International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance Analysis of Deep Neural Networks for Unimodal and Multimodal Biometric Authentication

[1]Pravin Jangid, [2]Geetanjali Nilesh Sawant, [3]Dr. Vinayak Ashok Bharadi. [4]Dr. Nupur Giri

Submitted: 26/04/2023 **Revised:** 25/06/2023 **Accepted:** 04/07/2023

Abstract: Biometric technology is a powerful tool that relies on distinct and measurable physiological or behavioral characteristics possessed by individuals. These traits serve as reliable means to verify and authenticate individuals. However, unimodal biometric systems encounter several challenges that hinder their effectiveness. These challenges include noisy data, variations within the same class, limited degrees of freedom, non-universality, susceptibility to spoof attacks, and high error rates. To address these challenges, researchers have turned to multimodal biometric systems. These systems leverage the use of two or more biometric modalities to enhance performance and overcome the limitations of unimodal systems. In the context of the discussed system, face and fingerprint modalities are utilized. To develop and evaluate the performance of the system, various deep learning algorithms such as VGG16, VGG19, CNN, and Inception are employed. These algorithms are trained, validated, and tested using the face and fingerprint data. Performance metrics such as accuracy, precision, and f1 score are calculated for both unimodal and multimodal configurations. The results of the performance evaluation demonstrate that CNN (Convolutional Neural Network) yields higher accuracy compared to the other models tested. This finding suggests that CNN is particularly well-suited for this biometric system, as it can effectively extract meaningful features from the face and fingerprint data and provide accurate classification results. By adopting a multimodal approach and utilizing deep learning algorithms such as CNN, the system successfully addresses the challenges faced by unimodal biometric systems. This implementation demonstrates improved accuracy, making it a promising solution for reliable and secure individual authentication in various applications.

Keywords*: Unimodal, Multimodal, Face recognition, Fingerprint recognition, CNN, VGG16, VGG19, Inception.*

1. Introduction

Biometrics refers to automated methods of identifying individuals by analyzing their unique biometric traits [1]. A biometric system or device captures these traits, extracts distinct features, and compares them to recognize or verify a person's identity. Fingerprint recognition is the most widely recognized and oldest form of biometrics. Other commonly used physiological characteristics include face, palm-prints, iris, retina, hand vein, and hand geometry [2]. Behavioral characteristics such as gait, voice, signature, and typing rhythm are also utilized for biometric identification. Due to its unparalleled uniqueness, biometrics is extensively employed for secure identification and personal verification purposes [1]. Fingerprint recognition is a well-established and widely used form of biometric identification. It relies on the distinct patterns and ridges found on an individual's fingertips, offering a reliable and unique biometric trait for personal verification and identification purposes [4].

Face recognition is a biometric technology that identifies and verifies individuals based on their unique facial features. It analyzes facial patterns, structures, and characteristics to match against stored data [4].

Multimodal biometrics involves the integration and combination of multiple biometric modalities to enhance the accuracy and reliability of identification and authentication systems. Instead of relying on a single biometric trait, such as fingerprints or facial recognition, multimodal biometric systems utilize two or more biometric modalities simultaneously [2].

The fusion of multiple biometric modalities allows for increased robustness, as it reduces the risk of false positives and false negatives

The present study focuses on correlation-based fingerprint recognition methods **[4]** that exhibit robustness but may lack accuracy. To enhance accuracy, the integration of another trait in a multimodal biometric system is proposed.

Additionally, multimodal systems can address challenges like intra-class variations, noisy data, non-universality, and restricted degrees of freedom. [6] Overall, multimodal biometrics provides a comprehensive and reliable

1Research Scholar, IT Department FAMT, Mumbai University Ratnagiri, India pravinjangid@gmail.com
2Research Scholar, IT Department FAMT, Mumbai University Ratnagiri, India geetanjalinileshsawant@gmail.com
3Professor, Head of Department-IT. FAMT, Mumbai University Ratnagiri,India vinayak.bharadi@famt.ac.in
4Professor, Head of Department-CS VESIT, Mumbai University Mumbai, India nupur.giri@ves.ac.in

approach to person verification and authentication, making it valuable in applications such as access control, border security, and identity management systems[7].

In this paper Multimodal Fusion of Face and Fingerprint as shown in figure 1 using CNN, VGG16, VGG19 and Inception are used. By combining different biometric traits, such as fingerprints, iris scans, voice patterns, or facial features, multimodal biometric systems can overcome the limitations of unimodal systems. They offer improved performance in terms of accuracy, security, and resistance to spoof attacks [3].

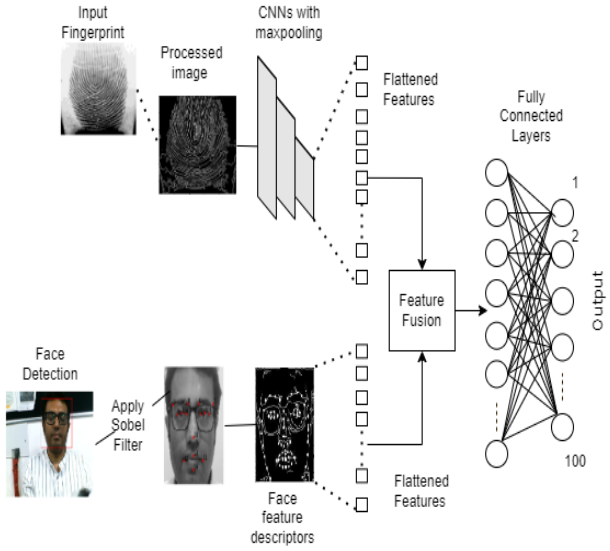


Fig .1 Multimodal biometrics system using two traits: face and fingerprint

2. Literature Survey

In paper [4] [10] ANN and KNN authors suggested that use of multimodal biometric increases the accuracy using raspberry pi IoT devices. Multimodal biometric system increases the accuracy. It uses KNN. Artificial neural networks and deep learning approaches were not used. Also the size of the dataset is significantly less.

In paper [5], extensive face and fingerprint data sets were analyzed using different normalization and fusion techniques. The findings of the study demonstrated that the performance of a multimodal biometric system surpasses that of a unimodal system in terms of accuracy and effectiveness.

M. N. Yaacob, S. Z. Syed et.a al. [6] proposed a Biometric Feature based Person Unique Identification System. The study focused on employing various image enhancement techniques such as Gaussian smoothing function and adjusting intensity values of pixels to enhance the quality of the images.

Different binarization methods were examined, and the most effective one was selected for the input thumb image. The study also highlighted the challenges of using ANN (Artificial Neural Network), including the difficulty in understanding its structure, the risk of overfitting with too many attributes, and the need for experimentation to determine the optimal network structure [7].

In paper [16] for fingerprint and face Normalization techniques [z-score, tanh, two quadrics] and fusion techniques [max score, match weighing] was used in paper [17] or fingerprint and face Neyman-Pearson based methodology was used by combining match scores provided by various biometric matchers.

3. Dataset Description

The dataset used in this study is curated specially for the work proposed. Section III (a) and (b) present the fingerprint and Face images collection and feature selection done. The preprocessing and training model is in section IV

1. III (a) FINGERPRINT RECOGNITION

The process of fingerprint recognition involves several key steps. The work proposed in the paper uses a fingerprint image captured using an Optical sensor R307 module shown in figure1. The captured image is then subjected to preprocessing (described in section IV a) to enhance its quality by eliminating noise and artifacts that could affect subsequent analysis [4].

Next, feature extraction is performed to identify the specific characteristics of the fingerprint. These features include ridge endings, bifurcations, and ridge orientations. Various algorithms, such as minutiae-based or ridge-based methods, are utilized to extract these unique features. Once the features are extracted, they are

transformed into a compact template, which serves as a mathematical representation of the fingerprint. During enrollment, the template is stored in a database along with the corresponding individual's identity [4].

For verification or identification, a captured fingerprint image undergoes preprocessing and feature extraction. The extracted features are then compared to the stored templates in the database using matching algorithms. Similarity between the presented fingerprint and the templates is calculated, and a decision is made based on predefined thresholds

In the current research 50 fingerprint images of 100 different users were used. In total 5000 images of fingerprints were used. Also R307 Optical Fingerprint Sensor Module was used for scanning fingerprint of the 100 different users.s Specifically, spectral features of a fingerprint are extracted.

The fingerprint Region of Interest (ROI) is chosen to have dimensions of 256 *288 pixels. This particular dimension is selected to ensure that the square ROI fits within the boundaries of the fingerprint image. The ROI serves as the input for feature extraction. This is shown in the Fig 2.

Fig.2 Fingerprint image

B) FACE RECOGNITION

The face recognition process begins with capturing an individual's face image using a camera. Facial landmarks, such as the eyes, nose, and mouth, are then detected to create a facial template [4] as shown in the fig 2.

Feature extraction techniques are employed to convert the facial template into a mathematical representation using CNN method for feature extraction.

During the identification or verification stage, the captured facial template is compared with pre-existing templates in a database. Matching algorithms calculate the similarity between the input template and the templates in the database, producing a confidence score or similarity measure. If the similarity exceeds a predetermined threshold, the person's identity is recognized or verified [8].

Facial keypoints detection using neural networks is a widely used technique in computer vision. It involves training a neural network model on a dataset of labeled facial images to accurately locate specific facial landmarks or keypoints.[9]

Fig.3 Face image and key points

It can be used to predict the key points on unseen images. By taking an input image, the model generates predictions for the coordinates or positions of the keypoints.

To train the model, a large dataset of 4850 facial images of 97 different users is used, where each image is annotated with the coordinates or pixel positions of the desired facial keypoints. The neural network learns to recognize patterns and features in the images that correspond to these key points through interconnected layers of neurons [10].

Convolutional neural networks (CNNs) are employed for facial keypoints detection due to their ability to capture spatial patterns effectively. During training, the model's parameters are optimized using techniques like backpropagation and gradient descent to minimize the difference between the predicted keypoints and the ground truth annotations.[11] **Fig 3** shows the Face image along with keypoints. of one of model

The complete Dataset consists of a total of 9700 images, 4850 each for fingerprint and face with dimensions of 96x96 pixels. Among these images, 2140 of them have

ground truth positions for all 15 facial keypoints, which serve as the dataset for training, validation, and testing the neural network

The dataset is divided randomly into three subsets: training set, validation set, and test set. From each user 40 images are taken for training and 10 for testing. Approximately 80% of the 9700 images (7760 images) are allocated to the training set. The remaining 20% of the images (1940 images) are further divided, with 50% of them (970 images) designated as the training set and the remaining 50% (970 images) used for validation.

The 15 facial keypoints that are annotated in the dataset include specific features on a person's face.

4. Model Development and Implementation

Proposed implementation steps –

Sensor level:

We merge biometric features collected from different sensors using Raspberry pi to create a combined biometric trait and process. This involves capturing an image of face using a digital camera and fingerprint using a module as shown in figure 1 both are attached to Raspberry pi, which then serves as the input for the system. Face images are examined by measuring the distances between different facial landmarks and analyzing the variations in their relative distances. Fingerprint images are typically analyzed at the pixel level to extract the unique features and patterns present in the fingerprints [4].

Fig. 4 Raspberry Pi

Preprocessing face images using Convolutional Neural Networks (CNN) involves several steps to prepare the data for analysis or recognition tasks. These steps include image resizing to ensure uniformity, face detection and alignment for consistency, image normalization to enhance quality, data augmentation for increased diversity, and feature extraction using pretrained CNN models like VGG, ResNet, or Inception. These preprocessing techniques aim to optimize the input data and improve the performance and accuracy of CNN-based face analysis or recognition systems. The specific implementation may vary depending on the application's requirements [10].

IV a Pre-Processing steps of Fingerprint recognition

1. **Image acquisition**: Fingerprint image is captured using R307 Optical Fingerprint Sensor Module connected to IoT device (Raspberry Pi)

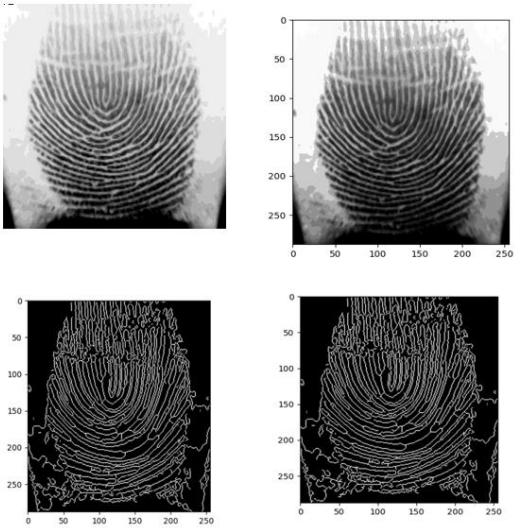


Fig. 5 FingerPrint Preprocessing - a) Enhanced Fingerprint b) Gray Scale and Histogram Equalization c) Ridge structure of Fingerprint d) Skeleton of Fingerprint

2. Image normalization and enhancement [Gray Scale and Histogram Equalization**]:** Captured fingerprint image may have noise, histogram equalization are used to enhance the image quality and remove unwanted noise.

3. Image binarization: The fingerprint image is divided into smaller regions to isolate the fingerprint ridges and valleys. Binarization is used to extract the ridge structure accurately.

4. Thinning or skeletonization [image size: 256 *288]: It extract the skeleton or core ridge structure from a fingerprint image [14][15].

IV a) Pre-Processing of Face Images

Using camera attached to IoT device,

1)Face image is detected using face detection algorithm convolutional neural networks

2)It is cropped and converted into grayscale image.

3)Face image is resized into (200 x 200) pixel.

4)Face key points are extracted as shown in the figure 6.

Fig. 6 Extracted Face Keypoints

5. Apply Sobel Filter to extract face features outline and convert face image into logical image

6. Plot Facial Keypoints into Logical Image

Fig. 7 Face Keypoints into Logical Image

2. V MODEL DEVELOPMENT AND IMPLEMENTATION

The primary objective of the proposed system is to minimize errors and enhance system performance by achieving a high acceptance rate during identification and authentication processes.

V Training Model:

Our dataset has been divided into three parts: Xtrain, Xval, and Xtest. We have assigned the ground truth values for these three sets as ytrain, yval, and ytest respectively.

Our model M will be trained using the training dataset (Xtrain, ytrain), and we will assess its performance by evaluating it on the validation set (Xval, yval) in order to fine-tune its parameters.

Once the training process is complete, we will evaluate our model M using the test set (Xtest, ytest) to obtain its final performance metrics. We assess the performance of our network models using two primary metrics. The first metric is detection accuracy, which is measured through regression loss. To elaborate, in our scenario, we quantify the loss by calculating the mean squared error **(MSE)** between the ground truth key points vector y and the predicted keypoints vector y-hat.

The second metric we consider is the time consumption during both the training and testing phases. We evaluate how much time is required for these processes [14].

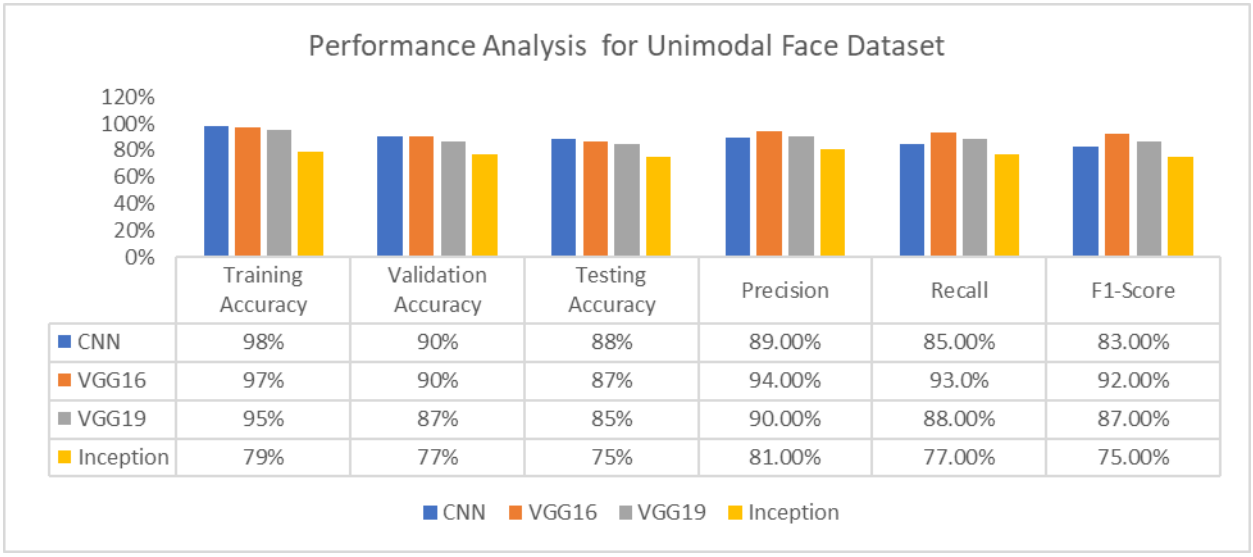
Epochs: Number of Epochs used to train models is 20.

5. Experimental Results

In this section the performance parameters like training accuracy, validation accuracy, testing accuracy, precision, recal and F1-score for unimodal and multimodal biometric system using CNN, VGG16, VGG19 and

Inception are discussed and performance matrix values are evaluated

A. Performance Analysis of CNN, VGG16 , VGG19, Inception- Based Architecture for Unimodal Face Dataset



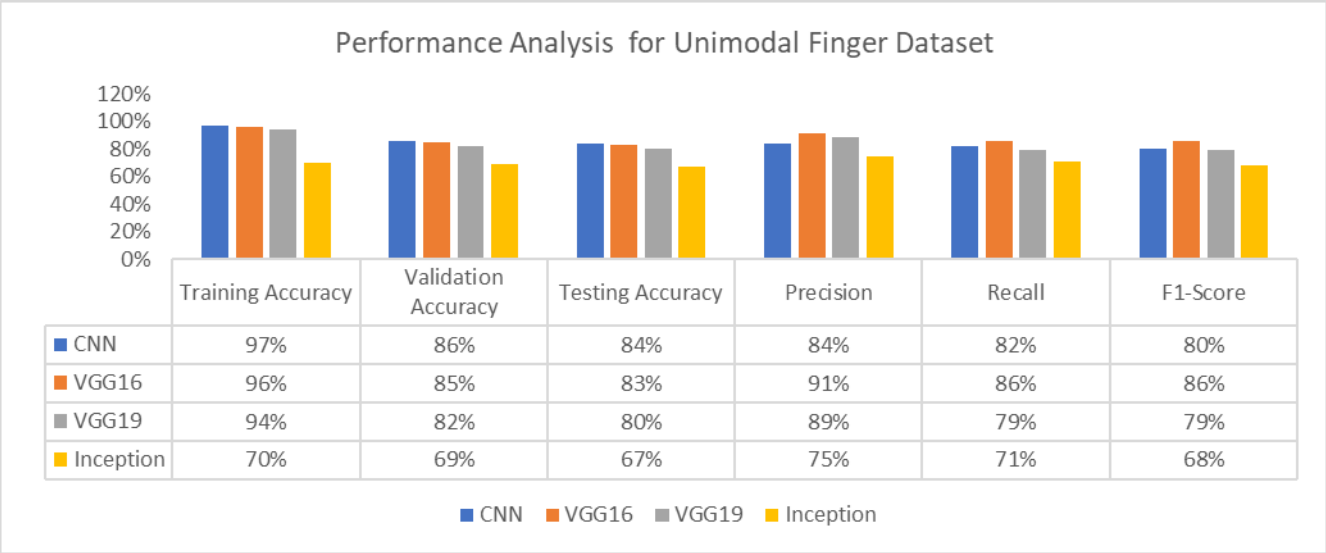
	Training Accuracy	Validation Accuracy	Testing Accuracy	Precision	Recall	F1-Score
CNN	98%	90%	88%	89.00%	85.00%	83.00%
VGG16	97%	90%	87%	94.00%	93.0%	92.00%
VGG19	95%	87%	85%	90.00%	88.00%	87.00%
Inception	79%	77%	75%	81.00%	77.00%	75.00%

Fig.8 Performance Analysis for Unimodal Face Dataset

The above figure 8 gives 98% training accuracy, 90% validation accuracy, 88% Testing accuracy, 89% Precision, 85% Recall and 83% F1-Score for CNN which is highest.

After evaluating the results of this technique, the conclusion derived that CNN technique gives the better performance as compared to VGG16, VGG19 and Inception.

B. Performance Analysis of CNN, VGG16 , VGG19, Inception- Based Architecture for Unimodal Finger Dataset



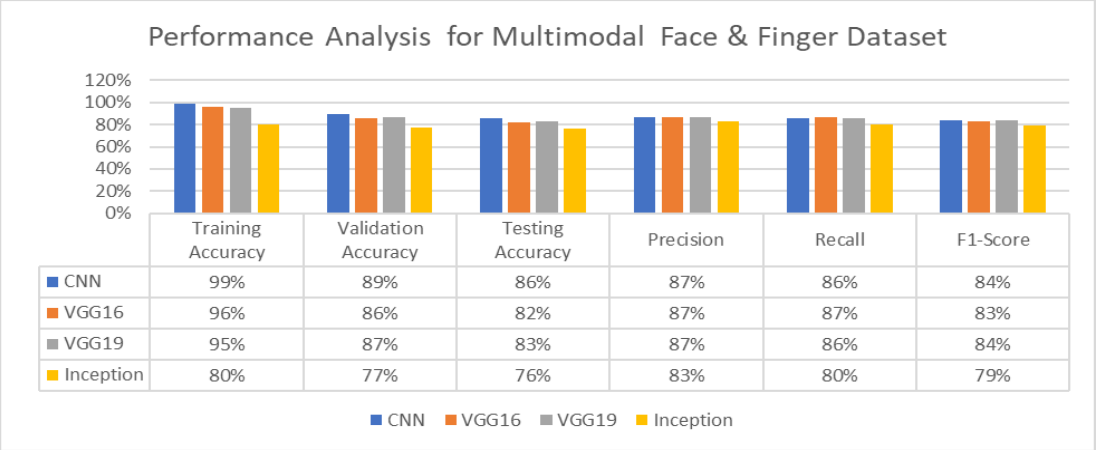
	Training Accuracy	Validation Accuracy	Testing Accuracy	Precision	Recall	F1-Score
CNN	97%	86%	84%	84%	82%	80%
VGG16	96%	85%	83%	91%	86%	86%
VGG19	94%	82%	80%	89%	79%	79%
Inception	70%	69%	67%	75%	71%	68%

Fig.9 Performance Analysis for Unimodal Finger Dataset

The above figure 9 gives 97% training accuracy, 86% validation accuracy, 84% Testing accuracy, 84% Precision, 82% Recall and 80% F1-Score for CNN which is highest.

After evaluating the results of this technique, the conclusion derived that CNN technique gives better performance as compared to VGG16, VGG19 and Inception.

C. **Performance Analysis of CNN, VGG16, VGG19, Inception- Based Architecture for Multimodal Face and Fingerprint Dataset**



	Training Accuracy	Validation Accuracy	Testing Accuracy	Precision	Recall	F1-Score
CNN	99%	89%	86%	87%	86%	84%
VGG16	96%	86%	82%	87%	87%	83%
VGG19	95%	87%	83%	87%	86%	84%
Inception	80%	77%	76%	83%	80%	79%

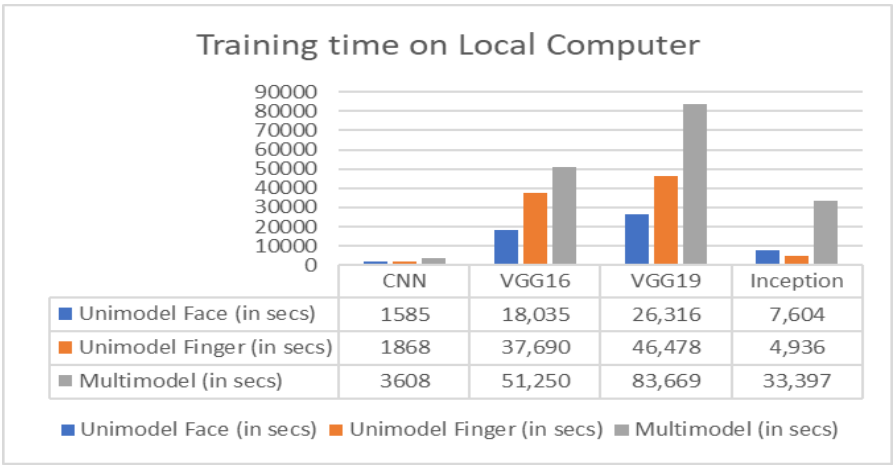
Fig.10 Performance Analysis for Multimodal Face and Finger Dataset

The above figure 10 gives 99% training accuracy, 89% validation accuracy, 86% Testing accuracy, 87% Precision, 86% Recall and 84% F1-Score for CNN which is highest.

After evaluating the results of this technique, the conclusion derived that CNN technique gives better performance as compared to VGG16, VGG19 and Inception.

II. DESIGN DETAILS

In this section the response training time for different architectures i.e. unimodal and multimodal biometric systems using CNN, VGG16, VGG19 and Inception are measured and recorded.



	CNN	VGG16	VGG19	Inception
Unimodel Face (in secs)	1585	18,035	26,316	7,604
Unimodel Finger (in secs)	1868	37,690	46,478	4,936
Multimodel (in secs)	3608	51,250	83,669	33,397

Fig. 11 Training time on Local Computer

Performance evaluation for different architectures shows the Training time for CNN which is lowest as compared to other deep learning models.

After evaluating the results of this technique, the conclusion derived that CNN technique gives the better performance as compared to VGG16, VGG19 and Inception.

Table 1.1 Comparison of unimodal and Multomodal using CNN

Factors	Biometric Sensing System		
	Unimodal		Multimodal
	Face	Fingerprint	Fingeprint and Face

Training Accuracy	98%	97%	99%
Validation Accuracy	90%	86%	89%
Testing Accuracy	88%	84%	86%

In the table 1.1 it is shown that training accuracy of multimodal biometric system is 99% which is better than the unimodal system for face and Fingerprint. Validation and testing accuracy for multimodal system is 89% and 86%.

Biometric data can contain **noise** if sensors are not well-maintained. For instance, dirt on a fingerprint scanner can result in a noisy fingerprint. Additionally, improper camera focusing can lead to unclear images of the face. In our research we overcome this factors.

Using multiple biometric traits for authentication, such as face and fingerprint, can be an effective deterrent against **spoofing attempts**. Since replicating the exact facial structure and texture, or creating an artificial replica of a fingerprint, is highly complex, it would be challenging for an impostor to spoof both biometric traits simultaneously. This multi-factor authentication approach can increase security and reduce the risk of successful spoofing attacks

A typical example is the issue of worn fingerprints that may result in **error**. In a multimodal biometric system such error or failure may not seriously affect an individual because other biometric technology like face recognition systems are employed in this research. Hence, the failure to **enroll rate is reduced in a multimodal system**, which is one of its major advantage.

Intra-class variations refer to the natural differences that exist within a group of individuals sharing the same biometric trait. A multimodal biometric system addresses the problem of **intra-class variations** by combining multiple biometric modalities (Face and Finger) to capture a broader range of individual characteristics.

A multimodal biometric system helps address the **restricted degree of freedom problem** by incorporating multiple biometric modalities for identification or authentication. By combining different biometric traits like fingerprint and face the system increases the robustness, mitigating the limitations of relying on a single biometric modality.

6. Conclusion

This research work focused on training, testing, and validating unimodal and multimodal systems using various deep learning architectures. The main objective was to explore the potential of multimodal biometric systems, which utilize two or more biometric modalities, in overcoming the limitations of unimodal biometric systems. The architectures employed in this work included CNN, VGG16, VGG19, and Inception.

A performance analysis was conducted on these architectures using a dataset comprising both face and fingerprint modalities. The analysis considered parameters such as Training Accuracy, Validation Accuracy, Testing Accuracy, Precision, Recall, and F1-Score. The experimental results revealed that CNN outperformed the other architectures across all these parameters.

For the unimodal systems (face and fingerprint), the training accuracies were found to be 98% and 97%, respectively. In the case of the multimodal system, the training accuracy was 99% which is better than the unimodal system. The training times for the unimodal face, unimodal fingerprint, and multimodal face and fingerprint systems were measured as 1585 seconds, 1868 seconds, and 3608 seconds, respectively. The performance evaluation demonstrated that CNN was the superior choice for training, validation, and testing purposes in this study.

References

[1] A. K. Jain, P.Flynn,A. Ross, “Handbook of Biometrics”, Springer, USA,ISBN-13:978-0-387-71040-2,pp.1-23,2007.

[2] A.K. Jain ,A. Ross.S.Prabhakar,"An introduction to biometric recognition", Circuits and Systems for Video Technology, IEEE Transactions on,Vol. 14,No. 1,pp.4-20,2004

[3] http://en.wikipedia.org/wiki/Biometrics , Accessed on 08.05.2023 , 10:42AM.

[4] O. Olazabal et al., "Multimodal Biometrics for Enhanced IoT Security," 2019 IEEE 9th Annual

Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2019, pp. 0886-0893, doi: 10.1109/CCWC.2019.8666599.

[5] N.Kavitha Devi, “Authentication using multimodal Biometric features”, International Journal of Computer Science and Mobile Computing, Vol.7 Issue.1, January- 2018, pg. 1-8

[6] M. N. Yaacob, S. Z. Syed Idrus, W. A. Wan Mustafa, M. A. Jamlos, M. H. Abd Wahab, Identification of the exclusivity of individual’s typing style using soft biometric elements. Annals of Emerging Technologies in Computing. 5, 10–26 (2021)..

[7] Jiya Tian,Yanqin Peng,"Research of the Matlab application in the f ing erprint identification system"IEEE explorer, 2012 International Conference on Image Analysis and Signal Processing.

[8] S. Selvarani; S. Jeba Priya; R. Smeeta Mary,"Automatic Identification and Detection of Altered Fingerprints' 'IEEE explorer, International Conference on Intelligent Computing Applications,March 2014.

[9] Chan, L.H. , S.H. Salleh and C. M. Ting, 2010. Face biometrics based on principal component analysis and linear discriminant analysis. J. Comput. Sci., 6: 693-699. DOl: I 0. 3844/jcssp. 20 10.693. 699.

[10] S. Thakre, A. K. Gupta and S. Sharma, "Secure reliable multimodal biometric fingerprint and face recognition," 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2017, pp. 1-4, doi: 10.1109/ICCCI.2017.8117796.

[11] Muhammad Imran Razzak et aI. , Multimodal face and finger veins biometric authentication, Scientific Research and Essays Vol. 5(17), pp. 2529-2534, 4 September, 2010.

[12] Dr Shubhangi D C, Manohar Bali: ―MultiBiometric Approaches to Face and Fingerprint Biometrics‖ International Journal of Engineering Research & Technology Vol. 1 Issue 5 ISSN: 2278-0181, July - 2012.

[13] Agrawal, Divyansh, et al. "A robust drug recall supply chain K.Sasidhar, Vijay L Kakulapati, Kolikipogu Ramakrishna: ―multimodal biometric systems – study to improve accuracy and performance‖ International Journal of Computer Science & Engineering Survey (IJCSES) Vol.1, No.2, November 2010.

[14] V. A. Bharadi and G. M. DSilva, "Online signature recognition using software as a service (SAAS) model on public cloud," in 2015 International Conference on Computing Communication Control and Automation, 2015, pp. 65-72

[15] B. Pandya, G. Cosma, A. A. Alani, A. Taherkhani, V. Bharadi, and T. M. McGinnity, "Fingerprint classification using a deep convolutional neural network," in 2018 4th International Conference on Information Management (ICIM), 2018, pp. 86-91

[16] B. E. Manjunathswamy, J. Thriveni, and K. R. Venugopal, ‘‘Bimodal biometric verification mechanism using fingerprint and face images(BBVMFF),’’ in Proc. IEEE 10th Int. Conf. Ind. Inf. Syst. (ICIIS),Dec. 2015, pp. 372–377

[17] M. D. J. Ghate and S. B. Patil, ‘‘Robust combination method for privacy protection using fingerprint and face biometrics,’’ in Proc. 4th Int.Conf. Rel., Infocom Technol. Optim. (ICRITO)(Trends Future Directions),Sep. 2015, pp. 1–6

[18] Perez-Siguas, R. ., Matta-Solis, H. ., Millones-Gomez, S. ., Matta-Perez, H. ., Cruzata-Martinez, A. ., & Meneses-Claudio, B. . (2023). Comparison of Social Skills of Nursing Students from Two Universities of Lima. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 14–19. https://doi.org/10.17762/ijritcc.v11i2.6105

[19] Dhablia, A. (2021). Integrated Sentimental Analysis with Machine Learning Model to Evaluate the Review of Viewers. Machine Learning Applications in Engineering Education and Management, 1(2), 07–12. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/12