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

FRMSDNET Classifier for Multimodal Feature Fusion Biometric Authentication

Parvathy J.[1], Dr. Poornima G. Patil[2]

Submitted: 26/09/2023 **Revised**: 16/11/2023 **Accepted**: 27/11/2023

Abstract: Biometric System (BS), which requires biometric information from people grounded on their physical characteristics and/or behavioral attributes like Fingerprint (FP), iris, face, or voice pattern, is a pattern recognition system. Unimodal BS has a number of issues in real-world applications. It may include the person's age factor, degradation of biological traits, and depending on just one trait that loses the authentication credits. A novel multi-BS grounded on the FP and iris has been proposed to get over the issue with unimodal systems. Two biometric characteristics, namely FP and iris are utilized in the proposed work. Initially, the FP image is pre-processed; in addition, by employing the Adaptive Ostu Mode segmentation (AOMS), the FP region is segmented. Next, the binarization and injection of segmented images into the ridge thinning procedure are done; next, the minutiae points are extracted. Afterward, the Iris image is pre-processed. Also, by deploying Kernel Snake Control Method (KSCM), the regions of the iris are segmented. The features are extracted from the segmented regions. By utilizing the Levy Good, the Bad, and the Ugly Optimizer (LGBUO), the significant features are selected as of the obtained features from the 2 phases. Lastly, the selected features are fused, which is then fed to the Fuzzy Residual Mean Squared Deviation Network (FRMSDNET) Classifier. The experiential outcomes exhibited that the proposed model attained an enhanced performance.

Keywords: *Adaptive Ostu Mode Segmentation Technique (AOMS), Kernel Snake Control Method (KSCM), Levy Good, The Bad, and the Ugly Optimizer (LGBUO), Fuzzy Residual Mean Squared Deviation Network (FRMSDNET) Classifier, deep learning.*

1. Introduction

Real-time applications were enabled by several advancements in technology via remote access in banking, medical, national security, trade, enterprise, law enforcement, along with various other sectors [1]. Recently, the major issue is security needs. In many ways, security could be managed. Prior to the secure resource's release, the correct person is identified by a process named authentication. To attain this, the person acquires counter-configuring unique information [2]. Human identity verification was provided with a Biometric Authentication System (BAS) to perform high-level security [3]. The physiological along with behavioural characteristics were utilized by BS for authentication [4]. The static mannerism of the human body is linked with the Physiological characteristics, which never alters regarding age; however, only to the behaviour of personnel, the behavioural style of biometrics is limited [5]. Face, Fingerprint (FP), iris, et cetera is the characteristics that come under the physiological characteristics, while signature, voice, and gait are included in the behavioural biometrics characteristics [6].

The Biometric System (BS) is named a verification system or identification system grounded on the applications utilized in the environment [7]. The system attempts to differentiate an individual as of a population when executing identification. It also seeks to verify a person's identity during verification [8]. It also has various applications like electronic data security, ATM, credit card, computer network login, e-commerce, mobile phones, distant learning, health records management, physical access control, internet access, et cetera [9]. Biometrics can't be lost, manipulated, copied, shared, forgotten, or else forged unlike other conventional methodologies, namely passwords and tokens [10].

Unimodal as well as multimodal are the 2 varieties of biometric recognition systems. The user is recognized by the unimodal system that utilized a single biometric trait [11]. Variations in biometric data, low recognition rate, spoofed easily, etcetera were the issues faced by the unimodal system [12]. By employing a scheme that merges numerous biometric input sources into a single decision, a combination of diverse BS could be employed to tackle the issues in unimodal BAS [13, 14]. Noise affects the multimodal BS slightly. The non-universality issue is tackled along with elevating the matching accuracy; thereby providing a secure storage environment [15]. When analogized to unimodal BS, a higher rate of security was provided by the Multi-BS [16]. Multi-sensors, -samples, -algorithms, and -instances are the four categories of multi-BS. Information fusion is defined as merging information as of multiple biometric sources [17]. Generating a proper function is the aim of fusion and it

[1] *Parvathy J, Research Scholar, VTU, Belagavi*
[2] *Dr Poornima G Patil, Assistant Professor, VTU, Belagavi*
** Corresponding Author Email: paru.jan@gmail.com*

optimally fuses the data gathered as of various BS [18]. Owing to a large number of issues like (1) intra-class variability (2) inter-class similarity (3) segmentation (4) noisy input (5) scalability, together with (6) template size, the BS design is a complex process.

The biometric input is processed by utilizing various algorithms. One amongst the most enriched as well as highly accurate approaches is the Convolutional Neural Network (CNN) , which has a high rate of recognition along with accuracy. However, noise, less investigation of features, and high complexity are the defects contained in the prevailing BAS. So, an efficient Multimodal Biometric Authentication (MBA) was proposed in this work utilizing a novel Fuzzy Residual Mean Squared Deviation Network (FRMSDNET) classifier.

1.1. Problem Definition

Despite developing several Machine learning (ML) and Deep Learning (DL) based models for biometric authentication, still, certain limitations are needed to be solved for better authentication. Some of the limitations of existing models are framed as follows:

• The biological traits get deteriorated owing to the person's age factor; also, the credit of authentication is lost as it relies on only one single trait. For instance, combinations of various traits are used for multimodal authentication.

• In the field of biometric recognition, the key research objects are FPs and iris, which are non-invasive and easier to accept. Zero error recognition could not be attained by the prevailing techniques in the field of FP and iris recognition. Thus, FPs and iris fusion recognition have become a research hotspot.

• In unimodal biometric systems, various issues are there regarding real-world applications. This operates on a single biometric modality, namely intra-class variations, non-universality, noise in sensed data, inter-class similarities, along with spoof attacks.

1.2. Objectives

To overcome these issues, the paper proposes an enhanced multimodal biometric authentication system using a novel FRMSDNET classifier. It consists of some research objectives, which are enlisted as follows:

 To propose novel Image Enhancement techniques to increase the structure and clarity of an image.

 To propose a novel Image Segmentation Technique that gives proper segmentation of an image

 To mount the detection accuracy and lower the computational complexity using the Novel Neural Network Classifier.

The proposed technique's major contributions are listed below:

• A Contrast Local Variance Adaptive Histogram Equalization (CLVAHE) method is proposed for increasing the contrasts of the images' contrasts; thus, the image details, which lie in the lower dynamic range, could be evaluated efficiently.

• An Levy Good, the Bad, and the Ugly Optimizer (LGBUO) is utilized for selecting the important features of both iris and FP for reducing the training time of the classifier.

• An iris region is segmented using Kernel Snake Control Method (KSCM) for avoiding the unwanted extractions of features.

• A FRMSDNET classifier is developed for authentication by removing the complexity

This remaining paper is arranged as: Section 2 exemplifies the associated work; Section 3 elucidates the proposed technique; Section 4 illustrates the proposed technique's results; lastly, the paper is winded up in Section 5.

2. Literature Survey

Mohammed Chachan Younis and Huthaifa Abuhammad [19] developed an MBA framework centered on a hybrid multi-phase feature fusion to deliver compact knowledge as of multiple model traits. Utilizing Discriminant Correlation Analysis (DCA) together with Canonical Correlation Analysis (CCA), the fusion was executed at each single as well as hybrid phase. Here, CCA helped to combine the multiple set of extracted features from the enhanced images' entirely connected vector layers. DCA incorporated the class association into correlation analysis; thus, the features not only had the maximum intrinsic correlation betwixt feature sets but also had class structure information. The outcomes exhibited that better accuracy was attained by the presented one than the prevailing models. The optimization need of CNN architecture to elevate its performance was the limitation of this presented model.

FerielCherifet et al. [20] engendered an effectual, unobtrusive, along with robust Multimodal BS (MBS) to validate smartphone users. Grounded on the score-level fusion strategy, ear shape along with arm gesture features-set was combined in this scheme. For computing ear features, an image fragmentation model was utilized for elevating the authentication system's performance. Random Forest (RF) was employed for performing classification. An enriched performance was attained by MBS than the prevailing ones. However, over fitting issues were created by RF utilization.

Rasha O. Mahmoud et al. [21] developed an MBA method for validating a person's identity grounded on the face along with iris features. MBA system overcame the shortcomings of both single systems, enhanced the overall system performance, and enriched the security. Two characteristics were combined in this system, one was physical (fingerprint), whereas the other was vital (ECG), which offered the advantage of liveness detection to the system that makes the system robust to spoof attacks. Here, the Rectangle Histogram of Oriented Gradient (R-HOG) extracted the facial features followed by Principal Component Analysis (PCA) performing the feature reduction. Gabor filters were utilized for extracting the iris features. The classification process was executed by employing a Deep Belief Network (DBN)-based classifier. The outcomes revealed that superior performance was attained by the presented scheme. Nevertheless, crucial information was lost by PCA utilization.

Ajai Kumar Gautam and Rajiv Kapoor [22] developed the Multi-Modal Biometrics (MMB) recognition system centered on the Feature Level and Score Level (FLSL) fusion technique along with Modified Deep Learning Neural Network (NN) (MDLNN) classifier. Entire inputted traits were enhanced by the Improved Plateau Histogram Equalization (IPHE) approach. The IPHE followed these steps: the images' pixel values were arranged in ascending order. Then, the histogram building was produced; afterward, the histogram median value was computed, which was rounded off to the nearest integer value, that is, the Threshold Value (TV). Thereafter, the EXOR operations of equivalent two histogram values were executed; also, the values are considered as Cumulative Distributions Function (CDF). After that, the gamma correction was performed for gathering the enhanced image by regulating the intensity level of the image. After completing all the steps of contrast enhancement, facial parts were segmented by Viola-Jones Algorithm (VJA). The unnecessary information in the ear along with finger traits was eliminated; also, the retina's blood vessel images were segmented by the Penalty and Pearson correlation-centric Watershed Segmentation (PPWS). The features were extracted later. After that, the Kernelized Linear Discriminants Analysis (KLDA) diminished the feature's dimensions. Then, the features were fused by the FLSL. The person is categorized as genuine or an imposter. A better performance was exhibited by the presented model. However, while diminishing the usable signal, IPHE utilization may lead to an elevation in contrast to background noise.

SurabhiHom Choudhury et al. [23] developed a scheme for adaptive rank-level biometric fusion. The optimum performance was provided by the rank-level fusion rules that utilized the Ant Colony Optimization technique. The plates deemed here were the Index, middle, together with ring fingernail plates. Also, the '3' customized pre-trained models, including Alex Net, ResNet-18, along with DenseNet-201 extracted the '3' nail plates' Deep Learning (DL) feature sets. The outcomes suggested that the optimal parameters for rank-level fusion for altering security levels were computed by this model, thereby contributing toward optimal performance accuracy. But, higher complexity was exhibited in this model.

Mohamed Hammad et al. [24] introduced a secure MBA system that employs a convolution NN (CNN) together with Q-Gaussian Multi-Support Vector Machine (QG-MSVM) grounded on a diverse level fusion. A feature along with decision-level fusion was the two-level fusion algorithms implemented. Later, CNN was utilized to extract the features. To elevate the performance, the classifier employed for authentication was QG-MSVM. The experiential outcomes exhibited that more efficiency, robustness, and reliability were attained by the generated multimodal systems than the prevailing MBA models. However, owing to the large feature size, the system's computation time was high.

Chanderkant and Sheetal Chaudhary [25] developed a watermarking-centered approach to guard templates in the MBA. Iris as well as face patterns were retrieved as of watermarked images at the authentication time; also, matching was executed with the query image. After the fusion of iris together with face matching scores at the match score level, the final decision was provided by the obtained matching score. The experiential outcomes demonstrated the developed system's higher security along with recognition accuracy. Owing to watermarking, the image's clarity was missed. Hence, learning the image for more time was led by the system.

M. Vijay and G. Indumathi [26] demonstrated the MBA model grounded on the Score Level Fusion (SLF). Pre-processing, Feature Extraction (FE), recognition score utilizing a Multi-Support Vector NN (Multi-SVNN) for all traits, SLF, together with recognition using a DBN were the five phases involved in the developed model's overall procedure. The experiential outcomes proved that maximal accuracy, sensitivity along with specificity was acquired by the developed models. However, for a small amount of data, the DBN was not appropriate.

B.Sreevidya and E.Chandra [27] introduced fresh security architecture meant for secured storage in the cloud. It provided robust authentication by wielding various biometric modalities as of users along with allowing/denying access accordingly. The Entropy-centric Local Binary Pattern (ELBP) was utilized for FE. For bearing the testimony to the authentication systems' performance, the Receiver Operating Characteristics Curve (ROC) was wielded. However, the optimal time period was not attained even though substantial improvement was

attained by the developed ELBP FE.

Chengsheng Yuan et al. [28] suggested a more robust and accurate Biometrics Liveness Detection (BLD) by considering the weighted Multimodal Convolution NN (MCNN)'s merit for extracting varied deep features. Initially, those FPs' invalid backgrounds were removed by executing the Region of Interest (ROI) operation. Then, for making full utilization of CNN's learning capacity devoid of human interactions, a multimodal feature fusion strategy was utilized. The experiential outcomes suggest that outstanding outcomes were attained by the presented approach in FP intra-database along with cross-material evaluations as well as in face anti-spoofing. However, the model was not cost-effective.

Arafat Rahman et al. [29] demonstrated an MBS combining Electro Encephalography (EEG) and keystroke dynamics. By utilizing multi-domain FE (time, frequency), Feature Selection (FS) (Gini impurity), classifier design, along with SLF, a machine-learning classification pipeline was engendered. Thus, the developed system's higher accuracy was demonstrated by the experiential outcomes. However, owing to the prevalence of vanishing gradient issues in a neural network, the data was learned slowly by the developed system.

Gurjit Singh Walia et al. [30] introduced an MBS centered on the optimal SLF. The Backtracking Search Optimization Algorithm (BSA) optimized the individual classifiers. By utilizing BSA, optimal values of confidence factors were obtained. BSA was an evolutionary technique to overcome the issues of slow computation as well as premature convergence. BSA comprised 4 steps, such as (1) initialization, (2) selection-I, (3) selection II, and (4) mutation and crossover. Proportional Conflict Redistribution (PCR-6) rules integrated the optimized belief masses of classifiers. Under several challenges, the experiential outcomes suggest that prevailing fusion methodologies were outperformed by the developed multimodal fusion method. However, the challenge here was to attain the score-centric fusion's adaptive nature.

VeeruTalreja et al [31] introduced a DL framework meant for feature-level fusion that engendered a secure multimodal template as of every single user's face along with iris biometrics. To combine more than two modalities, Fully Connected Layer (FCL) together with bilinear architecture were the DL-centered fusion architectures. The outcomes exhibited a low level of security even though the presented system attained higher efficacy.

Ayesha Tarannum et al. [32] developed a fresh multi-user-centered multi-modal authentication. For every user with sturdy integrity along with the encryption process, the dynamic keys were engendered here. To supply strong security to the huge data size in distributed applications, every single user's multi-modal features were utilized to implement fresh integrity verification. When analogized to the prevailing system, the experiential outcomes revealed the developed system's higher run time. However, owing to the over-fitting issue, the developed system had more complexity

Felix O. Babalola et al. [33] introduced the 1st multimodal vein database termed FYO. To be wielded as a defence against several attacks to BS, it had '3' biometric modalities; palmar vein, dorsal vein, along with wrist vein. The image was classified by developing Convolutional Neural Networks (CNN) with the model. Thus, superior performance was exhibited by the experiential outcomes. More time was taken for the classification owing to the system's narrow waist.

3. Proposed Multimodal Biometric Authentication System

For efficient detection along with a classification of authentication, a novel FRMSDNET classifier was proposed here. FP image and Iris image were the two inputs included in the proposed system. In this work, the iris and FPs are selected for constructing a multimodal biometric system for the given reasons: Iris modality is the most reliable biometric characteristic; it is a protected organ; also, it has a unique texture that remains unchanged during adult human life. For the identification process, the iris region is segmented as of the eye image. The FP is the most commonly used one that could not be lost or misplaced; also, they are always with the person. FP patterns, which could not be guessed, are non-transferable. Hence, for identity verification, a higher level of security is offered by FP biometrics. The input images are further derived under six common steps. From both inputs, the features are computed individually along with fused for classification. Figure 1 exhibits the proposed technique's block diagram.

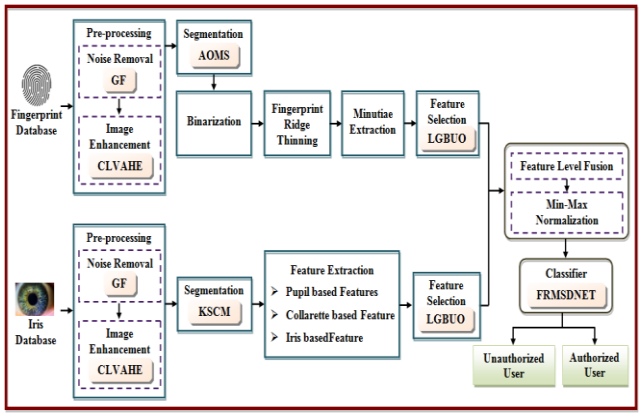


Fig 1. Block diagram of the proposed methodology

3.1 Fingerprint Image

The proposed methodology's initial phase has proceeded with the input FP image. The FP image in the BS aids in

elevating the security level with authentication. Pre-processing, segmentation, binarization, FP ridge thinning, minutiae extraction, and FS are the steps in which the input image is executed.

3.1.1 Pre-Processing

Because of the noise complexity along with the input image's fadedness, the input FP image is pre-processed. The Gabor filter is utilized to diminish the noise via pre-processing. The Contrast Local Variance Adaptive Histogram Equalization (CLVAHE) algorithm is employed to elevate the image's contrast.

Noise removal: Owing to the noise prevalence, the input image is filtered by executing this prominent step. Gabor filter also known as a linear filter is utilized in this proposed system. The noise is eliminated; in addition, the FP image's blurred regions are elevated. By multiplying a Gaussian envelope function with a complex oscillation, these filters' impulse responses are engendered that can be expressed as,

$$G = \exp\left(\frac{-\alpha^2 + \gamma^2\beta^2}{2\sigma^2}\right)\cos\left(2\pi\frac{\alpha}{\beta} + \Phi\right) \quad (1)$$

Where, the parameters utilized in the Gabor filter for better response are $\alpha, \lambda, \beta, \Phi$ the Gaussian envelope's sigma is signified as σ. The Gaussian envelope function restricts the impulse response of the input FP image in the spatial domain and also determines the temporal decay of impulse response and step size, which ensures that the filter is based on impulse response. In other words, the Gaussian envelope is used to specify either the spatial frequency or the orientation more precisely. Therefore, the very optimal convolution operation is performed over the input image, and the noises are eradicated efficiently. A Gabor filter modulated by Gaussian wave was viewed as a sinusoidal signal of a particular frequency and orientation. By utilizing the real parts of diverse Gabor filter kernels, the images are filtered to obtain the filtered image Y_f^{filter}

Contrast enhancement: Here, to elevate the valley's structure clarity and the image's ridges, the CLVAHE algorithm enhances the filtered image Y_f^{filter}.

For augmenting the image's contrast, a variant of adaptive histogram equalization is applied over all neighborhood pixels. Over-amplifying noise in relatively homogenous regions of an image was the tendency shown by the prevailing CLAHE technique. The clip limit value is chosen as of the local variance of an image in the proposed methodology to avoid over-amplification. The expectation of the squared deviation of random pixels from its sample mean is named local variance by which the pixel distribution in an image as of mean was determined. The distribution of pixels in an image was known by the standard deviation. Local variance helps to measure the texture of the image and the studies of spatial FP image structure for gaining insight information of images at a local scale. Therefore, the local variance is chosen here for gathering information regarding the spreading of pixels. Through this information, the correlation between each pixel of the image is identified by calculating the deviation of each pixel from their neighborhood pixel. The variance is calculated for every divided window in the image. From the variance of the windows, the best lower variance is considered as the clip limit. Thus, the contrast of the image is enhanced significantly, and also the over-amplification is reduced abundantly. Adding contrast makes it possible to detect even the smallest informative regions of FPs and provides information about the precise location of the ridges and edges.

CLVAHE works on the image's smaller regions named tiles instead of the whole image. By utilizing the bilinear interpolation, the surrounding tiles are blended for removing the false boundaries. This helps to improve image contrast. Also, it achieves enhancement by efficiently spreading out the most common intensity values, that is, stretching out the image's intensity range.

Primarily, the input FP is partitioned into contextual regions; then, the contextual region's histogram was computed. Each contextual in the image has been applied with a mapping function along with engendering the grey level. The grayscale mapping provides a similar enhancement for all regions of the image. By utilizing a weighted sum of the mappings of its four nearest regions, the mapping for a pixel is obtained. Grounded on the proximity of the pixel to the centers of the 4 nearest regions, the weights are computed. For each grey level in the histogram, a number of pixels are equally partitioned in the contextual region. The pixels are portioned based on taking the average of all the pixels in each gray level of the contextual region. The partitions help to efficiently analyze the pixel for enhancing and reducing the noise. Also, this partition is done manually based on the average pixel value. So, the average number of pixels in the grey level P_{avg} is computed as,

$$P_{avg} = \frac{N_k - N_j}{N_g} \quad (2)$$

Where, the pixels in the contextual region's direction are given as N_k the pixels in the contextual region's V direction are symbolized as N_j the number of grey levels was notated as N_g The user clip limit I_{CL} is calculated utilizing the local variance as,

$$I_{CL} = E[P_{avg}{}^2] - (E[P_{avg}])^2 \quad (3)$$

Where, the local mean of P_{avg} is denoted as. $E[P_{avg}{}^2]$

The pixels can be redistributed utilizing the threshold limit for each grey level. Therefore, the overabundant pixels' initial redistribution to the histogram's remaining bins is executed. For this, an auxiliary variable is determined. The value is added to every single bin for which this operation will not result in an exceedance of the clip limit. Afterward, the remaining excess is evenly distributed amongst all allowable bins. This is generally employed in a "while" loop. The pixels can be redistributed utilizing the threshold limit for each grey level. To eliminate the artificially induced boundaries, the neighboring tiles are combined after the execution of equalization using bilinear interpolation. Hence, Y_f^{pre} denotes the pre-processed image.

3.1.2 Segmentation

Utilizing the Adaptive Ostu Mode segmentation Technique (AOMS), the FP region is segmented from Y_f^{pre} after pre-processing. The technique of separating the foreground region signified by the FP region as of the background region as per the threshold level is named Otsu Thresholding. Otsu's method is adaptive thresholding for image binarization. Otsu Thresholding has simplicity and speed. Prior knowledge about the processed image is not required for Otsu's method. In Otsu Thresholding, the TV is not chosen but is determined automatically. The information taken as of the image histogram was utilized to determine the optimal threshold. An image histogram, which acts as a tonal distribution's graphical representation in a pre-processed image, is a sort of histogram. It plots the number of pixels for every single tonal value. Two peaks are enclosed in the histogram generated. Thus, a generic condition will select a TV, which lies in the middle of both the histogram peak values. However, this mechanism has limitations in distinguishing the object and background for diverse greyscale values to generate a binary image. The image is corrupted severely by additive noise along with diminished grey level histogram's sharp valley if the object area is small when analogized with the background area and the object's variance. The possibly incorrect threshold is then determined, which generates outcomes in the segmentation error. Rather than utilizing a mean value, the mode value of background pixels subtracted as of the standard deviation of object pixels is utilized to tackle these issues. This significantly reduces the additive noise and prevents the sharp valley of the gray level by selecting the optimal TV based on the mode value. Hence, the modified Otsu thresholding is called AMOS. The steps of AMOS are explained below.

Grounded on all thresholds' possibilities, the FP regions are separated by the proposed method. On each side of the threshold, a gauge of spread for the pixel levels, that is, the pixels that either fall in the foreground or else the background are also computed. By diminishing the intra-class intensity variance, or else equivalently augmenting the inter-class variance, the threshold is determined. Then, the threshold $\Im$ is expressed as,

$$\Im = \arg\max_{1<\Im<\Delta} \{\wp^2_{inter}(\Im) \mid \wp^2_{inter} = P_P(\lambda_P - \lambda_I)^2 + P_B(\lambda_B - \lambda_I)^2 \quad (4)$$

Where, the maximum pixel value in the input image is represented as Δ, cumulative probabilities of FP and background region are modeled as P_P and P_B, the adaptive mode of two regions is notated as λ_P and λ_B, the whole image's adaptive mode is specified as λ_I, the inter-class variance is denoted as $\wp_{inter}(\Im)$ which is independent of the optimal thresholds' choice. The occurrence probability of grey level G_i in the histogram for a given grey-level image is given by,

$$G_i = \frac{\hbar_i}{a \times b} \quad (5)$$

Where, the pixels plotted in the histogram are signified as $\hbar_i$, the image size is represented as $a \times b$. The probability distribution of two regions is computed by building a histogram utilizing the image's pixels as follows,

$$P_P = \sum_{i=1}^{\Im} G_i \quad (6)$$

$$P_B = \sum_{i=\Im+1}^{\Delta} G_i \quad (7)$$

Where, the FP region's pixel intensity range is proffered as $[1, \Im]$ and $[\Im+1, \Delta]$is the background region's pixel intensity range. The adaptive mode for both regions that aid to fix the optimal TV was obtained in the ensuing phase. Hence, it is formulated as,

$$\lambda_{L,B} = L_{L,B} + \frac{\mu_{L,B(c)} - \mu_{L,B(p)}}{2\mu_{L,B(c)} - \mu_{L,B(p)} - \mu_{L,B(s)}} * \delta_{L,B} \quad (8)$$

Wherein, the adaptive mode of two classes is represented by $\lambda_{L.B}$, the lower intensity value is notated as $L_{L.B}$, the frequency of two classes is proffered as $\mu_{L.B(c)}$, the frequency of class preceding and succeeding the modal class is symbolized as $\mu_{L.B(p)}$ and $\mu_{L.B(s)}$, the size of the class interval of '2' classes is indicated as $\delta_{L.B}$. Similarly, the whole image's adaptive node λ_I is measured. The TV

is computed as of the whole process along with analogized with the pixels in the input image is expressed as,

$$Y_f^{seg} = \begin{cases} Y_{fp} & i > \Im \\ Y_{back} & i < \Im \end{cases} \quad (9)$$

Wherein, the output of segmentation is denoted as Y_f^{seg}, the pixels in the input image are notated as i, the segmented FP region is modeled as Y_{fp}, along with the segmented background region is signified as Y_{back}

3.1.3 Fingerprint Ridge Thinning

By enhancing the contrast betwixt the ridges and valleys, binarization is executed to convert the greyscale FP image into a binary image after the FP region is segmented. This process has continued under the "FP Ridge Thinning," in which the image's digital shape is converted into a clear skeleton from which the proposed can extract efficient features. To diminish the thickness of the ridges to a single pixel, thinning is then applied to the enhanced FP image after being binarized. The process of reducing the lines' thickness as possible with minimal losses is referred to as a thinning process. This process is significant for identifying the FP image's exact pattern. Due to the thinning process, image redundancies and unwanted pixels can be eliminated. Also, the discontinuities are easily redrawn over the binarized image, and the pixels are changed into a single-pixel line. Therefore, accurate minute extraction over thinned images can be performed. The thinning operation aids to have comprehending that operator prior to reading on, as it is linked to the hit-and-miss transform. The image's output with the single pixel $Y_{f(thin)}^{seg}$ is expressed as,

$$Y_{f(thin)}^{seg} = Y_f^{seg} - \kappa \quad (10)$$

$$\kappa = \begin{cases} if\ Y^{str}{}_{FB} = Y_f^{seg}{}_{FP},\ then Y^{Fore} \\ if\ Y^{str}{}_{FB} \neq Y_f^{seg}{}_{FP},\ then Y^{BACK} \end{cases} \quad (11)$$

Where, hit-and-miss transform for each pixel is represented by κ. The segmented FP image's hit-and-miss transform is performed by translating the structuring element's origin to all points in the image and then contrasting the structuring element (Y^{str}) with the underlying FP image pixels. The pixel underneath the structuring element's origin is set to the foreground color (Y^{Fore}) (ridge region) if the foreground and background pixels in the structuring element $(Y^{str}{}_{FB})$ exactly match the foreground and background pixels in the FP image (Y_{FP}^{seg})

If it doesn't match, then that pixel is set to the background color (Y^{BACK}) (the gap betwixt two ridges). By combining this foreground region, the skeleton structure of the FP region is formed. This is the procedure for thinning the ridges of FP.

3.1.4 Minutiae Extraction

The point, where a ridge forks or diverges into branch ridges is named ridge bifurcation. These features are collectively named minutiae. Most FP extraction, as well as matching techniques, restricts the set of features to 2 sorts of minutiae, namely ridge endings and ridge bifurcations. Minutiae extraction aids in obtaining informative data about FP status. Crossing Number (CN) is the generally deployed technique of minutiae extraction. The usage of the thinning image in which the ridge flow pattern is 8-commented is encompassed in this technique. In the thinning image, by scanning the neighborhood of every single ridge pixel, the minutiae feature is extracted. CN, which is half the sum of the differences betwixt pairs of adjacent pixels in the input image, is estimated. If the crossing Number is 1, 2 and 3, or else greater than 3, then minutiae points are categorized into Termination, Normal ridge, and Bifurcation, correspondingly. The CN

$C_{Y_{f(thin)}^{seg}}$ is equated as,

$$C_{Y_{f(thin)}^{seg}} = \frac{1}{2}\sum_{j=1}^{8}\left|p^j - p^{j-1}\right| \quad (12)$$

Here, the neighboring pixel of P is depicted by P^j. As per the CN property, the features can be classified as bifurcations and ridge endpoints. The output of minutiae features $\aleph_{f(n)}^{ex}$ can be attained from the whole steps.

3.1.5 Feature Selection

The vital features are chosen from $\aleph_{f(n)}^{ex}$ utilizing LGBUO for decreasing the parameters after extracting the entire features of the FP. The selection of the "best" solution among the set of candidate solutions is involved in LGBUO. By utilizing an objective function, which is to be minimized or else maximized, the solution's degree of goodness is quantified. Generally, the Good, Bad, Ugly Optimizer (GBU) is motivated by the population moving toward the good and bad members. Also, it is centered on the impact of '3' members on the population updates. The restrictions of a random selection of ugly individuals and the absence of exploitation are possessed by the prevailing optimization technique and this engenders slow

convergence. In updating stage, Levy Flight is deployed to conquer these restrictions. A random walk-centered process that is characterized by a series of instantaneous jumps selected as of a probability density function that has a power law tail is called Levy flight. When weighed against the Brownian random walks, it is more effective in exploring unknown, large-scale search spaces. The Levy flight is employed to substitute the local searching for the optimal ugly solution, which engenders larger matching as well as optimization iterations in the global search process. The Levy flight provides better random solutions, where the step lengths have a stable distribution (a heavy-tailed probability distribution). When defining a random walk towards the current ugly member in a space of dimension more than one, the steps made are automatically in isotropic random directions for producing a new solution for each bat within the search space. So, the random solution does not go out from the search space. So, the new solutions for this selection of FP features are obtained accurately within the search space. This significantly eliminates the uncertainty caused by randomness and reduces the exploration complexity. Therefore, the iterations are reduced with higher selection accuracy. Hence, the modified existing GBU is known as LGBUO.

The population of members is initialized (the extracted features $\aleph^{ex}_{f(n)}$ are considered as a group of members) in this technique. It is represented in the matrix formation as,

$$\aleph^{ex}_{f(n)} = [\aleph^{ex}_{f(1)}, \aleph^{ex}_{f(2)}, ... \aleph^{ex}_{f(N)}]_{p\times q} \tag{13}$$

Where, the member's array representation is indicated by $p \times q$, and the number of populace members is depicted by N. After that, centered on the classification accuracy, the fitness is assessed for the populace's every member so that the value for optimizing their member is expressed by every member. The objective function O_n value is represented as a matrix representation is,

$$O_n = \begin{bmatrix} O_1\aleph^{ex}_{f(1)} \\ \vdots \\ O_i\aleph^{ex}_{f(i)} \\ \vdots \\ O_N\aleph^{ex}_{f(N)} \end{bmatrix}_{p\times 1}$$

Where, the objective function i^{th} of the populace's member is symbolized by O_i, the value objective function of the i^{th} member is expressed by $O_i\aleph^{ex}_{f(i)}$.

The best quasi-optimal solution is offered by the member with the minimum objective function together with the worst quasi-optimal solution is offered by the maximum objective function. As per the population's '3' particular members say good, bad, along with ugly, the population is updated in GBUO. In accordance with the objective function's value, the good is the population member that is the top quasi-optimal solution; in addition, the bad is the populace member that has exhibited the worst quasi-optimal key. A population member that directs the populace to situations in a contradictory direction is Ugly. After that, for the '3' members, the quasi-optimal solutions are articulated as,

$$M_g = \aleph_g | O_g \tag{15}$$

$$M_b = \aleph_b | O_b \tag{16}$$

$$M_u = \begin{Bmatrix} 1 & : \nabla < 1 \\ \nabla^{-\tau} & : \nabla \geq 1 \end{Bmatrix} \tag{17}$$

Where, the good member is M_g, the bad member is M_b; the ugly member is expressed by M_u. Centered on the minimum objective function O_g value, the good member is chosen. Centered on the maximum objective function O_b value, the bad member is chosen. The member's distributed step size is ∇, and the dimension of the member is τ. At last, the population member's position is updated in '3' stages. The population with the minimum fitness value that moves towards the good member is updated after updating the member's position. It is delineated as,

$$\aleph^{ex(D)}_{f(i)g(new)} = \aleph^{ex(D)}_{f(i)} + r(M_g{}^D - 2\times\aleph^{ex(D)}_{f(i)}) \tag{18}$$

$$\aleph_i = \begin{cases} \aleph_{ig(new)} & O_i^{g(new)} \leq O_i \\ \aleph_i & else \end{cases} \tag{19}$$

Here, the fresh value for the D^{th} variable of i^{th} member updated grounded on the good member is signified by $\aleph^{ex(D)}_{f(i)g(new)}$, the novel status of i^{th} member updated centered on the good member is delineated by $\aleph_{ig(new)}$, the value for D^{th} variable exemplified by i^{th} member is mentioned by $\aleph^{ex(D)}_{f(i)}$, the random value is indicated by r, the objective function's corresponding is denoted by $O_i^{g(new)}$. The population moves away as of a bad member.

$$\aleph^{ex(D)}_{f(i)b(new)} = \aleph^{ex(D)}_{f(i)} + r(2\times\aleph^{ex(D)}_{f(i)} - M_b{}^D) \tag{20}$$

$$\aleph_i = \begin{cases} \aleph_{ib(new)} & O_i^{b(new)} \leq O_i \\ \aleph_i & else \end{cases} \tag{21}$$

Here, the novel value for the D^{th} variable of i^{th} member updated centered on the bad member is depicted by $\aleph_{f(i)b(new)}^{ex(D)}$, the fresh status of i^{th} member updated grounded on the bad member is signified by $\aleph_{ib(new)}$, along with the objective's equivalent value is indicated by $O_i^{b(new)}$. In contrast to the population's movement, the ugly member causes the population to the position.

$$\aleph_{f(i)u(new)}^{ex(D)} = \aleph_{f(i)}^{ex(D)} + \frac{1}{5} * r(M_u{}^D - \aleph_{f(i)}^{ex(D)})\delta(O_u - O_i) \tag{22}$$

$$\aleph_i = \begin{cases} \aleph_{iu(new)} & O_i^{u(new)} \le O_i \\ \aleph_i & else \end{cases} \tag{23}$$

Here, the novel value for the D^{th} variable of i^{th} member updated grounded on the ugly member is signified by $\aleph_{f(i)u(new)}^{ex(D)}$, the fresh status of i^{th} member updated centered on the ugly member is depicted by $\aleph_{iu(new)}$, the related value's objective function is denoted by $O_i^{u(new)}$, the sign function is indicated by δ; the ugly member's objective function is expounded by O_u. Till the stopping criteria are met, that process is repeated. The obtained selected features are initialized from the entire steps.

$$\aleph_{f(n)}^{sel} = \{\aleph_{f(1)}^{sel}, \aleph_{f(2)}^{sel}, \ldots\ldots\ldots \aleph_{f(N)}^{sel}\} \tag{24}$$

Here, the number of selected features of the FP $I_{seg(n)}$ is depicted by $\aleph_{f(n)}^{sel}$. In Figure 2, the LGBUO's pseudo-code is delineated.

Input: Extracted features $\aleph_{f(n)}^{ex}$
Output: Selected features $\aleph_{f(n)}^{sel}$
Begin
  Initialize population $\aleph_{f(n)}^{ex}$, Maximum number of iteration $t_{\max}$
  Calculate objective function
    Set $t = 0$
    While $(t < t_{\max})$ do
      Update M_g, M_b, M_u
      For $i = 1$ /////number of populations
        Update $\aleph_i$ based on the M_g
        Update $\aleph_i$ based on the M_b
        Update $\aleph_i$ based on the M_u
      End for
      Save the best quasi-optimal solution
    End while
  Return selected features $\aleph_{f(n)}^{sel}$
End

Fig 2: Pseudo code of the LGBUO algorithm

3.2 Iris Image

Here, the iris image that is gathered from the openly accessible data is regarded as the input. The iris image is injected into the pre-processing in which similar steps done for the FP image are repeated for the iris image; in addition, for further processing, the pre-processed image Y_I^{pre} can be achieved.

3.2.1 Segmentation

By employing KSCM, the iris region is completely segmented from the pre-processed image Y_I^{pre}. A methodology, which could resolve a huge range of segmentation issues, is termed the snake model. Detecting and outlining the target object for iris region segmentation is the goal. The key steps of detecting and outlining the target are to initialize the boundary curve, the contour move, and the contour stops moving. However, while the snake curve flow is initiated at a long distance as of the minimal, there was a poor convergence performance of the contour meant for concave boundaries in the prevailing snake model. Logistic kernel value is multiplied for internal as well as external energy calculation for surpassing those demerits

For minimizing the energy function, the Snake propagates via the image area along with targets to find the iris image's edge. The snake S is equated as,

$$S_{Y_I^{pre}} = (g(\rho), h(\rho)) \tag{25}$$

Here, the coordinates of the 2-dimensional curve in the range of $[0,1]$ are signified by $(g(\rho), h(\rho))$. For maintaining the complete energy at the maximal level, the snake model moves with the influence of internal along with external energy.

$$\xi^{tot} = \xi^{\text{int}} + \xi^{ext} + \xi^{image} \tag{26}$$

Here, the total energy is depicted by ξ^{tot}, the internal energy is delineated by ξ^{int}, the external energy is mentioned by ξ^{ext}, and the image's energy under consideration is expounded by ξ^{image}. The piecewise smoothness in the curve is signified by the internal energy.

$$\xi^{\text{int}} = \left(\psi \left| \frac{\partial S}{\partial \rho} \right|^2 + \Omega \left| \frac{\partial^2 S}{\partial \rho^2} \right| \right) * K \tag{27}$$

Here, the degree of the resistance to stretching together with the bending of the contour is depicted by ψ and Ω. The external energy ξ^{ext} is typically defined such that the contour seeks the edges in the image Y_I^{pre}

$$\xi^{ext} = \left(\int_0^1 f((g(\rho), h(\rho))dS \right) * K \tag{28}$$

Here, the logistic kernels are indicated by K.

$$K = e^{\frac{-\|\varpi_1 - \varpi_2\|^2}{2\varphi^2}} \quad (29)$$

Here, the two vectors of the iris image are denoted by ϖ_1 and ϖ_2, and the sigma value is depicted by φ. The output Y_I^{seg} can be obtained from the contour-based.

3.2.2 Feature Extraction

One of the significant steps in the authentication of BS is FE. Moreover, it is the process of extracting Iris features from Y_I^{seg} to be wielded in selection and classification tasks. The iris's structure is more associated with the geometric shape; thus, the extraction of these features is also possible. The geometric feature of the iris plays a prominent role in biometric authentication because each human has a different iris structure. The geometric features provide more relevant information regarding the iris for the proposed model.

Thus, pupil-centric, Collarette-centric, and Iris-centric features, which are essential, are extracted. Pupil is defined to be circular or else round; yet, they are not really a circle. Thus, for identifying the pupil, the circle's geometrical properties could be utilized; also, these features could be wielded for recognition. Also, the Collarette and iris have different geometrical properties to represent the differentiation between humans. Therefore, FE is important to be considered here. Using these extracted geometrical features, the classifier has been trained and reduced the complexity by considering important features.

Pupil-based Feature: Here, from the iris image, the circle shape-centric features are extracted. Pupil Roundness, Largeness, along with Smoothness is the public-centric extraction. The gauge of sharpness at the corners and edges of a circle that is associated with the circle's sphericity and compactness is termed roundness. Grounded on the pupil's diameter, the pupil's diameter is estimated. Centered on the pupil's '8' angles, the diameter is computed. The pupil's smoothness is grounded on the pupil's curvature. Grounded on the diameter at different orientations, the roundness feature is computed in which the ratio betwixt the maximum and the minimum value is computed at horizontal, vertical, left, and right at 0, 45, 90 and 135°, correspondingly. Centered on the pupil's radius, the pupil's largeness is described. The iris's smoothness could be specified based on the curvature of the circle. Thus, the extracted features are equated as,

$$R_p = \frac{M(\ell_h, \ell_v, \ell_l, \ell_r)}{m(\ell_h, \ell_v, \ell_l, \ell_r)} \quad (30)$$

$$L_p = \frac{Y_I^{pre}(0^\circ_{\ell P} + 30^\circ_{\ell P} + 45^\circ_{\ell P} + 75^\circ_{\ell P} + 90^\circ_{\ell P} + 120^\circ_{\ell P} + 135^\circ_{\ell P} + 165^\circ_{\ell P})}{8} \quad (31)$$

Here, the FE based on the roundness is depicted by R_P, the maximum along with minimum value calculated in the entire angle are signified by $M(\bullet)$ and $m(\bullet)$, the extracted feature based on the largeness is delineated by L_P, the diameter of the pupil is mentioned by ℓP, the extracted feature based on the smoothness operation is indicated by P_s, the curvature points is depicted by C_z, the curvature means is specified by C_m, the number of curvature points is signified by Z

$$p_s = \frac{1}{Z}\sum_{z=1}^{z} C_z - C_m \quad (32)$$

Here, the FEbased on the roundness is depicted by R_P, the maximum along with minimum value calculated in the entire angle are signified by $M(\bullet)$ and $m(\bullet)$, the extracted feature based on the largeness is delineated by L_P, the diameter of the pupil is mentioned by ℓP, the extracted feature based on the smoothness operation is indicated by P_s, the curvature points is depicted by C_z, the curvature means is specified by C_m, the number of curvature points is signified by Z

Collarette-based Feature: In Collarette Roundness, Iris Ratio, along with Pupil Ratio steps, the features associated with the Collarette part that is concentric with the pupil are extracted. Grounded on the maximal along with minimal values of the collaret's '6' diameters, the collaret's roundness feature is extracted. Centered on the iris's diameter, the collaret iris ratio is calculated. The collaret iris ratio is calculated by summarizing the ratio of the collaret's diameter and the iris's diameter for all the angles. The iris's diameter is approximately 10-11 mm. The collaret, which is located about 1.5 mm away from the pupil, occupies one-third of the iris. Thus, at '8' points, the distance betwixt the collaret and iris diameter is measured. Regarding the polar coordinates, the iris is divided into equivalent spaces. The distance, as well as the collaret ratio, is estimated for every intersection point on the iris edge. Here, the distance between two intersection points is calculated based on the square root of the sum of the squares of the coordinates. Regarding the collaret and pupil, the collaret Pupil Ratio is estimated; in addition, equated similarly to Collarette Iris Ratio. The extracted features are equated as,

$$R_c = \frac{M(0^{\circ}\ell c, 30^{\circ}\ell c, 45^{\circ}\ell c, 75^{\circ}\ell c, 90^{\circ}\ell c, 120^{\circ}\ell c, 135^{\circ}\ell c, 165^{\circ}\ell c)}{m(0^{\circ}\ell c, 30^{\circ}\ell c, 45^{\circ}\ell c, 75^{\circ}\ell c, 90^{\circ}\ell c, 120^{\circ}\ell c, 135^{\circ}\ell c, 165^{\circ}\ell c)} \tag{33}$$

$$I_c = \frac{1}{8}\left(\frac{0^{\circ}\ell c}{0^{\circ}\ell I} + \frac{30^{\circ}\ell c}{30^{\circ}\ell I} + \frac{45^{\circ}\ell c}{45^{\circ}\ell I} + \frac{75^{\circ}\ell c}{75^{\circ}\ell I} + \frac{90^{\circ}\ell c}{90^{\circ}\ell I} + \frac{120^{\circ}\ell c}{120^{\circ}\ell I} + \frac{135^{\circ}\ell c}{135^{\circ}\ell I}\right) \tag{34}$$

$$P_c = \frac{1}{8}\left(\frac{0^{\circ}\ell c}{0^{\circ}\ell P} + \frac{30^{\circ}\ell c}{30^{\circ}\ell P} + \frac{45^{\circ}\ell c}{45^{\circ}\ell P} + \frac{75^{\circ}\ell c}{75^{\circ}\ell P} + \frac{90^{\circ}\ell c}{90^{\circ}\ell P} + \frac{120^{\circ}\ell c}{120^{\circ}\ell P} + \frac{135^{\circ}\ell c}{135^{\circ}\ell P}\right) \tag{35}$$

Here, the extracted features centered on the Collarette Roundness are depicted by R_c, the output of the Collarette Iris Ratio is signified by I_c, the output of Collarette Pupil Ratio is denoted by P_c, the collaret's diameter is mentioned by ℓc and the iris's diameter is indicated by ℓI.

Iris-based Feature: From the segmented image Y_I^{seg}, the Iris-centric features are extracted. Grounded on Iris's roundness as well as diameter, the features are extracted. The iris' roundness feature is computed similarly to the pupil. Centered on the average diameter at '8' points on the iris, the diameter-centric feature is appraised. The iris is not a definite circle; it is the usual distance diameter calculation as of the center to any point on the iris's outer circle isn't adequate.

$$R_I = \rho_{i\max} - \rho_{i\min} \tag{36}$$

$$d_I = \frac{(0^{\circ}\ell P + 30^{\circ}\ell P + 45^{\circ}\ell P + 75^{\circ}\ell P + 90^{\circ}\ell P + 120^{\circ}\ell P + 135^{\circ}\ell P + 165^{\circ}\ell P)}{8} \tag{37}$$

Here, the iris roundness feature is depicted by R_I, the Iris diameter-based feature is signified by d_I, and $\rho_{i\max}$ $\rho_{i\min}$ implies maximal along with minimal measured trace-points. The output of the extracted features Y_I^{ex} can be attained via the complete extraction process. Without losing the image information, the significant features are chosen. As same as the FP FS phase, the selection process is done by employing LGBUO. The output of the selected features is equated as,

$$\aleph_{I(n)}^{sel} = \{\aleph_{I(1)}^{sel}, \aleph_{I(2)}^{sel}, \ldots\ldots\ldots \aleph_{I(N)}^{sel}\} \tag{38}$$

Here, the number of selected Iris features is depicted by $\aleph_{I(n)}^{sel}$.

3.3 Feature Level fusion and Normalization

By deploying this, the feature is fused and normalized. By concatenating both features, the feature sets initiated from the FP and iris are consolidated into a single feature. Feature fusion attempts in extracting the most discriminative information as of the iris and FP, and eliminate redundant information. The stability along with reliability is enhanced by the combination of FP and iris features. The selected features' concatenation is equated as,

$$F_{Y_{fII(n)}^{sel}} = \{(\aleph_{f(n)}^{sel}), (\aleph_{I(n)}^{sel})\} \tag{39}$$

Here, the output of the concatenated features is depicted by $F_{Y_{fII(n)}^{sel}}$. By utilizing the min-max normalization that normalizes fused features within the range, the features are normalized.

$$\aleph_{fI(n)}^{nor} = \frac{F_{Y_{fII(n)}^{sel}} - \min(F_{Y_{fII(n)}^{sel}})}{\max(F_{Y_{fII(n)}^{sel}}) - \min(F_{Y_{fII(n)}^{sel}})} \tag{40}$$

Here, the output of the normalized features is indicated by $\aleph_{fI(n)}^{nor}$

3.4 Classification

By employing the FRMSDNET, the authentication can be done grounded on the normalized features $\aleph_{fI(n)}^{nor}$.Among the number of neural networks, the efficiency of deep neural networks is enhanced with more neural layers by the ResNet, which also minimizes the percentage of errors. In other words, the outputs from prior layers are added to the stacked layers' outputs by the skip connections, thus making it possible to train much deeper networks than previously possible. Therefore, it provides an accurate recognition outcome with a higher security level. One of the famous DL models, which consist of 18 convolution layer (1×1) connected with the standalone convolution layer is termed the Residual Network (ResNet). ResNet, which is a popular tool in the field of DL, is selected, because of its easier acceptability, high-security level, along with recognition accuracy than other prevailing systems. Their robustness is owing to their flexible architecture along with their capability in extracting features as of raw data. They learn from data to perform a task without being explicitly programmed, whereas non-machine learning models rely on the judgment of an expert, not a system that learns. Hence, the proposed machine learning model is optimum for successfully used in image classification. The convolution layer is also termed a residual block. Each residual block is subsequent to a Batch Normalization (BN) together with a Rectified Linear Unit (ReLu) activation function. To the activation

function called skip connection, the input is directly added. Since the gradient is back-propagated to previous layers and the repeated multiplication might make the gradient infinitely smaller, the disadvantages are harder for training owing to the vanishing gradient issue in generalResnet. The performance gets saturated or else even starts debasing quickly since the network goes deeper. For estimating the Loss Function in Neural Network, Mean Squared Deviation is employed. For updating the weights and bias, the Fuzzy layer in which several rules are generated is utilized. Thus, a considerable feature's learning process is developed. In Figure 3, the architecture of the FRMSDNET is depicted.

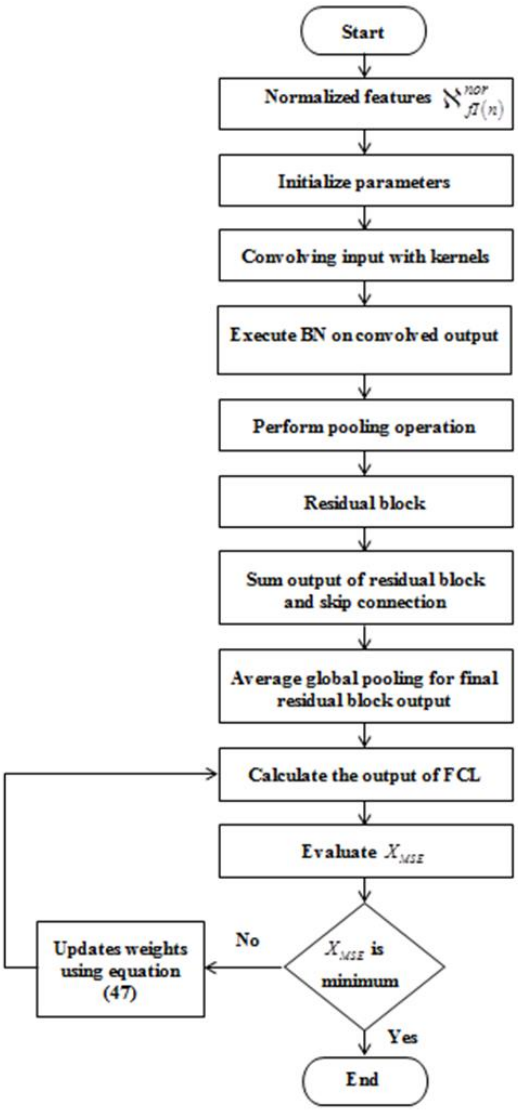


Fig 3: Architecture of FRMSDNET

The normalized features are represented in the matrix representation and the weight values ω_n are initialized. The kernel or Filter is the matrix of weights. The selected features' outputs are given to the convolution layer. By employing the kernels, the input is convolving. Then, the dot product has been calculated between them and updated to the output array.

The normalized features are represented in the matrix representation and the weight values ω_n are initialized. The kernel or Filter is the matrix of weights. The selected features' outputs are given to the convolution layer. By employing the kernels, the input is convolving. Then, the dot product has been calculated between them and updated to the output array. The

updated output can be equated as,

$$C(Y_{fI(n)}^{nor},\omega)=\sum_{d}\sum_{d}Y_{fI(n)}^{nor}(u-d,v-d)*\omega(d,d) \tag{41}$$

Here, the convolution operation's output is depicted by $C(\bullet)$, the kernels in the matrix are signified by $\omega(d,d)$, and the matrix representation of normalized features is indicated by u and v. BN is connected with the convolution layer's output for faster training and providing nonlinearity in the neural network. The gradient is also developed by the function. The BN

$B_{C(Y_{fI(n)}^{sel},\omega)}$ can be equated as,

$$B_{C(Y_{fI(n)}^{sel},\omega)}=\frac{C(Y_{fI(n)}^{nor},\omega)-E(C(Y_{fI(n)}^{nor},\omega))}{\sqrt{\text{var}\,C(Y_{fI(n)}^{nor},\omega)}} \tag{42}$$

Here, the mean value of $C(Y_{fI(n)}^{nor},\omega)^k$ is signified by $E(\bullet)$, and the variance of $C(Y_{fI(n)}^{nor},\omega)^k$ is depicted by $\text{var}\,C(\bullet)$. A pooling operation is performed to reduce the input image's parameters. The way of reduction is termed down-sampling. For detecting the maximum pixel value, the kernel matrix is mapped over the input array. The maximum value is updated to output and the pooling $\Re$ can be formulated as,

$$\Re=\frac{B_{C(Y_{fI(n)}^{sel},\omega)}-\omega}{S}+1 \tag{43}$$

Here, the strides of the kernel are depicted by S

Residual Block: The pooling layer's output is directly fed into the 1st residual block. Then, the convolution and BN are performed. The output is directly injected into the activation function of ReLu, which aids to study the intricate patterns in the data along with the residual block's output.

$$h(\Re)=f(\Re) \tag{44}$$

The output is modified to the below equation by introducing a skip connection.

$$h(\Re) = f(\Re) + \Re \tag{45}$$

By permitting the alternate shortcut path for the gradient to flow via, the skip connections resolve the vanishing gradient in deep NN. Till the final residual block, that process is continued. For generating one feature map for every related category, the output from the final residual block is directly fed into the average global pooling. Every feature map's average is considered along with the resulting vector and is fed directly into the FCL in which every input as of one layer is connected to the next layer's each activation unit. For classifying the input, the FCL is familiar with the softmax $\varepsilon(\bullet)$ activation function.

$$\varepsilon(h(\Re)) = \frac{e^{h'(\Re)}}{\sum_{r=1}^{\ell} e^{h'(\Re)}} \tag{46}$$

Here, the FCL's output is depicted by $h'(\Re)$.

In the end, the mean square error is computed, which determines

the current output's variations from the expected output. The mean square error is equated as,

$$X_{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left[\varepsilon\left(h(\Re)\right)_i - \varepsilon\left(h(\Re)\right)\right]^2 \tag{47}$$

Here, the current output is signified by $\varepsilon(h(\Re))_i$, and the expected output is depicted by $\bar{X}_i\,\varepsilon(h(\Re))$. Owing to the high variation of current and targeted output, the weight values should be generated using the fuzzy rules while the loss function is maximal. Fuzzy rules are generated around the Training data with huge estimation errors or else at the points where the Training Data group until the error limit is within the specified limit. By accumulating the simulation performance results of possible weights, the fuzzy rules are generated along with it operates on an 'if-then' principle where the 'if' is a vector of fuzzy premises; in addition, 'then' is a vector of fuzzy consequences. The example fuzzy rule equated as,

$$IF\ x\ is\ A\ THEN\ y\ is\ B \tag{48}$$

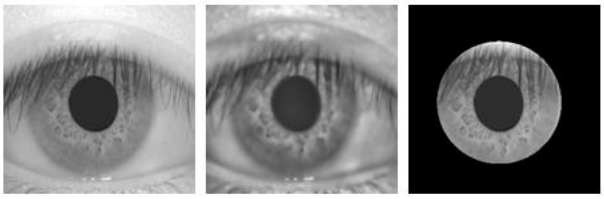
Similarly, grounded on the feature values, the fuzzy rules are generated and the weight values are updated. In the end, the output layer delivers its output as an authorized user or unauthorized user.

4. Results and Conclusion

The analysis of the proposed FRMSDNET by comparing its results with other existing models is done in this section. In the working platform of MATLAB, the proposed technique is employed. From CASIA-IrisV3 and Sokoto Coventry FP Dataset (SOCOFing), the data are amassed. The performance analysis and comparative analysis were carried out.

4.1 Dataset Description

3 subsets are encompassed in CASIA-IrisV3. They are labeled as CASIA-IrisV3-Interval, CASIA-IrisV3-Lamp, along with CASIA-IrisV3-Twins. A total of 22,035 iris images as of more when weighed against 700 subjects are comprised in it. Under near-infrared illumination, all iris images that are 8-bit grey-level JPEG files are amassed. A biometric FP database that is designed for academic research is SOCOFing. 6,000 FP images as of 600 African subjects together with unique attributes like labels for gender, hand, finger name, along with synthetically modified versions with '3' different levels of alteration for obliteration, central rotation, along with z-cut are encompassed in it. Figures 4 and 5 portray the sample input and output images of the iris.



(a) (b) (c)

Fig 4: Sample input and output image of the iris, (a) input image, (b) pre-processed image, and (c) segmented image

(a) (b) (c)

Fig 5: Sample input and output image of the fingerprint, (a) input image, (b) pre-processed image, and (c) segmented image

In Figure 4, (a) depicts the sample input image of the iris, (b) portrays the pre-processed image utilizing CLVAHE, and (c) showcases the segmented image utilizing KSCM. In Figure 5, (a) depicts the input image of the FP, (b) illustrates the pre-processed image utilizing CLVAHE, and (c) portrays the segmented image utilizing AMOS.

4.2 Performance Analysis of Classification

Regarding specificity, precision, f-measure False Positive Rate (FPR), sensitivity, accuracy, recall, Negative Prediction Value (NPV), False Negative Rate (FNR), Matthews Correlation Coefficient (MCC), along with False Rejection Rate (FRR), the performance analysis of the FRMSDNET is validated. The comparison of outcomes acquired from these is carried out with the prevailing models, namely Deep NN (DNN), Artificial NN (ANN), and Adaptive Neuro-Fuzzy Inference System (ANFIS), together with Recurrent NN (RNN). At last, the proposed system's Receiver Operating Characteristic (ROC) is also acquired.

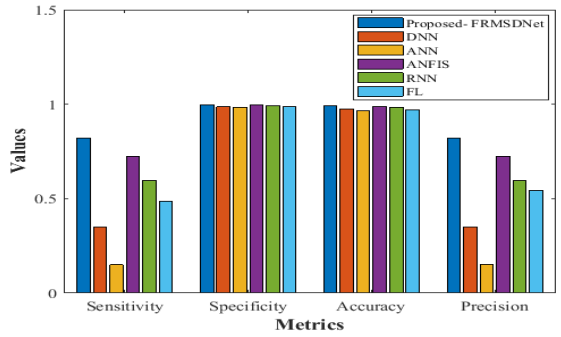


Fig 6: Performance analysis of the proposed FRMSDNET and the existing models in terms of sensitivity, specificity, accuracy, and precision

Figure 6 portrays the performance assessment of the proposed FRMSDNET and the conventional frameworks. The percentage of actual positive value is gauged by sensitivity, which is correctly identified. The proposed model's superior performance is exhibited by a higher sensitivity. A sensitivity of 0.82 is achieved by the proposed model, which is greater when analogized to the prevailing systems that accomplished sensitivity in the range of 0.15 to 0.725. Likewise, the proposed model's performance achieves greater specificity, accuracy, and precision values. When analogized to the conventional models, the proposed mechanism attains a specificity of 0.9964, accuracy of 0.99294, and precision of 0.82, which are higher. The proposed system's performance is attained by proper weight and bias updation for each neuron based on evaluating the loss function using Mean Squared Deviation. If the target and actual output shows higher variation, then the feature value-based fuzzy rules are generated for updating the weight and bias values. This significantly aids in decreasing the system complexity and provides better recognition outcomes.

Table 1: A comparative analysis of the proposed model and existing model in terms of recall, f- measure, NPV, and FPR

Techniques	***Recall***	***F-measure***	***NPV***	***FPR***
Proposed FRMSDNet	0.82	0.82	0.9964	0.0036
DNN	0.35	0.35	0.987	0.013
ANN	0.15	0.15	0.983	0.017
ANFIS	0.725	0.725	0.9945	0.0055
RNN	0.595	0.595	0.9919	0.0081
FL	0.5369	0.5698	0.984	0.0097

Regarding recall, f-measure, NPV, along with FPR, table 1 exhibits the comparative evaluation of the proposed system and the existing model. The proposed system's enhanced performance is displayed by higher recall along with f-measure values. The proposed model has the same recall as well as f-measure. It exhibits an improvement of 0.47 to DNN, 0.67 to ANN, 0.095 to ANFIS, 0.225 to RNN, and 0.5369 to FL. Likewise, when analogized to DNN, ANN, ANFIS, RNN, and FL the proposed system's improvement in NPV is 0.0094, 0.0134, 0.0019, 0.0045, and 0.0124 respectively. Similarly, when contrasted with the prevailing techniques, the FPR of the proposed approach is 0.0036, which is low. In FRMSDNET, the neurons remain active for a maximum number of iterations, thus it significantly learns the inputs. The proposed system's better performance is displayed by the lower FPR value.

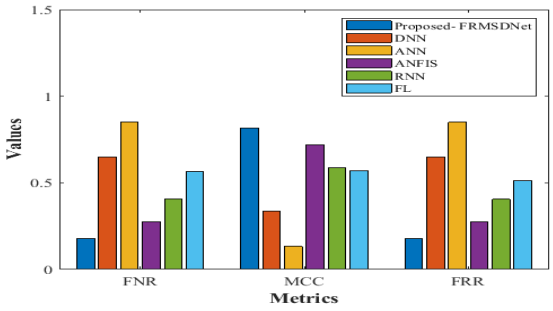


Fig 7: Performance analysis of proposed and existing models in terms of FNR, MCC, and FRR

Regarding FNR, MCC, as well as FRR, the appraisal of the proposed and prevailing systems is illustrated in Figure 7. Better performance is brought about by lower FNR and FRR values. The proposed model has the same FNR and FRR (0.18), which is lower analogized to the prevailing models; while the existing models attain higher FNR and FRR in the range of 0.65 to 0.405. Likewise, better performance is achieved by higher MCC. When analogized

to prevailing models, the proposed model achieves a higher MCC of 0.8164. Hence, it is inferred from the metrics that the proposed framework exhibits better performance for multimodal biometric authentication. This is because the fusion of FP along with iris traits has the benefit of integrating information extracted from different biometrics and identifying the compact set of features that enhances the stability as well as reliability of the recognition performance.

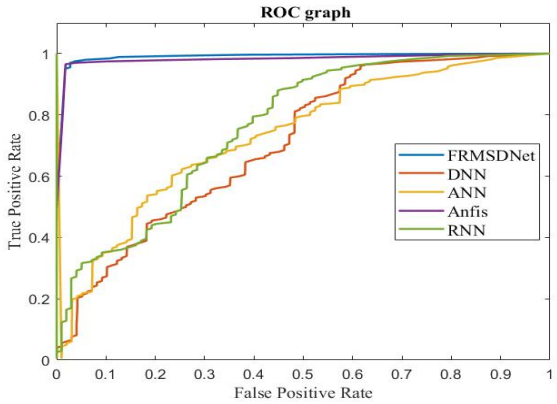


Fig 8: ROC curve of proposed FRMSDNET and existing model

In Figure 8, the proposed FRMSDNET and the prevailing system's ROC curve are delineated. An effectual technique of analyzing the performance quality of diagnostic tests is termed ROC. By plotting the True Positive Rate (TPR) against the FPR, a ROC curve is engendered. The ROC curve is a measure of recognition accuracy for the particular features that gave rise to that curve. This measure can range in value from 0.5 to 1.0. The model with a higher ROC achieves even higher accuracy while simultaneously achieving a higher TPR and lower FPR. Thus, the area under the ROC curve is higher as the decision-making accuracy increases. When weighed against the prevailing DNN, ANN, ANFIS, and RNN, the sensitivity and specificity of the FRMSDNET are superior.

Table 2: Recognition Rates of the proposed model compared to unimodal systems

Bio-Metric Systems	***Precision***	***Recall***	***FPR***	***FRR***
Iris (unimodal)	0.81	0.81	0.0038	0.19
Fingerprint (unimodal)	0.735	0.735	0.0053	0.265
Iris&Fingerprint (Multimodal)	0.5435	0.5369	0.0097	0.513

The performance outcomes of feature-level fused systems in comparison with unimodal systems of Iris and FP are displayed in Table 2. By enhancing the recognition accuracy, strong and stable performance is exhibited by the fused feature descriptor. The outcomes exposed that the recognition rate is mounted by the fusion of these two biometric traits rather than utilizing a single biometric trait in any authentication system. Thus, the Feature level fusion could derive the discriminatory information; also, it encloses the richest information as of multiple features. Thus, for the final recognition, it has a considerable benefit.

4.3 Performance Analysis of Pre-Processing

Regarding Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), together with Structural Similarity Index (SSIM), the proposed CLVAHE performance evaluation is done. The outcomes are analogized to the current CLAHE, HE, and Contrast Stretching (CS).

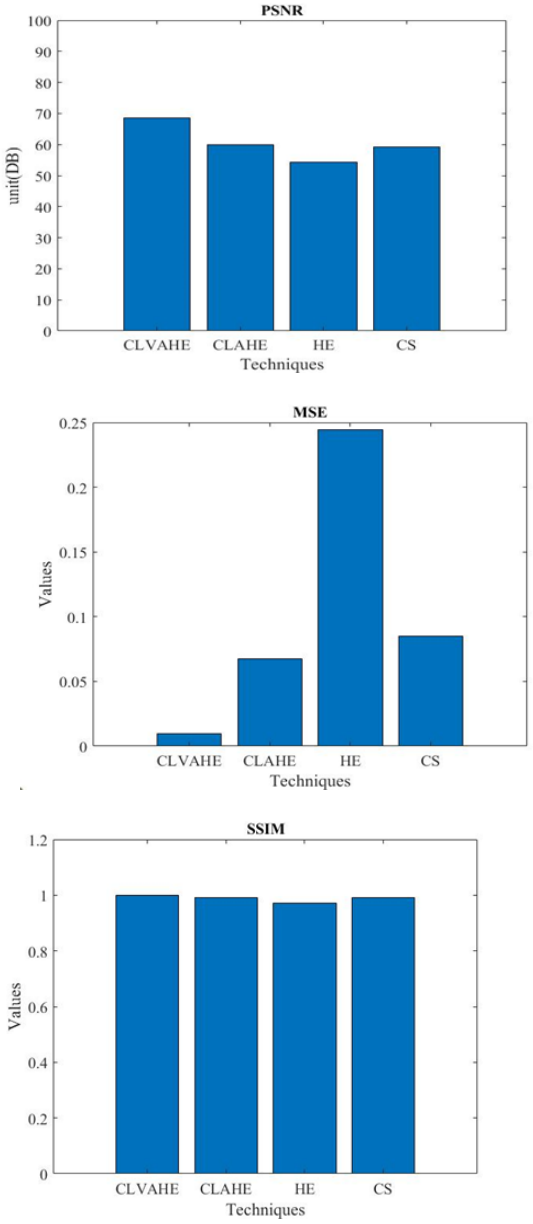


Figure 9: Performance analysis of proposed model and existing model, (a) shows in terms of Peak Signal to Noise Ratio(PSNR), (b) shows in terms of Mean Square Error(MSE), and (c) shows in terms of Structural Similarity Index(SSIM)

The performance evaluation of the proposed along with the current techniques is depicted in Figure 9. To measure the image's quality, PSNR is wielded. The system's efficacy will be proved by the high PSNR value. For PSNR, the proposed system attains 68.519db; while the prevailing CLAHE, HE, and CS achieved 60.011db, 54.426db, and 59. 318db.The system's performance is high if the MSE value is low. The proposed system's MSE is 0.0094, which is low when analogized to the prevailing CLAHE, HE, and CS. 0.9989 is the SSIM of the proposed technique; whereas, the current methodology attained in the range of 0.992 to 0. 9906. The proposed system's efficient performance in the pre-processing stage is attained by setting the appropriate clip limit based on the overall intensity information using a local variance. Therefore, inducing the best clip limit in the existing CLAHE increases and provides a higher-clarity image. When weighed against the prevailing system, the proposed technique attained superior performance.

4.4 Performance analysis for the CASIA dataset

Here, the performance of the proposed FRMSDNET is contrasted with the Multiclass Support Vector Machine (MSVM), Support Vector Machine (SVM), along with the Optimized Source Classifier (OSC).

Table 3: Comparative analysis of the proposed method with the existing method

Methods	***Accuracy***
proposed FRMSDNET	99.4%
MSVM[27]	89%
SVM [19]	96.6%
OSC[30]	98.43%

Table 3 illustrates the proposed technique's performance with the prevailing approaches regarding accuracy. When analogized with the conventional techniques, better accuracy (99.4%) was achieved by the proposed one. However, the conventional techniques like MSVM and SVM attained 89% and 96.6%, correspondingly. The outcomes prove that the authentication performance could be enhanced by the multimodal fusion of FP and iris biometrics. For protection enrichment in phrases of high overall performance along with the inherent aliveness property of biometrics, the proposed system is beneficial. The proposed technique aims to design a multi-biometric system, which could incorporate various fusion strategies, such as FE, feature selection, feature fusion using max-min, and decision level using FRMSDNET under a dynamic system for meeting varying security demands. The presented model lies in the fact that it engages 2 biometric traits concurrently with efficient segmentation and feature selection steps in the domain of feature-level fusion, for testing the system's efficacy in authenticating identities. Hence, it is concluded that the proposed one outperforms the prevailing frameworks.

5. Conclusion

An efficient Multimodal Biometric Authentication (MBA) utilizing a novel Fuzzy Residual Mean Squared Deviation Network(FRMSDNET) classifier is proposed in this paper. The biometric traits, namely iris, and FPs that are acquired from publically available sources are combined by the proposed model. Here, by utilizing feature-level fusion, the features are extracted and fused. Next, by employing FRMSDNET, the classification was carried out. Then, with respect to diverse metrics, the experimental analysis is done where the proposed technique's performance and comparative analyses were performed. The proposed system accomplishes an accuracy of 0.99294 as per the final finding. Likewise, for all other metrics, namely specificity, precision, f-measure, FPR, sensitivity, accuracy, recall, NPV, FNR, MCC, along with FRR, the model accomplishes superior performance. Hence, when weighed against other prevailing models, it is deduced that the proposed system exhibits better performance. With a few advanced methods for security purposes, the work could be expanded in the future.

References

[1] Arafat Rahman, Muhammad E H Chowdhury, AmithKhandakar, SerkanKiranyaz, KhShahriya Zaman, Mamun Bin IbneReaz, Mohammad Tariqul Islam, MaymounaEzeddinand Muhammad Abdul Kadir, "Multimodal EEG and keystroke dynamics based biometric system using machine learning algorithms", IEEE Access, vol. 9, pp. 94625-94643, 2021.

[2] El-Sayed A El Dahshan, Mahmoud M Bassiouni, SeptaveraSharvia and Abdel Badeeh M Salem, "PCG signals for biometric authentication systems an in-depth review" Computer Science Review, vol. 41,pp. 1-25, 2021.

[3] Po Ya Hsu, Po Han Hsu, Tsung-Han Lee and HsinLi Liu, "Motion artefact resilient scg-based biometric authentication using machine learning" 43rdAnnual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 01-05 November 2021, Mexico, 2021.

[4] Shaymaa M Hamandi,AbdulMonem S Rahma and Rehab F Hassan, "Design a multi biometric system for safe access to buildings", Al Qadisiyah Journal of

Pure Science, vol. 26, no. 4, pp. 275-287, 2021.

[5] Swati KChoudhary and AmeyaKNaik, "Multimodal biometric authentication with secured templates a review", 3rdInternational Conference on Trends in Electronics and Informatics, 23-25 April 2019, Tirunelveli, India, 2019.

[6] Arjun BenagatteChannegowda and H N Prakash, "Multimodal biometrics of fingerprint and signature recognition using multi-level feature fusion and deep learning techniques", Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 1, pp. 185-197, 2021.

[7] Venkatesh R, Uma MaheswariNandJeyanthi S, "Multiple criteria decision analysis based overlapped latent fingerprint recognition system using fuzzy sets", International Journal of Fuzzy Systems, vol. 20, no. 2, pp. 1-27, 2018.

[8] Jorge Blasco and Pedro Peris-Lopez, "On the feasibility of low-cost wearable sensors for multi-modal biometric verification", Sensors, vol. 18, no. 9, pp. 1-20, 2018.

[9] ShashankTripathi, Jay Murgi, Kalpana Rai SnehaSoniandRajde, "A literature survey on multi model bio metric system", Journal of Computing Technologies, vol. 10, no. 2, pp. 1-5, 2021.

[10] Mohamed Hammad, Yashu Liu and Kuanquan Wang, "Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint", IEEE Access, vol. 6, pp. 26527-26542, 2018.

[11] Nada Alay and Heyam H Al-Baity, "Deep learning approach for multimodal biometric recognition system based on fusion of iris face and finger vein traits", sensors, vol. 20, no. 19, pp. 1-17, 2020.

[12] Vandana and Navdeepkumar, "A study of biometric and identification and verification system", International Conference on Advance Computing and Innovative Technologies in Engineering, 04-05 March 2021, Greater Noida, India, 2021.

[13] AnterAbozaid, Ayman Haggag, Hany Kasban and Mostafa Eltokhy, "Multimodal biometric scheme for human authentication technique based on voice and face recognition fusion", Multimedia Tools and Applications, vol. 78, pp. 16345–16361, 2019.

[14] Gavisiddappa, ShivakumarMahadevappa, Chandrashekar Mohan Patil, "Multimodal biometric authentication system using modified relieff feature selection and multi support vector machine", International Journal of Intelligent Engineering and Systems, vol. 13, no. 1, pp. 1-12, 2020.

[15] Tariq M Khan, "Hardware implementation of multimodal biometric using fingerprint and iris", 2022,https://doi.org/10.48550/arXiv.2201.05996

[16] Arthi R, ManojkumarD, Aksa Abraham, Allada Rahul Kishan and AlekhyaSattenapalli, "Deep learning based multi-modal biometric security system using visible light communication", WSEAS Transactions on Systems and Control, vol. 17, pp. 34-41, 2022.

[17] BasmaAmmour, LarbiBoubchir, ToufikBoudenandMessaoudRamdani, "Face iris multimodal biometric identification system", Electronics, vol. 8, pp. 1-18, 2020.

[18] Pooja Akulwar and NatarajAVijapur, "Secured multi modal biometric system a review", 3rdInternational Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2019), 12-14 December, Palladam, India, 2019.

[19] Mohammed ChachanYounis and HuthaifaAbuhammad, "A hybrid fusion framework to multi-modal bio metric identification", Multimedia Tools and Applications, vol. 80, pp. 25799–25822, 2021.

[20] FerielCherifi, Kamal Amroun and Mawloud Omar, "Robust multimodal biometric authentication on IoT device through ear shape and arm gesture", Multimedia Tools and Applications, vol. 80, pp. 14807–14827, 2021.

[21] Rasha O Mahmoud, Mazen M. Selim and Omar A. Muhi, "Fusion time reduction of a feature level based multimodal biometric authentication systems", International Journal of Sociotechnology and Knowledge Development, vol. 12, no. 1, pp. 67-83, 2020.

[22] Ajai Kumar Gautam and Rajiv Kapoor, "Multi modal biometric recognition system based on FLSL fusion method and mdlnn classifier", Turkish Journal of Computer and Mathematics Education, vol. 12, no. 12, pp. 241-256, 2021.

[23] SurabhiHom Choudhury, Amioy Kumar and ShahedulHaqueLaskar, "Adaptive management of multimodal biometrics a deep learning and metaheuristic approach", Applied Soft Computing, vol. 26, no. 8, pp. 1-20, 2021.

[24] Mohamed Hammad, Yashu Liu and Kuanquan Wang, "Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint", IEEE Access, vol. 7, pp. 26527-26542, 2018.

[25] Chander Kant and SheetalChaudhari, "A

watermarking based approach for protection of the templates in multimodal biometric system", Procedia Computer Science, vol. 167, pp. 932-941, 2020.

[26] Vijay M and Indumathi G, "Deep belief network-based hybrid model for multimodal biometric system for futuristic security applications", Journal of Information Security and Applications, vol. 58, no. 3, pp. 1-14, 2021.

[27] Sreevidya B and Chandra E, "Entropy based local binary pattern (ELBP) feature extraction technique of multimodal biometrics as defence mechanism for cloud storage", Alexandria Engineering Journal, vol. 58, no. 1, pp. 103-114, 2019.

[28] Chengsheng Yuan, Shengming Jiao, Xingming Sun and Jonathan Wu Q. M, "MFFFLD a multi-modal feature fusion based fingerprint liveness detection", IEEE Transactions on Cognitive and Developmental Systems, vol. 14, no. 2, pp. 648-661, 2022.

[29] Arafat Rahman, Muhammad E. H Chowdhury, AmithKhandakar, SerkanKiranyaz, Shahriya Zaman K. H, Mamun Bin IbneReaz, Mohammad Tariqul Islam, MaymounaEzeddin and Muhammad Abdul Kadir, "Multimodal EEG and keystroke dynamics based biometric system using machine learning algorithms", IEEE Access, vol. 9, pp. 94625-94643, 2021.

[30] Gurjit Singh Walia, Tarandeep Singh, Kuldeep Singh and NeelamVerma, "Robust multimodal biometric system based on optimal score level fusion model", Expert Systems with Applications, vol. 116, pp.364-376, 2018.

[31] VeeruTalreja, Matthew C Valenti and Nasser M Nasrabadi, "Deep hashing for secure multimodal biometrics", IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1306-1321, 2020.

[32] Ayesha Tarannum, Zia Ur Rahman, Koteswara Rao L, Srinivasulu T and Aime Lay Ekuakille, "An efficient multi-modal biometric sensing and authentication framework for distributed applications", IEEE Sensors Journal, vol. 20, no. 24, pp. 15014-15025, 2020.

[33] Onsen Toygar, Felix O Babalola and Yiltan Bitirim, "FYO a novel multimodal vein database with Palmar, dorsal and wrist biometrics", IEEE Access, vol. 8, pp. 82461-82470, 2020.