Convolutional Neural Network Approach for Bimodal (Face and Fingerprint) Biometric Identification System

Aditi Sangle
Department of Computer Engineering
A. P. Shah Institute of Technology
Thane (M.H.), India, 400615
aditisangle2@gmail.com

Zeal Punamiya
Department of Computer Engineering
A. P. Shah Institute of Technology
Thane (M.H.), India, 400615
zealpunamiya15@gmail.com

Vedang Gore
Department of Computer Engineering
A. P. Shah Institute of Technology
Thane (M.H.), India, 400615
vedang.nilesh.gore@gmail.com

Soumyojyoti Dutta
Department of Computer Engineering
A. P. Shah Institute of Technology
Thane (M.H.), India, 400615
soumyojyoti.datta@gmail.com

Prof. Sachin Malave
Department of Computer Engineering
A. P. Shah Institute of Technology
Thane (M.H.), India, 400615
shmalave@apsit.edu.in

***Abstract*—Sometimes the process of machine learning is found to be difficult for analysing and computing the desired result. This generally happens when the dataset is relatively small. Therefore, we have considered taking images for faces as well as fingerprints. This is keenly effective security wise and training wise too. In this paper we focus on working on the learning ability of the Siamese neural network which is one of the most famous networks among the other convolutional networks. In this study, we investigate a technique for training Siamese neural networks, which use a special structure to prioritize input similarity. When the training is done, the model then has the ability to extend its prediction qualities not only to new classes, but also to unknown distribution sets as well. The base being a convolutional architecture, we are therefore able to get a much higher and stronger results in comparison to other deep neural networks. The purpose of this paper is affected by the inefficient accuracy provided when only one biometric system is used. For example, in the case of twins the face might be considered as same by the model; this would be a huge breach in the security of the whole state. In order to rectify it we take in face as well as fingerprint images which, for a twins case which works perfectly with high accuracy.**

***Keywords*—convolutional networks, biometric system,Siamese neural network, security, accuracy.**

I. Introduction

Devices that use biometric identification rely on physical traits such a fingerprint, face patterns, iris or retinal patterns to confirm a user's identity. Network logon is just one application where biometric authentication is gaining popularity. For the device to compare a fresh sample provided during the logon process to a biometric template or identifier (a sample known to be from the authorized user), it is necessary to store the information in a database. The two most often used techniques for biometric authentication are face recognition and fingerprint recognition. There are some locations where a higher level of protection is necessary that can be equipped with a face recognition and fingerprint recognition system. When you offer a physical aspect for authentication, biometric authentication is carried out by comparing it to a copy that has been stored, biometric authentication is carried out by comparing it to a copy that has been stored.

Fingerprint authentication is a type of biometric authentication that instantly verifies a user's identification by comparing their fingerprint to a template that has been stored. Due to the fact that each person has a distinct fingerprint, fingerprint scans are an inherent characteristic or "something you are," making them impossible to predict and challenging to change or falsify. By analyzing and comparing the dermal ridges on a person's fingers, fingerprint recognition enables verification or identification of that person. One of the earliest methods for automatically recognising individuals was fingerprint recognition, which is still one of the most widely used and reliable biometric methods today. To ascertain whether the reported identity of the person is accurate, it performs a one-to-one comparison. The submitted claim of identification is either rejected or accepted by a verification system.

Face recognition is a biometric identification technique that relies on body measurements—in this case, the face and head—to confirm a person's identity using their facial biometric pattern and data. The technique uses a collection of specific biometric information about each person's face and expression to identify, confirm, and/or authenticate them. Finding several instances of the same face in a database of training photos from an input image is the goal of face recognition. Face recognition is one of the areas of pattern recognition and computer vision study because it has so many real-world applications in the fields of biometrics, information security, access control, law enforcement, smart cards, and surveillance.

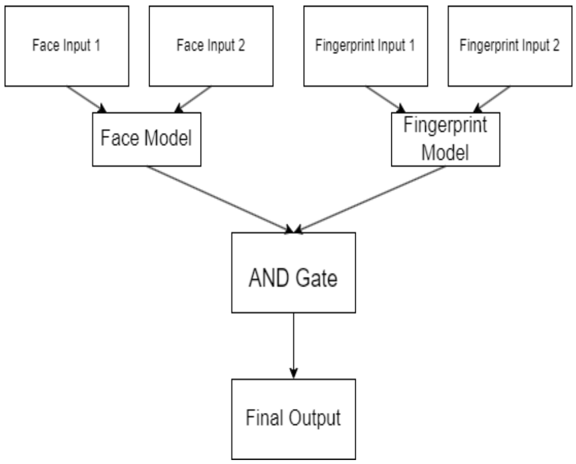

Fig. 1. Flow of the System [2]

Multimodal biometric systems boost performance and resilience against imposters' attacks and environmental changes in order to get beyond the limitations associated with systems based on a single modality of biometric data. This system falls within the categories of hybrid, multi-instance, multi-sensor, multi-algorithm, and multi-modal systems. The system for multi-biometric recognition incorporates data from many biometric sources. The key benefit of a multimodal system over a standard single biometric is that it makes recognition more precise and safe. The remaining portion of the paper is divided as follows: A literature review is a brief summary of related studies. Our suggested algorithm's supporting procedures are referred to as the proposed system. Conclusion brings the proposed work to a close while Experimental results and discussion clarify the experimental results.

II. Literature Survey

There are many ways to implement face and fingerprint recognition. After discussing about several ways, we have chosen the Siamese Neural network. Many studies have been done on these biometric systems.

According to Cherrat et al., 2017[1], The biometrics modalities are often unique, measurable or automatically validated or permanent.

According to Borra, Reddy and Reddy, 2018[2], Fingerprint have become an essential biometric trait due to its uniqueness and invariant to every individual. This biometric modality is more used and more acceptable by users as the acquiring device is comparatively small. Moreover, the recognition accuracy is relatively much higher than the others biometric recognition system based on the retina, ear shape, iris, etc. Face recognition is a biometric recognition technology based on facial feature (human of course) information (feature vectors) for identification or verification.

According to Mane and Shah, 2019[2], human and computer performance on facial identification is a research topic with both scientific research value and widely application prospects. The bimodal biometric recognition system combines any two biometric recognition system. The main advantage of multimodal system against traditional single biometric is the recognition process we achieve is more secure and accurate (Unar, Seng and Abbasi, 2014)[3]. The advantage of using face and fingerprint is these things are more natural for recognizing and their deployment cost is relatively low.

Ross and Jain (2003)[4] presented information fusion in a multimodal system at various levels, the performance of different normalization techniques and fusion rules in the context of a multimodal biometric system based on the face, fingerprint and hand -geometry features of a user. They also observed that multimodal systems utilizing user-specific weights perform better compared to systems that assign the same set of weights to the multiple biometric traits of all users.

Yang and Zhang (2012)[5] made experimental systems with a fusion of fingerprint and finger vein. These biometric features were extracted using a unified Gabor filter method. The feature level fusion is generated based on supervised local preserving correlation analysis framework. This work is evaluated using a database only of 600 fingerprints and finger veins respectively.

Son and Lee (2005)[6] have been subjected a fusion of face and iris using reduced joint feature vector method. However, it is not verified on a large amount of data.

Ross and Govindarajan (2005)[7] presented multimodal biometric system that uses hand and face feature level fusion for biometric recognition purposes. Moreover, the experiments have been tested on both intra-modal and inter-modal fusion with R, G, B channels. But it does not accord eigen-coefficients of face and minutiae points of fingerprints, so still lacks significant accuracy.

Ma, Popoola and Sun (2015)[8] achieved a good fingerprint and finger vein identification system by concatenating the feature vectors. But, the accuracy of this technique is so less, it does not satisfy the requirements of real-world applications.

Huang et al. (2015)[9] used adaptive face and ear recognition system(bimodal) based on sparse coding. Talking about Siamese neural network, Koch, Zemel and Salakhutdinov[10] used Siamese neural network for one shot image recognition. Here, image representations were learnt via a supervised metric-based approach using Siamese neural network then reuse the network's feature for one short learning without any re-training.

Lake et al.[11] approached the problem of one-shot learning by inverting a composition casual process, addressing one-shot learning for character recognition with a method called Hierarchical Bayesian Program Learning

III. Problem Statement

To propose a hybrid system of Neural Network: Siamese Convolutional neural network (CNN) based multi-biometric Fingerprint and face identification system. The score provided by these systems is combined for improving Human identification not only in companies but also in other industries too where manual attendance or recognition is of utmost importance. The need of authentication is the need of the hour. For security systems to work correctly identifying the person is very important. So in order to propose a hybrid bimodal model of siamese convolutional neural network to be used for the purpose of verifying the identity of a person using face and fingerprint data , this project is deployed. To inaugurate a system which enhances the current traffic management by developing intelligent traffic lights based upon machine learning and artificial intelligence.

IV. Scope

• This system is classified as multi-modal and hybrid systems.
• In order to overcome the limitation concerned unimodal biometric system, the multimodal biometric system increase the robustness and performance against the imposter‘s attack and environment variations.
• The advantage of combining the fingerprint and face is its devices are less expensive and easier to deploy.
• Moreover, the face is one of the most natural methods to identify an individual, it does not restrict the movement of the person and its deployment cost is relatively low.
• The user will give an input image of face or fingerprint, the model will compare it with the predefined verification images and predict if the two images of the user are same or not.

V. Proposed System

A. Proposed System Overview

The Siamese nets were first introduced in the early 1990s by Bromley and LeCun to solve signature verification as an image matching problem (Bromley et al., 1993). A Siamese neural network consists of twin networks which accept distinct inputs but are joined by an energy function at the top. This function computes some metric between the highest level feature representation on each side. The parameters between the twin networks are tied. Weight tying guarantees that two extremely similar images could not possibly be mapped by their respective networks to very different locations in feature space because each network computes the same function. We propose the use of a siamese convolutional neural network for the purpose of bimodal biometric verification. The two modes we use here are face and fingerprint.

B. Architecture of the model

An architectural diagram is a visual representation that maps out the physical implementation for components of a software system. It shows the general structure of the software system and the associations, limitations, and boundaries between each element. The input images and anchor(verification) images of face are taken as input by the face model. The model then does feature extraction and encodes the images. The difference layer of the model then calculates by how much do these images differ which is then flattened by the fully connected flatten(dense) layer and then if the difference is less then we get a score greater than the threshold value and the face is verified as Yes(True). If the difference is greater then we get a score less than the threshold value and the face is verified as No(False). The same process happens with the fingerprint input and anchor images and we get a value of either yes(True) or No(False). Then we apply the logical AND gate to verify if both the outputs are True or not and output it as the final result.

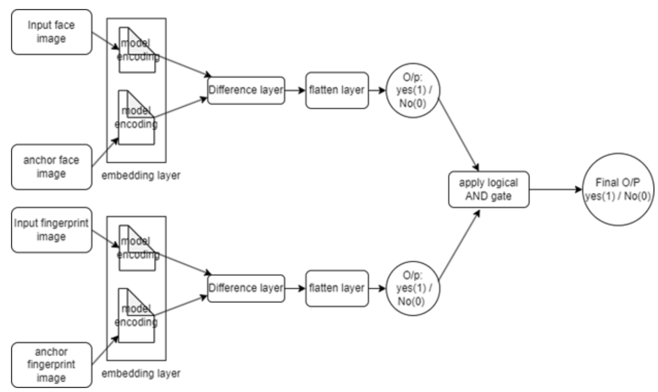

Fig. 2. Architectural Diagram

C. Data Flow of the model

A Data Flow Diagram is a graphical representation of the “flow” of data through an information system, modeling its process aspects. Often it is a preliminary step used to create an overview of the system that can later be elaborated.
Level-0 Data collected from components that is face input 1, face input 2, fingerprint input 1 and fingerprint input 2 are sent to the respective Models, then the model gives the prediction and then AND gate is applied on the predictions of the models giving the final output . Level-1 Data is collected from user in the form of images. This data is then sent to the respective Neural Network Model for intermediate processing. The verification status is predicted for both face and fingerprint data. AND gate is applied between the two predictions and the final output is recieved.

D. About the model

Using L layers and Nl units per layer, our typical model is a siamese convolutional neural network, where h1,l stands for the hidden vector in layer l for the first twin and h2,l for the second twin. In the first L 2 layers, rectified linear (ReLU) units are our only choice. The final layers contain sigmoidal units.

The model is composed of a series of convolutional layers, where each layer employs a single channel, filters of various sizes, and a fixed stride of 1. To maximise performance, a multiple of 16 is specified for the number of convolutional filters. The output feature maps are subjected to a ReLU activation function by the network, which may be followed by maxpooling with a filter size and stride of 2. As a result, the kth filter map for each layer takes the following form:

$$a_{1,m}^{(k)} = \max - \text{pool}\left(\max\left(0, \mathbf{W}_{l-1,l}^{(k)} \star \mathbf{h}_{1,(l-1)} + \mathbf{b}_l\right), 2\right)$$
$$a_{2,m}^{(k)} = \max - \text{pool}\left(\max\left(0, \mathbf{W}_{l-1,l}^{(k)} \star \mathbf{h}_{2,(l-1)} + \mathbf{b}_l\right), 2\right)$$

where $\mathbf{W}_{l-1,l}$ is the 3-dimensional tensor representing the feature maps for layer l and we have taken $\star$ to be the valid convolutional operation corresponding to returning only those output units which were the result of complete overlap between each convolutional filter and the input feature maps.

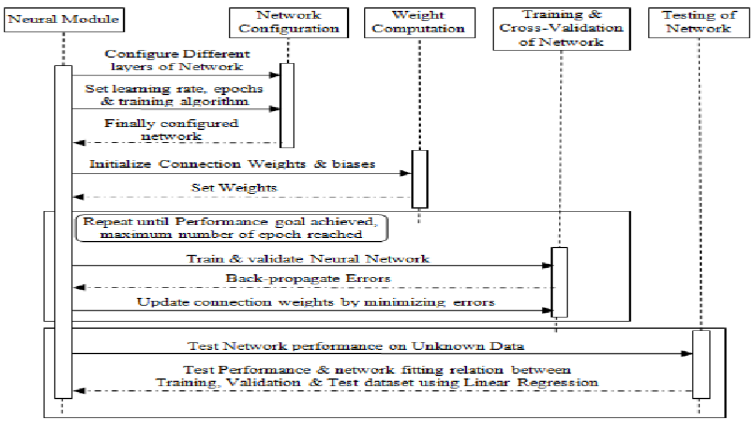

Fig. 3. Sequence diagram

VI. Design Details

A. General Dataset

For implementation of this project we have made use of two datasets. The first one is the labelled faces in the wild dataset and the second dataset is our own dataset that we have created.

I. Labelled faces in the wild dataset: The data set contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured. 1680 of the people pictured have two or more distinct photos in the data set. The only constraint on these faces is that they were detected by the Viola-Jones face detector.

B. Personal Biometric Dataset

Our biometric dataset consists of a total of 5151 images out of which 2410 images are face images and 2741 images are fingerprint images. The face images are of three subjects with subject1 having 729 face images, subject2 having 867 face images and subject3 having 814 face images. The fingerprint images contain fingerprint images of index finger of three subjects with at least 600 images each. There is also a set of 900 fingerprint images of random fingers from these subjects.

Fig. 4. Labelled faces in wild dataset

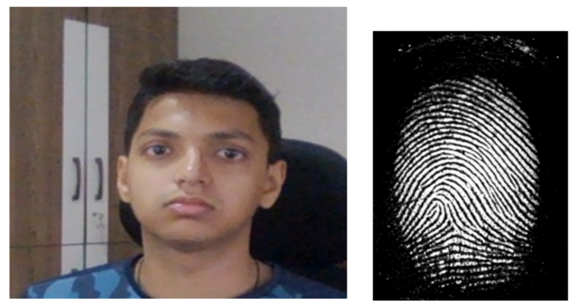

Fig. 5. our own dataset sample image 1

Fig. 6. our own dataset sample image 2

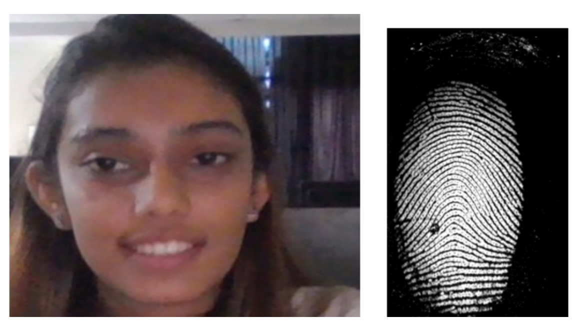

Fig. 7. our own dataset sample image 3

C. Model Details

The following figure is the summary of the layers of the embedding for the model. We have used a 2d convolution layer to filter and enhance the features of the image. Then the max pooling layer gives the feature map of the maximum prominent features. We then again use this same combination twice to enhance and filter and extract the most prominent features of the image. Then we use a flatten layer to bring all the features in the same dimension. Then we use a dense layer as the fully connected layer.

```
embedding.summary()
Model: "embedding"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_image (InputLayer)    [(None, 100, 100, 3)]     0

 conv2d (Conv2D)             (None, 91, 91, 64)        19264

 max_pooling2d (MaxPooling2D  (None, 46, 46, 64)       0
 )

 conv2d_1 (Conv2D)           (None, 40, 40, 128)       401536

 max_pooling2d_1 (MaxPooling  (None, 20, 20, 128)      0
 2D)

 conv2d_2 (Conv2D)           (None, 17, 17, 128)       262272

 max_pooling2d_2 (MaxPooling  (None, 9, 9, 128)        0
 2D)

 conv2d_3 (Conv2D)           (None, 6, 6, 256)         524544

 flatten (Flatten)           (None, 9216)              0

 dense (Dense)               (None, 4096)              37752832

=================================================================
Total params: 38,960,448
Trainable params: 38,960,448
Non-trainable params: 0
_________________________________________________________________
```

Fig. 8. Layers

VII. Experimental Setup

A. Software Requirements

- Windows 7 or higher OS
- Jupyter Notebook: Project Jupyter is a project with goals to develop open-source software, open standards, and services for interactive computing across multiple programming languages. It was spun off from IPython in 2014 by Fernando Pérez and Brian Granger. Project Jupyter's name is a reference to the three core programming languages supported by Jupyter, which are Julia, Python and R. Its name and logo are an homage to Galileo's discovery of the moons of Jupiter, as documented in notebooks attributed to Galileo. Project Jupyter has developed and supported the interactive computing products Jupyter Notebook, JupyterHub, and JupyterLab. Jupyter is financially sponsored by NumFOCUS.
- Tensorflow : TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow was developed by the Google Brain team for internal Google use in research and production. The initial version was released under the Apache License 2.0 in 2015. Google released the updated version of TensorFlow, named TensorFlow 2.0, in September 2019. TensorFlow can be used in a wide variety of programming languages, including Python, JavaScript, C++, and Java. This flexibility lends itself to a range of applications in many different sectors.
- Python: Python is one of the widely used programming languages for building systems that indulge in Image Processing as well as Machine Learning. Python provides amazingly powerful libraries and tools that help us in achieving the tasks efficiently.
- Scikit Learn: Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. Scikit-learn is a NumFOCUS fiscally sponsored project.
- Software needed to activate the GPU: CUDA (or Compute Unified Device Architecture) is a parallel computing platform and application programming interface (API) that allows software to use certain types of graphics processing units (GPUs) for general purpose processing, an approach called general-purpose computing on GPUs (GPGPU). CUDA is a software layer that gives direct access to the GPU's virtual instruction set and parallel computational elements, for the execution of compute kernels. CUDA is designed to work with programming languages such as C, C++, and Fortran. This accessibility makes it easier for specialists in parallel programming to use GPU resources, in contrast to prior APIs like Direct3D and OpenGL, which required advanced skills in graphics programming. CUDA-powered GPUs also support programming frameworks such as OpenMP, OpenACC and OpenCL; and HIP by compiling such code to CUDA. CUDA was created by Nvidia. When it was first introduced, the name was an acronym for Compute Unified Device Architecture, but Nvidia later dropped the common use of the acronym.
- Google Colab(if gpu is unavailable): Colaboratory, or "Colab" for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing access free of charge to computing resources including GPUs.
- Matplotlib: Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK. There is also a procedural "pylab" interface based on a state machine (like OpenGL), designed to closely resemble that of MATLAB, though its use is discouraged.[3] SciPy makes use of Matplotlib.

B. Hardware Requirements

List of items used	
Hardware	Specification
Fingerprint Scanner	Futronic FS 80
GPU	Nvidia RTX 3060
Processor	AMD Ryzen 7
Software	Specification
Nvidia CUDA	11.2
Nvidia CuDNN	8.1.0
FTR Scan software	-
Tensorflow	2.9

VIII. Results

Fig. 9. Verification of face

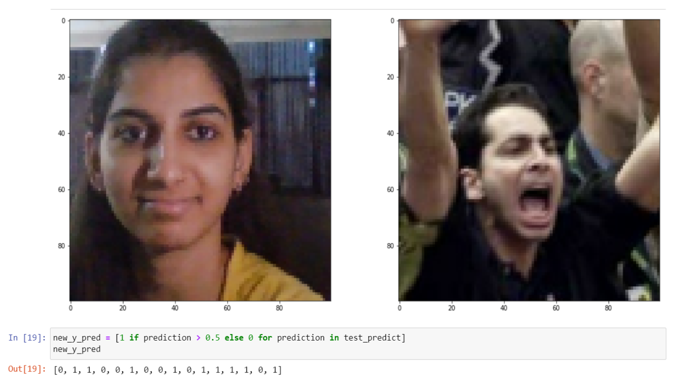

Fig. 10. Verification of face

This is how we get the predicted values for a test batch of face data. The images are of index 0 and if you see the output in the next block we get a 0 at the 0th index as the face images are not of the same person.

This is how we get the predicted values for a test batch of fingerprint data. The images are of index 7 and if you see the output in the next block we get a 1 at the 7th index as the face images are not of the same person.

Fig. 11. Verification of fingerprint

Fig. 12. Verification of face

Here we run the model to verify if the faces are same,if they are same we get the result true or else false.

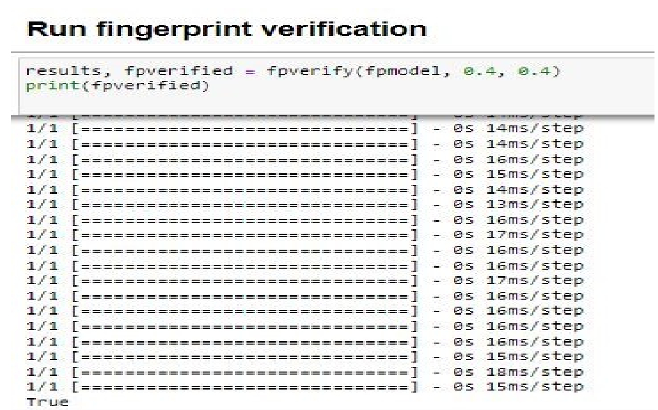

Fig. 13. Verification of fingerprint

Here we run the model to verify if the finger print images are same,if they are same we get the result true or else false. After obtaining the two results of face and fingerprint

```
final_verified = AND(faceverified, fpverified)
print('The verification status is:', final_verified )
```

```
The verification status is: True
```

Fig. 14. Final verification status

respectively ,they are then processed in AND gate to get the final verification status

IX. Conclusion

In this study, we have introduced a system for human recognition , which is based on combination of face and fingerprint.We have used Siamese neural network and CNN for this recognition system. We have also shown why this is better than unimodal biometric systems based on Siamese neural network. Overall, biomodal biometric recognition systems represent an exciting and rapidly advancing field with tremendous potential for improving security and convenience in a wide range of applications. With continued research and development, these systems are poised to become a critical component of our digital lives in the years to come.

References

[1] Cherrat EM, Alaoui R, Bouzahir H, Jenkal W. 2017. High-density salt-and-pepper noise suppression using adaptive dual threshold decision based algorithm in fingerprint images. In: 2017 Intelligent Systems and Computer Vision (ISCV). Fez: IEEE, 1–4.

[2] Borra SR, Reddy GJ, Reddy ES. 2018. An efficient fingerprint identification using neural network and BAT algorithm. International Journal of Electrical Computer Engineering 8(2):1194–1213DOI 10.11591/ijece.v8i2.pp1194-1213

[3] Mane S, Shah G. 2019. Facial recognition, expression recognition, and gender identification. In: Balas V, Sharma N, Chakrabarti A, eds. Data Management, Analytics and Innovation. Singapore: Springer, 275–290.

[4] Unar JA, Seng WC, Abbasi A. 2014. A review of biometric technology along with trends and prospects. Pattern Recognition 47(8):2673–2688 DOI 10.1016/j.patcog.2014.01.016

[5] Ross A, Jain A. 2003. Information fusion in biometrics. Pattern Recognition Letters 24(13):2115–2125 DOI 10.1016/S0167-8655(03)00079-5.

[6] Yang J, Zhang X. 2012. Feature-level fusion of fingerprint and finger-vein for personal identification. Pattern Recognition Letters 33(5):623–628 DOI 10.1016/j.patrec.2011.11.002.

[7] Son B, Lee Y. 2005. Biometric authentication system using reduced joint feature vector of iris and face. In: International Conference on Audio- and Video-Based Biometric Person Authentication. Berlin, Heidelberg: Springer, 513–522

[8] Ross AA, Govindarajan R. 2005. Feature level fusion of hand and face biometrics. In: Biometric Technology for Human Identification II, Orlando, Florida, United States. Vol. 5779. International Society for Optics and Photonics, 196–205 DOI 10.1117/12.606093.

[9] Ma H, Popoola OP, Sun S. 2015. Research of dual-modal decision level fusion for fingerprint and finger vein image. International Journal of Biometrics 7(3):271–285DOI 10.1504/IJBM.2015.071949.

[10] Huang Z, Liu Y, Li X, Li J. 2015. An adaptive bimodal recognition framework using sparse coding for face and ear. Pattern Recognition Letters 53:69–76 DOI 10.1016/j.patrec.2014.10.009

[11] Koch, Zemel Salakhutdinov Siamese neural network for one shot image recognition Lake, Brenden M, Salakhutdinov, Ruslan R, and Tenenbaum,Josh. One-shot learning by inverting a compositionalcausal process. In Advances in neural information processing systems, pp. 2526–2534, 2013