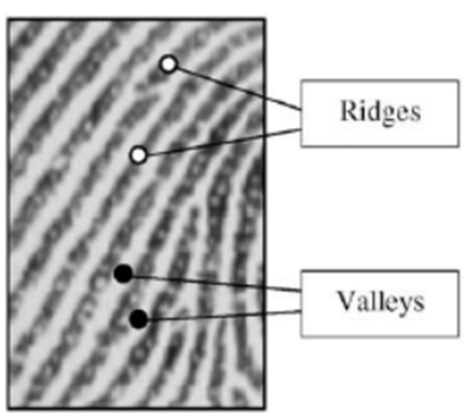
1. ABSTRACT

The paper discusses the application of deep learning, particularly convolutional neural networks (CNNs), in biometrics. It introduces a gender recognition system solely based on CNNs, using fingerprint images as input to determine an individual's gender. Prior to using neural networks, a number of research works were already carried out by numerous researchers on gender determination through fingerprint analysis. Some of these techniques include DWT+PCA and Support Vector Machine (SVM). We also emphasised on data preprocessing techniques to improve the quality of input images. Furthermore, it compares the obtained results with previous research findings in this area incorporating neural networks.

2. INTRODUCTION

Gender detection through fingerprints utilizes biometric technology to accurately determine an individual's gender based solely on fingerprint images. By analysing the unique patterns in fingerprints, this method has diverse applications in fields like law enforcement and healthcare, enhancing security and enabling personalized services.

Fingerprints are graphical patterns of ridges and valleys on fingertips, with ridge density showing gender disparities [1].



Mark A. Acree [2], in his investigation of "gender disparities in fingerprint ridge density," established that women exhibit notably higher ridge density, leading to finer ridge detail, compared to men. Despite the marginal differences in mean ridge densities observed between Caucasian females and males (2.18 ridges/25 mm2) and between African American females and males (1.71 ridges/25 mm2), this consistent magnitude of distinction across the extensive sample population rendered it statistically significant.

Fingerprints gets established at nine months of foetal development and remain stable throughout an individual's life, barring accidents like cuts which can affect the quality of fingerprint images. However, injuries such as cuts, burns, and bruises are temporary and do not alter the underlying ridge configurations permanently [3]. Once fully healed, the original fingerprint patterns are restored, making fingerprints a reliable and resilient biometric identifier [4].

2.1. Analysis of Fingerprint

Fingerprints possess various features that are crucial for identification purposes. Some of these features include:

1. **Ridge Patterns:** The most prominent feature of a fingerprint is its ridge pattern, which consists of ridges (raised areas) and furrows (depressed areas). These patterns are unique to each individual and are used for identification.

2. **Minutiae:** These are small ridge characteristics that occur where the ridge structure changes, such as ridge endings, bifurcations (where a ridge splits into two), and dots (short ridges). Minutiae are used extensively in automated fingerprint recognition systems for matching and comparison.

3. **Ridge Count:** The total number of ridges present in a fingerprint within a specific area, typically used in manual fingerprint analysis for identification purposes.

4. **Ridge Density:** The number of ridges per unit area, which can vary between individuals and is used as a distinguishing feature in fingerprint analysis.

5. **Ridge Width:** The width of individual ridges, which can vary in size and shape, contributing to the uniqueness of a fingerprint.

6. **Ridge Thickness:** The thickness of the ridges in a fingerprint, which can also vary and be used as a distinguishing feature.

7. **White Lines:** These are non-ridged areas that appear as white lines on a fingerprint due to the presence of sweat or oils on the skin, and they can be used as additional features for identification.

These features collectively contribute to the uniqueness and distinctiveness of each fingerprint, enabling accurate identification and classification in forensic and biometric applications.

3. LITERATURE REVIEW

Research in this field traces back to 1999, when early investigations involved manual ridge counting on specific areas of the fingerprint epidermis for gender determination. Since then, a multitude of methods have emerged solely based on fingerprint analysis for gender detection.

Noteworthy contributions include Kralik and Novotny's introduction of mean epidermal ridge breadth in 2003 [5], followed by Ahmed Badawi et al.'s manual extraction of ridge counts, ridge thickness to valley thickness ratio (RTVTR), and white lines count in 2006, achieving accuracies of 80.39%, 86.5%, and 88.5% using Fuzzy C-Means (FCM), Linear Discriminant Analysis (LDA), and Neural Network, respectively [6].

Subsequent advancements saw Manish Verma et al. utilizing ridge width, RTVTR, and ridge density as features in 2008, reaching a remarkable 91% accuracy for both genders using Support Vector Machine (SVM) [7]. In 2012, Kaur and Mazumdar employed Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), and Power Spectral Density (PSD) for ridge feature analysis, achieving 90% accuracy for females and 79% for males [8].

In the same year, Gnanasivam et al. [9] introduced a fingerprint-based method using Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) for feature extraction, coupled with K-Nearest Neighbors (KNN) as a classifier, yielding 91% accuracy for males and 84% for females.

Ting Tang in 2012 proposed image enhancement techniques before applying feature extraction methods, resulting in an 86% gender classification accuracy [10].

Further innovations include Rijo Jackson Tom's use of frequency domain analysis and DWT+PCA for classification in 2013, achieving 70% accuracy [11]. Heena Agarwal et al. in 2014 employed RTVTR, ridge density, and ridge width as determining features, with Multiclass SVM as classifier, achieving 91% accuracy in gender classification [12]. Samta Gupta et al. in the same year utilized frequency domain analysis for feature extraction, followed by DWT and Backpropagation Neural Network for classification, achieving 91% successful classifications [13].

In 2018, Shehu et al. introduced a Resnet34 model based on Convolutional Neural Networks (CNN), achieving 75% accuracy in gender prediction [14].

More recently, in 2022, Chung Ting Hsiao et al. experimented with VGG16, Resnet50, and Inception-V3 models, yielding accuracy rates of 79.2%, 79.1%, and 77.1%, respectively [15].

These cumulative efforts illustrate the evolution and advancements in gender classification through fingerprint analysis over the past few years.

4. METHODOLOGY

4.1. Data Acquisition

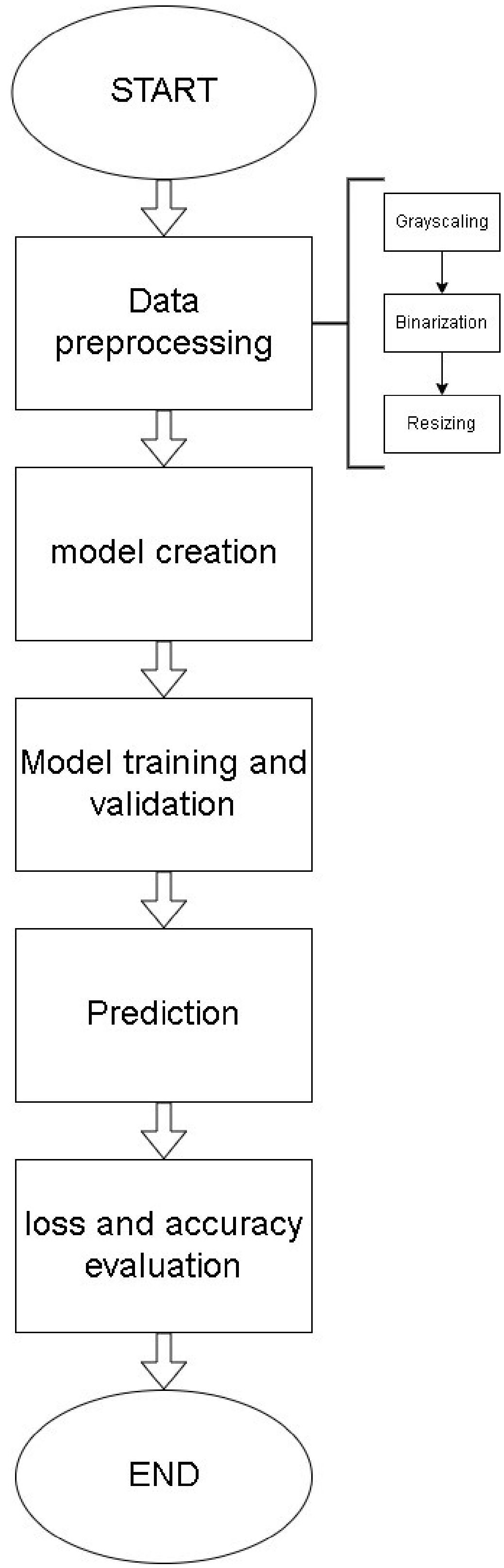
The dataset used for our research was Sokoto Coventry Fingerprint Dataset (SOCOFing). Sokoto Coventry Fingerprint Dataset (SOCOFing) is a biometric fingerprint database proposed by Yahaya Isah Shehu et al [16]. SOCOFing is made up of 6,000 fingerprint images from 600 African subjects and contains unique attributes such as labels for gender, hand and finger name as well as synthetically altered versions with three different levels of alteration for obliteration, central rotation, and z-cut.

4.2. Data Preprocessing

The data preprocessing phase comprises grayscaling, binarization, and resizing of all the fingerprint images, upon which the deep learning model will be trained and evaluated. The preprocessing of fingerprint images is conducted to enhance the visibility of fingerprint features relevant to gender identification, especially the ridges and valleys.

Since our prime objective is to articulate the ridge structure from the fingerprints, grayscaling is applied on the images to lower the color channels down to one, reducing unnecessary computations. In the next stage, the grayscaled fingerprint images get binarized in order to reduce noise, enhance the contrast between the fingerprint ridges and valleys, and to increase computational efficiency as binary images are easier to process in comparison to grayscale images. Finally, it was observed that a relatively small number of images in the dataset vary in their dimensions. Therefore, all images were resized to a fixed dimension.

After processing all the fingerprints, the dataset is partitioned into training and testing sets in a ratio of 10:1, respectively.



4.3. Model Creation

Convolutional Neural Network (CNN) was used for training and validation of the fingerprints. The model consists of two Convolutional layers followed by two max pooling layers and finally an Artificial neural network (ANN) comprising 4 layers. Activation function used in the ANN except the last layer was ReLU, while Sigmoid is used in the last layer, as the model is a binary classifier. Within the ANN, dropouts were added to help the model generalize better to new, unseen data by preventing overfitting and encouraging robust learning of features.

The CNN design implemented is summarized as follows:

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_4 (Conv2D)            (None, 101, 94, 32)       320

max_pooling2d_4 (MaxPooling  (None, 50, 47, 32)        0
2D)

conv2d_5 (Conv2D)            (None, 48, 45, 32)        9248

max_pooling2d_5 (MaxPooling  (None, 24, 22, 32)        0
2D)

flatten_2 (Flatten)          (None, 16896)             0

dense_8 (Dense)              (None, 64)                1081408

dropout_4 (Dropout)          (None, 64)                0

dense_9 (Dense)              (None, 16)                1040

dropout_5 (Dropout)          (None, 16)                0

dense_10 (Dense)             (None, 4)                 68

dense_11 (Dense)             (None, 1)                 5

=================================================================
Total params: 1,092,089
Trainable params: 1,092,089
Non-trainable params: 0
```

4.4. Model Training and Validation

The dataset used to train the CNN comprises of 10,000 fingerprints of both women and men, summing up to 20,000 fingerprint images for training. Consequently, 1000 fingerprints of both women and men were sent for validation to the trained model. The CNN undergoes training for 12 epochs with a batch size of 32.

Throughout the training and validation phase, a record of accuracy and loss was kept for both training and validation data at each epoch. This record provides insight into the learning process of the model.

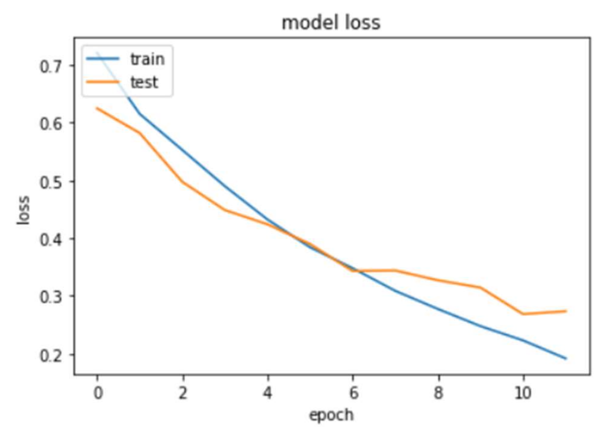
4.5. Model Prediction

The trained model is made to run on same 2000 fingerprint images used for validation in order to examine its performance on unseen data. The results of the prediction phase are concisely illustrated via the confusion matrix.

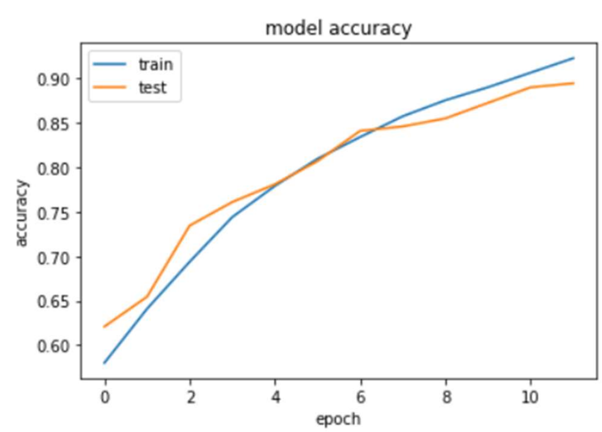
4.6. Loss and Accuracy Evaluation

During model training and validation phase, we tracked the loss and accuracy changes with respect to epochs in order to determine the performance of the model.

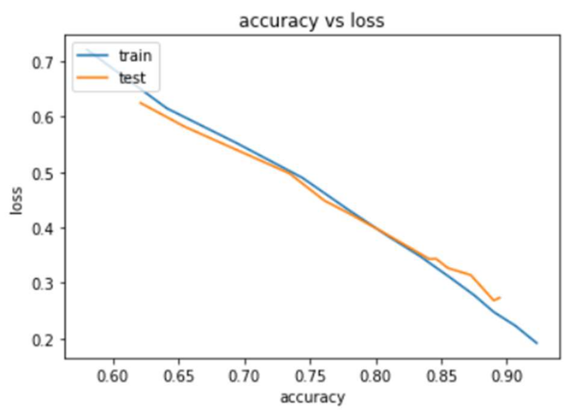
The following graph depicts the difference in loss with respect to epoch while the model was in training phase and in validation phase:



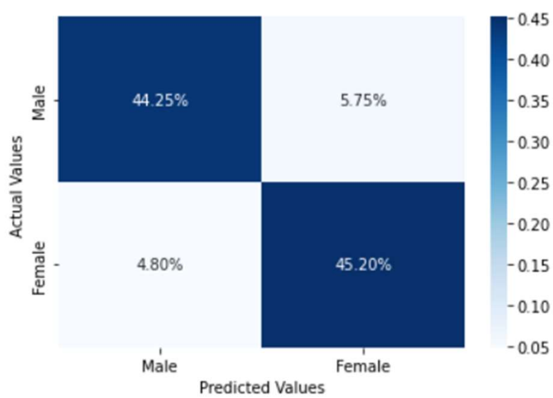
The graph below represents the variation in accuracy for training and validation phases with respect to epoch:



Loss and Accuracy forms the two major metrics to determine the performance of any machine learning or deep learning model. The following graph depicts the performance of the model during training and validation phase:



5. RESULT ANALYSIS



2000 fingerprints were used in model prediction stage comprising 1000 fingerprints of male and 1000 of female.

As per the results,

Correct prediction for male: 885 i.e. 885 fingerprints actually belonged to male and they were predicted as male.

Incorrect prediction for male: 115 i.e. 115 fingerprints actually belonged to female but were predicted as male.

Correct prediction for female: 904 i.e. 904 fingerprints actually belonged to female and were correctly predicted.

Incorrect prediction for female: 96 i.e. 96 fingerprints actually belonged to male but were predicted as female.

Overall, the model achieved an accuracy of 89.45 % during the evaluation phase.

Particularly for males, the model turns out to be 88.50 % accurate, while for females, the accuracy reached 90.4 %.

Previously, Chung-Ting Hsiao et al. [15] carried an experiment using popular Convolutional neural networks such as VGG16, Inception-v3 and Resnet50. The following table describes how well our proposed CNN compare to the other CNN models used previously in this field of research:

Model	Accuracy
VGG16	79.2 %
Inception-v3	77.1 %
Resnet50	79.1 %
CNN proposed by the paper	89.45 %

6. CONCLUSION

This paper introduces a system employing Convolutional Neural Network (CNN) to predict or confirm an individual's gender on the basis of fingerprint characteristics. It provides an overview of fingerprint recognition and details the recognition process step by step, including data acquisition, image pre-processing, model training, and evaluation. Python programming language along with its popular machine learning python packages such as numpy, keras, and matplotlib were utilised to reach the objective of this paper. While the comparison with other CNN models adds depth to our analysis, it also highlights avenues for further research and model refinement.

As we continue to explore the potentials of CNNs in biometric applications, this study contributes valuable insights that pave the way for future advancements in the field.

7. REFERENCES

[1] Ali, Mouad & Mahale, Vivek & Yannawar, Pravin & Gaikwad, Ashok. (2016). Overview of Fingerprint Recognition System. 10.1109/ICEEOT.2016.7754902.

[2] Acree M.A. 1999. Is there a gender difference in fingerprint ridge density? Forensic Science International, 102: 35–44.

[3] A.s, Falohun & Fenwa, Olusayo & Funmilola Alaba, Ajala. (2016). A Fingerprint-based Age and Gender Detector System using Fingerprint Pattern Analysis. International Journal of Computer Applications. 136. 975-8887. 10.5120/ijca2016908474.

[4] Babler. 1991 Handbook of fingerprint recognition. Springer Verlag.

[5] Kralik, Miroslav & Novotný, Vladimír. (2003). Epidermal ridge breadth: An indicator of age and sex in paleodermatoglyphics. Variability and Evolution. 11. 5-30.

[6] Badawi, Ahmed & Mahfouz, Mohamed & Tadross, Rimon & Jantz, Richard. (2006). Fingerprint-Based Gender Classification. Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV'06. 1. 41-46.

[7] Verma, Manish & S., Agarwal. (2008). Fingerprint Based Male-Female Classification. Advances in Soft Computing. 53. 251-257. 10.1007/978-3-540-88181-0_32.

[8] Kaur, Ritu & Mazumdar, Susmita & Tech, M & Bhilai, Rcet & India. (2012). Fingerprint based gender identification using frequency domain analysis. International Journal of Advances in Engineering and Technology. 3.

[9] Gnanasivam, P & Muttan, S. (2012). Estimation of Age Through Fingerprints Using Wavelet Transform and Singular Value Decomposition. International Journal of Biometric and Bioinformatics.

[10] T. Tang, "Fingerprint recognition using wavelet domain features," 2012 8th International Conference on Natural Computation, Chongqing, China, 2012, pp. 531-534, doi: 10.1109/ICNC.2012.6234738.

[11] Tom, Rijo. (2018). Fingerprint based gender classification using PCA and Wavelet Transforms.

[12] Heena Agrawal, Prof. Siddhartha Choubey." Fingerprint Based Gender Classification using multi- class SVM."International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 3, pp 47-52, August 2014.

[13] Gupta, Samta & Rao, A. Prabhakara. (2014). Fingerprint Based Gender Classification Using Discrete Wavelet Transform & Artificial Neural Network. 10.13140/RG.2.2.15292.69769.

[14] Shehu, Yahaya & Ruiz-Garcia, Ariel & Palade, Vasile & James, Anne. (2018). Detailed Identification of Fingerprints Using Convolutional Neural Networks. 1161-1165. 10.1109/ICMLA.2018.00187.

[15] Hsiao, Chung-Ting & Lin, Chun-Yi & Wang, Po-Shan & Wu, Yu-Te. (2022). Application of Convolutional Neural Network for Fingerprint-Based Prediction of Gender, Finger Position, and Height. Entropy. 24. 475. 10.3390/e24040475.

[16] Shehu, Y. I., Palade, V., & James, A. (2018). Sokoto Coventry Fingerprint Dataset. ArXiv. /abs/1807.10609