COMPUTER METHODS IN MATERIALS SCIENCE
Informatyka w Technologii Materiałów

Publishing House
A K A P I T

Vol. 20, 2020, No. 2

A NOVEL FEATURE BIOMETRIC FUSION APPROACH FOR IRIS, SPEECH AND SIGNATURE

MAMTA GARG[1*], AJAT SHATRU ARORA[2], SAVITA GUPTA[3]

[1]*Department of Computer Engineering, Govt Polytechnic College For Girls, Jalandhar, India, 144001.*
[2]*Department of Electrical and Instrumentation Engineering, Sant Longowal Institute of Engineering & Technology, Longowal, India, 148106.*
[3]*Department of Computer Science and Engineering, University Institute of Engineering & Technology, Panjab University, Chandigarh, India, 160016.*
[*]*Corresponding author: mamta2050@yahoo.co.in*

Abstract

With an ever-increasing emphasis on security and the new dimensions in security challenges facing the world today, the need for automated personal identification/verification system based on multimodal biometrics has increased. This paper addresses the issue of multiple biometric fusion to enhance the security of recognition. The paper utilizes iris, speech,and signature for the novel fusion. A segregated classification mechanism for each biometric is also presented. The fusion is done on the base of features extracted at the time of individual classification of biometrics. Different feature extraction algorithms are applied for different biometrics. The paper has utilized 2-Dimensional Principle Component Analysis (2DPCA) for Iris, Scale Invariant Feature Transform (SIFT) for signature and Mel-frequency cepstral coefficients for speech biometric. This paper utilizes Genetic Algorithm for the optimization of the evaluated features. The classification is done using Artificial Neural Network (ANN).

Key words: Biometric Fusion, Scale Invariant Feature Transform, 2-Dimensional Principle Component Analysis, Mel-Frequency Cepstral Coefficient, Genetic Algorithms, Artificial Neural Networks

1. INTRODUCTION

Biometric Fusion is a combination of two or more than two biometrics. There can be N number of biometrics which can be combined to form a fusion. A fused biometric is better than a unimodal biometric classification due to high sophistication in the structure of formation. Figure 1 demonstrates the single biometric classification mechanism.

This paper utilizes the feature based fusion approach for Iris, Signature and Speech biometric. Separate feature extraction algorithms were utilized for different biometrics. Optimization of the extracted features is not a compulsory step but it may enhance the relevance of the features extracted. A brief

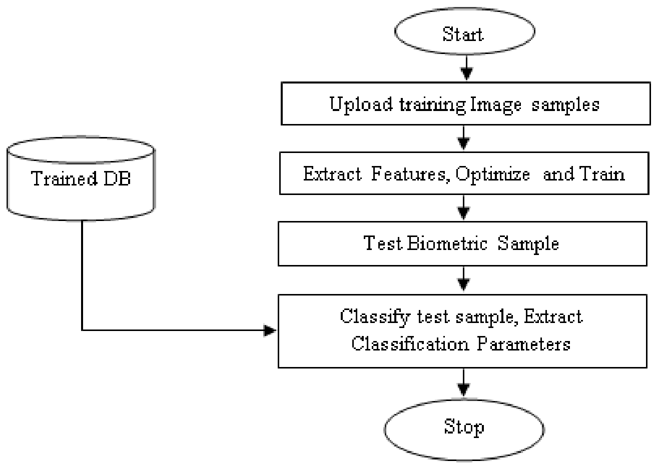

Fig. 1. *General Unimodal biometric classification mechanism.*

survey of literature for feature extraction, optimization and classification is as follows.

Cui et al. (2004) extracted PCA as key features for their proposed research work. They utilized CASIA Irıs dataset for training and classification and applied inter- and intra-class distance distribution for the classification process. Haung and Wang (2006) applied Genetic Algorithm (GA) for the optimization of the feature sets. The authors followed the architecture similar to figure 1 for training and classification and employed Support vector Machine (SVM) as the classification algorithm. SVM is a binary classifier (Yu & Kim, 2012) and hence it can only classify only true and false. Lin et al. (2008) implemented Swarm Intelligence as key optimization and used SVM for classification. They used Radial Basis Function (RBF) as the kernel to SVM. Karouni et al. (2011) presented the same type of architecture but for signature verification. Instead of SVM, classification was performed through Neural Network (NN), which is a multiclass classifier (Ou & Murphey, 2007) and uses a three-layer architecture. It surely provides better efficiency and classification accuracy as compared to SVM. Malode and Sahare (2017) used MFCC for feature extraction and Hidden Markov Model (HMM) for classification in speech recognition system. It has been empirically proven in many publications (Chetty & Wagner, 2005; Chen & Chu, 2006; Rattani et al., 2007; Zhang et al., 2007; Rattani & Tistarelli, 2009; Almayyan et al., 2011; Liau, & Isa, 2011; Bokade & Sapkal, 2012; Park & Kim, 2013; Nadheen & Poornima, 2013; Dhameliya & Chaudhari, 2013; Eskandari et al., 2014; Saleh & Alzoubiady, 2014, Veluchamy & Karlmarx, 2016; Haghighat, 2016; Sarhan et al., 2017; Leghari et al., 2018; Carol & Fred, 2018; Supreetha Gowda et al., 2018) that multimodal biometrics systems improve the recognition accuracy by integrating complementary information over unimodal biometrics systems. Features represent rich information about biometrics; fusion at feature level is believed to be bestowing better performance (Veluchamy & Karlmarx, 2016).

2. PROPOSED METHODOLOGY

One of the major driving forces behind the development of the proposed framework was to demonstrate that it is possible to design an effective deployable multimodal biometric system with reduced subset of features. The proposed work uses multimodal biometric fusion in order to enhance the security. The proposed algorithm is done in two phases.

- Unimodal classification
- Multimodal fusion

Figure 2 demonstrates the framework of the proposed algorithm. The different unimodal (Iris, Signature, and Speech) are utilized to form a multimodal biometric fusion. The iris unimodal classification uses 2DPCA for feature extraction followed by Genetic Algorithm (GA) as feature optimization. In the similar fashion, the signature takes Scale Invariant Feature Transform (SIFT) as key feature extraction algorithm followed by the GA for the optimization. Speech authentication utilizes Mel-Frequency-Cepstrum for feature extraction followed by GA for feature optimization.

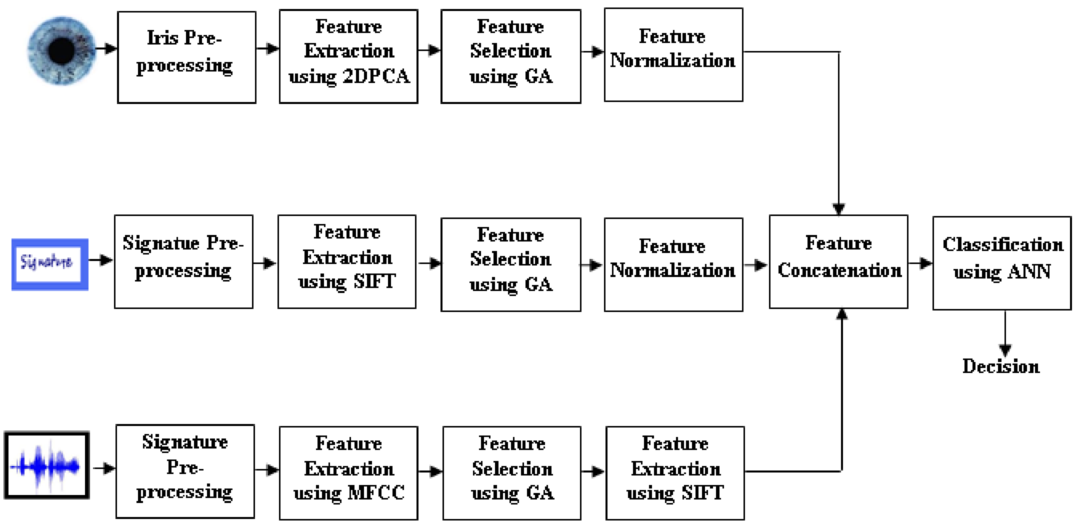

Fig. 2. General framework of work flow architecture.

In this research, we have applied 2DPCA to extract features and to represent iris image with low dimensionality features. 2DPCA is used to lower the dimensionality of the data set containing large number of interrelated variables, while maintaining the variations in the data sets to the great extent. It is a typical statistical technique based on orthogonal transform to change a set of values of correlated variables into a set of values of uncorrelated variables called principal components. The main function of 2DPCA is to reduce the large dimensionality of the observed variables to the smaller dimensionality of independent variables. The extraction of the best parametric representation of acoustic data is an important task in the design of any speaker recognition system. The speech features must provide a sufficient representation of the speech signal. The Mel-Frequency Cepstral Coefficients (MFCC) based on Mel scale is a foremost approach used for extraction speech features.

Mel-Scale operates frequencies below 1000 Hz with a linear behavior; whereas frequencies over a 1000 Hz are handled logarithmically. It gives more importance to lower frequency changes. Hence MFCC is based on the known fact that the low-frequency components of the speech signal carry information which is phonetically more significant for human perception than carried by high-frequency components. It means that humans are more stimulated by changes at lower frequencies than changes at higher ones.

The main steps for finding the MFCC coefficients are taking the log magnitude spectrum of the windowed waveform and then passed through triangular filters and finally computing the DCT of the waveform to generate the MFCC coefficients. After localization the signature, the main focus is on to obtain features from the image that are invariant to scale, position and orientation. To obtain transformation invariance between the features, local feature extraction algorithms like SIFT are used.

It generates descriptors representing the texture around the key-points. The SIFT algorithm has broken down to four steps. In the first step, scale space extrema using the Difference of Gaussian (DoG) is constructed. After that, significant key point candidates are localized and refined by eradicating the low contrast points. Thirdly, a key point orientation assignment based on local image gradient is performed and in the last step, a descriptor is generated to compute the local image descriptor for each key point based on image gradient magnitude and orientation.

Each biometric section is trained by Artificial Neural Network. If the ANN is adopted without applying feature selection technique, then the large dimensionality of the extracted features would cause to degrade performance of the ANN. So robust feature selection technique is implemented to reduce irrelevant data and maintain discriminating power of data. The three layer architecture of Neural Network takes the optimized feature set of each biometric at the input layer. The target set for each trained sample is the identity number for which the Neural is trained. As for example if there are 5 samples, the target set would be {1,2,3,4,5} for the feature vectors. Algorithm 1 demonstrates the working of training of each individual identity for each biometric section.

Algorithm 1. Train Biometrics (Iris_Set, Signature_Set, Speech_Set)

```
Foreach BiometricS in Biometrics // Taking individual biometric section
Foreach Bv in BioMetric// For each sample in Biometric
F_value(Bv)=ExtractFeature(Bv);
Target(Bv)=Bv;// Taking identity value as the target set of the biometric
End For
Initialize NeuralNetwork( F_value, Target, X) // The input layer takes the
feature vector,
//Target set and X number of neurons as the input  layer
Train(); // Training the Neural Network
Save to DB( );
End
```

Every neural network propagates the information with the help of neurons, mentioned as X in algorithm 1. The way of propogation can be either liner, polynomial or quadratic.

Table 1 represents the values of extracted and optimized set range for training for different biometrics. The range listed in table 1 is for a single image of each biometrics. A total of 500 samples are trained in the proposed architecture.

***Table 1**. Range of single image for each biometric.*

Iris 2DPCA	GA	Signature SIFT	GA	Speech MFCC	GA
20×280	6×280	20×128	6×128	12-dimensional	6-dimensional

Table 1 illustrates that when 2DPCA is applied to an image, it returns a 20×280 Eigen based PCA points and after the optimization through GA, it reduces to 6×280. In the similar fashion, for signature, SIFT is applied and a 20×128 key point set is evaluated followed by optimization technique GA which results in a 6×128 optimized vector set. When the speech goes through feature vector extraction through MFCC, it returns a Mel – Coefficient of size 12-dimensional MFCCs and after optimization, it reduces to 6-dimsensional MFCCs. The architecture of GA is presented in table 2.

***Table 2**. Architecture of GA.*

Parameter	Values
Population Size	Total Feature Count for each biometric
Selection Function	Roulette wheel selection
Mutation Type	Intermediate
Iterative Proceeding Fitness Function	1 if Fs>Ft otherwise Fs is a current selected feature value Ft is a threshold of fitness

The proposed algorithm utilizes a mixture of standard and real-time data set. The iris dataset is collected from http://biometrics.idealtest.org. As the standardization of the iris data set is very difficult and it requires high-end equipment to pre-process the image for the required processing, the proposed work took from the online repositories. Rest two biometrics are from real people near the development place. The classifier is trained using 70% of data while it is tested for remaining 30% data. Figure 3 represents a glimpse of the databases.

For self-developed signature database, 500 signatures from 50 persons (10 signatures each person) with different educational backgrounds and age are collected. Everyone is asked to sign non-overlapping signatures using pen on a white sheet of A4 sized paper. The signatures are collected in two sessions over a period of three months to consider intra-class variations in the signatures with time. The purpose of this exercise is to have considerable variability in the signature database. All these sample signatures are scanned using HP ScanJet 200 Flatbed Photo Scanner with a resolution of 300 dpi and stored as genuine signatures in jpg format for further processing.

Ten speech segments between 5-10 sec duration are recorded for 50 speakers in two sessions. The sampling frequency is originally set at 44.1 KHz for all recordings in order to preserve acoustical quality of sound signals. The recordings are captured by using low cost microphone and Audacity software in an environmentally controlled room to reduce acoustical interferes and stored with 16 bits/sample resolution. The silence parts of the recorded utterance at the beginning and end are trimmed off from the recorded file. The speech data recorded utterances for the English language digits (1 to 10).

The extracted features are passed to ANN which is again a two-phase processing algorithm in the training section:

- Feeding forward
- Validating backward

As figure 4 demonstrate that there are 35 neurons in the processing of data from the input layer to hidden layer structure.The training layer uses the Levenberg-Marquardt algorithm for training. The feeding of the data takes validation parameters like Gradient, time etc for the stopping criteria.

Once the feeding is complete, the back propagation plays its role (figure 5). The performance parameter for backpropagation is mean square error. The back propagation ensures the training validation so that the classification process gets high accuracy.

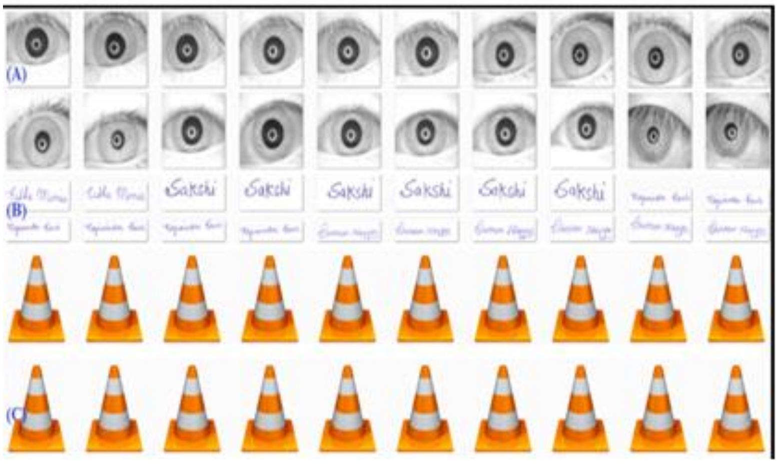

Fig. 3. Database samples of iris, speech and signature

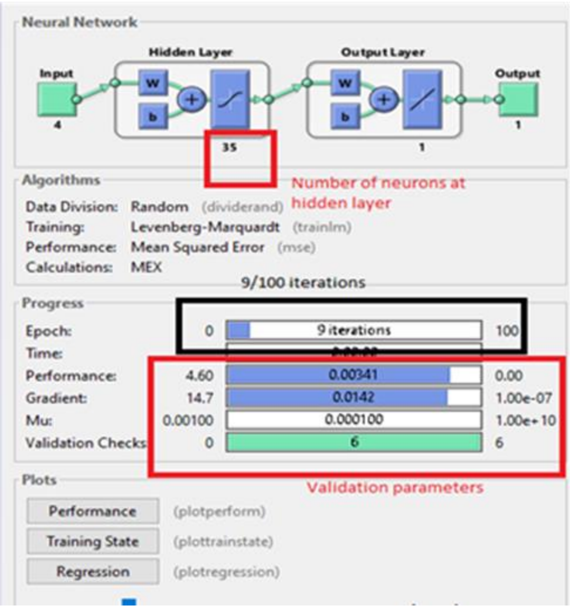

Fig. 4. Architecture of the training of ANN.

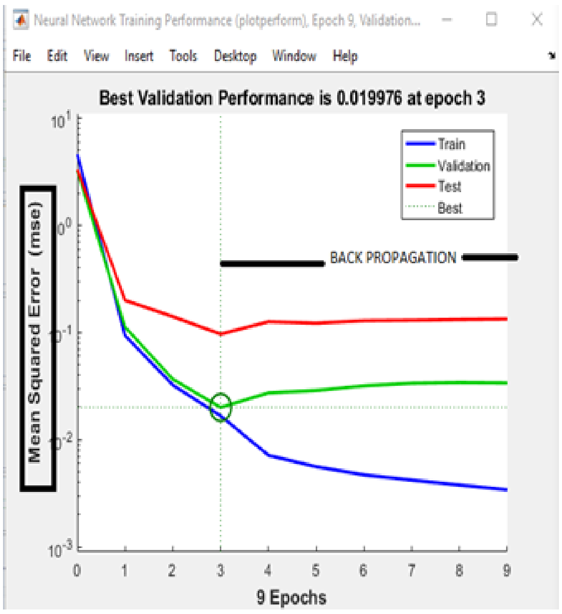

Fig. 5. Backpropagation of training.

Each biometric section has its own classification sections to make the prediction more accurate and precise. The applied NN also performs regression analysis while training the samples. In any training algorithm, the selection of the features for training depends upon the regression model and the layout of propagation.

Close R (Regression) value leads to an optimal training mechanism. As shown in figure 6 (Regression value for Iris) the initial regression is 0.54319 whereas final regression is 0.51703. The value of R changes as the iteration of propagation is increased. A close regression value represents perfection in the training architecture.The difference between the initial regression value and final regression value is pretty less and it will result in a good classification accuracy.This regression model is for iris verification mechanism. The other regression models for speech and signature stand in the same range.

Table 3 represents the classification accuracy and Mean Square Error (MSE) of each biometric classification set. A total of 500 samples for each category is tested. The MSE is calculated as:

$$MSE = \frac{\sum \text{Non Classified Bits}}{\text{Bit Count}} \tag{1}$$

Table 3. Range of single image for each biometric.

Iris		Signature		Speech	
Classification accuracy	MSE	Classification accuracy	MSE	Classification accuracy	MSE
98.32	0.26	97.93	0.31	98.12	0.38

The final fusion and fusion classification check is presented in algorithm 2.

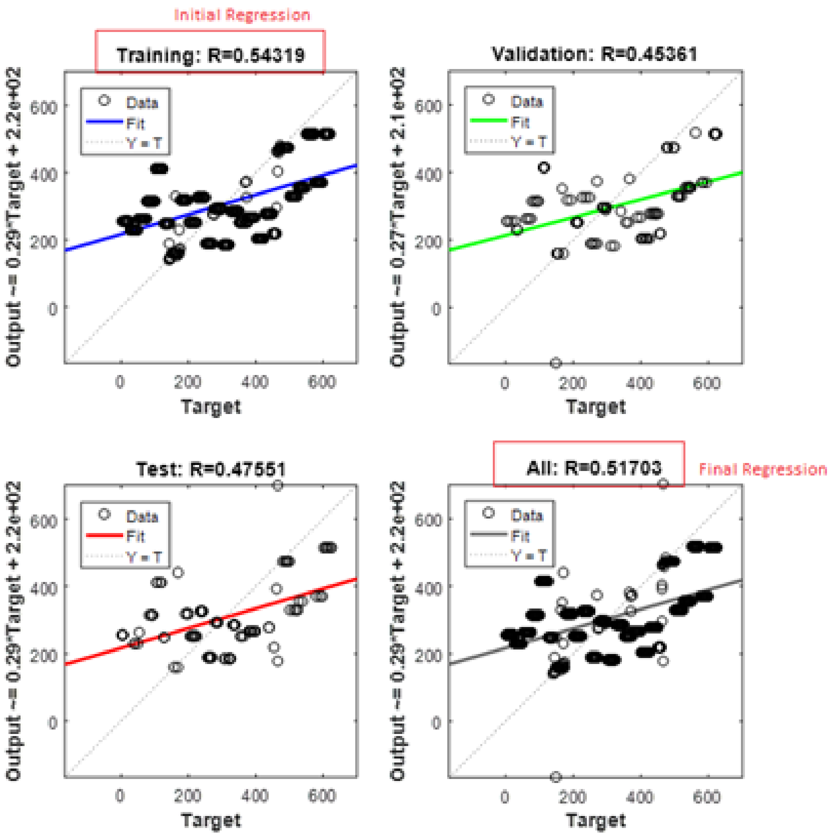

Fig. 6. Final regression model of training.

Algorithm 2. Fusion_classification (Test_Samples, Totalsample, Feature_Value1, Feture_Value2, Feature_Value3)

```
For each sample in Totalsamples  // For each sample in Totalsamples
Fusion_Value(sample)=Feature_Value1(sample)+Feature_Value2(sample)
+Feature_Value3 (sample)

//Summing up the scores of each Classified feature value

End For

Test_Feature_Set=Extracted_Fetures(Test_Samples);  // Taking the feature
set of test samples

Fused_Test= Fusion(Test_Feature_Set)

Match_Found=0; // Initializing the match value to 0;

For i=1:Totalsamples

Current= Fusion_Value(i);

If  Current==Fused_Test

Match_found=i; // storing the matching sample number

End If

End For

End Algorithm
```

Algorithm 2 represents the fusion and classification structure of the proposed algorithm. The optimized feature set of each segment is passed to Algorithm 2. The algorithm generates the score of each feature vector and sum them up with other scores. When it comes to identification, the test sample of each biometric segment follows the same process of feature extraction and individual classification as shown in figure 2. Initially the match score value is set to be 0. The iteration where the fusion score of test sample matches with the training score, comes out to be the classified set. If no match is found, then the authentication is failed.

3. RESULTS AND DISCUSSION

The evaluation of the results is done using the parameters listed below.

3.1. False Acceptance Rate (FAR)

It is the total number of falsely accepted samples by the proposed algorithm with respect to the total number of supplied samples as expressed in equation:

$$FAR = \frac{\text{Total number of false accepted samples}}{\text{Total number of suplied samples}} \tag{2}$$

3.2. False Rejection Rate (FRR)

It is the ratio of total number of falsely rejected samples to the total number of supplied samples. It is calculated as:

$$FRR = \frac{\text{Total number of false rejected samples}}{\text{Total number of suplied samples}} \tag{3}$$

3.3. Accuracy

It is the measure of correct identification and calculated according to equation:

$$Accuracy = \left[1 - \left(FAR + FRR\right)\right] \times 100 \tag{4}$$

The average accuracy of fusion classification lies between 97-99% where as FAR and FRR lies in between 0.01 and 0.06. The result section has compared the proposed results with other biometric fusion works (Rattani et al., 2007; Bokade & Sapkal, 2012; Nadheen & Poornima, 2013; Dhameliya & Chaudhari, 2013; Veluchamy & Karlmarx, 2016). It is not necessary that each researcher has taken the same biometric sample or same dataset. Hence studies compared in figure 7 can only serve as a qualitative comparison. Since the experiments are conducted under different conditions by researchers, it is difficult to extract absolute conclusions through the comparisons. An overall accuracy is compared in this segment.

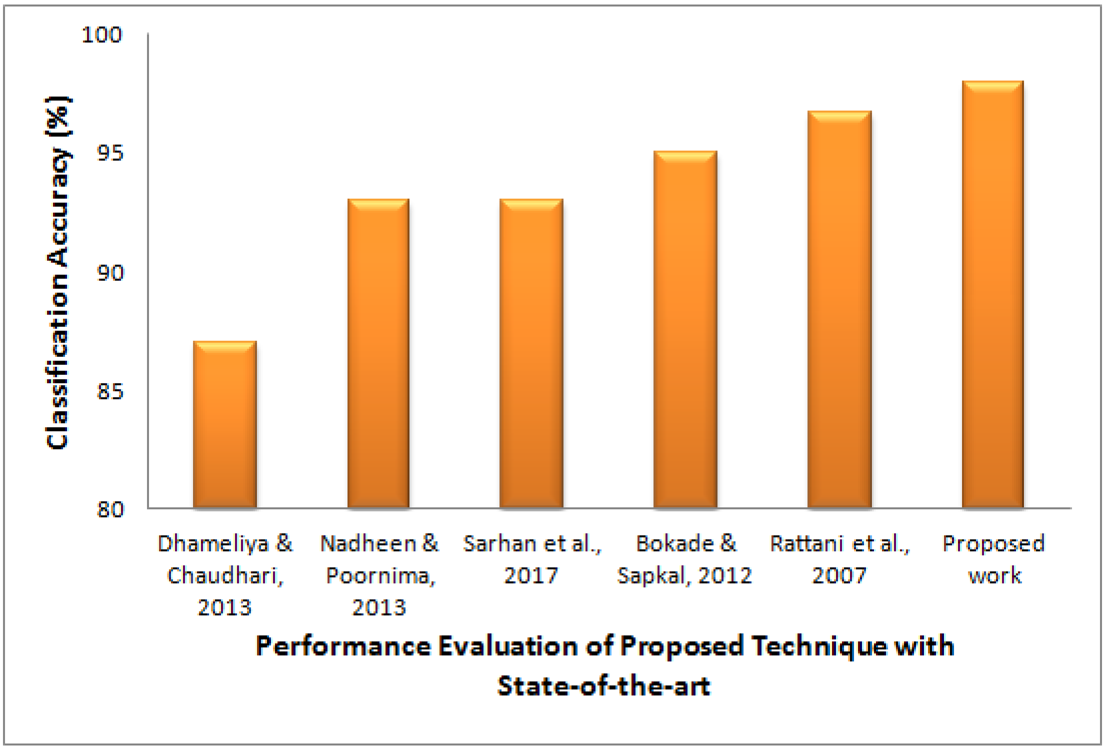

Fig. 7. *Backpropagation of training*

.

4. CONCLUSIONS

The research carried out in this paper aimed to improvise the fusion process. To attain the desired goal, a novel framework is presented which classifies each section first. The usage of Genetic Algorithm has made the extracted feature vector quite precise and the role of NN is significant. Feature fusion method is utilized for the fusion purpose. Overall classification accuracy of the proposed algorithm lies between 96-98% and a comparative analysis is also presented. Implementing GA before the training and testing stage tackles the 'curse of dimensionality' challenge, thereby increasing the computing speed and puts low constraint on classifier. The presented research work has areas of improvement. Other optimization algorithm from Swarm Intelligence series can also be tried.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The Authors received no funding.

REFERENCES

Almayyan, W., Own, H., Abel-Kader, R., Zedan, H., 2011, A multimodal biometric fusion approach based on binary particle optimization,in: *Research and Development in Intelligent Systems XXVIII*, SGAI, Eds, Bramer, M., Petridis, M.,. Nolle, L., Springer, London, 139-152.

Bokade, G.U., Sapkal, A.M., 2012, Feature level fusion of palm and face for secure recognition, *International Journal of Computer and Electrical Engineering*, 4(2), 157-160.

Carol, J., Fred, A.L., 2018, Feature level fusion using physical biometric traits, *International Journal of Biomedical Engineering and Technology*, 26(1), 54-70.

Chen, C.H., Chu, C., 2006, Fusion of face and iris features for multimodal biometrics, in: *Advances in Biometrics ICB*, Eds, Zhang, D., Jain, A.K,. vol. 3832, Springer, Berlin, Heidelberg, 571-580.

Chetty, G.,. Wagner, M., 2005, Investigating feature-level fusion for checking liveness in face-voice authentication, in: *Proc. 8th IEEE International Symposium on Signal Processing and its Applications*, Sydney, 66-69.

Cui, J., Wang, Y. , Huang, J. , Tan, T., Sun, Z., Ma, L. 2004, An Iris Image Synthesis Method Based on PCA and Super-Resolution, in: *Proc. 17th Int. Conf. on Pattern Recognition*, Cambridge, 471-474.

Dhameliya, M.D, Chaudhari, J.P., 2013, A multimodal biometric recognition system based on fusion of palmprint and fingerprint, *International Journal of Engineering Trends and Technology*, 4, 1908-1911.

Eskandari, M., Toygar, Ö., Demirel, H., 2014, Feature extractor selection for face–iris multimodal recognition, *Signal, Image and Video Processing*, 8(6), 1189-1198.

Haghighat, M., Abdel-Mottaleb, M., Alhalabi, W., 2016, Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition, *IEEE Transactions on Information Forensics and Security*, 11(9), 1984-1996.

Huang, C.L., Wang, C.J., 2006, A GA-based feature selection and parameters optimization for support vector machines, *Expert Systems with Applications*, 31(2), 231-240.

Karouni, A, Daya, B., Bahlak, S., 2011, Offline signature recognition using neural networks approach, *Procedia Computer Science*, 3, 155-161.

Leghari, M., Memon, S., Chandio, A.A., 2018, Feature-level fusion of fingerprint and online signature for multimodal biometrics," in *Proc: IEEE International Conference on Computing, Mathematics and Engineering Technologies*, Sukkur, 1-4.

Liau, H.F., Isa, D., 2011, Feature selection for support vector machine-based face-iris multimodal biometric system, *Expert Systems with Applications*, 38(9), 11105-11111.

Malode, A.A., Sahare, S.L. , An improved speaker recognition by using VQ & HMM, in: *Proc. IET Chennai 3rd International on Sustainable Energy and Intelligent Systems*, Tiruchengode, 377-383.

Nadheen, M.F., Poornima, S., 2013, Feature level fusion in multimodal biometric authentication system, *International Journal of Computer Applications*, 69(18), 36-40.

Ou, G., Murphey, Y.L., 2007, Multi-class pattern classification using neural networks, *Pattern Recognition*, 40(1), 4-18.

Park, G., Kim, S., 2013, Hand biometric recognition based on fused hand geometry and vascular patterns, *Sensors*, 13(3), 2895-2910.

Rattani, A., Kisku, D.R., Bicego M., Tistarelli, M., 2007, Feature level fusion of face and fingerprint biometrics, in: *Proc. First IEEE International Conference on Theory, Applications and Systems*, Crystal City, 1-6.

Rattani, A., Tistarelli, M., 2009, Robust multi-modal and multi-unit feature level fusion of face and iris biometrics, in: *Advances in Biometrics*, Eds, Tistarelli, M., Nixon, M.S., vol. 5558, Springer, Berlin, Heidelberg, 960-969.

Saleh, A., Alzoubiady, L.M., 2014, Decision level fusion of iris and signature biometrics for personal identification using ant colony optimization, *International Journal of Engineering and Innovative Technology*, 3(11), 35-42.

Sarhan, S., Alhassan, S., Elmougy, S., 2017, Multimodal biometric systems: a comparative study, *Arabian Journal For Science And Engineering*, 42(2), 443-457.

Supreetha Gowda, H.D., Hemantha Kumar, G., Imran, M., 2018, Multi-modal biometric system on various levels of fusion using LPQ features, *Journal of Information and Optimization Sciences*, 39(1), 169-181.

Veluchamy, S., Karlmarx, L.R., 2016, System for multimodal biometric recognition based on finger knuckle and finger vein using feature-level fusion and k-support vector machine classifier, *IET Biometrics*, 6(3), 232-242.

Yu, H., Kim, S., 2012, SVM Tutorial — Classification, Regression and Ranking, in: *Handbook of Natural Computing*, Eds, Rozenberg, G., Bäck, T., Kok J.N., Springer, Berlin, Heidelberg, 479-506.

Zhang, Z., Wang, R., Pan, K., Li, S.Z.,. Zhang, P., 2007, Fusion of near infrared face and iris biometrics, in: *Advances in Biometrics*, Eds., Lee, S.W., Li, S.Z., vol. 4642, Springer, Berlin, Heidelberg, 172-180.

NOWATORSKIE POŁĄCZENIE CECH BIOMETRYCZNYCH DLA TĘCZÓWKI, MOWY I PODPISU

Streszczenie

W związku z ciągłym wzrostem wymagań dotyczących bezpieczeństwa i nowymi wyzwaniami stojącymi dzisiaj w tym zakresie przed światem istnieje potrzeba tworzenia ststremów wykorzystujących biometrię multimodalną do automatycznej identyfikacji/weryfikacji osób. Artykuł opisuje problem zastosowania multimodalnej fuzji biometrycznej do poprawy bezpieczeństwa rozpoznawania osób. Do nowej fuzji wykorzystano tęczówkę, mowę i podpis. Zaprezentowano oddzielny mechanism dla każdego czynnika biometrycznego. Fuzję przeprowadzono wykorzystując cechy wybrane w danej chwili czasu indywidualnie dla każdego czynnika. Dla różnych czynników zastosowano różny algorytm wyboru cech biometrycznych. Zastosowano 2-wymiarową analizę podstawowych składników (ang. 2-Dimensional Principle Component Analysis - 2DPCA) dla tęczówki, skaloniezmiennicze przekształcenie cech (ang. Scale Invariant Feature Transform - SIFT) dla podpisu oraz parametry melcepstralne (ang. Mel-Frequency Cepstral Coefficients) dla mowy. W artykule wykorzystano metodę Algorytmów Genetycznych do optymalizacji oceny poszczególnych cech. Klasyfikację przeprowadzono wykorzystując sztuczne sieci neuronowe.

Received: August 23, 2020.
Received in a revised form: November 30,2020.
Accepted: December 24, 2020.