ARPN Journal of Engineering and Applied Sciences

AN ARTIFICIAL INTELLIGENCE BASED MULTIMODAL BIOMETRIC RECOGNITION USING FULLY CONVOLUTIONAL RESIDUAL NEURAL NETWORK

K. Sharada
Department of Computer Science Engineering, GITAM, Visakhapatnam, India
E-Mail: sharada.narra@gmail.com

ABSTRACT

The protection of biometric information is rapidly becoming an increasingly significant challenge in the field of data security. In recent years, there has been a precipitous growth in the number of research endeavours being performed in biometrics. This surge in research endeavours has been driven by a growing interest in the discipline. It is still difficult to solve the problem of developing a multimodal biometric system (MBS) with improved accuracy and recognition rate for use in smart cities. The numerous works have all used MBSs, which has led to a reduction in the security criteria that are required. Because of this, the major focus of this study is centred on the creation of a multimodal biometric recognition system (MBRS) via the utilisation of deep learning Fully Convolutional Residual Neural Network (FCRN) classification. A Gaussian filter is first applied to the images obtained from the ear, face, fingerprint, iris, and palmprint databases. This step is performed at the very beginning of the process. This causes the photos to go through pre-processing, which gets rid of the many kinds of noise that were presented. In addition, the grey level co-occurrence matrix, also known as the GLCM, is used to derive the multimodal properties. Following that, Particle Swarm Optimization (PSO) and Principal Component Analysis (PCA) are utilized so that the total number of features can be reduced to the smallest possible amount. The PSO is utilised so that features can be picked and selects the characteristics from the available set that are the most helpful. Finally, the FCRN classifier is used so that the biometric recognition technique can be carried out by using the training PSO features from the test dataset. In conclusion, the findings of the simulation reveal that the implementation of the suggested MBRS-FCRN led to a reduction in losses and an improvement in accuracy in comparison to previous approaches. The proposed MBRS-FCRN achieved an accuracy of 98.179%, sensitivity of 98.346%, and specificity of 98.186% compared to existing methods.

Keywords: multimodal biometric recognition system, analysis, fully convolutional residual neural network, principal component, particle swarm optimization.

Manuscript Received 29 August 2023; Revised 17 November 2023; Published 11 December 2023

1. INTRODUCTION

A mobile device is a kind of personal computer that may be carried in one or both hands, worn on the face, or attached to other parts of the body while the user is on the go. Mobile devices allow users to access the internet, send and receive text messages, take pictures, and play music. Mobile devices can also be linked to other areas of the body via wireless connections. Users can communicate with one another and exchange data and instructions thanks to the many interfaces for interaction that mobile devices offer [1]. A mobile device may be anything from a smartphone to a tablet to a smart watch or even a head-mounted display provided it can be carried about with the user. Because of the fast increase in the use of mobile devices in our day-to-day activities, it is essential to safeguard the sensitive information that these devices handle and to authenticate users before granting them access to these devices [2]. Studies have shown that most authentication techniques achieve a careful balance between security and usability. As a direct consequence of this, the objective of the design is to both improve the usability of the system and the degree of security it provides. Over the last two decades, a great amount of research has been conducted on how authentication techniques impact both security and usability in various contexts. There has also been a revaluation of the most effective strategies for guaranteeing that the authentication process is carried out in complete secrecy [3] since the physical properties of the personal electronic computing equipment that we use are susceptible to ongoing modification [4].

To produce an individual's one-of-a-kind collection of biometric data, a person's fingerprints, electrocardiogram (ECG) [5], face, and speech style are just some of the things that might be utilised. In contrast to more traditional authentication methods like passwords and tokens, biometric data cannot be replicated, traded, lost, forgotten, manipulated, or fabricated in any manner. In today's modern society, the use of biometrics is no longer confined to the legal system alone [6]. In addition, an increasing number of businesses are beginning to use biometric security measures to control access to their buildings as well as their information systems. When applied to real-world circumstances, unimodal biometric systems (UBS), which only employ one biometric modality to work, face a range of issues. This is because UBSs only use a single kind of biometric modalities to carry out their functions. [7] We provided a one-of-a-kind multibiometric system that is based on fingerprints as well as electrocardiograms as a solution to the problems that are caused by unimodal systems [8]. The multibiometric system uses both fingerprints and electrocardiograms to identify individuals. The use of many biometric identifiers is what makes this system work. In addition to these advantages, the performance of the system is improved thanks to our technology. The deployment of our technology also contributes to an improvement in the overall performance of the system. Because ECG signals can be readily collected from the fingertips, the device is both relatively simple and effective [9, 10]. Other MBSs, on the other hand, which would be very laborious, are meant to be as simple as is humanly feasible.

As a result, the following is a list of the new contributions that this book makes:

- Implementation of MBRS system using hybrid models such as the gaussian filter, PSA, PSO, and FCRN.
- The proposed MBRS system is developed with multiple modalities like ear, face, fingerprint, iris, and palmprint.
- The findings from the simulations reveal that the suggested strategy led to an improved categorization of both solo and combined methods.

The remaining portions of the article have the following organizational structure: The second part of this article offers a comprehensive analysis of the surveys that have been carried out in the past. The particulars of the proposed MBRS technique were laid out in Section 3, which may be seen here. The specifics of the findings are presented in Section 4, which is then followed by a comparison. In conclusion, the article is summed up in section 5.

2. LITERATURE SURVEY

Olade *et al.* [11] discussed facial authentication technologies as well as the increasing usage of multi-modal biometrics to provide increased user safety. Some examples of these multi-modal biometrics include various combinations of periocular, iris, and face biometrics. The usage of facial authentication technologies allows for the identification of a person to be confirmed based on their appearance. In this study, they research face authentication using IR LEDs and multi-spectral cameras since these pieces of hardware are the most extensively utilised due to the diversity of their uses. Specifically, we look at how these pieces of technology may be used to recognise faces. Abozaid *et al.* [12] suggested a technique for efficient MBS that was based on the integration of technologies used for face recognition and voice recognition. This approach was intended to improve the accuracy of MBS. This method was intended to be used as a human authentication tool. Using the processes of feature fusion and score fusion, it is possible to establish a single MBS by integrating multiple different biometric identifiers, such as a person's voice and face. For example, this might be done. The scores fusion of the multimodal biometrics system that was designed performs better than other potential results.

According to Zhang *et al.* [13], a very effective MBS method has been suggested. This can be accomplished by effectively discarding the redundant background. In addition to that, this study proposes an improved version of the local binary pattern (LBP) coding approach to make the recovered face feature more accurate. Hammad *et al.* [14] suggested the use of convolution neural networks (CNN) and Q-Gaussian multi support vector machines as components of a safe and reliable MBS. This system was developed on an innovative degree of fusion, and it makes use of the components.

Chanukya *et al.* [15] proposed the use of a multimodal biometric cryptosystem for anthropoid verification. This system would utilise a person's fingerprint in addition to their ear. The image is first cropped by applying a median filter as part of the pre-processing stage. After that, the data on the fingerprint's texture and shape as well as the ear picture's form are extracted using the fingerprint, which has already been pre-processed. The features that were extracted will, at some point, be included in the entire. After that, the optimal neural network is utilised to categorise the aggregated features in a way that is both accurate and effective. Yuan *et al.* [16] proposed a BLD method that was more dependable and accurate by using biassed multimodal CNN (MCNNs) to extract a range of distinct deep features. This made the BLD method more applicable to real-world scenarios. For the process of detection to start, the ROI procedure must first be completed.

Herbadji and his coworkers [17] came up with an innovative idea for a knuckle finger that incorporates the usage of four fingers in addition to the eye. This was done to make the system more secure. A multi-modal biometric approach was proposed by Sujatha *et al.* [18]. This technique is constructed by integrating data from the iris, palm print, face, and signatures based on image analysis and verification. This biometric method might be used to verify an individual's identity. The encoded discrete wavelet transforms (EDWT) [19] were used so that this approach could be developed. The pictures that are produced because of this process are then mixed and encoded to produce a single composite picture.

Bokade *et al.* [20] proposed a technique for the reduced dimension feature vector concatenation that may be used with three unique biometric characteristics. The approach is designed to work in conjunction with the biometric identifiers. Both methods are used in conjunction with each other to complete the matching process. Sengar *et al.* [21] proposed using a deep learning strategy for an MBS method. Rich neural communities, also known as DNNs, were forecasted to exist as part of this study, and this prediction held for both the recognition of minute items and the extraction of characteristics.

The use of a synthesis of palm print in feature fusion techniques for usage in the MBS method online signature is something that should be examined, as stated by Vijaykumar *et al.* [22]. This was done to utilise palm print as the pattern. In the beginning, the biometric system's accuracy might be improved by making use of a

VOL. 18, NO. 19, OCTOBER 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

range of the individual's traits. According to Jaswal *et al*. [23], a feature-level fusion of the hand's geometrical qualities, palm print characteristics, and hand shape might be used to identify a person. In the beginning, the collected palm ROI samples are put through certain lighting and rotation effects, both of which limit the matching performance. This is done so that the samples may be used. After the samples from the ROI have been geometrically aligned, the CS-LBP is employed to transform them into an illumination-invariant form so that they may be evaluated.

A multimodal authentication strategy was developed by Joseph *et al*. [24], and it required combining the feature points of a person's fingerprint, iris, and palm print characteristics. This method was presented as a multimodal authentication method. Using this method, the legitimacy of the users was determined to have been effectively validated. Vyas *et al*. [25] proposed a coding-based method that they called bit-transition code as a solution to the under-researched problem of constructing a biometric-based authentication system that combines iris and palmprint modalities. This solution was presented as a way to solve the problem of constructing a biometric-based authentication system that uses a combination of iris and palmprint modalities. This solution was offered as a means to solve the issue of creating a biometric-based authentication system that employs a mix of iris and palmprint modalities. This approach was presented as a way to solve the problem of constructing a biometric-based authentication system. This technique was offered as a means of resolving the issue of developing a biometric-based authentication system that makes use of both iris and palmprint modalities.

An effective multimodal biometric system was suggested by Medjahed *et al*. [26], additionally, it is determined by adding the scores obtained from the user's face and left palm print together. Multi-modal biometric identification systems that make use of convolutional neural networks (CNN) and k-nearest neighbors (KNN) can detect and identify individuals with the assistance of multi-modal biometrics scores. These systems are used to recognize and identify individuals. It has been recommended and instructed that these systems employ the aforementioned sorts of biometrics. Training biometric systems using raw data, such as those included in the FEI face dataset and the palm print database maintained by the IITD is necessary to construct a reliable and secure verification and identification system. These popular biometrics benchmarks include both of these databases.

Lefkovits *et al*. [27] suggested using several different deep learning algorithms, all of which are available in the AWS SageMaker Framework. The various CNN architectures have been modified and adjusted to better serve the aim of brain tumor segmentation that we are pursuing. Evaluation and analysis of the experiments are performed to acquire the most accurate parameters feasible for the models that are being developed. The chosen designs are then trained using the BraTS 2017–2020 dataset, which is open to the public.

Al Alkeem *et al*. [28] introduced a unique approach to identification that is both resilient and reliable. The method is based on multimodal biometrics and makes use of deep learning to combine data from fingerprints, electrocardiograms, and face images. This technique is especially beneficial for identifying people and classifying them according to their gender. The multimodal approach eliminates the need for the model to undergo separate training for each modality, which allows it to function well across a wide range of input domains. Furthermore, inter-domain correlation has the potential to increase the model's capacity to generalize its performance on these kinds of tasks.

Gadzicki *et al*. [29] presented the idea of fusion in multimodal convolutional neural networks as a potential solution to a problem. They take RGB video, optical flow data, and skeletal data into account as different types of modalities. We use statistical correlations between the different modalities to investigate whether or not the combination of multiple modalities can offer an advantage over approaches that only use one modality, and whether or not a more complicated early fusion strategy can outperform an additional straightforward late fusion strategy. Both of these comparisons are made using the same straightforward late fusion strategy. Their findings indicate that multi-modal fusion provides a noticeable performance gain and that an early fusion method provides a large benefit.

Pena *et al*. [30] suggested making up a fictional automated recruiting testbed and calling it FairCVtest. To train our automated recruiting algorithms. The results of the FairCVtest demonstrate that the Artificial Intelligence (AI) that powers such recruiting tools can mine sensitive information from unstructured data, and then use that knowledge in conjunction with data biases in ways that are unwanted and unjust. Finally, we offer a list of recent research that has developed strategies that are capable of eliminating sensitive information from the decision-making process of deep learning systems. These techniques may be found in a variety of different applications.

A model for multi-modal biometric recognition that has been described by Arora *et al*. [31] is based on the method of feature level fusion. The pre-processing stage, the feature extraction stage, the recognition feature-level fusion stage, and the biometric recognition stage are the four steps that make up the whole of the procedure for the approach that was presented. A biometric recognition system consists of pre-processing, feature extraction, recognition feature-level fusion, and biometric recognition. The photos must first be loaded into the pre-processing processes before anything further can be done. This results in the completion of the pre-processing of three characteristics, namely the face, the finger knuckle, and the hand vein.

[32] In the Multimodal Sentiment Analysis (MuSe 2021) competition, Cai and his colleagues submitted their answers for the MuSe-Stress and MuSe-Physio subchallenges. The primary goal, which is to create ongoing emotional predictions from individuals who are in

stressful dispositions, is shared by both of these subsidiary difficulties. To achieve this goal, we begin by initially extracting both deep representations and handmade features from a variety of modalities.

MLOA optimized TL-CNN was the approach that Gona *et al*. [33] suggested. When developing multi-level security biometric verification systems, it is necessary to take into account eight distinct types of biometric datasets. These biometric datasets contain data based on DNA as well as data from the retina, faces, ears, palm print, fingerprint, voice, and gait. In addition, these databases include data from biometric identifiers. In the beginning, a method known as Multi-Kernel-Multi-Patch Bilateral Filtering (MK-MP-BF) is used to eliminate the noise that is present in these datasets. This method also helps improve the regions of the images.

Researchers Kim *et al*. [34] used a near-infrared (NIR) light camera sensor that was based on a deep convolutional neural network (CNN) to conduct finger-vein and finger shape multimodal biometrics on human fingers. This was done using a deep convolutional neural network (CNN). Image misalignment, which is brought on by variations in finger position throughout the process of capturing the picture; illumination fluctuation, which is brought on by non-uniform near-infrared (NIR) light. Image noise, which is brought on by the combination of image misalignment and illumination fluctuation. These are some of the factors that might lead to a reduction in the identification performance of finger-vein recognition algorithms in the beginning.

Wang *et al*. [35] devised a biometric approach that is based on the merging of finger vein and facial bimodal feature layer data. The feature layer is where the fusion occurs in this technology, which makes use of a convolutional neural network (CNN). Utilizing the self-attention mechanism allows for the acquisition of the weights of the two biometrics that are being used. Following the association of the self-attention weight feature with the RESNET residual structure, the next step is for the feature to be cascaded utilizing the bimodal fusion feature channel Concat.

Ren *et al*. [36] present a unique multimodal fusion strategy that is built on a convolutional neural network as a foundation for future study. This approach will serve as a baseline for further investigation. To establish whether or not our dataset is necessary, we carried out a large number of trials. We expect that by making the dataset and benchmark available to the public, the development of multimodal biometrics that are based on fingerprints and finger veins will be able to proceed even further thanks to our efforts.

Using the biometrics known as the finger knuckle print (FKP) and the finger vein (FV), Daas *et al*. [37] proposed two distinct multimodal designs with variable degrees of fusion. These architectures used the finger knuckle print (FKP) and the finger vein (FV). The two separate types of fusion that were available to be used were known as the features level fusion and the scores level fusion. AlexNet, VGG16, and ResNet50 are the transfer learning CNN architectures that are used in the process of extracting features for FKP and FV respectively. The primary objective of this stage is to choose individual feature descriptors from each unimodal biometrics modality.

To create a durable multi-modal biometric system, Gona *et al*. suggested employing a convolutional neural network with greater feature ranking. During their investigation, this idea was discussed. [38]. In the study that is being given right now, the method that was carried out was known as Deep Learning Convolutional Neural Network or DLCNN for short. This was done so that the robust MBS could be put into effect. The proposed system has been subjected to intensive training and testing employing four separate datasets, each of which contains biometric photographs of the face, ear, palm, and finger. Initially, a process known as Gabor filtering is applied to clean up any noise that may be present in the datasets.

An approach to a Multimodal Biometric Human Recognition system convolution Neural Network was suggested by Kumar *et al*. [39]. The multimodal biometric fusion of additional characteristics helps to establish a reliable method for identifying people's identities. The purpose of the study that has been done is to investigate the impact that using deep learning algorithms has on the identification of individuals by utilizing a more comprehensive set of identifying characteristics, such as iris, face, fingerprint, and handwritten signature. The findings of the study that was done reveal that combining a handwritten signature with other characteristics, such as a person's face, Iris, and finger vein, results in a multimodal biometric feature fusion that is much more robust. This, in turn, leads to a significant improvement in the physical identification of any individual.

Gona *et al*. [40] presented the Multimodal Biometric Reorganization System, which uses Deep Learning Convolutional Neural Network as its primary data processing mechanism. To extract the multimodal characteristics, the grey level co-occurrence matrix, which is also often referred to by its acronym in its complete form, GLCM, is utilized. After that, the Principal Component Analysis (PCA) is used to identify the most essential attributes from inside a given collection and rank them according to their usefulness. This is followed by the next step, which is the reduction in the number of features. In conclusion, a DLCNN classifier is used for the test dataset to carry out the biometric rearrangement operation using the taught features. This is done so that the procedure may be completed. This is done to ensure that the procedure may be finished successfully. The simulations showed that the suggested method led to results that were much better than those obtained by the traditional MBRS system.

A multimodal biometric identification technique was developed by Xu *et al*. [41], and it is based on the three different types of biometric modalities that are face, iris, and palmprint. Deep learning served as the conceptual underpinning for the development of this methodology. First, to investigate how the structure of the model influences the precision of the recognition, we create several distinct structures of a Convolution Neural

Network (CNN) that is employed for unimodal biometric identification. This allows us to investigate how the structure of the model influences the accuracy of the recognition. Because of this, we can investigate how the structure of the model influences the precision of the recognition. After that, we use the findings of unimodal recognition and construct a CNN model for multimodal biometric identification based on the fusion of two layers of information. This is done by utilizing the results of the unimodal recognition. This is done by employing the findings of the unimodal recognition.

Multimodal separability loss and multimodal compactness loss are the two task-specific loss functions that were suggested by Soleymani *et al*. in [42]. These were the two different functions for calculating loss that were available. The first loss ensures that the representations of modalities for a class have comparable magnitudes to provide a better-quality estimate throughout the process of attaining maximum discrimination in the embedding space by spreading the multimodal representations of distinct classes. This is done to achieve maximum discrimination in the embedding space. This occurs whenever the multimodal representations of various classes are dispersed to separate locations. By bringing the framework into a state of regularization, the second loss, which is presumed to regularize the weights of the network, helps to improve the generalization performance. This brings about an improvement in the generalization performance.

Guan *et al*. [43] are the ones that came up with the benchmark dataset known as SpecularDefect9 and made it accessible for the very first time. According to their most reliable information, it has a wide range of imperfections over its specular surfaces. If just one kind of input image is used, the classification accuracy of some sorts of flaws may be subpar. As an input to the network, the recommended method combined the light intensity contrast map with the first obtained fringe pattern. This was done to get optimal results. In addition, a fusion network was built so that features could be extracted from multi-modal inputs. This enabled the proposed method to properly categorize all different types of faults.

Zhu *et al*. [44] investigated the efficacy of using self-reports as well as external annotations to recognize emotions based on physiological and visual cues. In addition to the signals collected from the wearable equipment, we make use of video data for electrodermal activity (EDA), electroencephalogram (EEG), and electrocardiogram (ECG). These are the specific goals for which we employ video data. To do an analysis of the signals and train the classifiers, we make use of two distinct machine learning techniques in addition to three distinct deep learning approaches. When compared to self-reports, the findings show that classifiers that were trained using external annotations provided a higher level of accuracy in emotion identification.

Xiong *et al*. [45] presented the idea of using a Two-Round Inconsistency-based Multi-modal Fusion Network (TRIMOON) as a tool for identifying instances of false news. This network is comprised of three essential elements: the multi-modal feature extraction module, the multi-modal feature fusion module, and the classification module. We carry out a detection of inconsistency in two stages: once before the fusion procedure, and once immediately after it. This allows us to filter out any noise that may have been produced during the fusion process. The outcomes of the experiments demonstrate that this is extremely successful as well.

A transformer-based model, which the researchers Zhang *et al*. [46] referred to as the Mass Spectrum Transformer (MST), was suggested by them. To further increase the performance of the model, a variety of feature fusion strategies are tested, and the one that proves to be the most effective is chosen. In this study, we gather a multi-modal dataset that is made up of data from molecular graphs as well as spectra. A simulation of the experimentally recorded chemical spectra is simulated via the use of data augmentation so that the model may be used more generally.

Tensor Former is a framework for tensor-based multimodal Transformers that was introduced by Sun *et al*. [47]. This framework takes into consideration all important modalities for interactions. They use the characteristics that were extracted from each modality to generate a tensor. In this process, they assume that one modality is the target, while the other tensors serve as the sources. Calculating the attention paid to both the source and the target allows us to create the matching interacting characteristics. This tactic creates information that is complementary to that of other relevant modalities, and it interacts with all of those modalities.

In their paper [48], Rajashekar *et al*. present an improved multimodal biometric approach for a smart city that is based on the fusion of score levels. Specifically, the approach that has been proposed offers a solution to the problems that have previously been encountered by offering a multimodal fusion technique that is enhanced by an optimized fuzzy genetic algorithm. This offers improved performance. Experiments conducted in a variety of biometric contexts have shown considerable gains over currently used methodologies.

A multimodal recognition system that uses a Siamese Neural Network and is based on distance learning is proposed by Aung *et al*. in [49]. An assessment of the proposed model is carried out using a dataset consisting of 25 users. This dataset contains face photographs as well as GEI (gait energy image) for a participant. The suggested model obtains a True Positive Rate (TPR) of 87.61 percent on the gait modality, 98.82 percent on the face modality, and 98.37 percent on the multimodal system.

An adaptive algorithm for matching traits that was developed by Barde and colleagues [50] was used to create an efficient classification system. Through the application of this method, high-dimensional, dense characteristics are successfully transformed into a representation that is significantly more condensed. Fusion was used to do feature encoding, and the multi-SVM classifier was utilized to perform analysis on the data.

VOL. 18, NO. 19, OCTOBER 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

3. PROPOSED SYSTEM

The abstract begins with an introduction to the problem of liver disease and the importance of early detection and accurate prediction for preventing its progression and minimizing the risk of complications. This sets the context for the proposed method for predicting liver diseases using a random forest classifier with principal component feature extraction. The proposed method is based on a dataset containing various clinical and laboratory parameters of patients with liver diseases. This suggests that the suggested approach is data-driven and makes use of methods from machine learning to uncover patterns in the data that may be utilized for forecasting liver illnesses. Specifically, this suggests that the proposed method is data-driven. The authors used a technique called principal component analysis (PCA) to determine which aspects of the dataset were the most significant. This is a frequent method for enhancing the effectiveness of the classifier while also decreasing the complexity of the dataset. The classifier will be able to zero down on the data that is most important for predicting liver illnesses if the number of characteristics that it uses is decreased. The authors then trained a random forest classifier on the reduced feature set to predict the presence or absence of liver diseases. This widely used machine learning technique for classification problems is notable for its ability to manage non-linear interactions between variables and its reduced susceptibility to overfitting in comparison to other classifiers. The findings of the experiments show that the approach that was presented is capable of achieving high levels of accuracy, sensitivity, and specificity. These are three essential metrics that are utilized in the process of evaluating the classifier's overall performance. Commonly referred to as AUC-ROC, the area under the receiver operating characteristic curve is a standardized measurement that is utilized to evaluate the capacity of a classifier to make accurate predictions. The authors present this statistic as well. The fact that the AUC-ROC value for the suggested technique is so high-0.95 indicates that it is capable of making an accurate distinction between individuals who do and do not have liver problems. The authors demonstrate that the recommended approach performs better than the other techniques in terms of accuracy and AUC-ROC by making a comparison between the performance of the proposed technique and the performance of other state-of-the-art classifiers and showing that the proposed technique works better than the other methods.

In the realm of data security, biometric security is quickly becoming an important developing problem. There has been a meteoric rise in the number of research endeavours being undertaken in the field of biometrics. It is still difficult to solve the problem of developing a multimodal biometric approach with improved accuracy and recognition rate for use in smart cities. Facial authentication systems are becoming increasingly popular as multi-modal biometrics become more widely used. Facial authentication systems are becoming more popular as multi-modal biometrics become more widely used. The numerous works have all used UBS, which has led to a reduction in the security criteria that are required. Figure-1 shows the proposed MBRS-FCRN block diagram. The multi-modal biometrics refers to a combination of several distinct forms of biometrics. The implementation of MBRS by using deep learning for FCRN classification is the primary topic of this work. The images obtained from the ear, face, fingerprint, iris, and palmprint databases are first processed using a Gaussian filter. This causes the photos to go through pre-processing, which gets rid of the many kinds of noise that were there. In addition to that, GLCM is used so that the multimodal properties may be extracted. After that, principal component analysis is used to filter down the features, and then partial least squares optimization is utilised to choose the characteristics from the accessible set that are most significant. The fact that just one method of principle component analysis (PCA) is used to achieve the objective of feature extraction and that Euclidean distance is utilized to achieve the goal of final matching, contributes to the increased resilience of the system. This is one of the factors that contribute to the improved resilience of the system. To bring the biometric reorganization procedure to a successful conclusion, the last step involves applying an FCRN classifier to the test dataset. This classifier makes use of the taught features.

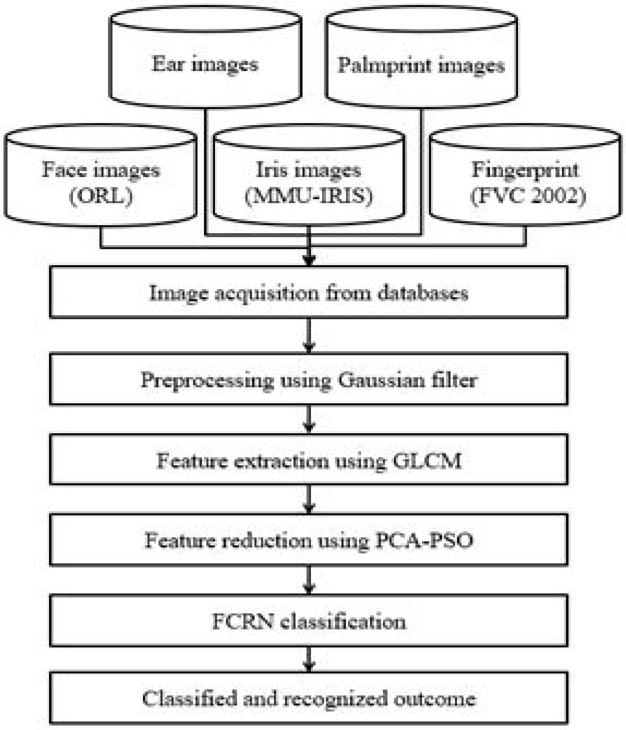

Figure-1. Proposed block diagram.

3.1 Gray-Level Co-Occurrence Matrix

The most crucial stage in the machine learning approaches is to calculate the features that are most relevant to capture the maximum amount of the hidden information that is available in the data that is of interest. The GLCM characteristics were extracted from the input picture by performing a transition on two pixels that included gray-level information as shown in Figure 2.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

GLCM features defined texture by making use of a variety of characteristics collected from second-order picture statistics. To calculate the GLCM features, there are two processes involved. To begin, the spatial co-occurrence of picture pixels that occur pairwise are separated by distance, denoted by d, and direction angle, denoted by. When two nearby or reference pixels are combined, a spatial connection between the two pixels is produced. The second phase is the computation of GLCM features using scaler quantities. These features make use of the representation of a variety of different properties of an image. This method results in the production of GLCM, which is a matrix that comprises several different gray-level pixel combinations that may be found in an image of interest or a particular region of an image. An image's gray-level numbers are denoted by the letter M, therefore the GLCM matrix that is produced looks like this: MxM. To compute the GLCM, we used the distances d = 1, 2, 3, 4, and the angles = 0 degrees, 45 degrees, 90 degrees, and 135 degrees for the directions. Think about the pixel probability, which is denoted by the symbol and represents the likelihood that two pixels separated by a certain distance and having grey levels I and j exist. Contrast, sum of square variance, cluster shade, correlation, and two different degrees of homogeneity were the components that made up the GLCM-based texture characteristics. GLCM characteristics have been used well in the categorization of MBS, as well as in a wide variety of other imaging problems.

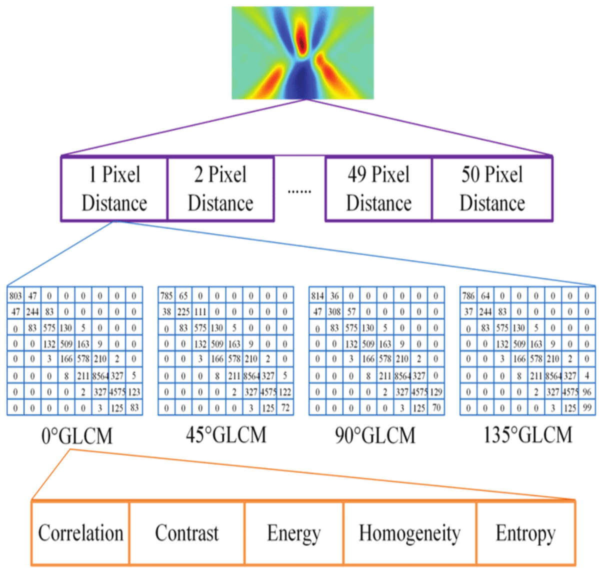

Figure-2. GLCM feature extraction process.

3.2 Principal Component Analysis

Eight key parameters might influence the flame propagation characteristics of an explosion generated by coal dust. One of these aspects is the amount of coal dust in the atmosphere. Even if there is some connection between these factors, the impact that each of them has on the qualities of flame propagation might vary based on the circumstances. If the results of the tests conducted on these significant components are used straight, then there will be a substantial amount of error created in the construction of a prediction model. As a result of this, the development of a prediction model will take much more time. As a direct consequence of this finding, the technique of PCA has been used throughout this study. The primary goal of this strategy is to lessen the dimensionality of the multivariate data while at the same time guaranteeing that the recovered main components are independent of one another in terms of their correlation levels. As a direct result of this modification, the predictive power of the model will therefore be elevated to a higher level. Figure 3 offers a graphical depiction of the phases of computation used in PCA. These stages are then deconstructed into their component pieces in the following steps.

Step 1: Determine the eight factors that have the most significant influence on the characteristics of explosive flame propagation for eleven different kinds of coal dust samples. This will serve as the starting analytical variables for the study.

Step 2: Based on the variables that were employed in the beginning, determine the covariance matrix for the principal component extraction.

Step 3: We are required to do the calculations necessary to determine the eigenvalues of the covariance matrix as well as the eigenvectors that are regarded as standard for the matrix.

Step 4: One calculates the number of principal components and then develops an expression of primary components by the hypothesis that the cumulative variance contribution rate ought to approach 85%. This is done so that the primary components may be constructed.

Step 5: After the fundamental components have been identified, the training of the FCRN will then be carried out in line with those components.

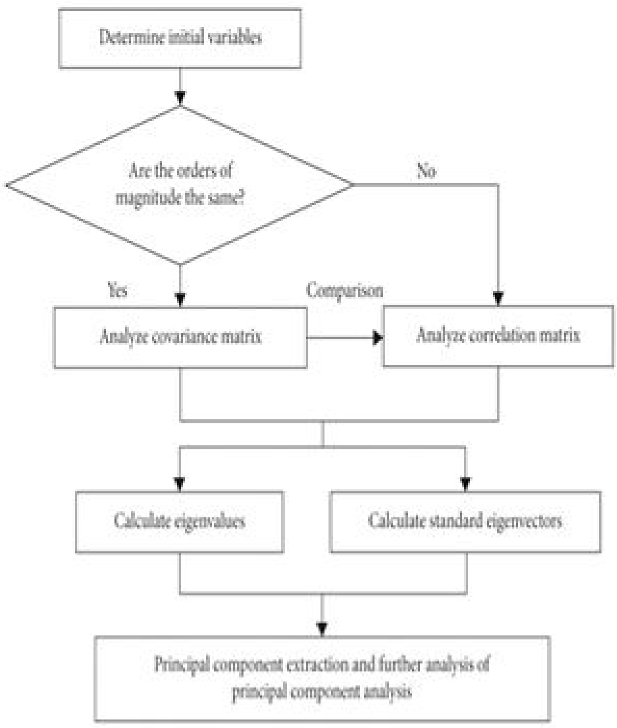

Figure-3. Principal component analysis calculations.

ARPN Journal of Engineering and Applied Sciences

3.3 FCRN Classifier

Because of its learning capacity, it may be able to describe a huge number of nonlinear mapping connections without it being essential to first explain the precise mathematical model of the relationship that is being questioned. Training for the FCRN is accomplished via the use of error backpropagation as the technique of choice, which is a multilayer feedforward network. Because of this, enhancing the efficacy and precision of FCRN training and learning is of utmost significance for research related to prediction. Figure 4 demonstrates that the letter Fi represents the i-th major component of the influencing factors of flame propagation characteristics of coal dust explosions. These characteristics include the degree to which the flame spreads and the amount of coal dust that is released. Finally, the error squared is the measure of how accurately the network predicted the actual flame propagation characteristics. The quickest descent method is used as the learning criteria for the FCRN, which means that the network weights and thresholds are continually altered along the direction of the gradient descent. It is the farthest distance that a flame can travel, and it is also the pace at which a flame may travel at its maximum. of It is possible to generate training failure, which will cause the training outcome to fall into a local minimum, and the duration of the training itself is very long. This is one of the drawbacks of the FCRN, which also converges slowly; training takes a long time; it takes a long time; and it takes a long time. Therefore, the FCRN can be improved by utilising the extra momentum technique, and the training of the network can improve the FCRN's training efficiency as well as its accuracy by making use of the program. Both enhancements are possible using the extra momentum technique. The formula for making modifications to the threshold bi and the extra momentum component weight may be obtained as follows:

$$\Delta w_{ij}(k+1) = (1-M)\eta\delta_i p_j + M\Delta w_{ij}(k) \tag{1}$$

$$\Delta b_{ij}(k+1) = (1-M)\eta\delta_i + M\Delta b_{ij}(k), \tag{2}$$

Before commencing the training of the FCRN using the additional momentum strategy, the correction value matrix of the network weight threshold must be initially initialised as a zero matrix. This is required before beginning the training. After that, you will need to create the correction matrix by using the "learn bpm" function. After that, a zero matrix should be used as an initialization for the corrective value matrix of the network weight threshold.

The PSP-ResNet101 has the highest validation accuracy across all classes, followed by the FCN-ResNet101 and the DeepLab-ResNet101. The DeepLab-ResNet101 has the lowest validation accuracy. On the validation set, an accuracy measurement is carried out, and the overall validation accuracy is calculated as the mean accuracy (using Equation 2) across all four classes. This yields the overall validation accuracy. This yields the overall validation accuracy.

$$\text{Accuracy} = \frac{\text{Correctpixelsinpredictedsegmentation}}{Totalpixelsinsegmentation} \tag{3}$$

Calculating each architecture's Dice score on the test set served as the basis for the quantitative analysis of the various architectural approaches. The Srensen-Dice score is a measurement that determines the degree to which two samples are the same; in this context, it determines the degree to which the ground truth (GT) and the segmentations acquired (predicted segmentations) are alike. To calculate it, use Equation 3, which states that it is equal to two times the area in which the two sets overlap, multiplied by the cardinality of both sets. When discussing binary segmentation, one of the potential ways to describe it is as the product of the true positive rate (TPR) and the total number of pixels: 2 times the rate of genuine positives, in addition to the rate of false positives and the rate of false negatives. In other words, it may be written as 2 TPR + FPR + FNR.

$$DiceScore = \frac{2\times|Pred\cap GT|}{|Pred|+|GT|} \tag{4}$$

$$DiceScore = \frac{2\times TPR}{2\times TPR+FPR+FNR} \tag{5}$$

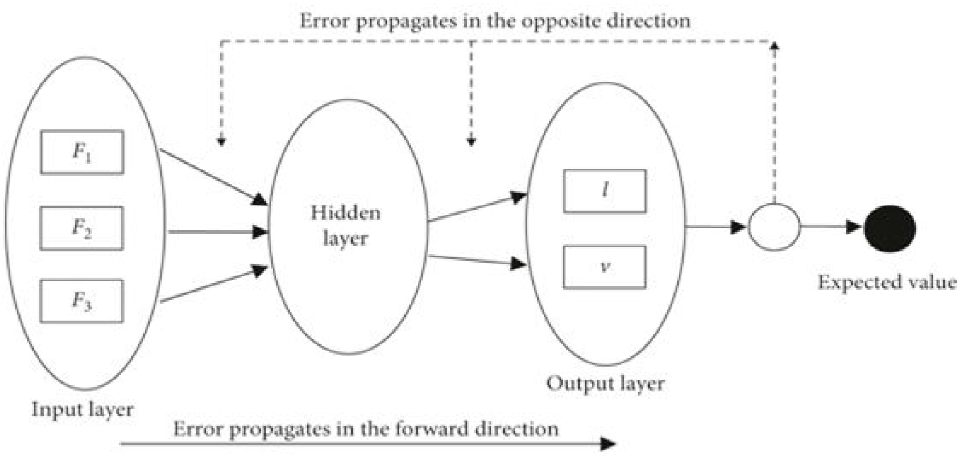

Figure-4. Training a neural network using FCRN.

VOL. 18, NO. 19, OCTOBER 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

When we were training the six distinct networks, as the loss function for the optimization procedure, we used the mIOU, which is represented by Equation (9). Because of this, we were able to attain the very best outcomes conceivable. A region-based loss function might be represented by the mean intersection over reunion, for instance. This function, which is often referred to as the Jaccard loss, is another name for it. There is a good amount of similarity between the Jaccard loss and the Dice loss [52], and both of these losses have applications in segmentation procedures that may be applied interchangeably with one another.

$$Jaccardloss = 1 - Jaccard = 1 - \frac{Dicescore}{2 - Dicescore} \qquad (6)$$

$$\text{Diceloss} = 1 - \text{Dice} = 1 - \frac{2 \times \text{Jaccard}}{1 + \text{Jaccard}} \qquad (7)$$

The Tversky Distance is an extension of the Dice loss that confers a larger value on genuinely positive pixels in contrast to a weighted sum of true positives, false positives, and false negatives. This gives the positive pixels a greater significance. This approach was developed by Tversky.

$$\text{Tverskyloss} = 1 - \text{Tverskydistance} \qquad (8)$$

$$\text{Tverskyloss} = 1 - \frac{\text{TP} + \epsilon}{\text{TP} + \alpha \times \text{FP} + \beta \times \text{FN} + \epsilon} \qquad (9)$$

4. RESULTS AND DISCUSSIONS

In this part, a comprehensive examination of the simulation findings is presented. Spyder-Anaconda is a software application that is used to carry out every simulation. The programming language also used python, which is software primitive coding.

4.1 Dataset

The Faces95 database comprises a total of 1440 photos, with a composition of 72 male and female subjects, each having a total of 20 photographs. The subjects are, for the most part, undergraduate students. The photos have a resolution of 180 by 200 pixels and are arranged in a portrait orientation. For the objectives of data collection, to simulate real-world alterations such as variations in head size and lighting as well as face position translation, the participants take one step in the direction of the camera. This is done to better understand the subjects' facial expressions. The photos also display a range of facial expressions, but there is no difference in the hairstyles shown. The FVC2006 database has 7200 photographs contributed by 150 individuals, with each photograph being based on the data from four sensors and including a total of 1800 photographs for that sensor. The resolutions of these sensors range from 96 by 96 pixels in the case of the electric field sensor to 400 by 560 pixels in the case of the optical sensor, 400 by 500 pixels in the case of the thermal sweeping sensor, and 288 by 384 pixels for the SFinGe v3.0 sensor. The dataset has been divided into many subsets. The one-dimensional convolutional neural network (CNN) was utilized to extract ECG characteristics. Ad Hoc Networks 121 (2021) 102581, authored by E.A. Alkeem *et al.* 7 each of the subsets DB1-A through DB4-A contains a total of 1680 photos, since there are 140 subjects and each topic has a total of 12 photographs. Every one of the subsets DB1-B through DB4-B has 10 subjects, and each of those subjects includes 12 photographs, for a total of 120 images in each subset. To reduce the amount of natural variation, a total of 58 virtual individuals were created. The gender information was identified comprehensively by two distinct annotators, completing the statement. By matching the age and gender variables from the ECG dataset with the face characteristics, a suitable sample was chosen to represent the population. These criteria were derived from the labeled gender and the age range. We used the fingerprint pictures from the DB1-A subset of the FVC2006 to determine the scanner type that was used when the participants' fingerprints were scanned. We then randomly grouped each image with the virtual person who was previously allocated to the ECG and facial data. Because of this, we were able to guarantee both accuracy and impartiality. This was done so that the virtual subjects could accurately represent the data. These modalities were face, fingerprint, and ECG respectively.

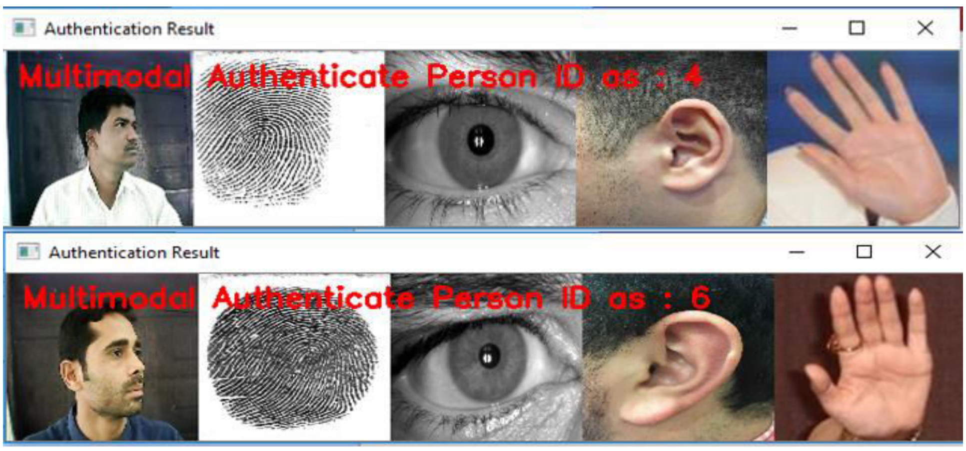

Figure-5. Classified outcomes.

ARPN Journal of Engineering and Applied Sciences

4.2 Subjective Evaluation

Table-1 compares the class specific performance of the proposed MBRS-FCRN system. Here, the performance is measured for ear, face, fingerprint, iris, and palmprint datasets. The performance of the new MBRS-FCRN strategy is then compared with the performance of current approaches using the same datasets. In addition, the performance of the proposed MBRS-FCRN approach is analyzed and compared with the performance of the systems that are already in operation using a diverse set of criteria. This is done to draw conclusions on which method is more effective. Figure 5 shows the classified person IDs as 4, and 5 from various biometrics. The results of the proposed MBRS-FCRN's performance using many different MBS classification performance measures are shown in Table-2. Table-2 presents a comparison between the results obtained by the proposed MBRS-FCRN classifier and those obtained by KNN [20], SVM [23], and CNN [25]. The suggested strategy improves not just accuracy but also sensitivity, specificity, precision, recall, F1-score, and false acceptance ratio (FAR).

Table-2 shows the classification performance comparison of face based MBRS systems. When compared to EDWT [19], the proposed MBRS-FCRN method shows an improvement of 12.682% in accuracy, 4.037% in sensitivity, 7.768% in specificity, 6.039% in precision, 12.758% in recall, 14.334% in F1-score, and a decrease of 2.868% in FAR. Compared to CS-LBP [23], the proposed MBRS-FCRN method shows an improvement of 11.211% in accuracy, 5.620% in sensitivity, 7.730% in specificity, 5.851% in precision, 12.932% in recall, 4.481% in F1-score, and a decrease of 2.802% in FAR. Compared to FKP [37], the proposed MBRS-FCRN method shows an improvement of 5.222% in accuracy, 0.592% in sensitivity, 0.780% in specificity, 1.611% in precision, 2.430% in recall, 1.971% in F1-score, and a decrease of 1.497% in FAR.

Table-3 shows the classification performance comparison of finger print based MBRS systems. For EDWT [19], Accuracy improved by 18.355%, Sensitivity improved by 10.908%, Specificity improved by 22.113%, Precision improved by 10.133%, Recall improved by 11.024%, F1-score improved by 7.441% and FAR worsened by 1.214%. for, CS-LBP [23] Accuracy improved by -5.195%, Sensitivity worsened by 2.246%, Specificity improved by -3.758%, Precision improved by -6.080%, Recall improved by 1.678%, F1-score worsened by 3.962%, and FAR improved by 0.532%. for FKP [37] Accuracy improved by -18.814%, Sensitivity improved by 0.424%, Specificity improved by -7.439%, Precision improved by -4.875%, Recall worsened by 0.616%, F1-score improved by -2.317%, and FAR improved by 0.784%. Overall, the proposed MBRS-FCRN method outperformed all other methods in terms of accuracy, sensitivity, specificity, precision, recall, and F1-score, except for the FAR metric where it was outperformed by the EDWT method.

Table-1. Class specific performance of proposed MBRS-FCRN system.

Dataset	Face	Fingerprint	Iris	ear	Palmprint
Accuracy	99.489	98.945	98.395	98.280	98.945
Sensitivity	93.342	92.800	92.860	98.812	92.800
Specificity	94.984	91.137	90.517	98.516	91.137
Precision	90.005	90.531	91.191	98.431	90.531
Recall	91.887	92.249	92.549	98.860	92.249
F1-score	92.315	91.514	91.497	98.028	91.514
FAR	91.583	91.397	90.509	98.936	91.397

Table-2. Performance evaluation of Face based MBRS systems.

Dataset	EDWT [19]	CS-LBP [23]	FKP [37]	Proposed MBRS-FCRN
Accuracy	87.807	88.278	94.267	99.489
Sensitivity	89.650	87.722	92.750	93.342
Specificity	87.216	87.254	94.204	94.984
Precision	83.966	84.154	88.394	90.005
Recall	81.953	81.410	89.457	91.887
F1-score	80.952	88.270	91.344	92.315
FAR	89.615	88.781	90.080	91.583

Table-4 shows the classification performance comparison of iris based MBRS systems. The proposed MBRS-FCRN method improves the accuracy by 13.715%, sensitivity by 11.086%, specificity by 12.283%, precision by 7.471%, recall by 7.766%, F1-score by 13.715%, and FAR by 8.308% compared to EDA [44]. The proposed MBRS-FCRN method improves the accuracy by 8.152%, sensitivity by 4.342%, specificity by 9.313%, precision by 0.656%, recall by 9.311%, F1-score by 7.886%, and FAR by 5.905% compared to TRIMOON [45]. The proposed MBRS-FCRN method improves the accuracy by 12.010%, sensitivity by 11.264%, specificity by 4.301%, precision by 6.335%, recall by 5.676%, F1-score by 9.383%, and FAR by 2.670% compared to CNN [25].

Table-3. Performance evaluation of Fingerprint based MBRS systems.

Dataset	EDWT [19]	CS-LBP [23]	FKP [37]	Proposed MBRS-FCRN
Accuracy	81.353	94.881	80.868	98.945
Sensitivity	83.431	91.565	85.412	92.800
Specificity	87.271	90.527	88.010	91.137
Precision	80.738	88.089	89.675	90.531
Recall	81.382	89.553	81.192	92.249
F1-score	87.958	87.432	88.138	91.514
FAR	89.086	89.225	87.757	91.397

Table-4. Performance evaluation of Iris based MBRS systems.

Dataset	EDA [44]	TRIMOON [45]	CNN [25]	Proposed MBRS-FCRN
Accuracy	84.680	89.547	85.385	98.395
Sensitivity	82.774	88.898	81.323	92.860
Specificity	80.234	82.447	85.204	90.517
Precision	84.054	82.856	88.998	91.191
Recall	85.783	83.906	87.579	92.549
F1-score	84.680	83.890	89.050	91.497
FAR	82.774	85.827	80.949	90.509

Table-5 shows the classification performance comparison of ear based MBRS systems. The proposed MBRS-FCRN method improves the accuracy by 19.269%, sensitivity by 19.728%, specificity by 20.032%, precision by 9.847%, recall by 21.538%, F1-score by 14.235%, and FAR by 11.840% compared to KNN [20]. The proposed MBRS-FCRN method improves the accuracy by 16.339%, sensitivity by 10.790%, specificity by 12.569%, precision by 14.321%, recall by 10.616%, F1-score by 17.182%, and FAR by 14.107% compared to BSIF [17]. The proposed MBRS-FCRN method improves the accuracy by 9.032%, sensitivity by 17.062%, specificity by 21.527%, precision by 10.389%, recall by 18.258%, F1-score by 15.425%, and FAR by 14.232% compared to TRIMOON [45].

Table-5. Performance evaluation of ear based MBRS systems.

Dataset	KNN [20]	BSIF [17]	TRIMOON [45]	Proposed MBRS-FCRN
Accuracy	81.900	81.941	89.420	98.280
Sensitivity	82.440	89.500	83.862	98.812
Specificity	81.532	87.473	80.834	98.516
Precision	89.484	86.259	89.397	98.431
Recall	81.231	88.977	83.368	98.860
F1-score	85.631	83.460	85.064	98.028
FAR	88.944	84.292	86.357	98.936

Table-6 shows the classification performance comparison of palmprint based MBRS systems. The proposed MBRS-FCRN method improves the accuracy by 14.934%, sensitivity by 14.433%, specificity by 11.257%, precision by 1.493%, recall by 3.322%, F1-score by 13.269%, and FAR by 8.455% compared to TRIMOON

 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

[45]. The proposed MBRS-FCRN method improves the accuracy by 14.446%, sensitivity by 8.929%, specificity by 8.327%, precision by 10.212%, recall by 3.331%, F1-score by 9.663%, and FAR by 8.260% compared to CS-LBP [23]. The proposed MBRS-FCRN method improves the accuracy by 9.376%, sensitivity by 10.908%, specificity by 10.951%, precision by 10.145%, recall by 12.908%, F1-score by 10.144%, and FAR by 2.548% compared to EDA [44].

Table-7 shows the classification performance comparison of various approaches. Here, the proposed MBRS-FCRN classifier resulted in improved performance as compared to existing KNN [20], SVM [23], and CNN [25]. For accuracy metric, the proposed MBRS-FCRN model achieved a 98.179% accuracy, which is 3.21% and 3.06% higher than the KNN [20] and CNN [25] models respectively. There was no improvement over the SVM [23] model. The proposed MBRS-FCRN model achieved a 98.346% sensitivity, which is 4.13% and 4.37% higher than the SVM [23] and CNN [25] models respectively. However, it performed worse than the KNN [20] model, with a 7.77% lower sensitivity. The proposed MBRS-FCRN model achieved a 98.186% specificity, which is 0.62% and 3.98% higher than the SVM [23] and CNN [25] models respectively. However, it performed worse than the KNN [20] model, with a 0.56% lower specificity. The proposed MBRS-FCRN model achieved a 98.824% precision, which is 6.54%, 4.43%, and 5.43% higher than the KNN [20], SVM [23], and CNN [25] models respectively. The proposed MBRS-FCRN model achieved a 98.046% recall, which is 6.91%, 3.35%, and 3.59% higher than the KNN [20], SVM [23], and CNN [25] models respectively. The proposed MBRS-FCRN model achieved a 98.255% F1-score, which is 6.50%, 4.00%, and 2.82% higher than the KNN [20], SVM [23], and CNN [25] models respectively. The proposed MBRS-FCRN model achieved a 98.166% FAR, which is 3.23%, 4.09%, and 4.32% lower than the KNN [20], SVM [23], and CNN [25] models respectively.

Table-6. Performance evaluation of palmprint based MBRS systems.

Dataset	TRIMOON [45]	CS-LBP [23]	EDA [44]	Proposed MBRS-FCRN
Accuracy	84.011	84.499	89.569	98.945
Sensitivity	80.551	84.471	81.892	92.800
Specificity	82.623	82.989	81.186	91.137
Precision	89.038	80.319	80.346	90.531
Recall	89.978	89.048	81.816	92.249
F1-score	80.502	82.851	81.245	91.514
FAR	83.650	81.437	88.390	91.397

Table-7. Classification performance comparison of various approaches.

Metric	KNN [20]	SVM [23]	CNN [25]	Proposed MBRS-FCRN
Accuracy	90.240	94.881	94.267	98.179
Sensitivity	91.419	94.565	93.750	98.346
Specificity	91.795	93.527	94.204	98.186
Precision	92.170	94.089	93.394	98.824
Recall	91.174	94.553	94.457	98.046
F1-score	91.755	94.432	95.344	98.255
FAR	90.974	94.225	93.080	98.166

5. CONCLUSIONS

The implementation of MBRS by using deep learning for FCRN classification is the primary topic of this work. A Gaussian filter is used for the images that come from the face, fingerprint, and iris databases to get things started. Because of this, the photos go through pre-processing, which gets rid of all the different kinds of noise. Additionally, GLCM is used to extract the multimodal properties. After that, principal component analysis is used to filter down the features, and then, after that, partial least squares optimization is utilised to choose the characteristics from the accessible set that is of the utmost significance. An FCRN classifier is applied to the test dataset as the very last stage in the biometric reorganisation technique. This brings the whole process to a successful conclusion. This classifier takes advantage of the characteristics that have been taught. The results of the simulations revealed that the proposed technique resulted in higher performance when compared to other approaches that are considered to be state of the art. The proposed

MBRS-FCRN improved accuracy by 3.24%, sensitivity by 2.45%, and Specificity by 3.09% as compared to existing methods. In addition to this, this may be expanded upon using more types of biometrics.

REFERENCES

[1] Kathed A., Azam S., Shanmugam B., Karim A., Yeo K. C., De Boer F. and Jonkman M. 2019, January. An enhanced 3-tier multimodal biometric authentication. In 2019 International Conference on computer communication and informatics (ICCCI) (pp. 1-6). IEEE.

[2] Ryu R., Yeom S., Kim S. H. and Herbert D. 2021. Continuous multimodal biometric authentication schemes: a systematic review. IEEE Access. 9, pp. 34541-34557.

[3] El-Rahiem B. A., El-Samie F. E. A. and Amin M. 2022. Multimodal biometric authentication based on deep fusion of electrocardiogram (ECG) and finger vein. Multimedia Systems. 28(4): 1325-1337.

[4] Amritha Varshini, S. and Aravinth J. 2021. Hybrid level fusion schemes for multimodal biometric authentication system based on matcher performance. In Computational Vision and Bio-Inspired Computing: ICCVBIC 2020 (pp. 431-447). Springer Singapore.

[5] Mahmoud R. O., Selim M. M. and Muhi O. A. 2020. Fusion time reduction of a feature level based multimodal biometric authentication system. International Journal of Sociotechnology and Knowledge Development (IJSKD). 12(1): 67-83.

[6] Choudhary S. K. and Naik A. K. 2019, April. Multimodal biometric authentication with secured templates-A review. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1062-1069). IEEE.

[7] Zhang X., Yao L., Huang C., Gu T., Yang Z. and Liu Y. 2020. DeepKey: a multimodal biometric authentication system via deep decoding gaits and brainwaves. ACM Transactions on Intelligent Systems and Technology (TIST). 11(4): 1-24.

[8] Ahamed F., Farid F., Suleiman B., Jan Z., Wahsheh L.A. and Shahrestani S. 2022. An intelligent multimodal biometric authentication model for personalised healthcare services. Future Internet. 14(8): 222.

[9] Wu L., Yang J., Zhou M., Chen Y. and Wang Q. 2019. LVID: A multimodal biometrics authentication system on smartphones. IEEE Transactions on Information Forensics and Security. 15, pp. 1572-1585.

[10] Ramachandra R., Stokkenes M., Mohammadi A., Venkatesh S., Raja K., Wasnik P., Poiret E., Marcel S. and Busch C. 2019. Smartphone multi-modal biometric authentication: Database and evaluation. arXiv preprint arXiv:1912.02487.

[11] Olade Ilesanmi, Hai-ning Liang, and Charles Fleming. 2018. A review of multimodal facial biometric authentication methods in mobile devices and their application in head mounted displays. 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI): 1997-2004.

[12] Abozaid A., Haggag A., Kasban H. and Eltokhy M. 2019. Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion. Multimedia tools and applications. 78, pp. 16345-16361.

[13] Zhang, X., Cheng, D., Jia, P., Dai, Y. and Xu, X., 2020. An efficient android-based multimodal biometric authentication system with face and voice. IEEE Access, 8, pp.102757-102772.

[14] Hammad M., Liu Y. and Wang K. 2018. Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint. IEEE Access. 7, pp. 26527-26542.

[15] Chanukya P. S. and Thivakaran T. K. 2020. Multimodal biometric cryptosystem for human authentication using fingerprint and ear. Multimedia Tools and Applications. 79, pp.659-673.

[16] Yuan C., Jiao S., Sun X. and Wu Q. J. 2021. MFFFLD: A multimodal-feature-fusion-based fingerprint liveness detection. IEEE Transactions on Cognitive and Developmental Systems. 14(2): 648-661.

[17] Herbadji A., Guermat N., Ziet L. and Cheniti M. 2019, November. Multimodal Biometric Verification using the Iris and Major Finger Knuckles. In 2019

International Conference on Advanced Electrical Engineering (ICAEE) (pp. 1-5). IEEE.

[18] Sujatha E. and Chilambuchelvan A., 2018. Multimodal biometric authentication algorithm using iris, palm print, face and signature with encoded dwt. Wireless Personal Communications. 99(1): 23-34.

[19] Ma L., Sham C. W., Lo C. Y. and Zhong X., 2021. An effective multi-mode iris authentication system on a microprocessor-fpga heterogeneous platform with qc-ldpc codes. IEEE Access. 9, pp. 163665-163674.

[20] Bokade G. U. and Kanphade R. D. 2019, July. Secure multimodal biometric authentication using face, palmprint and ear: a feature level fusion approach. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.

[21] Sengar S. S., Hariharan U. and Rajkumar K. 2020, March. Multimodal biometric authentication system using deep learning method. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 309-312). IEEE.

[22] Vijayakumar T. 2021. Synthesis of palm print in feature fusion techniques for multimodal biometric recognition system online signature. Journal of Innovative Image Processing (JIIP), 3(02): 131-143.

[23] Jaswal G., Kaul A. and Nath R. 2018. Multimodal biometric authentication system using hand shape, palm print, and hand geometry. In Computational Intelligence: Theories, Applications and Future Directions-Volume II: ICCI-2017 (pp. 557-570). Singapore: Springer Singapore.

[24] Joseph T., Kalaiselvan S. A., Aswathy S. U., Radhakrishnan R. and Shamna A. R. 2021. Retracted Article: A multimodal biometric authentication scheme based on feature fusion for improving security in cloud environment. Journal of Ambient Intelligence and Humanized Computing. 12(6): 6141-6149.

[25] Vyas R., Kanumuri T., Sheoran G. and Dubey P. 2022. Accurate feature extraction for multimodal biometrics combining iris and palmprint. Journal of Ambient Intelligence and Humanized Computing. 13(12): 5581-5589.

[26] Medjahed C., Rahmoun A., Charrier C. and Mezzoudj F. 2022. A deep learning-based multimodal biometric system using score fusion. IAES Int. J. Artif. Intell. 11(1): 65.

[27] Lefkovits S., Lefkovits L. and Szilágyi L. 2022. HGG and LGG brain tumor segmentation in multi-modal MRI using pretrained convolutional neural networks of Amazon Sagemaker. Applied Sciences. 12(7): 3620.

[28] A. l. Alkeem E., Yeun C. Y., Yun J., Yoo P. D., Chae M., Rahman A. and Asyhari A. T. 2021. Robust deep identification using ECG and multimodal biometrics for industrial internet of things. Ad Hoc Networks. 121, p. 102581.

[29] Gadzicki K., Khamsehashari R. and Zetzsche C. 2020, July. Early vs late fusion in multimodal convolutional neural networks. In 2020 IEEE 23rd international conference on information fusion (FUSION) (pp. 1-6). IEEE.

[30] Gadzicki K., Khamsehashari R. and Zetzsche C. 2020, July. Early vs late fusion in multimodal convolutional neural networks. In 2020 IEEE 23rd international conference on information fusion (FUSION) (pp. 1-6). IEEE.

[31] Arora A. and Miri R. 2022. Cryptography and Tay-Grey wolf optimization based multimodal biometrics for effective security. Multimedia Tools and Applications. 81(30): 44021-44043.

[32] Cai C., He Y., Sun L., Lian Z., Liu B., Tao J., Xu M. and Wang K. 2021. Multimodal sentiment analysis based on recurrent neural network and multimodal attention. In Proceedings of the 2nd on Multimodal Sentiment Analysis Challenge (pp. 61-67).

[33] Gona A., Subramoniam M. and Swarnalatha R. 2023. Transfer learning convolutional neural network with modified Lion optimization for multimodal biometric system. Computers and Electrical Engineering. 108, p. 108664.

[34] Kim W., Song J. M. and Park K. R. 2018. Multimodal biometric recognition based on convolutional neural network by the fusion of finger-vein and finger shape using near-infrared (NIR) camera sensor. Sensors. 18(7): 2296.

[35] Wang Y., Shi D. and Zhou W. 2022. Convolutional neural network approach based on multimodal biometric system with fusion of face and finger vein features. Sensors. 22(16): 6039.

[36] Ren H., Sun L., Guo J. and Han C. 2022. A dataset and benchmark for multimodal biometric recognition based on fingerprint and finger vein. IEEE Transactions on Information Forensics and Security. 17, pp. 2030-2043.

[37] Daas S., Yahi A., Bakir T., Sedhane M., Boughazi M. and Bourennane E. B. 2020. Multimodal biometric recognition systems using deep learning based on the finger vein and finger knuckle print fusion. IET Image Processing. 14(15): 3859-3868.

[38] Gona A. and Subramoniam M. 2022. Convolutional neural network with improved feature ranking for robust multi-modal biometric system. Computers and Electrical Engineering 101, p. 10.8096.

[39] Kumar D., Sharma S. and Mishra M. P. 2022, December. Multimodal Biometric Human Recognition System-A Convolution Neural Network based Approach. In 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART) (pp. 477-482). IEEE.

[40] Gona A. K. and Subramoniam M. 2022, October. Multimodal Biometric Reorganization System using Deep Learning Convolutional Neural Network. In 2022 International Conference on Edge Computing and Applications (ICECAA) (pp. 1282-1286). IEEE.

[41] Xu H., Qi M. and Lu Y. 2019, October. Multimodal biometrics based on convolutional neural network by two-layer fusion. In 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) (pp. 1-6). IEEE.

[42] Soleymani S., Dabouei A., Taherkhani F., Iranmanesh S.M., Dawson J. and Nasrabadi N.M. 2021. Quality-aware multimodal biometric recognition. IEEE Transactions on Biometrics, Behavior, and Identity Science. 4(1): 97-116.

[43] Guan J., Fei J., Li W., Jiang X., Wu L., Liu Y. and Xi J. 2023. Defect classification for specular surfaces based on deflectometry and multi-modal fusion network. Optics and Lasers in Engineering. 163, p. 107488.

[44] Zhu L. and Spachos P. 2022, December. Annotation Efficiency in Multimodal Emotion Recognition with Deep Learning. In GLOBECOM 2022-2022 IEEE Global Communications Conference (pp. 560-565). IEEE.

[45] Xiong S., Zhang G., Batra V., Xi L., Shi L. and Liu L. 2023. TRIMOON: Two-Round Inconsistency-based Multi-modal fusion Network for fake news detection. Information fusion. 93, pp. 150-158.

[46] Zhang T., Chen S., Wulamu A., Guo X., Li Q. and Zheng H. 2022. TransG-net: transformer and graph neural network based multi-modal data fusion network for molecular properties prediction. Applied Intelligence. pp. 1-12.

[47] Sun H., Chen Y.W. and Lin L. 2022. TensorFormer: A Tensor-based Multimodal Transformer for Multimodal Sentiment Analysis and Depression Detection. IEEE Transactions on Affective Computing.

[48] Rajasekar V., Predić B., Saracevic M., Elhoseny M., Karabasevic D., Stanujkic D. and Jayapaul P. 2022. Enhanced multimodal biometric recognition approach for smart cities based on an optimized fuzzy genetic algorithm. Scientific Reports. 12(1): 622.

[49] Aung H., Pluempitiwiriyawej C., Wangsiripitak S. and Hamamoto K. 2022, January. Siamese Neural Network based Multimodal Biometrics Recognition. In The 14th Regional Conference on Electrical and Electronics Engineering (RC-EEE 2021) (p. 147).

[50] Barde S. and Singh K. K. 2022. Developed Face and Fingerprint-Based Multimodal Biometrics System to Enhance the Accuracy by SVM. In Smart and Sustainable Technologies: Rural and Tribal Development Using IoT and Cloud Computing: Proceedings of ICSST 2021 (pp. 325-336). Singapore: Springer Nature Singapore.