Multimodal Deep Learning Based Score Level Fusion Using Face and Fingerprint

Krishna Shinde[1](✉) and Charansing Kayte[2]

[1] Department of Computer Science and IT, Dr. B.A.M.U. Aurangabad, Aurangabad, India
shreekriss@gmail.com

[2] Department of Digital and Cyber Forensic, Government Institute of Forensic Science, Dr. B.A.M.U. Aurangabad, Aurangabad, India

Abstract. In the previous decade, biometrics referred to the automatic recognition of persons based on their physiological or behavioural traits but unimodal biometrics have their limitations. Due to its potential to overcome some of the inherent limitations of single biometric modalities while simultaneously enhancing overall recognition rates, multimodal biometrics has recently gained prominence. In this research, we offer a multimodal biometric person authentication system based on pre-train transfer learning VGG16 with CNN and CNN models, that uses the user's face and fingerprint biometric traits. We have used the own collected samples of same person KVKR face and fingerprint dataset for experimental work. First, we have applied pre-processing data augmentation technique on face and fingerprint data then image enhancement techniques on fingerprint data. In the features extraction, we have extraction the hidden feature of the face and fingerprint images using pre-train VGG16 with CNN and CNN models. The hstack method has been used to combine the features and SoftMax classifier use for features classification. The fusion score is calculated using the fixed-rule-based maximum rule technique, finally we have done comparative analysis of the unimodal and multimodal biometric recognition system.

Keywords: VGG16 · CNN · Deep learning · Score Fusion

1 Introduction

In recent years, biometric-based authentication systems have grown in popularity in several applications that require a reliable verification/identification method. Biometric authentication is a pattern-recognition system that identifies a person using a feature vector obtained from a physiological or behavioural characteristic [1]. The technique is currently used in a variety of high-security identity and individual verification systems. Multimodal biometrics has increased in importance in recent years, due to single modal biometrics' limitation, which includes non-universal, noisy sensor data, substantial intra-user variability and vulnerability to spoofing attacks [2]. The drawbacks of unimodal biometric systems are addressed by multi-modal biometric systems. For example, consider the issue of non-universality: a fraction of users may lack a certain biometrics

R. Manza et al. (Eds.): ACVAIT 2022, AISR 176, pp. 140–152, 2023.
https://doi.org/10.2991/978-94-6463-196-8_13

attribute [3]. The multimodal biometric system can be used for combination of multiple modality properties. In the multimodal system sensor-level fusion, feature-level fusion, score-level fusion and decision-level fusion are the four forms of fusion. In a variety of ways, multimodal biometric systems surpass unimodal biometric systems, making them a particularly appealing secure recognition solution [4, 5]. Deep learning is based on an artificial neural network that constructs feature hierarchies using statistical machine learning techniques. There is made up of three layers such as input, hidden layer and output layer and the layers' nodes are all to inter connected. The raw data is transferred to the input layer, where each node stores part of the information they encounter and the information is then passed on to the next layer nodes, which develop abstract knowledge about the data. Deep learning has lately had a significant influence on biometrics systems, generating surprising outcomes. Many of the shortcomings in traditional machine learning methods, particularly in feature extraction procedures, have been solved by deep learning algorithms. Deep neural networks can adapt to changes in biometric images and extract features from raw data [6, 7]. In this article, the Sect. 1 introduced of biometric and multimodal biometric, face and fingerprint recognition. In the next section have related work and existing work of multimodal biometric recognition system. Section 3 about database benchmark, Sect. 4 experimental setup and hyperparameter, Sect. 5 have introduce proposed methodology, features extraction, feature classification and fusion techniques. Section 6 have performance analysis of proposed system, Sect. 7 has result and discussion and last section have conclusion, contribution and acknowledge.

2 Literature Survey

The following Table 1 shows existing work of multimodal biometric system review starting from the author, biometric traits, databases, algorithm, level of fusion, and lastly the recognition result.

Table 1. Literature Survey

Sr. No	Author & year	Biometrics Traits	Database	Techniques	Fusion Level	Result
1	Arun Ross 2003 [8]	Face, Fingerprint, and Hand Geometry	NA	PCA	Score	FAR 0.03%
2	Shi-Jinn Horng 2009 [1]	Face, Fingerprint, and Finger Vein	NIST	SVM	Score	99.8%
3	Mendu. Anusha 2016 [9]	Fingerprint, Face, and Iris	HMM	Doughman's, WLD, and Decision Tree	Features	90.00%

(*continued*)

Table 1. (*continued*)

Sr. No	Author & year	Biometrics Traits	Database	Techniques	Fusion Level	Result
4	E. Sujatha 2017 [10]	Iris, Palm Print, Face, and Signature	CASIA	DWT	Features	99.90%
5	Supreetha Gowda H D 2018 [7]	Face and Iris	CASIA and ORL	CNN	Feature	99.00%
6	Veeru Talreja 2019 [11]	Face and Iris	NA	CNN	Features	99.00%
7	Nada Alay 2020 [6]	Iris, Face and Finger Vein	SDUMLA -HMT	CNN	Features Score	99.22 100
8	EI Mehdi 2020 [12]	Finger Vein and Face	DB1 and DB2	CNN, RF, SVM	Score	99.98
9	EI Mehdi 2020 [13]	Fingerprint, Finger Vein and Face	SDUMLA -HMT	CNN, SVM, LR, and RF	score	99.49
10	Priti Shende 2020 [14]	Face, Palm Veins, and Fingerprint	Self-Created	CNN and SVM	Features	99.10%
11	Arjun Benagatte 2021 [15]	Fingerprint and Signature	SDUMLA -HMT and MCYT	HOG and CNN	Features	93.33%
12	Mehwish Leghari 2021 [16]	Fingerprint and Online Signature	NA	CNN	Features	99.10%

3 About Database

In this study, we have utilized the KVKR face and fingerprint database for the same person. Under the supervision of Prof. Dr. K. V. Kale, Programme Coordinator, UGC SAP (II) DRS Phase-I, we gathered data in the Multimodal Biometrics Research Laboratory at the Dept. of Computer Science & Information Technology, Dr. B. A. M.U., Aurangabad. Face and fingerprint data were obtained from Research Scholars and PG students aged 21 to 40. The face KVKR database collects 10 various positions of the face such as the frontal face, left 90, left 60, Right 90, Right 60, chin up, Down, small smiling, big smiling, and close eyes, etc. position set and neutral and smiling facial expression. The operation distance 1 m, Resolution of the image is 640*480, the following Fig. 1 show that the sample of face data.

The fingerprint KVKR database collection begins with the left-hand little finger and progresses to the right hand's little finger one by one.

Fig. 1. Face Data sample

Fig. 2. Fingerprint Data Sample

Each dry and natural finger has been gathered and the image resolution is 320 * 480 grayscale images. The sample of fingerprint data is shown in Fig. 2 show that the sample of fingerprint data.

4 Experimental Setup

The proposed method is constructed deep learning with Python 3.6 and other open-source library tools such as Keres, TensorFlow, CUDA and image processing libraries like OpenCV, matplotlib and scikit-learn, among others. The method runs on a laptop window 11 with an Intel Core-i5 CPU, NVidia 2 GB of memory and 8 Gb ram. The training models used Jupyter notebook IDE and the pre-processing was done using spider IDE.

In this study, we have used own collected same person samples of 30 subject KVKR face and fingerprint databases. The features extraction has done using transfer learning pre-train VGG16 with CNN models and CNN model. These extracted features have classified using SoftMax classifier. In the VGG16 model the default input size of the ImageNet wights images is $224 \times 224 \times 3$ but we are allowed to modify this size as per the requirement. Thereby, we are modifying the size as $64 \times 64 \times 3$ as per the experimental requirements. The hyperparameters was kept same for the all-transfer learning as well as CNN based experiments such as Epoch is 8, Batch Size is 64, learning rate is 0.001, Dropout rate is 0.1, activation function ReLU, Optimizer function as Adam, Loss function as categorical cross entropy (as it is multi-class classification) and SoftMax classifier is use as a classifier. The unimodal biometric face and fingerprint-based person identification system, we have used 960 images for training and 240 images for validation (80:20%).

5 Proposed Methodology

In multimodal face and fingerprint score fusion system, we have first applied pre-processing techniques on KVKR dataset, then divided the dataset in to different percentages for models training and used CNN and VGG16 with CNN models for calculating recognition rate. In the multimodal biometric person identification system, first time we have used 960 images for training and 240 images for validation (80:20%), second time used 720 images for training and 480 for validation (60:40%) and last time used 600 images for training and 600 images for validation (50:50%). In Fig. 3 shows general structure of multimodal biometric face and fingerprint fusion system. This system is divided into two phase those are training and testing. In training phase, first load KVKR dataset and resize in to 64 * 64, then used hstack method to combine the resize images features, then to train the networks using the proposed structure. The trained model wights will be stored and use this model wights in to testing phase for test images. In the testing phase, KVKR test dataset image lode and resized into 64 * 64, combine the face and fingerprint images features using hstack technique, then test image using previously stored model wights. Lastly, shown in the screen as the result of person recognition in multimodal biometric system as the person ID belonging to which class and fusion score.

5.1 Pre-processing

In this study, we have used two pre-processing techniques those are data augmentation and image enhancement. The data augmentation technique has used to artificially increase the size of training dataset by making multiple copies of the images. The more data may lead to more proficient deep learning neural network models. The augmentation technique can provide picture variations that help fit models generalize as well as overcome the overfitting problem. This study KVKR face data has applied rotation,

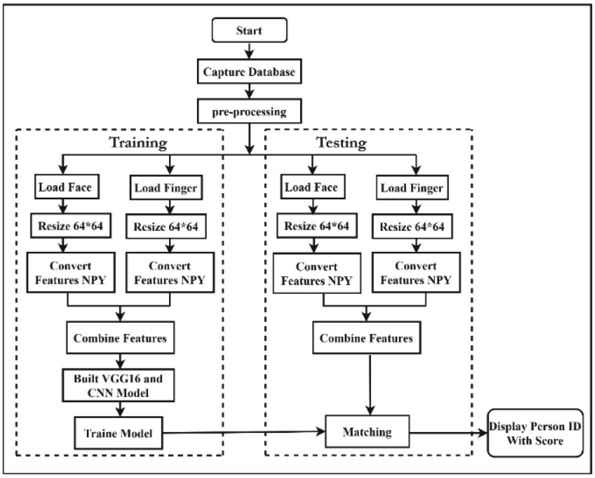


Fig. 3. Proposed Methodology

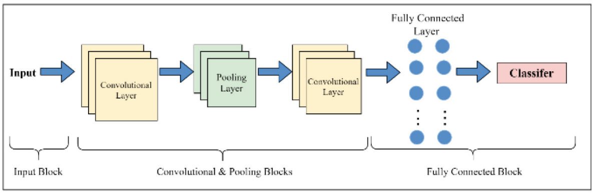


Fig. 4. Convolutional Neural Network Architecture

zoom and horizontal flip operation in the augmentation and KVKR fingerprint data have used cropping and brightness level operation in the augmentation technique. After augmentation the KVKR fingerprint data, we have applied image enhancement technique on KVKR fingerprint database. The image enhancement technique has some ridgelines in fingerprint images flow in various directions by applying Gabor-type filters on these ridgelines no noise between them is introduced and the resultant image is cleaner than the original [19].

5.2 CNN

CNNs are a deep learning approach that consists of numerous layers and is inspired by the biological visual cortex. CNNs are one of the most prevalent forms of neural networks used to detect and classify images and objects. These CNN models are used in a range of applications and domains, but they're especially popular in image and video processing. The building blocks of CNNs are filters and kernels. Using the convolution technique, the Kernels extract the appropriate information from the input. To train and evaluate deep learning CNN models, each input image is passed through a sequence of convolution layers with filters (Kernels), Pooling layers, fully connected layers (FC), and the SoftMax function, which uses probabilistic values to identification an item. These models have used convolutional layers for extracting the image features. Each convolutional contains a set of weighted matrices known as filters or kernels that slide over the input image to identify specific information. Colours and basic patterns are detected by the CNN's initial layers of filters. Then, as they progress through the levels, they begin to see more intricate patterns. Each filter uses a convolution operation to generate a feature map to discover features [17, 20]. In Fig. 4 these three blocks are used to constrict a CNN architecture by varying the size of blocks, addition or removing a block. This model has used three hidden layers and four convolution and max pooling layer.

5.3 VGG16

The VGG models were developed by the Visual Geometry Group at Oxford University. VGG16, one of the most common models, has 16 layers and the ImageNet database was used to train the VGG16 model extensively. This massive database has over 14 million photos divided into 20000 categories Five convolution blocks make up a VGG16 model. The VGG16 model consists of five convolution blocks. Each convolution block has two

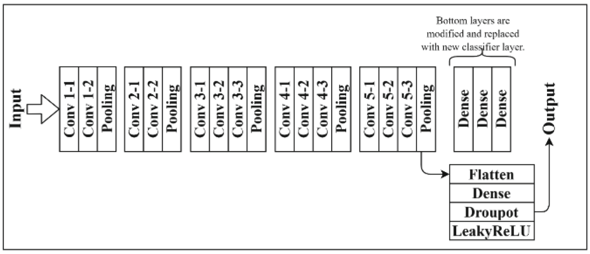


Fig. 5. Modified of VGG16 Model layers using CNN.

convolutional layers (size: 3 × 3) and one max-pooling layer (size: 2 × 2). The prediction and classification tasks are handled by the fully connected (FC) layers [20].

The Fig. 5 shows the architectural block diagram of the convolutional layer's configuration for VGG16 using transfer learning.

The initial layers are frozen layers and we cannot modify them. However, in order to perform transfer learning using this convolution architecture, the bottom layers are modified and replaced with new classifier layer using CNN model. This architecture used eight hidden layer of CNN model.

5.4 Features Classification

SoftMax: Classification of object is very important task normally different literature uses SVM based classifier, which works on the classification score based on hyper plane that separate the data into two categories but, SoftMax classifier predict the class label on the basis of calculated probabilities we used the fully connected layer for final calculation of score, as we calculating probabilities in previous layer and got the value for calculating label. The SoftMax function is the most often utilised activation function at the output layer for multi class classification. When using real values, the SoftMax function calculates the probability distribution and returns a value in the range of 0 to 1, with the total of the probabilities equating 1. The output layer in multiclass issues would have 'n' neurons, where n is the number of classes. Each neuron would provide a probability value for each class, with the predicted class being the neuron with the greatest value.

5.5 Score Level Fusion

In this approach, we have used pre-train transfer learning VGG16 with CNN and CNN model for the fusion of face and fingerprint. In this study, we have calculated similarity score, the fully connected layer is entered into the SoftMax classifier for measuring similar scores, then each subject score is fused using fixed-rule-based maximum rule technique for biometric score fusion. This technique has normalized final score vector and in the output identity of the subject shows who have the largest fusion score. This is a most popular method for fusion, in the maximum rule defined, 'f' is fusion score, 'xm' is the number of modalities used for fusion [17, 18]. This technique has selected the largest value score as fusion score. The following equation shows mathematical description:

$$f = \max(x1, x2, x3 \ldots\ldots x_m)\ldots\ldots\ldots \quad (1)$$

6 Performance Analysis

6.1 Classification and Confusion Matrix

The KVKR face and fingerprint database were used to test the face and fingerprint score fusion multimodal biometric identification system and the results were assessed using several evaluation metrics, including confusion matrix, Precision, Recall, Support, Micro and weighted average, and F1 Score. Precision depicts the model's positive predictive value, whereas recall depicts its sensitivity and true positive rate.

We utilized micro-averages to integrate the findings across the thirty categories to get the overall accuracy and recall. Figure 6 classification report of precision (P) and recall (R) rate of overall classes of VGG16 with CNN data split (80:20%). In the Fig. 7 show that the normalize and without normalize confusion matrix of the face and fingerprint score fusion result of VGG16 with CNN data split (80:20%). The confusion matrix vertical X axis shows that the true labels id of each class and horizontal axis Y show that the predicted labels id of each class.

	precision	recall	f1-score	support
ID 0 (ID_1)	1.00	1.00	1.00	40
ID 1 (ID_2)	1.00	1.00	1.00	40
ID 2 (ID_3)	1.00	1.00	1.00	40
ID 3 (ID_4)	1.00	1.00	1.00	40
ID 4 (ID_5)	1.00	1.00	1.00	40
ID 5 (ID_6)	1.00	1.00	1.00	40
ID 6 (ID_7)	1.00	1.00	1.00	40
ID 7 (ID_8)	1.00	1.00	1.00	40
ID 8 (ID_9)	1.00	1.00	1.00	40
ID 9 (ID_10)	1.00	1.00	1.00	40
ID 10 (ID_11)	1.00	1.00	1.00	40
ID 11 (ID_12)	1.00	1.00	1.00	40
ID 12 (ID_13)	1.00	1.00	1.00	40
ID 13 (ID_14)	1.00	1.00	1.00	40
ID 14 (ID_15)	1.00	1.00	1.00	40
ID 15 (ID_16)	1.00	1.00	1.00	40
ID 16 (ID_17)	1.00	1.00	1.00	40
ID 17 (ID_18)	1.00	1.00	1.00	40
ID 18 (ID_19)	1.00	1.00	1.00	40
ID 19 (ID_20)	1.00	1.00	1.00	40
ID 20 (ID_21)	1.00	1.00	1.00	40
ID 21 (ID_22)	1.00	1.00	1.00	40
ID 22 (ID_23)	1.00	1.00	1.00	40
ID 23 (ID_24)	1.00	1.00	1.00	40
ID 24 (ID_25)	1.00	1.00	1.00	40
ID 25 (ID_26)	1.00	1.00	1.00	40
ID 26 (ID_27)	1.00	1.00	1.00	40
ID 27 (ID_28)	1.00	1.00	1.00	40
ID 28 (ID_29)	1.00	1.00	1.00	40
ID 29 (ID_30)	1.00	1.00	1.00	40
accuracy			1.00	1500
macro avg	1.00	1.00	1.00	1500
weighted avg	1.00	1.00	1.00	1500

Fig. 6. Face and Fingerprint Classification Matrix of VGG16 with CNN (80:20)

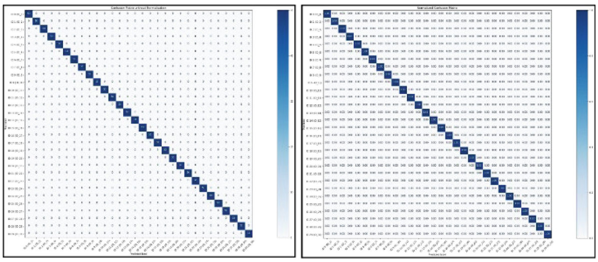


Fig. 7. Confusion Matrix of Face and Fingerprint VGG16 with CNN (80:20)

7 Result and Discussion

Here, the accuracy of the proposed VGG16 with CNN and CNN model. In this study, we have used the KVKR face and fingerprint same person database and calculated unimodal as well as multimodal biometric Accuracy. Figures 8 and 9 shows that the face and fingerprint recognition models accuracy and loss.

Figures 10, 11 and 12 show that the face and fingerprint score fusion model accuracy and loss.

In Table 2 shows, unimodal and multimodal recognition accuracy. In the face recognition we have got 98.86% accuracy and 1.14% equal error rate, fingerprint recognition got 87.08% accuracy and 12.92% equal error rate. In multimodal score fusion first time in VGG16 with CNN got 99.65% accuracy in (80:20%) split, second time got 99.30% accuracy (60:40%) split and third time got 99.25% accuracy (50:50%) split. The CNN model using first time got 99.50% accuracy (80:20%) split, second time got 99.10% accuracy (60:40%) split and third time got 98.89% accuracy (50:50%) split.

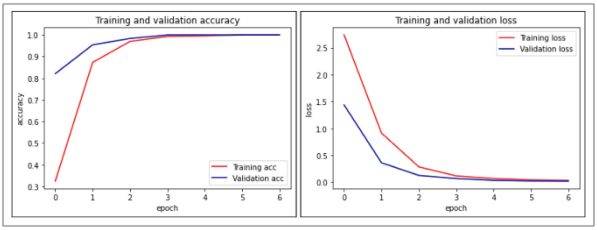


Fig. 8. Face Recognition Model Accuracy and Loss

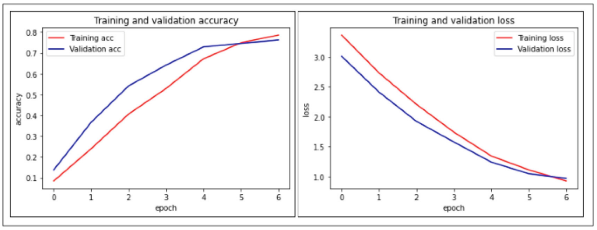


Fig. 9. Fingerprint Recognition Model Accuracy and Loss

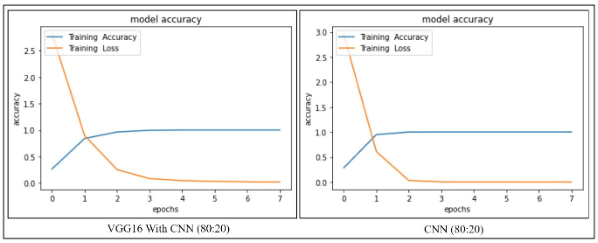


Fig. 10. Face & Fingerprint Fusion (80:20) Accuracy and Loss

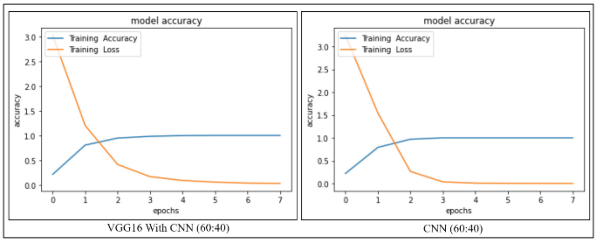


Fig. 11. Face & Fingerprint Fusion (60:40) Accuracy and Loss

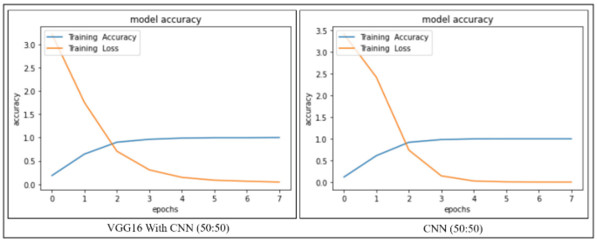
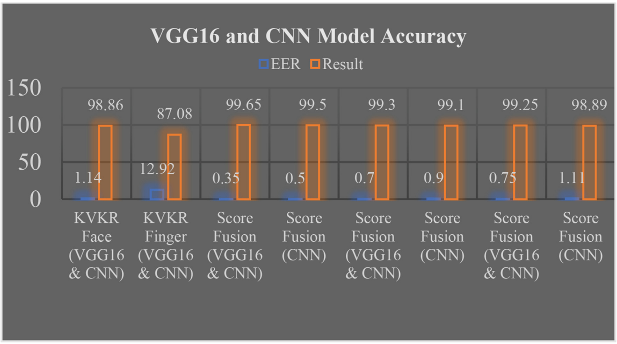


Fig. 12. Face & Fingerprint Fusion (50:50) Accuracy and Loss

In Graph 1 show that the comparative analysis of unimodal and multimodal biometric recognition accuracy. We have got in VGG16 with CNN model split (80:20%) database good recognition accuracy than unimodal biometric.

Table 2. Comparative Analysis of Unimodal and Multimodal Biometric Recognition Accuracy

Data	Techniques/ Algorithm	No. of Subject and data split percentages	EER	Accuracy
KVKR Face	VGG16 and CNN	30 (80:20)	1.14%	98.86%
KVKR Fingerprint	VGG16 and CNN	30 (80:20)	12.92%	87.08%
Face & Finger Score Fusion	VGG16 and CNN	30 (80:20)	**0.35%**	**99.65%**
Face & Finger Score Fusion	CNN	30 (80:20)	0.50%	99.50%
Face & Finger Score Fusion	VGG16 and CNN	30 (60:40)	0.70%	99.30%
Face & Finger Score Fusion	CNN	30 (60:40)	0.90%	99.10%
Face & Finger Score Fusion	VGG16 and CNN	30 (50:50)	0.75%	99.25%
Face & Finger Score Fusion	CNN	30 (50:50)	1.11%	98.89%



Graph 1. Comparative Analysis of Unimodal and Multimodal Biometric Recognition Accuracy

8 Conclusion

In this paper, we have used a deep learning technique to create a face and fingerprint based unimodal and multimodal person identification system. This work, we have used KVKR face and fingerprint own collected samples of same person database. First, we have applied data augmentation technique on the face and fingerprint database for artificially incurring the size of the database then fingerprint has applied image enhancement for improving ridgelines that run in different directions and the resulting image will be clearer than the original. In this study we have uses 30 subject databases and proposed VGG16 with CNN based face recognition got 98.86% accuracy and fingerprint recognition got 87.08% accuracy. In multimodal biometric person identification system, which uses the user's face and fingerprint using VGG16 with CNN model have got good recognition accuracy in (80:20%) split 99.65%. We have got in multimodal score fusion good recognition accuracy than the unimodal biometric system.

9 Contributions

- Most of the multimodal biometric recognition systems have used traditional methods and very few researchers have attempted the work in deep learning based.
- Our proposed system can be achieved good recognition accuracy in limited as well as larges datasets.
- We have designed a robust multimodal biometric system for person identification using face and fingerprint modality in a deep learning approach.

Acknowledgments. I would like to thank Dr. K.V. Kale for permission for me to collection of databases in the Multimodal Lab. I would acknowledge and thank CSMNRF, Pune for financial support SARTHI an Autonomous Institute Government of Maharashtra. Authors would like to extends thanks to Dr. Babasaheb Ambedkar Marathwada University for the publication support.

References

1. Shi Jinn H., Yuan Chen, Ray-Shine R., Rong-Jian C., Jui-Lin L., and Kevin Octavius S., An Improved Score Level Fusion in Multimodal Biometric Systems, International Conference on Parallel and Distributed Computing, Applications and Technologies, 2009.
2. M. Ahmad, W.L. Woo and S.S. Dlay, Multimodal Biometric Fusion at Feature Level: Face and Palmprint, IEEE, 2010.
3. S. V. and Jules R. Tapamo, Integrating Iris and Signature Traits for Personal Authentication Using User-Specific Weighting, Sensors, 2012.
4. Arun A. Ross, Karthik Nandakumar and Anil K. Jain, Handbook of Multibiometric, springer, 2006.
5. Krishna S. and Sumegh T., Development of Face and Signature Fusion Technology for Biometrics Authentication, International Journal of Emerging Research in Management &Technology, 2017.
6. Nada A. and Heyam H., Deep Learning Approach for Multimodal Biometric Recognition System Based on Fusion of Iris, Face, and Finger Vein Traits, Sensors, 2020.
7. Supreetha G. H D, Mohammad I., and Hemantha Kumar G., Feature level fusion of Face and Iris using Deep Features based on Convolutional Neural Networks, IEEE, 2018.
8. A. Ross and A. K. Jain, Information fusion in biometrics, Pattern Recognition Letters, 2003.
9. M. Anusha and T.V.V. Krishna, Multimodal Biometric System Integrating Fingerprint Face and Iris, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 10, October 2016.
10. E. Sujatha and A. C., Multimodal Biometric Authentication Algorithm Using Iris, Palm Print, Face and Signature with Encoded DWT, Springer, 2017.
11. Veeru T., Sobhan S., Matthew C. Valenti, and Nasser M. Nasrabadi, Learning to Authenticate with Deep Multibiometric Hashing and Neural Network Decoding, arXiv:1902.04149v3 [cs.CV] 7 Mar 2019.
12. El M. Cherrat, Rachid Alaoui, and Hassane Bouzahir, SCORE FUSION OF FINGER VEIN AND FACE FOR HUMAN RECOGNITION BASED ON CONVOLUTIONAL NEURAL NETWORK MODEL, International Journal of Computing, 2020.
13. El M. C., Rachid Alaoui, and Hassane Bouzahir, Convolutional neural networks approach for multimodal biometric identification system using the fusion of fingerprint, finger-vein and face images, Peer J Computer Science, 2020.
14. Priti S. and Yogesh H. D., Convolutional Neural Network Based Multimodal Biometric Human Authentication using Face, Palm Veins and Fingerprint, International Journal of Innovative Technology and Exploring Engineering (IJITEE) Volume-9 Issue-3, January 2020.
15. Arjun B. Channegowda and H N Prakash, Multimodal biometrics of fingerprint and signature recognition using multi-level feature fusion and deep learning techniques, Indonesian Journal of Electrical Engineering and Computer Science Vol. 22, No. 1, pp. 187–195, April 2021
16. Mehwish L., Shahzad Memon, Lachhman Das Dhomeja, Akhtar Hussain Jalbani and Asghar Ali Chandio, Deep Feature Fusion of Fingerprint and Online Signature for Multimodal Biometrics, MDPI, 2021.
17. Nada A. and Heyam H. Al-Baity, A multimodal biometric system for personal verification based on different level fusion of iris and face traits, Biosci. Biotech. Res. Comm. 12(3): 767-778, 2019.
18. Connor S. and Taghi M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning, Springer 2019.
19. Muhammad U. Munir and Dr. M. Y. Javed, Fingerprint Matching using Gabor Filters, National Conference on Emerging Technologies, 2004.

20. Karen S. and Andrew Zisserman, VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, Published as a conference paper at ICLR, 2015.
21. U. Gawande and Yogesh G., Biometric security system: a rigorous review of unimodal and multimodal biometrics techniques, Int. J. Biometrics, Vol. 10, No. 2, 2018.