Implementation of Convolutional Neural Networks in Image Recognition," *J. Phys. Conf. Ser.*, vol. 1087, no. 6, p. 062018, Sep. 2018, doi: 10.1088/1742-6596/1087/6/062018.

[17] C. Tang, Q. Zhu, W. Wu, W. Huang, C. Hong, and X. Niu, "PLANET: Improved Convolutional Neural Networks with Image Enhancement for Image Classification," *Math. Probl. Eng.*, vol. 2020, pp. 1–10, Mar. 2020, doi: 10.1155/2020/1245924.

[18] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Dec. 2021, doi: 10.1186/s40537-021-00444-8.

[19] J. Kim, B. Rim, N. J. Sung, and M. Hong, "Left or Right Hand Classification from Fingerprint Images Using a Deep Neural Network," *Comput. Mater. Contin.*, vol. 63, no. 1, pp. 17–30, Mar. 2020, doi: 10.32604/CMC.2020.09044.

[20] S. Jalali, R. Boostani, and M. Mohammadi, "Efficient fingerprint features for gender recognition," *Multidimens. Syst. Signal Process.*, vol. 33, no. 1, pp. 81–97, Mar. 2022, doi: 10.1007/S11045-021-00789-6/METRICS.

[21] A. Mishra and S. Jain Asst Professor, "A review on identification of gender using fingerprints," *Int. J. Health Sci. (Qassim).*, vol. 6, no. S2, pp. 9624–9634, May 2022, doi: 10.53730/IJHS.V6NS2.7514.

[22] B. K. Oleiwi, L. H. Abood, and A. K. Farhan, "Integrated Different Fingerprint Identification and Classification Systems based Deep Learning," in *2022 International Conference on Computer Science and Software Engineering (CSASE)*, Mar. 2022, pp. 188–193, doi: 10.1109/CSASE51777.2022.9759632.

[23] G. Singh, "Determination of Gender Differences from Fingerprints Ridge Density in Two Northern Indian Population of Chandigarh Region," *J. Forensic Res.*, vol. 03, no. 03, p. 3, 2012, doi: 10.4172/2157-7145.1000145.

[24] C. T. Hsiao, C. Y. Lin, P. S. Wang, and Y. Te Wu, "Application of Convolutional Neural Network for Fingerprint-Based Prediction of Gender, Finger Position, and Height," *Entropy*, vol. 24, no. 4, p. 475, Apr. 2022, doi: 10.3390/E24040475/S1.

[25] P. Sudharshan Duth and M. P. Mirashi, "Fingerprint Based Gender Classification using ANN," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5, pp. 1779–1782, Jun. 2019, doi: 10.23883/IJRTER.2018.4099.CWM02.

[26] H. Chotimah, H. Susilo, M. Henie, and I. Al, "Fingerprint Analysis and Gender Predilection among medical Students of Nepal Medical College and teaching Hospital," *Int. J. Res. Rev.*, vol. 4, no. June, pp. 6–13, 2017, (accessed Jan. 02, 2020). Available: https://www.ijrrjournal.com/IJRR_Vol.4_Issue.7_July2017/Abstract_IJRR0010.html.

[27] P. Terhorst, N. Damer, A. Braun, and A. Kuijper, "Deep and Multi-Algorithmic Gender Classification of Single Fingerprint Minutiae," in *2018 21st International Conference on Information Fusion (FUSION)*, Jul. 2018, pp. 2113–2120, doi: 10.23919/ICIF.2018.8455803.

[28] T. Nataraja Moorthy, S. Rajathi, and A. K. Sairah, "Gender determination from fingerprint patterns distribution among Malaysian tamils," *Int. J. Med. Toxicol. Leg. Med.*, vol. 22, no. 1–2, pp. 34–37, Jan. 2019, doi: 10.5958/0974-4614.2019.00009.3.

[29] O. N. Iloanusi and U. C. Ejiogu, "Gender classification from fused multi-fingerprint types," *Inf. Secur. J. A Glob. Perspect.*, vol. 29, no. 5, pp. 209–219, Sep. 2020, doi: 10.1080/19393555.2020.1741742.

[30] B. Pandya, G. Cosma, A. A. Alani, A. Taherkhani, V. Bharadi, and T. . McGinnity, "Fingerprint

classification using a deep convolutional neural network," in *2018 4th International Conference on Information Management (ICIM)*, May 2018, pp. 86–91, doi: 10.1109/INFOMAN.2018.8392815.

[31] J. Xiong *et al.*, "Application of Histogram Equalization for Image Enhancement in Corrosion Areas," *Shock Vib.*, vol. 2021, pp. 1–13, Jan. 2021, doi: 10.1155/2021/8883571.

[32] A. Narayanan and S. K, "Gender Detection and Classification from Fingerprints Using Pixel Count," *SSRN Electron. J.*, Aug. 2019, doi: 10.2139/ssrn.3444032.

[33] Y. Qi, Y. Li, H. Lin, J. Chen, and H. Lei, "Research on Gender-related Fingerprint Features," Aug. 2021, doi: 10.48550/arxiv.2108.08233.

[34] Y. I. Shehu, A. Ruiz-Garcia, V. Palade, and A. James, "Sokoto Coventry Fingerprint Dataset," Jul. 2018, doi: 10.48550/arxiv.1807.10609.