

Article

A Dual Multimodal Biometric Authentication System Based on WOA-ANN and SSA-DBN Techniques

Sandeep Pratap Singh * and Shamik Tiwari 

School of Computer Science, University of Petroleum & Energy Studies, Dehradun 248007, India

* Correspondence: sandeep102209@gmail.com

Abstract: Identity management describes a problem by providing the authorized owners with safe and simple access to information and solutions for specific identification processes. The shortcomings of the unimodal systems have been addressed by the introduction of multimodal biometric systems. The use of multimodal systems has increased the biometric system's overall recognition rate. A new degree of fusion, known as an intelligent Dual Multimodal Biometric Authentication Scheme, is established in this study. In the proposed work, two multimodal biometric systems are developed by combining three unimodal biometric systems. ECG, sclera, and fingerprint are the unimodal systems selected for this work. The sequential model biometric system is developed using a decision-level fusion based on WOA-ANN. The parallel model biometric system is developed using a score-level fusion based on SSA-DBN. The biometric authentication performs preprocessing, feature extraction, matching, and scoring for each unimodal system. On each biometric attribute, matching scores and individual accuracy are cyphered independently. A matcher performance-based fusion procedure is demonstrated for the three biometric qualities because the matchers on these three traits produce varying values. The two-level fusion technique (score and feature) is implemented separately, and their results with the current scheme are compared to exhibit the optimum model. The suggested plan makes use of the highest TPR, FPR, and accuracy rates.

Keywords: multimodal biometric authentication system; electrocardiograph (ECG); sclera; fingerprint; whale-optimization-algorithm-based artificial neural network (WOA-ANN); salp-swarm-algorithm-based deep belief network (SSA-DBN)

**Citation:** Singh, S.P.; Tiwari, S. A

Dual Multimodal Biometric

Authentication System Based on

WOA-ANN and SSA-DBN

Techniques. *Sci* **2023**, *5*, 10.<https://doi.org/10.3390/sci5010010>

Academic Editors: Asok Ray and

Chandrachur Bhattacharya

Received: 4 December 2022

Revised: 6 February 2023

Accepted: 23 February 2023

Published: 1 March 2023

**Copyright:** © 2023 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license ([https://creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)[https://creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/)

4.0/).

1. Introduction

Confirmation is the most common way of recognizing a person or thing, providing access control to frameworks through a coordinating course of client information, with the information stored in an approved dataset. Customary confirmation frameworks, in view of logins and passwords, are definitely more vulnerable to attacks than biometric validation frameworks. Biometric-recognizable proof is the craft of utilizing physical or actual attributes (i.e., iris, hand vein, face, finger vein, palm print, fingerprints, hand, tooth shape, ECG, ear shape, and so on) and conduct qualities (i.e., voice, walk, mark, keystroke elements, and so on) to distinguish clients [1]. The fundamental properties of biometric frameworks are exceptionality (the biometric should be one of a kind for every individual, in any event, even for twins), comprehensiveness (the biometric should be moved by the individual), permanency (the biometric ought not to be impacted by age), quantifiability (the biometric ought to be quantifiable with basic specialized devices), and ease of use (the framework ought not to be difficult to utilize). The use of biometrics is far-reaching, yet the frameworks are in danger as the climate and their utilization can influence estimations and require reconciliation, as well as extra equipment, and biometrics cannot be reset once compromised. Biometric frameworks depend on specific information about extraordinary organic qualities [2]. Many works ordinarily endeavor to utilize three biometric characteristics: ECG sign, sclera, and unique marks.

An electrical sign transmitted by the heart is referred to as an ECG signal. It tends to be estimated in a sophisticated manner by sensors attached to the chest of a human [3,4]. Recently, it has been utilized for validation, as it encounters the vast majority of biometric determinations. One of the upsides of ECG-based validation is aliveness, which is dissimilar to other verification strategies, for example, finger impressions and secret key confirmation, as it tends to be used exclusively by living people. Additionally, ECG can perhaps be given by a more extensive range of individuals, for example, handicapped people and the incapacitated, who cannot give regular biometrics; this includes, for example, unique marks, iris scans, or palm prints. Moreover, ECG information can perhaps be taken from different parts of the human body, for example, the fingertip, making it pertinent for a great many people [5,6].

The sclera is the white portion of eye, which acts as a strong wall for the eye. The sclera is covered with a bodily fluid layer required for the smooth development of the eye [7,8]. It is encircled by the optic nerve and is the thickest layer. The different inside pieces of the sclera are the episcleral, present beneath the conjunctiva; the sclera legitima, which is a thick white tissue responsible for the white shade of the sclera; and the lamina fusca, which is a flexible fiber. The sclera contains vein designs that are unique for all individuals. Indeed, even twins have different vein designs. The example of vein design can be utilized for recognizing people. This example is noticeable and will not ever change over the lifespan of an individual [9,10]. In a unimodal biometric framework, just the removed sclera can be utilized for distinguishing between individuals.

Unique finger-impression-based confirmation is one of the main advances in biometric verification and is acquiring prominence in everyday life [11,12]. The development of an acknowledgment framework for confirming unique marks is critical, and has drawn the interest of a variety of specialists. Fingerprint acknowledgment frameworks are currently comprehensively utilized for 1:1 coordination (confirmation) or 1:N identification (ID) worldwide. They are able to perform dependably when used to recognize people for delicate certifiable applications, such as financial exchanges, PC/cell security, criminology, and global boundary crossing. Principally, unique finger impression acknowledgment frameworks coordinate finger impressions in order to be validated with the current database of fingerprints [13,14]. Then, at that point, it might likewise be utilized as a gadget for security. Taking into account finger-impression-based verification frameworks, a finger impression is an association of a number of edges and valleys on the surface of a picture [15,16].

There are two sorts of biometric framework: the unimodal biometric framework and the multimodal biometric framework. Unimodal biometric frameworks utilize only one biometric quality for acknowledgment, which frequently causes problems; this includes, for example, biometric information variability, the absence of peculiarity, low acknowledgment precision, and parody attacks. To address these issues, multimodal biometric frameworks are utilized. Multimodal biometric frameworks, combining various attributes, address the limits of unimodal biometric frameworks in matching precision, trouble mimicry, comprehensiveness, attainability, and so forth [17]. When contrasted with single-modular biometric frameworks, a multimodal biometric framework increases acknowledgment precision, security, and framework dependability.

However, the information degeneration of any of the biometric modalities in the multimodal biometric framework corrupts the results. This is because of the way that most existing multimodal approaches also employ combination rules that are fixed or cannot successfully adapt to the diversities of biometric attributes and ecological changes. Regardless, the degree of combination turns into an issue of concern; at that point, the element portrayal and matching procedures are too muddled to even consider performing. From that point, the interest becomes divided between the verification execution, calculations, and cost. Subsequently, to overcome the current difficulties, this work has fostered an original multimodal biometric validation framework whose goal is as follows:

- To make an equal combination multimodal biometric framework that utilizes three biometric qualities, specifically ECG, finger impression and sclera.

- To make a successive combination multimodal biometric framework that utilizes three biometric qualities, particularly ECG, unique mark and sclera.

The forthcoming segment describes the format of the article. The ensuing section presents a review of the literature and the related examinations. Section 3 discusses the proposed technique and Section 4 shows the outcomes of this study's evaluation of consistent association traffic. Section 5 concludes and proposes future endeavors.

2. Literature Survey

El-Rahiem et al. [18] introduced a multimodal biometric confirmation technique in light of the profound merging of ECG and finger vein. This framework has three primary parts: biometric pre-handling, profound component extrication, and verification. At the time of pre-handling, standardization and separation strategies were adjusted for every biometric. The highlights were extracted during the element extraction process using a proposed deep CNN model [19–21]. Then, at that point, the confirmation cycle was performed on the removed elements utilizing five notable AI classifiers: SVM, KNNs, RF, NB, and ANN. Likewise, to address the profound elements in a low-layered highlight space and accelerate the verification task, the creators embraced MCCA (multi-set canonical correlation analysis). Exploratory outcomes stated an improvement as far as the validation execution process is concerned, with EERs of 0.12% and 1.40% highlighting the combination and score combination, respectively.

Valsaraj et al. [22] dissected the EEG signals for the trademark highlights evoked by the development and creative mind of four unique upper appendage developments. Similar development symbolism undertakings were looked at for their execution and legitimacy in fostering a powerful multi-modular biometric framework for people with engine handicaps. The review included ten subjects who implemented the envisioned raising of the right and left hands and the gripping of the right and left clenched hands. Alongside the envisioned development (engine symbolism), information for genuine appendage development was gathered, and the presentation was analyzed for both the nonexistent and real development. This framework pipeline achieved a misleading acknowledgment rate of under 2% for every one of the fanciful and real activities. An innovative multimodal technique for consolidating diverse MI activities was effectively executed with 98.28% precision. In addition, both fanciful and real developments showed a similarly great capacity for biometric purposes, recommending the convenience of the introduced biometrics framework for individuals with lost engine capacities or individuals with unfortunate engine symbolism abilities.

Cherifi et al. [23] developed a completely hidden and strong multimodal validation framework that naturally verifies a client based on how he or she answers the phone, after both arm and ear signal biometric modalities were removed from this single activity. To address the issues that ear and arm signal confirmation frameworks face in real-world applications, the creators proposed another strategy, based on image discontinuity, that makes ear acknowledgment more powerful in corresponding to impediment. The ear highlight extraction process was made locally by utilizing Neighborhood Stage Quantization (LPQ), in order to obtain heartiness, as well as posture and brightening variety. They also developed a set of four factual measurements to distinguish elements from arm motion signals. The multimodal biometric framework achieved an EER of 5.15%.

Gavisiddappa et al. [24] introduced a successful element determination calculation in order to decide the ideal component values for additional working on the exhibition of multimodal biometric confirmation. At first, the info pictures were gathered from the CASIA dataset. Then, at that point, highlight extraction was performed by utilizing the Neighborhood Double Example (LBP), seemingly trivial details that include extraction, Histogram of Arranged Angle (Hoard), and Dark Level Co-Event Grid (GLCM); this highlights bunch noticeable quality, Reverse Distinction Second Standardized (IDMN), and autocorrelation. After highlight extraction, the adjusted help includes the determination calculation that was utilized for dismissing the insignificant elements or for picking the ideal highlights.

3. Proposed Strategy

In this research work, two multimodal biometric systems are developed by combining three uni-modal biometric systems, including ECG, sclera and fingerprint. The two multi model biometric systems are known to be the following: (1) a sequential modal biometric system, and (2) a parallel modal biometric system.

As shown in Figure 1, the parallel model biometric system is developed by using score-level fusion, which is further discussed in detail.

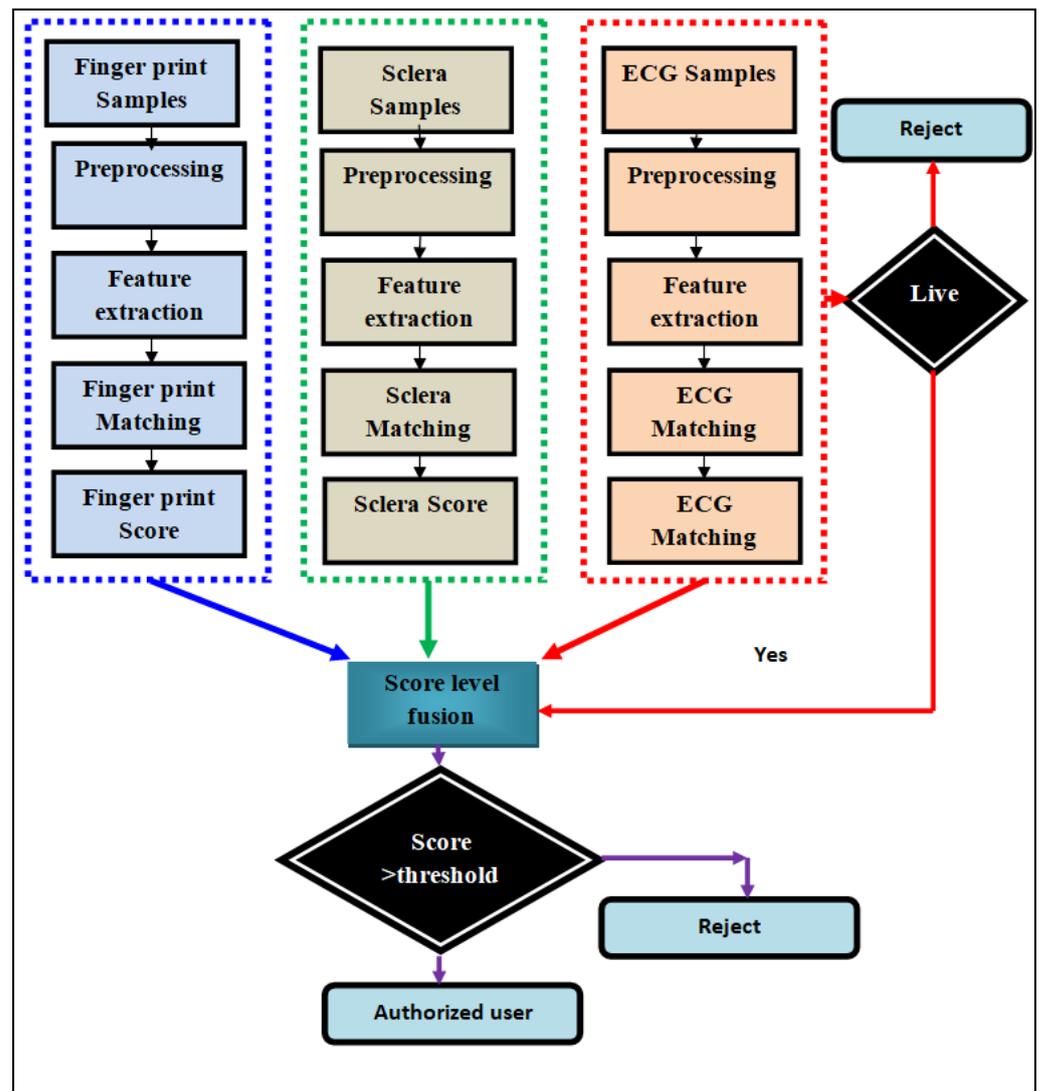


Figure 1. Parallel modal architecture.

As shown in Figure 2, the sequential model biometric system is developed by using decision-level fusion, which is further discussed in detail.

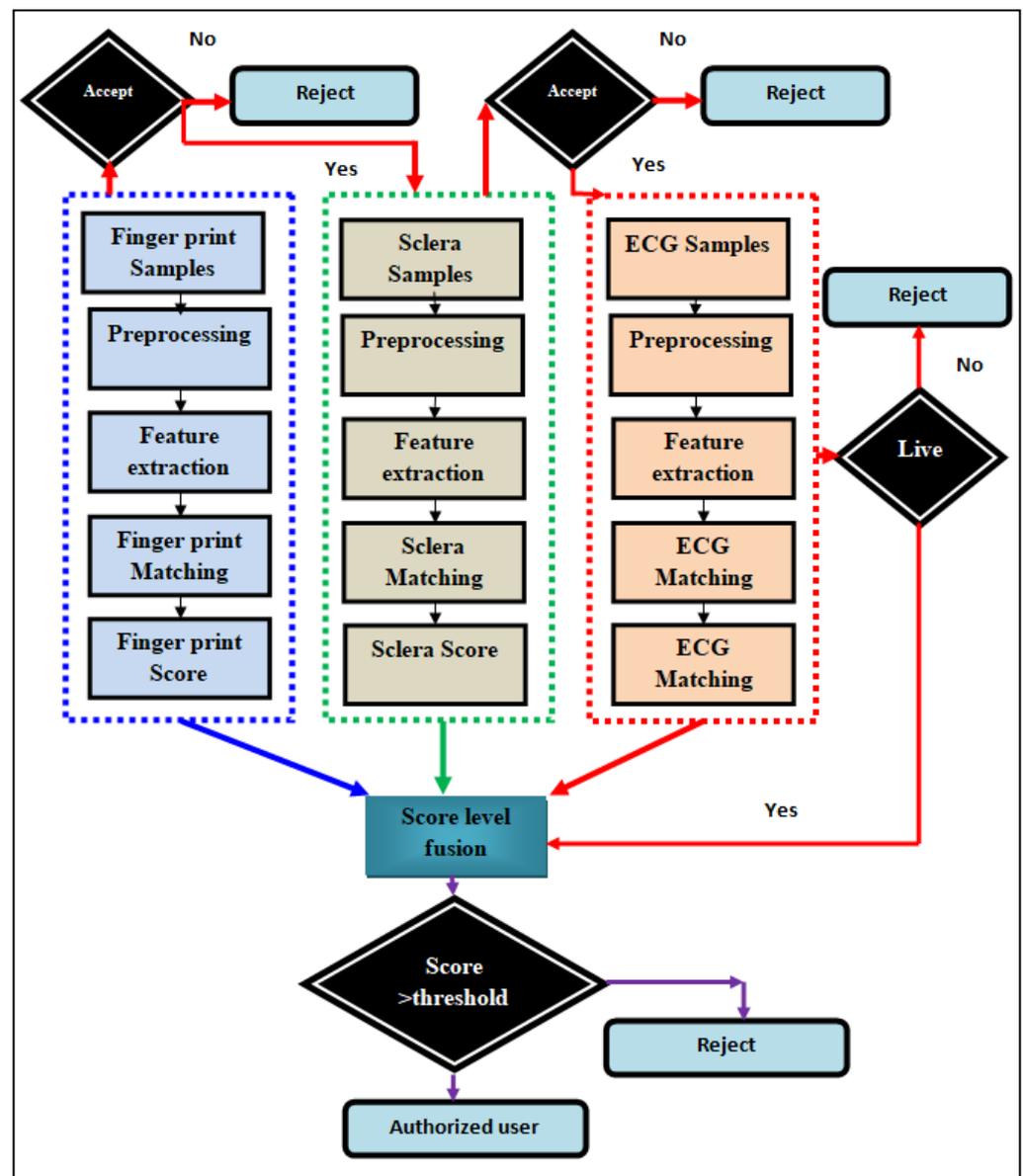


Figure 2. Sequential modal architecture.

3.1. Parallel and Sequential Modal Common Methodology

The parallel model architecture comprises a decision-making box for ECG, which states that if a person is alive, then fusion is proceeded to; otherwise, it is rejected.

The sequential model architecture comprises a decision-making process for each component, i.e., fingerprint, sclera and ECG. It evaluates whether the finger print and the sclera is matching, then moves further or rejects the user; the same process is performed for the ECG as for the parallel model.

3.1.1. Fingerprint

For fingerprint, the preprocessing step includes noise removal through binarization, fingerprint enhancement through the Gabor filter, histogram equalization, and the extraction of the Region of Interest (ROI). Finally, normalization is applied. Minutia marking, thinning, and the removal of breaks and spikes are carried out on the fingerprint to extract the features. A Convolutional Neural Network (CNN) is used for matching and generating the fingerprint score.

Binarization

The first grayscale picture is changed over into a double picture, which presents the picture as a 2D dark-level force capability with values from 0 to $L - 1$, where L indicates all the singular dim levels. Allow η to signify the all-out number of pixels in a picture and be the quantity of pixels with dim level i ; the likelihood that dim level i might happen is characterized as

$$\mathfrak{R}_i = \frac{\eta_i}{\eta} \tag{1}$$

The unique finger impression picture dark level is found using the middle value of

$$\forall_T = \sum_{i=0}^{L=1} i\mathfrak{R}_i \tag{2}$$

After averaging, the unique mark picture pixels are characterized into two unmistakable gatherings; here, t is as the limit value. The objects of interest in the forefront and foundation of a given picture relate to the w_1 and w_2 , respectively. Conditions (3) and (4) are the individual probabilities:

$$w_1(t) = \sum_{i=0}^{L=1} \mathfrak{R}_i \tag{3}$$

$$\text{and } w_2(t) = \sum_{i=i+1}^{L=1} \mathfrak{R}_i \tag{4}$$

The typical dark level qualities for w_1 and w_2 are determined, respectively, with the accompanying conditions:

$$\mathfrak{N}_1(t) = \sum_{i=i+1}^{L=1} \frac{i\mathfrak{R}_i}{w_1(t)} \tag{5}$$

$$\text{and } \mathfrak{N}_2(t) = \sum_{i=i+1}^{L=1} \frac{i\mathfrak{R}_i}{w_2(t)} \tag{6}$$

Thereafter, the enhancement process is carried out in order to improve the quality of the impression.

Enhancement

(a) Gabor filter

Finger impressions have a few general qualities, such as wrinkles and edges. Other palm print qualities include standard lines and kinks. A bank of 2D changed Gabor channels is utilized to channel palm print and unique mark pictures every which way in order to feature these qualities and eliminate clamors. In altered Gabor channels, rather than cosine capability $\cos(a; S)$, another occasional capability $f(a, S_1, S_2)$ is utilized. It is formed from two cosinusoidal useful bends with the various periods S_1 and S_2 . The parts over the x-pivot comprise a cosinusoidal useful bend with a period S_1 , and the parts beneath the x-hub comprise another cosinusoidal utilitarian bend with the various period S_2 . A 2D changed Gabor channel has the accompanying structure in the picture space (a, b) , as displayed in (7) and (8), where a and b are pixel organizers, and ϕ is the nearby direction of the current pixel:

$$\mathfrak{S}(a, b, S_1, S_2, \phi)\mathfrak{S}(a, b, S_1, S_2, \phi) = \mathfrak{h}_a''(a, b, S_1, S_2, \phi)\mathfrak{h}_b''(b, \phi) \tag{7}$$

$$\mathfrak{S}(a, b, S_1, S_2, \phi) = \left[\exp\left(\frac{-a\phi^2}{\alpha_a^2}\right) f(a, \phi, S_1, S_2) \right] \left\{ \exp\frac{-b\phi^2}{\alpha_b^2} \right\} \tag{8}$$

where

$$a\phi = a \cos \phi + y \sin \phi, \quad b\phi = -a \sin \phi + y \cos \phi \tag{9}$$

(b) Histogram Equalization

For tuning the force appropriation that comprises those pixel values, the phase that needs to be accomplished is histogram leveling. Consolidating the force of fluffy histogram balance with combined histogram evening out favors twofold histogram adjustment. Histogram leveling is utilized to uniformly spread the pixel power histogram in order to increase the dynamic scope of the pixels; this consequently maximizes the differentiation. Combined histogram adjustment is completed before fluffy histogram equalization. Combined histogram evening out moves towards the execution for total histogram leveling of a picture. The picture histogram is acquired first. Then, a histogram of combined conveyance capability is acquired. For each dimension that is worth of unique picture, another relating value is found using the histogram adjustment method.

Absolute number of pixels = ℓ

Complete number of conceivable power levels = ξ , then

$$\wp_k = \frac{\xi}{\ell} \times c_R(k) - 1 \tag{10}$$

Now the enhanced image of the fingerprint, $\wp_k = [\wp_{1x1}, \dots, \wp_{1xn}]$, is followed by feature extraction $\rho(\wp_k) = \rho_k[\wp_{1x1}, \dots, \wp_{1xn}]$.

Feature Extraction

Feature extraction helps to extract knowledgeable features in order to avoid a high error rate, which leads to the inaccurate authorization of the user.

(a) Minutiae Extraction

The construction of a Crossing number (CN) is widely utilized for the separation of the random data. The crossing number procedure limits the misleading acknowledgment rate (FAR). As edges have almost 150 choices and peculiarity, it is more straightforward to look through all the different random data assortments related to the unique mark. As a general rule, for each 3×3 window, on the off chance that the focal picture component is 1 and explicitly has 3 1-value neighbors, then, at that point, the focal picture component is an edge branch. In the event that the focal picture component is 1 and has just a single 1-value neighbor, then the focal picture component might be an edge finish (i.e., for an image component ρ , in the event that $\lambda_n(\rho) = 1$, it is an edge finish; assuming that $\lambda_n(\rho) = 3$, it is an edge bifurcation point).

$$\lambda_n = 0.5 \sum_{i=1}^8 |\rho_i - \rho_{i+1}| \tag{11}$$

Matching is material just during distinguishing proof or confirmation after they gave unique mark. This step finishes up the finger impression acknowledgment process.

Normalization

In this process, the scores from the unique finger impression module are matched while utilizing the Suprema range, from 0 to 1. Thus, prior to continuing with combination, the match scores from the finger vein module should be standardized, since these scores range from 0 to ~500 (note that these address distances). We utilize twofold sigmoid capability for score standardization, which maps the stretching of the obtained scores [0, 1]. The standardized score that utilizes the twofold sigmoid is then given as follows:

$$\forall'_k(\lambda_n) = \begin{cases} \frac{1}{1 + \exp(-\frac{2(\forall_k - t)}{R_1})} & \text{if } \forall_k < t \\ \frac{1}{1 + \exp(-\frac{2(\forall_k - t)}{R_2})} & \text{otherwise} \end{cases} \tag{12}$$

where t is the reference working point, and R_1 and R_2 signify the left and right edges of the area where the capability is direct, i.e., the twofold sigmoid capability shows straight attributes in the stretch $(t - R_1, t - R_2)$. Factor t is, for the most part, decided to be of

some worth compared to the district of cross-over between the veritable and fraud score conveyance; meanwhile, R_1 and R_2 are made equivalent to the degree of cross-over between the two disseminations toward the left and right of t , respectively.

3.1.2. Sclera

Sclera is the second component used by the multimodal system in this work. The sclera image is also preprocessed to avoid a high error rate. Preprocessing comprises the following steps:

Normalization

Standardization accomplishes the straight change of the picture so that it can squeeze into a specific space. Here, a min–max standardization strategy is used for the normalization of the picture, which straightly changes the data. Min–max standardization is completed by using the following condition:

$$Q = \frac{\tilde{X} - \tilde{X}_{\min}}{\tilde{X}_{\max} - \tilde{X}_{\min}} \tag{13}$$

where \tilde{X}_{\min} and \tilde{X}_{\max} are the base and greatest qualities in picture \tilde{X} , where Q is the standardized picture.

Bilateral Filter

The two-sided channel takes a weighted amount of pixels in a nearby neighborhood; the loads rely upon the spatial distance and the power distance. Unequivocally, at a pixel area x , the result of a two-sided channel is figured as follows:

$$F(X) = \frac{1}{C} \sum_{Y' \in N(X')} e^{-\frac{\|Y'-X\|^2}{2\sigma_R^2}} e^{-\frac{\|F(Y')-F(X)\|^2}{2\sigma_K^2}} \tag{14}$$

where σ_R^2 and σ_K^2 are boundaries controlling the tumble off of loads in spatial and power spaces, respectively, $N(X')$ is a spatial neighborhood of pixel $F(X)$, and C is the standardization steady. This reciprocal channel is generally used for smoothing the picture in a space of low variety that would further develop division.

3.1.3. ECG

ECG is the third component used by the multimodal system in this work. The ECG image is also preprocessed in order to avoid a high error rate. The preprocessing comprises the following steps:

Median Filter

The main idea of the middle channel is to have the information signal as input and then have it supplant with the middle of the adjoining passages. The neighbors' style is termed the "window", which slides, section by passage, alongside the whole sign. For one-layered (1D) signals, the central window is obvious in the essential few that are placed in the previous and succeeding sections; in the interim, two-Layered (2D) or higher-layered plans are feasible. It is vital to note that, the windows have passages with odd qualities, the middle is assessed effortlessly and the sections can be attempted by the explored window mathematically. In order to obtain a simple scope of the sections, there is more than one potential middle. The result of a non-recursive channel to some degree is that the middle worth of the information inside the window is centered at the point.

If " $X(k) \leq k \leq L$ and $Y(k) \leq k \leq L'$ ", respectively, the info and result of the one-layered (1D) SM channel of window size is $2N + 1$; then

$$Y(k) = \text{mid}\{X(k - N), \dots, X(k - 1), X(k), X(k + 1), \dots, X(k + N)\} \tag{15}$$

For account start and last impact, $X(N)$ and $X(1)$, respectively, are rehashed N times toward the beginning and toward the completion of the information.

QRS Extraction

Profound learning procedures demonstrated their effectiveness in an excess of space. The stacked auto encoder brain network can be used in our identifier; it is a connection of heterogeneous and homogeneous brain organizations. The encoder layers exist within homogeneous organizations. The softmax classifier that is later put into the last encoder layer is termed the heterogeneous part. A fundamental (shallow) autoencoder comprises three layers; specifically, the principal layer is the information, the subsequent layer is the secret layer, and the final layer is the reproduction layer (yield). For the most part, it is prepared covetously if the autoencoder involves more than one secret layer. Figure 1 illustrates the average stacked autoencoder structures. Provided information X , the relating yield Y of a brain network is as follows:

$$Y = wX + b \tag{16}$$

In which X is the columns of the framework, w indicates the loads connecting the information hubs to the corresponding hub from the secret layer, f is the enactment capability, and b is the predisposition vector. The point of enhancement comprises tracking down the weight grid w_e that maps the contribution to the secret layer (encoder), and the weight framework w_d that recreates the contribution from the secret layer (decoder), that is

$$\operatorname{argmin}_{w_e w_d} (f_d(w_d f_e(w_e X + b_e) + b_d) - X)^2 \tag{17}$$

Equation (17) demonstrates that the outcome from every layer is the consequence of a network increase in the information vector with the encoder weight framework; afterwards, an inclination vector is added. Essentially, the result of every decoder layer is a grid increase in the result of the point of reference layer and the decoder weight lattice; afterwards, an inclination vector is added. The consequence of every hub in the design is reassessed in terms of enactment capability. The goal capability that is utilized for driving the streamlining of the autoencoder is the squared contrast among the information vector, X , and the result of the encoder–decoder phases. Basically, the autoencoder brain network is a solo learning framework that attempts to imitate a duplicate of its contribution once the results are obtained, with no requirement for marked examples. The principal objective of building autoencoder engineering is to look for an inborn portrayal of the information that cannot be found by hand-created highlights. The bottleneck secret layer is taken advantage of as a productive arrangement of programmed highlights. As referred to, every secret layer is pre-prepared alone, and afterward, every layer is stacked and organized in order to construct the acknowledgment framework. Practically speaking, only the encoder layers are utilized for the stacked autoencoder. After linking both the encoder layers and the softmax classifier layer, the entire framework is adjusted in a supervised

$$err = \frac{1}{2} \sum_{i=1}^N (Y_i - t_i)^2 \tag{18}$$

where Y_i represents the anticipated results, t_i represents the objectives (wanted yields) and N is the quantity of results. The scaled-form slope improvement calculation is utilized for the preparation stages. It was picked for its straightforwardness of execution and low algorithmic intricacy. Furthermore, the scaled form inclination is more powerful against introductory supposition decisions than a straightforward slope plunge technique. The slope update of the loads is evaluated using the following accompanying conditions:

$$w_{k+1}^i = w_k^i + \beta \nabla g^i \tag{19}$$

In Equation (19), the loads of layer i at the time of the k^{th} emphasis are refreshed by the inclination plunge ∇g , succeeding a stage size β . The slope drop is figured, beginning at the result layer in the regressive course until the information layer is arrived at. This calculation is rehashed a few times until a combination is reached. It has been demonstrated in many examinations that the back spread is exceptionally effective in preparing multi-facet structures. The value of the inclination that limits the mistake is established using the following form:

$$\nabla g^i = \min_{w^i} e^i = - \frac{de^i}{dw^i} \tag{20}$$

Without a doubt, the encoder layers pre-preparing exactness is not significant. Continuously, an adjustment stage is performed, including all the beforehand and pretrained brain network layers. Notwithstanding, for the last adjustment stage, the Levenberg–Marquardt algorithm is instead utilized for its greater exactness, intermingling strength, and semi-independency from starting speculation. Practically speaking, the back spread way to deal with limiting the mean squared blunder accomplishes great execution, although it is to the detriment of a rapid union. To speed up the assembly, cross-entropy capability is utilized to appraise the blunder, resulting in the softmax layer; this is the last highlight that comes about from the encoder layers.

3.1.4. Convolutional Neural Network

A CNN is opted for in order to eliminate the loss of information and simultaneously deal with high dimensionality [18]. The CNN is applied in all the three components for extracting the features in the sclera, and also for matching the scores generated for the fingerprint, sclera and ECG systems. The basic 2D CNN architecture is explained below.

Basic Working

For the information tests, the work strategy takes on a two-layered convolutional activity, and the last objective is to extricate and create matching scores. The two-layered convolutional activity adds an aspect to the first one-layered case, which is a profound convolutional brain organization that is utilized in order to include location information. The principal element of two-layered convolutional networks is the profound convolutional brain organization. The system that uses sclera for identification is separated into 7 parts, which are the information input layer, convolution layer (same aspect), pooling layer (same facial inclination), full association layer, and order yield layer. For two-layered convolution tasks, a one-layered convolution piece is shaped. This cycle needs to utilize two aspects in order to concentrate and shape the last combination highlights. The two-layered convolution highlight extraction calculation is displayed in Equation (21), and the result cycle model chart of the two-layered convolution layer is displayed in Figure 3.

The two-layered convolution highlight extraction calculation is as follows:

$$X_i^{l,j} = f \left(b_f + \sum_{a=1}^m w_a^j x_{i+a-1}^{l-1,j} \right) \tag{21}$$

In view of the 2D CNN model preparation, the particular preparation process is as follows:

Stage 1: The Fingerprint and ECG information is imported, and one-hot encoding is performed on the mark. Then, at that point, the datasets are partitioned after the cluttered request, different boundaries of the 2D CNN network structure are instated, and the weight and the arbitrary number with an offset close to zero are set.

Stage 2: The forward engendering computation is understood. Furthermore, pooling procedures are performed on all the data and the matching score is located. The thick neurons that result are determined and associated with the full-interface layer. The softmax

capability is utilized in order to ascertain the likelihood of yielding a prepared neuron, the formula for which is as follows:

$$\alpha_j^L = \frac{e^{z_j^L}}{\sum_k e^{z_k^L}} \tag{22}$$

where L indicates the number of contributions of all the neurons in the layer.

Stage 3: The converse preparation is finished. The misfortune capability can be determined by determining the result likelihood and genuine mark of every neuron. The computation formula is as follows:

$$Loss = -\frac{1}{N} \sum_1^N \sum_{j=1}^K H_j \times \log(\alpha_j^L) \tag{23}$$

Among them, Loss is the misfortune capability, N is the preparation set example clump number, k is the classification number, H_j is the classification j^{th} enuine result mark, and α_j^L is the yield likelihood of the j^{th} neuron in the L^{th} layer. It is decided whether the Loss value fulfills the requirements. Both the Loss value and the accuracy value are steady. Assuming that the prerequisite is met, the preparation is finished and the weight and offset are saved; otherwise, Stage 4 is moved to.

Stage 4: The weight and offset of the preparation network are refreshed, which gives the features of the sclera. The equation is as follows:

$$w_{ij}^L = w_{ij}^L - \alpha \frac{\sigma}{\sigma w_{ij}^L} L(w, b) \tag{24}$$

$$b_i^L = b_i^L - \alpha \frac{\sigma}{\sigma b_{ij}^L} L(w, b) \tag{25}$$

where w_{ij}^L addresses the heaviness of the L^{th} layer, α addresses the learning rate, i is the middle of the road layer neuron, and j addresses the resulting neuron. The refreshed loads and counterbalances are acted in sync 2 in order to enter the 2D CNN model for preparation.

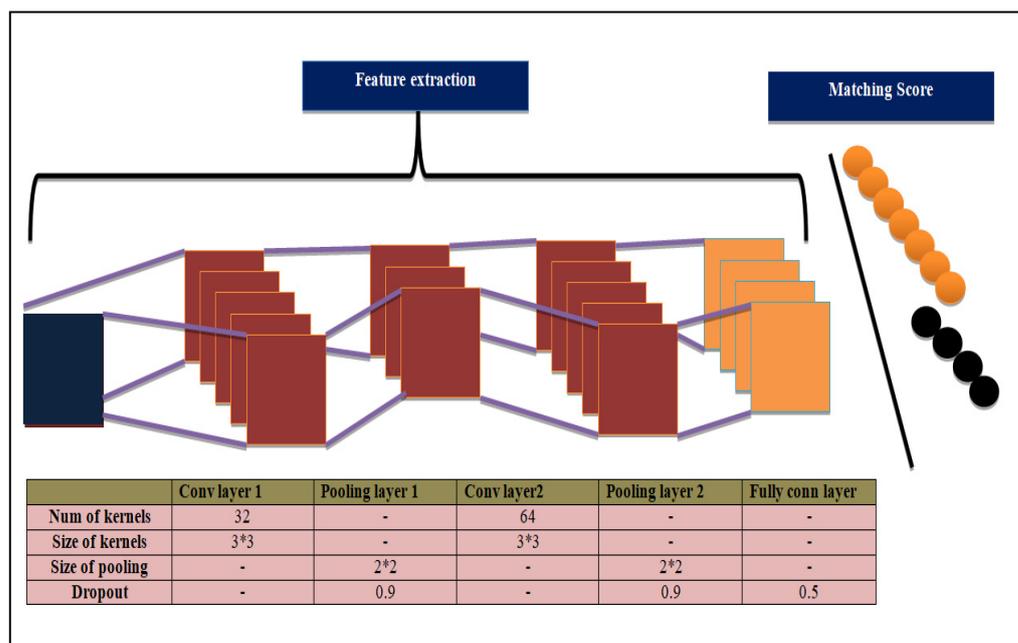


Figure 3. The 2D CNN architecture.

3.1.5. Parallel Fusion

Parallel fusion uses an optimized Deep Belief Network (DBN). The ReLU activation function of the DBN is optimized using the salp swarm algorithm (SSA). The optimized DBN classifier is utilized for obtaining the scores of every biometric modal. After that, the fusion rule is utilized among the ECG scores, sclera scores and the fingerprint scores in order to obtain the end score.

The DBN classifier is created by utilizing one multi-facet Perceptron (MLP) layer and two limited Boltzmann machine (RBM) layers, as shown in Figure 4. The associations are put among the covered up and the apparent neurons in the DBN classifier, and there are comparisons made between the secret neurons and the apparent neurons. The score-level combination yield is provided as the contribution to the primary RBM. The result obtained from the secret layer of the primary RBM is exposed to the contribution of the apparent layer in the subsequent RBM, and the result from the subsequent RBM is taken care of in the information layer within the MLP layer. The noticeable layer that has the element vector as its feedback, and that has the secret layer of the principal RBM, can be evaluated as follows:

$$z^I = \{z_j^e, z_j^f, z_j^i\} \tag{26}$$

$$F^I = \{F_1^I, F_2^I, \dots, F_n^I, \dots, F_r^I\}; 1 \leq n \leq R \tag{27}$$

where M_1^Q addresses the Q^{th} of the apparent RBM 1 neuron, F_n^I signifies the N^{th} secret layer, and R is the aggregate sum of the secret neurons. Every neuron in the secret layer and noticeable layer has a predisposition. By looking at the band, their predispositions in the covered up and noticeable layer can be surmised. The two predispositions connected with the neurons in the two layers for RBM 1 are given by the following:

$$X^I = \{X_1^I, X_2^I, \dots, X_Q^I, \dots, X_Z^I\} \tag{28}$$

$$Y^I = \{Y_1^I, Y_2^I, \dots, Y_Q^I, \dots, Y_Z^I\} \tag{29}$$

where M_1^Q is the inclination towards the Q^{th} noticeable layer, and Y^I intensifies the predisposition connected with the n^{th} secret layer. For the primary RBM, the weight vector is given as follows:

$$w^I = \{w_{QN}^I\}; 1 \leq Q \leq Z; 1 \leq 1 \leq R \tag{30}$$

where w_{QN}^I signifies the N^{th} hidden neuron's weight and Q^{th} signifies the noticeable neuron, and the weight vector component is indicated as $z \times r$. Subsequently, the result of the secret layer is processed in view of its loads and predisposition relating to each apparent neuron, as follows:

$$F_N^I = \sigma \left[b_n^1 + \sum z^I w_{QN}^I \right] \tag{31}$$

where the actuation capability is addressed as σ . Subsequently, the result acquired in the main RBM is evaluated as follows:

$$F^I = \{F_n^I\}; 1 \leq n \leq Rh \tag{32}$$

From that point onward, the educational experience of RBM 2 starts utilizing the result of the secret layer. The apparent neuron's aggregate sum is equivalent to the aggregate sum of the secret RBM 1 neurons, and is formulated as follows:

$$z^2 = \{z_1^2, z_2^2, \dots, z_N^2\} = \{F_n^I\}; 1 \leq n \leq R \tag{33}$$

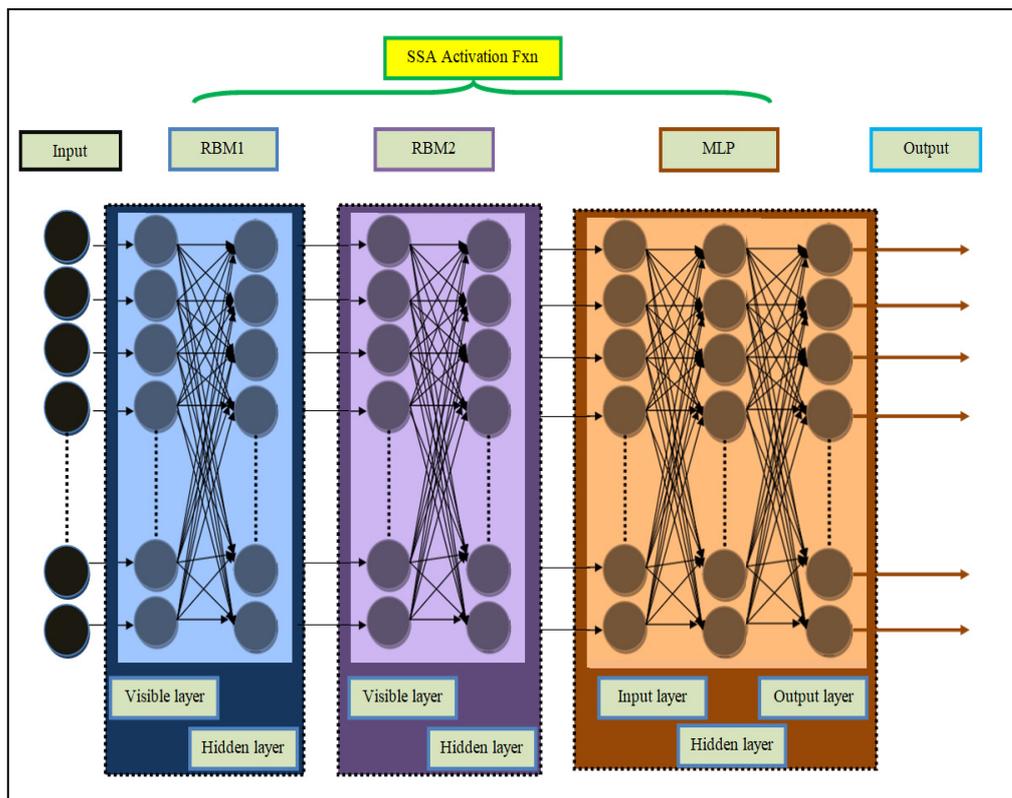


Figure 4. Proposed SSA_DBN architecture.

The second RBM secret layer is portrayed as follows:

$$F^2 = \{F_1^2, F_2^2, \dots, F_n^2\}; 1 \leq n \leq R \tag{34}$$

The secret layer and the apparent layer predisposition have similar portrayals, which are given in Equations (33) and (34). The load in the second RBM layer is communicated as follows:

$$w^2 = \{w_{NN}^2\}; 1 \leq Q \leq Z; 1 \leq 1 \leq R \tag{35}$$

where w_{NN}^2 is the load between the N^{th} secret neuron and the N^{th} noticeable neuron in the RBM 2 layer. The element of the weight vector is signified. The result of the secret neuron is resolved in light of the principal case, as follows:

$$F_N^2 = \varphi \left[b_n^2 + \sum z^2 w_{QN}^2 \right] \forall M_Q^2 = F_N^1 \tag{36}$$

Consequently, the result obtained from the secret layer is communicated as follows:

$$F^2 = \{F_n^2\}; 1 \leq n \leq P \tag{37}$$

Equation (38) is evaluated as the contribution to MLP, which addresses the aggregate sum of the neurons existing in the information layer. The input of the MLP layer is evaluated using the following:

$$\tilde{h} = \{\tilde{h}_1, \tilde{h}_2, \dots, \tilde{h}_N\} = \{F_n^1\}; 1 \leq n \leq P \tag{38}$$

which addresses the aggregate sum of the neurons that are offered to the secret layer as a result of the subsequent RBM. The secret layer of the MLP is communicated as follows:

$$t = \{t_1, t_2, \dots, t_p, \dots, t_U\}; 1 \leq P \leq U \tag{39}$$

where U indicates the quantity of secret neurons. YP signifies the predisposition of P^{th} secret neurons, in which $P = 1, 2, \dots, U = 1, 2, \dots, Q$. The MLP result is communicated as follows:

$$o = \{o_1, o_2, \dots, o_z, \dots, o_U\}; 1 \leq z \leq U \tag{40}$$

In which U represents the complete neurons in the result. The MLP comprises two weight vectors, one among the covered up and included layer, and the other among the yield and secret layer. Considering that w^j is the weight vector among the information and secret layers, it is described as follows:

$$w^j = \{w_{N^P}^J\}; 1 \leq K \leq m; 1 \leq r \leq x \tag{41}$$

where $w_{N^P}^J$ is the load among the j^{th} input neuron and the P^{th} secret neuron so that the size of w^j is $r \times x$. In view of loads and predisposition, the result of the secret layer is figured as follows:

$$h_r = X_r \left[\sum_{n=1}^r h_n w_{jr}^J \right] \forall h = F_N^2 \tag{42}$$

where φ_r indicates the inclination of the secret neurons, which is optimized by SSA and $h_n = F_N^2$, so that the contribution to the MLP is the aftereffect of RBM 2. The loads among the result layer and secret layer are communicated by the following:

$$w^L = \{w_{rz}^L\}; 1 \leq K \leq m; 1 \leq r \leq x \tag{43}$$

The result vector is determined in view of the result of the secret layer, and the weight w^L , as shown below:

$$O_z = X \left[\sum_{r=1}^Q h_r w_{rz}^L \right] \tag{44}$$

where w_{rz}^L is the load between the P^{th} secret and the z^{th} output neuron, and the secret layer yield is represented by h_r .

Now, the ReLU activation function is replaced by the SSA. The Salp has a place within the Salpidae family. Similar to how fish swim in schools [18], the Salp structures its chains in order to benefit its outcome when taking care of security, motion, and multiplication. The SSA's conduct is determined by working this out with the Salp chain, which looks for ideal food sources. In the SSA, in light of the person's (that is, the Salp's) position in the chain, they are divided into supporters or pioneers. The chain is started by the pioneer, while devotees submit to bearings for their development.

The proposed strategy outlines the SSA, where its closeness to the one more multitude smart calculation and also its effortlessness are shown. At the point at which the SSA began with the Salp populace instatement, the multitude of the Salp is epitomized as a 2D network. Then, the Salp's wellness can be in order to decide upon the Salp that utilizes the ideal wellness (the pioneer). The pioneer area can be overhauled by the following:

$$O_t = \left[O_1^1 O_2^1, \dots, O_d^1 O_1^2 O_2^2, \dots, O_d^2 : O_1^n : O_2^n, \dots, O_d^n \right] \tag{45}$$

$$O_1^1 = \lfloor Y_i + E_1((U_i - L_i))E_2 + L_i \rfloor E_3 \geq 0 Y_i + E_2((U_i - L_i))E_2 + L_i \rfloor E_2 \geq 0 Y \tag{46}$$

Here, the area of the starting Salp in the i^{th} boundary and the mean food area in the i^{th} boundary are addressed. Furthermore, the lower and upper limits of the i^{th} boundary are addressed, and, correspondingly, the coefficient is assessed by the accompanying condition. E_1 and E_2 also signify irregular qualities inside [0, 1].

$$E_1 = 2e^{-(\Delta l/L)^2} \tag{47}$$

Here, L addresses the maximal emphasis and l shows the current cycle. It is clear that the coefficient is critical in SSA, since it adjusts the investigation of and the abuse in the entire looking through technique.

$$X_i^j = \frac{1}{2}vt^2 + \Phi_0t \tag{48}$$

Here, $j \geq 2$, Φ_0 signifies an underlying pace, addresses the area of the j^{th} Salp in the i^{th} aspect, and addresses the time and duration. During improvement, E_1 , the time, shows the cycle. Consequently, the inconsistency among the cycles is identical to 1. Consider that $\Phi_0 = 0$, the resulting equation is used:

$$X_i^j = \frac{1}{2}(X_i^j - X_i^{j-1}) \tag{49}$$

When $j \geq 2$, assuming that some Salp moves beyond the looking through space condition, (6) demonstrates how to return them to the looking through space condition:

$$X_i^j = \left\{ \begin{matrix} l^j & \text{if } (X_i^j \leq l^j s^j) \\ X_i^j & \text{if } (X_i^j > l^j s^j) \end{matrix} \right\} \tag{50}$$

3.1.6. Sequential Fusion

An optimized Artificial Neural Network (ANN) is utilized for the sequential fusion. The Sigmoid activation function of the ANN is optimized using the Whale Optimization Algorithm (WOA), as shown in Figure 5. By using the effective employing of OR rule, the decision output combination is attained for achieving the superlative performance.

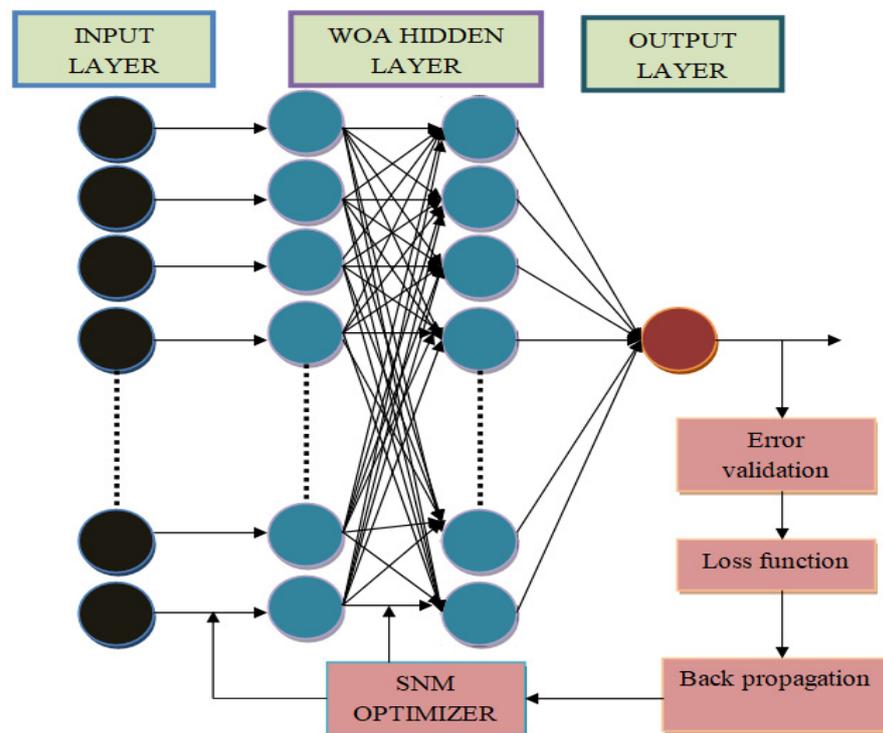


Figure 5. Proposed WOA-ANN architecture.

Building forecasting models is the same as determining the output variable that is the best estimate of the target value, given the known inputs. Typically, for environmental monitoring, a large number of parameters influence the target, and the connection between the inputs and outputs is not linear. Different forecasting strategies are accessible, with ANN being chosen in a comparative analysis presented in a previous study on various

forecasts. Creating a prediction model is a component of machine learning techniques, which requires a large dataset for training.

For pattern p (ζ_{net}), the net input to the hidden layer neuron is determined as the average of every input neuron emission ($\Gamma_{p,ji}$: input value), multiplied by weight ($O_{p,ji}$). For pattern p , an activation function is used for calculating the neuron outcome j of the hidden layer ($act_{\Gamma,j}$) and the neuron outcome k of the output layer ($o_{p,k}$), as follows:

$$T_{WOA} = Lw_{t+1} \tag{51}$$

where the activation function coefficient of WOA is T_{WOA} , and NET is described as $act_{\Gamma,j}$ or $o_{p,k}$, as per the following:

$$act_{\Gamma,j} = T \left(\sum_i \Gamma_{p,i} O_{p,ji} \right) \tag{52}$$

$$o_{p,k} = T \left(\sum_j act_{p,j} w_{p,kj} \right) \tag{53}$$

where $O_{p,ji}$ and $w_{p,kj}$ are the weights of the connections between the input layer neuron and the hidden layer neuron j , and between the hidden layer neuron and the output layer neuron k , respectively. Weights are set to modest random numbers at the start. Because of the sigmoid, nonlinearity functions are commonly employed. To minimize error, the learning algorithm updates the weights ($O_{p,ji}$ and $w_{p,kj}$).

The error's sum (e_p) in every neuron in pattern p is estimated as per the following:

$$e_p = \frac{1}{2} \sum_k (g_{p,k} - o_{p,k})^2 \tag{54}$$

$$te = \sum_p e_p \tag{55}$$

where $g_{p,k}$ represents the target value for the sequence at synapse k , and te is the overall error. With a forward process, the activation level computations pass over the hidden layer to the output layer (s). Every neuron in every subsequent layer combines its data and then performs a frequency response to generate its outcome. The channel's activation function subsequently generates the conclusive result, which is the projected goal value.

Each neuron's error value includes the level of mistake connected with that neuron. As a result, the neurons are ordered to refresh the proper weight modification. Output neuron weights are deployed as per the following:

$$\delta_{p,k(o)} = o_{p,k} (1 - o_{p,k}) (g_{p,k} - o_{p,k}) \tag{56}$$

A period control instrument is utilized to distinguish the sort of movement that occurs over the course of time. It does not only manage type A and type B swarm ways of behaving, but also the movements of jellyfish towards another approaching tide. The obligation cycle component is described in full in the accompanying subsection. In order to achieve an optimum result, the work used Sutskever Nesterov momentum to update the weights and learning rate.

$$w_{t+1} = w_t + \varphi_t v_t - \theta_t \nabla T(w_t + \varphi_t v_t) \tag{57}$$

At long last, the work proposes a shrewd activation determination model using the different features of the components.

Presently, the Whale calculation is utilized in order to enact the capabilities expressed in Equation 51. The use of this method is motivated by observing the natural way whales behave, which compares to the average way the Whale calculation accomplishes an upgraded arrangement. There are two stages to the calculation: the abuse and the investigation stage.

In the abuse stage, the whale creates a trail in a twisting way in order to encompass its prey. In the subsequent stage, the prey is arbitrarily investigated.

Numerical Clarification of the Abuse Stage:

Prey circling deals with two presumptions: the first is that the objective prey is the possibility of accomplishing the greatest arrangement; and the second is that other arbiters of pursuit consistently change their situations when a search specialist exists. This conduct is addressed as follows:

$$Lw_{t+1} = Lw_t - A_t \times D_{vect} \tag{58}$$

$$D_{vect} = \left| b_{coeffvect} \times Lw_{(t)} - L_{\alpha(t)} \right| \tag{59}$$

where Lw_t is the whale’s final position at iteration ‘ t ’, which is also proportional to the prey position; Lw_{t+1} is the whale’s existing position; and the distance between the prey and the whale is represented by D_{vect} , A_t and $b_{coeffvect}$, whose coefficient vectors represented as follows:

$$A_t = 2 \times A_{vect} \times r_{vect} + A_{vect} \tag{60}$$

$$b_{coeffvect} = 2 \times A_{vect} \tag{61}$$

For shrinking the search space corresponding to the spiral path trailed by the whale, the value of A_{vect} is minimized; this, in turn, decreases the oscillating range of A . Updating Spiral Path Position: The coordinates of the whale’s position are represented by (Lw, qw) and the coordinates of the prey’s position are represented by (Lt, qt) . The whale that trails the spiral path is represented by the following:

$$Lw_{t+1} = \exp^{cR} \times \cos(2\pi r) \times D_{Lvect} + L_{tvect} \tag{62}$$

$$D_{vect} = \left| b_{coeffvect} \times Lw_{t+1} - Lw_t \right| \tag{63}$$

where c denotes the constant that finds the logarithmic spiral path shape and R is the arbitrary variable that lies between -1 and $+1$. Thus, finally, the D_{vect} determines the values of the input features in ANN.

4. Results and Discussion

The proposed work is implemented in the Python programming language. The proposed model’s performance is compared to other existing models in terms of the ROC (the Receiver operating characteristic, which uses a sensitivity, specificity, efficiency, likelihood ratio).

For the fingerprint system, the dataset used was the FVR 2004-FVC 2004, which is a fingerprint database that consists of 4 separate databases, each of which has 80 photos of fingers. An optical sensor is used for DB1 and DB2, a thermal sweeping sensor is used for DB3, and the synthetic generator is used for DB4 to capture the 8-bit gray-level fingerprint images: link: <http://bias.csr.unibo.it/fvc2004/databases.asp>, (accessed on 1 July 2022).

The ECG system used The MIT-BIH Arrhythmia database, which is well-known in ECG-based biometrics research and is available through the Physionet repository. It has 48 signals, derived from ambulatory 2-lead measurements, each lasting 30 min. A wide variety of arrhythmias were represented by the 47 individuals; link: <http://physionet.org/physiobank/database/mitdb/>, (accessed on 1 July 2022).

For the sclera system, the SBVPI dataset was used. The freely accessible SBVPI dataset is intended primarily for sclera recognition research, although it can also be used for studies applying iris and periocular recognition methods. There are 55 subjects represented by 1858 high-resolution RGB photographs of their eyes; link: <https://sclera.fri.uni-lj.si/datasets.html>, (accessed on 18 July 2022). These datasets are all publicly available [25,26].

4.1. Performance Analysis of Parallel Modal Architecture

The proposed parallel SSA-DBN model is inspected, focused on measurements such as accuracy, sensitivity, specificity, F-measure, precision, FPR, MCC, FNR, FRR, NPV and computation time; it is furthermore analogized with existing techniques, such as alexNet-CNN, Resnet50, DBN, and ANN. Table 1 arranges the proposed strategy’s assessment, along with the common techniques that are focused on authorized user recognition.

Table 1. Evaluation of classification metrics for the proposed parallel SSA-DBN on recall, f-measure, NPV, specificity, accuracy and precision.

Techniques	Recall	f-Measure	NPV	Specificity	Accuracy	Precision
Proposed Parallel SSA-DBN	98.12	93.55	95.62	96.11	97.13	96.65
Existing alexNet-CNN	94.44	90.76	90.65	90.00	93.93	91.38
Existing ResNet50	91.22	88.99	89.78	86.74	92.29	88.57
Existing DBN	90.52	88.21	84.88	87.45	91.06	87.55
Existing ANN	86.12	83.73	82.54	81.86	88.69	83.38

Table 1 displays the proposed parallel SSA-DBN investigation, focuses on different execution measurements, including accuracy, sensitivity, specificity, F-measure, precision, FNR, FPR, NPV, FRR, etc. The measurements’ estimations are accomplished by being focused on 4 fundamental boundaries, similar to genuine TP, TN, FP and FN. The previously mentioned boundaries are focused on the shown measurements. TP shows that the real worth is not an assault and that the worth anticipated additionally yields something similar; TN shows that the genuine worth is an assault and that the anticipated worth yields something similar; FP establishes that the genuine worth is an assault, yet the anticipated worth does not show an assault; and FN shows that the genuine worth is not an assault anyway, and that the anticipated worth is expressed as an assault. Thus, an assault’s (unauthorized) assessment is totally dependent on the 4 boundaries.

Figure 6 shows the proposed parallel SSA-DBN graphical examination, along with the assorted existing techniques; these include alexNet-CNN, Resnet50, DBN, and ANN, which are focused on measurements, i.e., specificity, accuracy, recall, NPV, precision and F-measure. A review of the measurements’ accuracy shows the work’s adequacy in terms of the assorted dataset; in other words, the implemented strategy has unwavering quality. The proposed method achieves a more noteworthy specificity, accuracy, recall, NPV, precision and F-measure, of 98.12%, 97.13%, 98.12%, 95.62%, 96.65% and 93.55%, respectively; in any case, the existent method accomplishes the metrics value, ranging from 81.86 to 94.44%, which shows the lower viability of this plan compared to the parallel SSA-DBN procedure proposed. In addition, the parallel SSA-DBN proposed is broken down, and it fixates on F-measure measurements that depict the exactness of lopsided dispersion probability. With respect to quantifiable measurements, the proposed strategy yields a higher value in terms of recognizing an authorized user and avoids a high false detection rate, as shown in Table 1. Thus, the parallel SSA-DBN strategy proposed yields proficient unwavering quality and, furthermore, sidesteps the misclassification of assaults, compared to the existing procedures.

Table 2 discusses the metrics of the biometric authentication model, including FRR, FPR, Computation, MCC and FNR. The proposed model achieved an FRR, FPR, Computation, MCC and FNR value of 0.03%, 0.02%, 31.117%, 94.49% and 0.03%, respectively, which are better metrics values compared to the existing methods. This outcome mainly demonstrates the analysis between the predicted and the actual class values. A low FNR, FRR and FPR value illustrates a good model. MCC is the proportion of subjects genuinely analyzed as negative compared to every one of the individuals who had negative experimental outcomes (counting data that were erroneously analyzed as non-assault). This

characteristic can anticipate how likely it is for information be sought after, in the event of a negative experimental outcome.

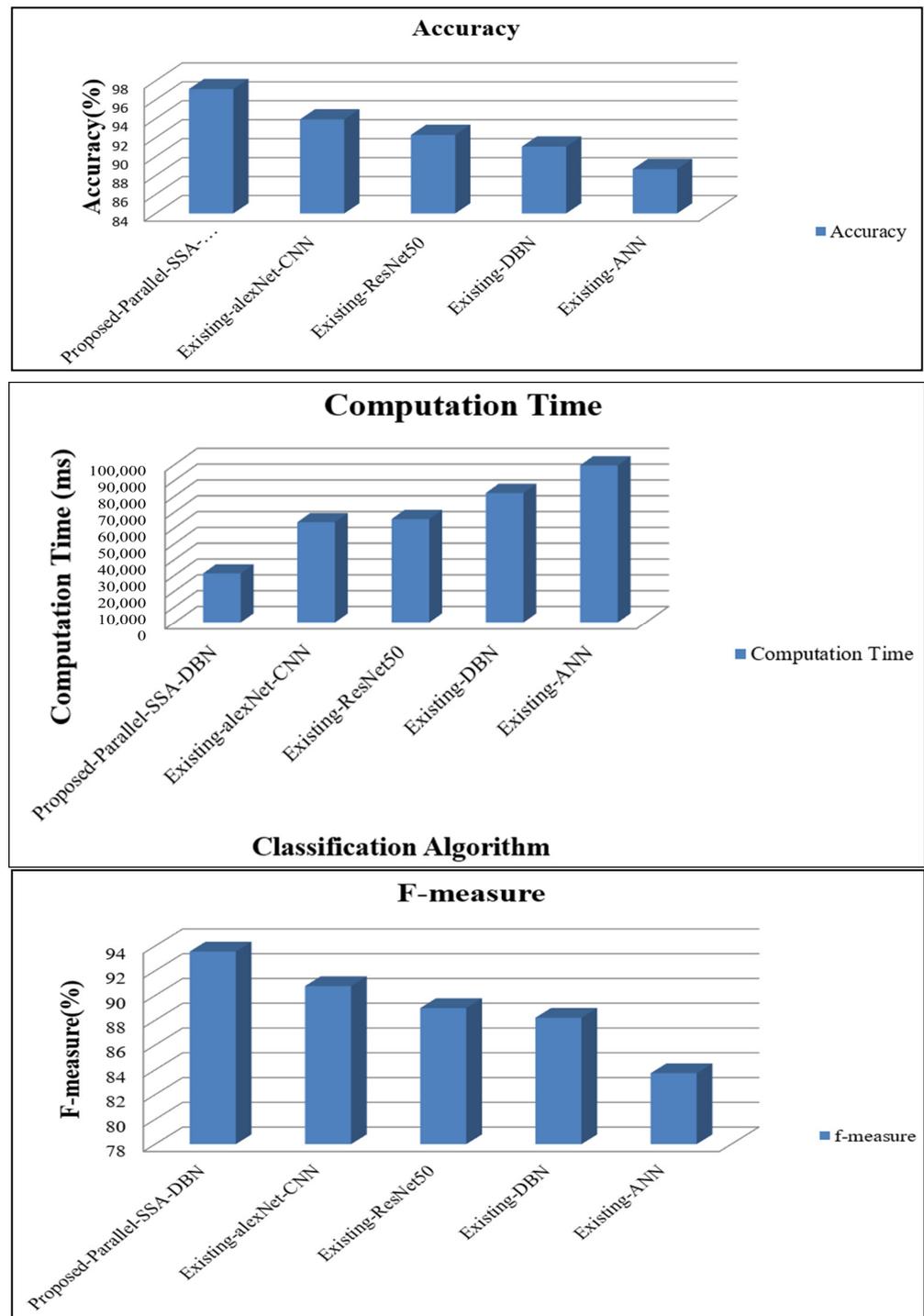


Figure 6. Cont.

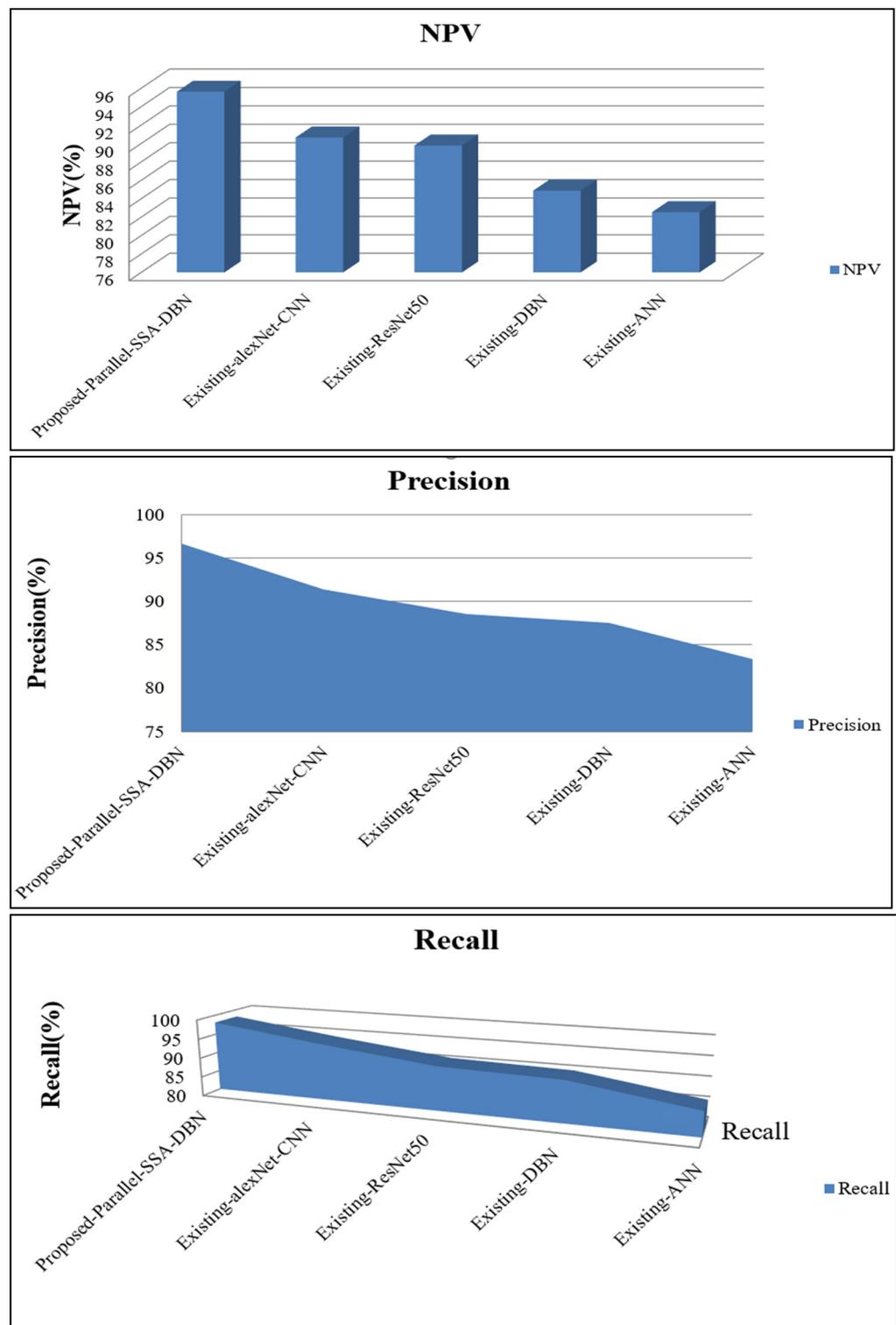


Figure 6. Cont.

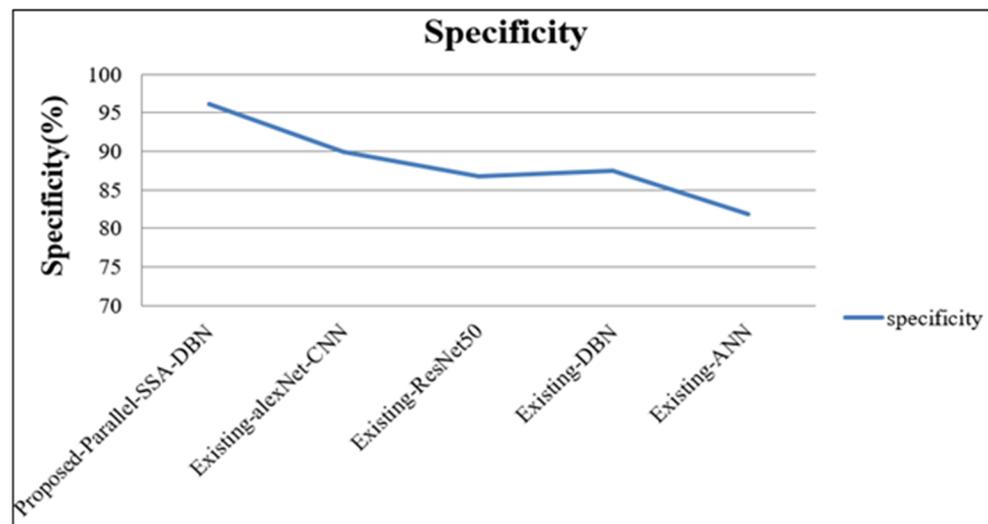


Figure 6. Graphical demonstration of classification metrics for proposed method along with existing methods.

Table 2. Evaluation of classification metrics for proposed parallel SSA-DBN on FPR, MCC, FRR, computation time and FNR.

Techniques	FPR	MCC	FRR	Computation Time	FNR
Proposed Parallel SSA-DBN	0.02	94.49	0.03	31,117.00	0.03
Existing alexNet-CNN	0.37	92.00	0.29	63,474.00	0.29
Existing ResNet50	0.53	91.00	0.51	65,454.00	0.51
Existing DBN	0.61	88.37	0.72	81,986.00	0.72
Existing ANN	0.96	84.89	0.95	99,384.00	0.95

Figure 7 displays the graphical assessment focused on measurements such as FRR, FPR, Computation, MCC and FNR, which are focused on the parallel SSA-DBN proposed. Then, the measurements accomplished are compared with different existing philosophies, such as alexNet-CNN, Resnet50, DBN, and ANN. Focusing on a plan to work successfully, it should involve the possibility of yielding a low FRR, FPR, and FNR value. The convention proposed yields of an efficient exactness, explicitness and responsiveness. The proposed strategies' accomplished values surpassed the existing techniques' accomplished measurement values, ranging between 0.29 and 92%; this is somewhat inefficient, compared to the method proposed. The parallel SSA-DBN procedure proposed yields productive measurements that are valuable to finding the assault and are shown to be efficacious compared to the existing strategy.

Figure 8 demonstrates the fitness vs. iteration for the proposed method, along with the existing methodology. The proposed method tends to perform well for activating the neurons, with a high level of fitness and a low number of iterations. The proposed method tends to achieve a fitness value that ranges between 66 and 194, whereas the existing technique achieves values between 30 and 185, which is low compared to the proposed method.

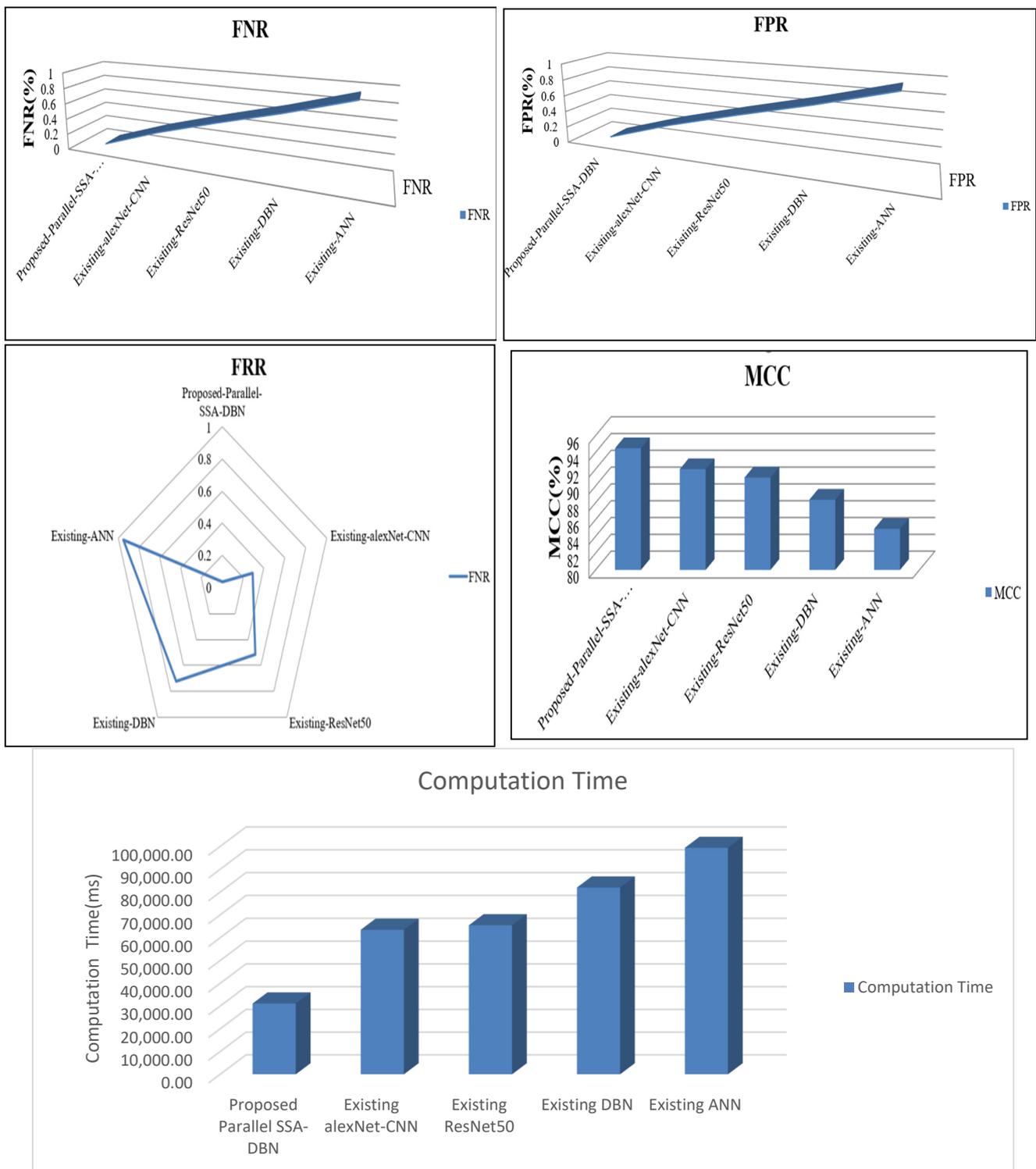


Figure 7. Classification metrics graphical demonstration.

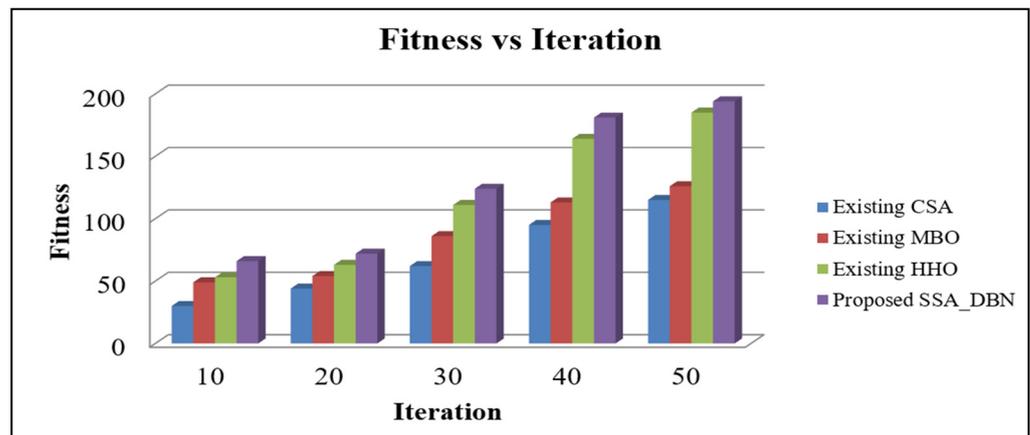


Figure 8. Fitness vs. iteration for proposed parallel SSA-DBN.

4.2. Performance Analysis of Proposed Sequential Modal Architecture

The sequential WOA-ANN proposed is inspected, focused on measurements such as accuracy, sensitivity, specificity, F-measure, precision, FPR, MCC, FNR, FRR, NPV, and computation time and, furthermore, is compared to the existent techniques, including alexNet-CNN, Resnet50, DBN, and ANN. Table 1 arranges the proposed strategy’s assessment, along with the common techniques focused on authorized user recognition.

Table 3 shows the proposed sequential WOA-ANN graphical examination, along with the assorted existing techniques, such as alexNet-CNN, Resnet50, DBN, and ANN, which are focused on measurements, i.e., specificity, accuracy, recall, NPV, precision, MCC and F-measure. A review of the measurements’ accuracy shows the work’s adequacy in terms of the assorted dataset; in other words, the implemented strategy has unwavering quality. The proposed plan achieves more noteworthy specificity, accuracy, NPV, precision, MCC and F-measure values of 95.54%, 98.00%, 95.63%, 95.23%, 94.56% and 93.79%, respectively;; in any case, the existing method accomplishes a metrics value that ranges between 80.18 and 91.85%, and that embodies the plan’s lesser viability compared to the sequential WOA-ANN procedure proposed. In addition, the sequential WOA-ANN proposed is broken down, focused on F-measure measurements that depict the exactness of lopsided dispersion probability. With respect to quantifiable measurements, the proposed strategy yields a higher value in terms of recognizing an authorized user and avoids a high false detection rate, as shown in Table 3 Thus, the sequential WOA-ANN strategy proposed yields proficient unwavering quality and, furthermore, sidesteps the misclassification of assaults compared to the existent procedures.

Table 3. Evaluation of classification metrics for proposed sequential WOA-ANN on specificity, accuracy and others.

	Specificity	Accuracy	Precision	F-Measure	NPV	MCC
Proposed Sequential WOA-ANN	95.54	98.00	95.23	93.79	95.63	94.56
Existing alexNet-CNN	91.46	94.42	90.85	91.54	91.85	92.95
Existing ResNet50	87.00	91.67	88.86	89.10	88.13	91.89
Existing DBN	87.69	91.12	86.11	87.56	83.00	86.46
Existing ANN	80.18	87.35	82.94	84.93	82.11	82.05

Figure 9 demonstrates the classification metrics for the proposed method, along with the existing methodology. The proposed method tends to perform well with a low error rate, whereas the existing techniques achieve a low metrics value, which leads to a low error rate.

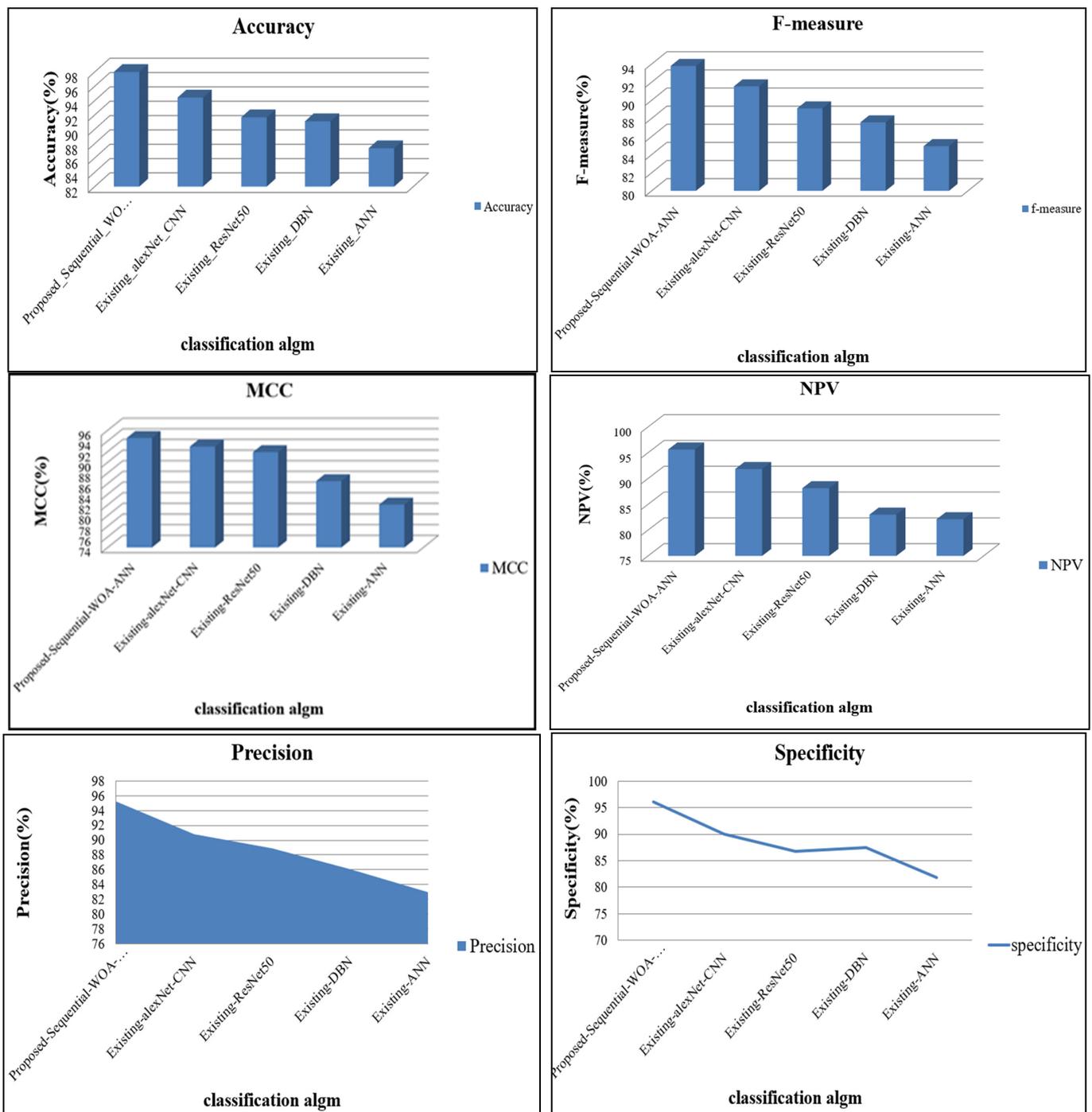


Figure 9. Classification metrics for sequential model architecture.

Table 4 shows the proposed sequential WOA-ANN graphical examination, along with the assorted existing strategies, such as alexNet-CNN, Resnet50, DBN, and ANN, which are focused on several measurements, specifically FRR, FPR, recall, computation time and FNR. The measurements' recall shows the work's viability in terms of the assorted dataset; in other words, the implemented grouping strategies have unwavering quality. Therefore, the proposed plan accomplishes a more prominent recall value of 98.46%; nonetheless, the existent strategy accomplishes a recall value ranging from 86.74% to 95.90%, which that shows that this method has a lower viability compared to the sequential WOA-ANN procedure proposed. Additionally, the sequential WOA-ANN plan proposed is examined focused on FPR, FRR and FNR measurements, which depict the probability

of misclassification. Concerning the FNR measurements, the implemented procedure yields a minimum worth of 0.03 FPR, 0.024 FRR and 0.02 FNR; in any case, the existing strategies yield FPR, FRR and FNR values ranging between 0.39 and 0.95%, which causes misclassification. Thus, the sequential WOA-ANN strategy proposed yields effective unwavering quality and, furthermore, sidesteps the misclassification of assaults compared to the existing procedures.

Table 4. Evaluation of classification metrics for proposed sequential WOA-ANN on FPR, FRR and others.

	FPR	FRR	Computation Time (ms)	FNR	Recall
Proposed Sequential WOA-ANN	0.0311	0.024	27,717	0.02	98.46
Existing alexNet-CNN	0.4295	0.388	57,464	0.39	95.90
Existing ResNet50	0.6112	0.644	77,814	0.64	92.22
Existing DBN	0.6966	0.861	89,986	0.86	89.35
Existing ANN	0.9572	0.986	99,114	0.99	86.74

Figure 10 discusses metrics such as FRR, FPR, recall, computation time and FNR for the biometric authentication model. The proportion of subjects genuinely analyzed as negative to every one of the individuals who had negative experimental outcomes (counting data that were erroneously analyzed as non-assault). This characteristic can anticipate how likely it is for information to be sought, in the event of a negative experimental outcome.

Figure 11 demonstrates the fitness vs. iteration for the proposed method, along with the existing methodology. The proposed method tends to perform well in terms of activating the neurons that have a high fitness value and a low number of iterations. The proposed method tends to achieve a fitness value that ranges between 68 and 198, whereas the existing techniques achieve values between 32 and 192, which is low compared to the proposed method.

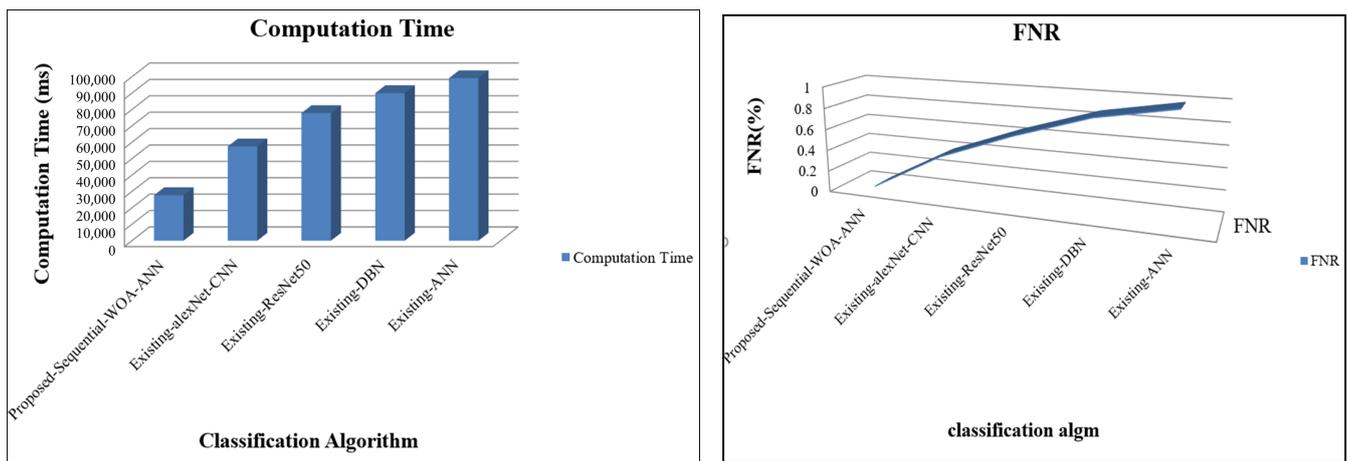


Figure 10. Cont.

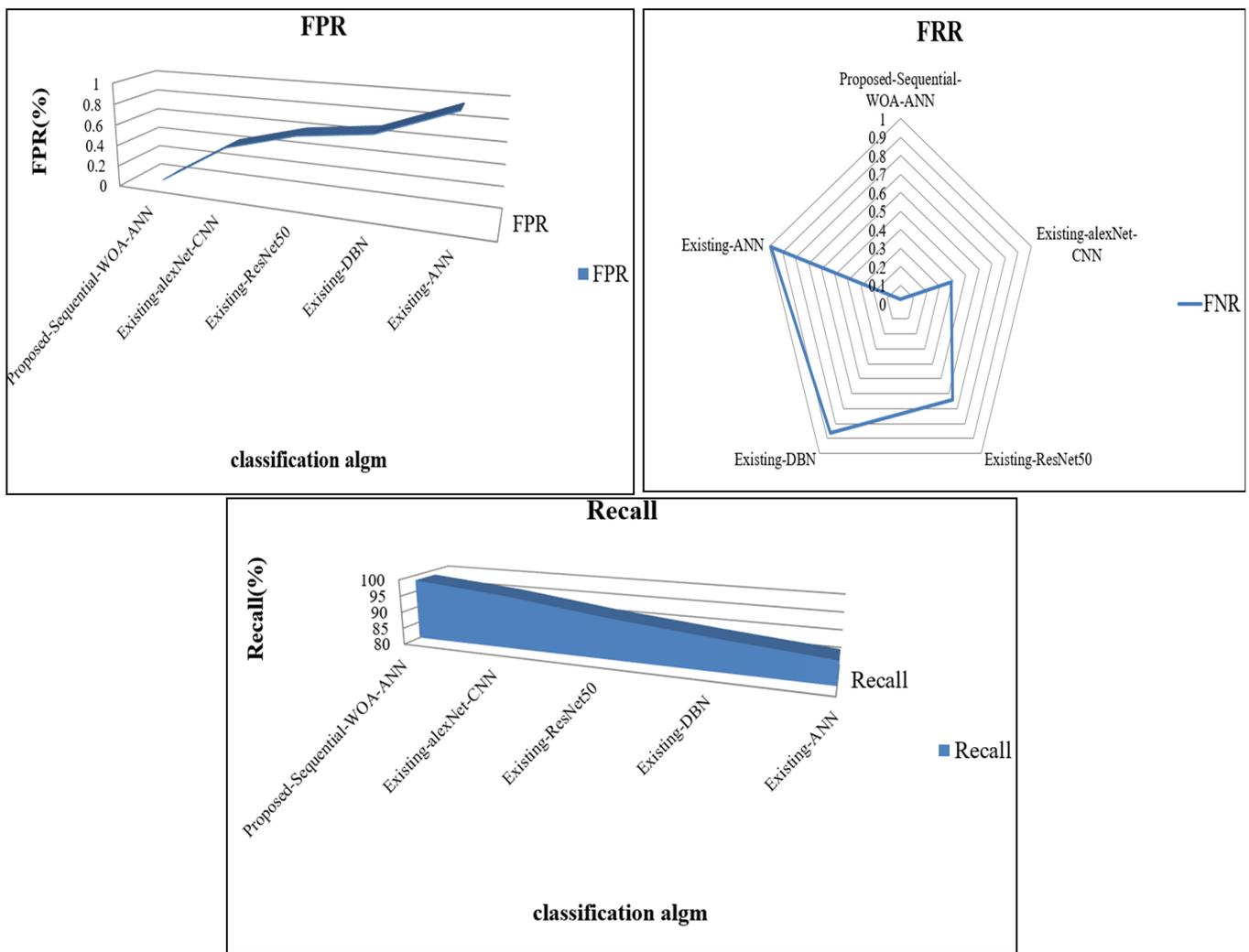


Figure 10. Graphical demonstration of classification metrics for sequential model architecture.

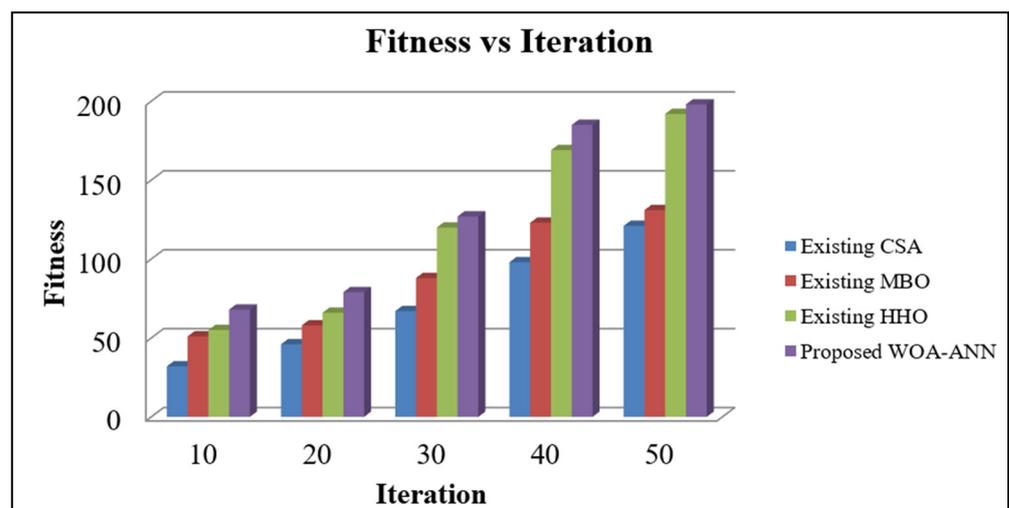


Figure 11. Fitness vs. iteration for proposed sequential WOA-ANN.

5. Conclusions

Authentication is a significant element in guaranteeing security for different applications. A multimodal biometric framework improves the vigour of the verification in-

strument as a result of its intrinsic robustness, uniqueness, and universality across various modalities. Appropriately, the work has created double multimodal biometric validation, which works on the strength of the framework, for example, by obtaining human hereditary codes and data for future reference. The proposed multimodal biometric system avoids the data degeneration of any one of the biometric models, which would deteriorate the system's performance. In the proposed work, the ECG supports liveness, and fingerprint and sclera support in the case of arrhythmia-like conditions. The proposed method gets adapted to variations in biometric traits and environmental changes. The ensemble fusion method is adapted with highly efficient feature representation and matching techniques. The model does not compromise between authentication performance, computation, and cost. The experiential assessment proved that the parallel model plan achieves 97.13% accuracy, 96.11% specificity, 93.55% f-measures and an FNR of 0.03%. In addition to that, the sequential model plan achieves 98.00% accuracy, 95.54% specificity, 93.79% f-measures and an FNR of 0.02%. Overall, the sequential model tends to obtain a better outcome compared to the parallel model, and remains highly secure compared to the existing techniques.

Author Contributions: Conceptualization, S.P.S.; software, S.P.S.; validation, S.P.S. and S.T.; formal analysis, S.P.S.; investigation, S.P.S.; resources, S.P.S.; data curation, S.P.S.; writing—original draft preparation, S.P.S.; writing—review and editing, S.P.S. and S.T., visualization, S.P.S.; supervision, S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Fingerprint: FVR 2004-FVC 2004 is a fingerprint database, available at link: <http://bias.csr.unibo.it/fvc2004/databases.asp>, (accessed on 1 July 2022), ECG: MIT Arthmyia—The MIT-BIH Arrhythmia database is available at the link: <http://physionet.org/physiobank/database/mitdb/> (accessed on 1 July 2022). Sclera: SBVPI dataset is available at link: <https://sclera.fri.uni-lj.si/datasets.html> (accessed on 18 July 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sudhamani, M.J.; Sanyal, I.; Venkatesha, M.K. Artificial Neural Network Approach for multimodal biometric authentication system. In *Proceedings of Data Analytics and Management*; Springer: Singapore, 2021; pp. 253–265.
2. Balaji, S.; Rahamathunnisa, U. A Review on Multimodal Biometric Authentication in Healthcare. In *Proceedings of the 8th International Conference on Advanced Computing and Communication Systems, Coimbatore, India, 25–26 March 2022*; Volume 1, pp. 1–5.
3. Bala, N.; Gupta, R.; Kumar, A. Multimodal Biometric System Based on Fusion Techniques: A Review. *Inf. Secur. J. A Glob. Perspect.* **2022**, *31*, 289–337. [[CrossRef](#)]
4. Gayathri, M.; Malathy, C.; Prabhakaran, M. A Review on Various Biometric Techniques, Its Features, Methods, Security Issues and Application Areas. In *Computational Vision and Bio-Inspired Computing*; Springer: Cham, Switzerland, 2020; pp. 931–941.
5. Shalini, P. Multimodal biometric decision fusion security technique to evade immoral social networking sites for minors. *Appl. Intell.* **2023**, *53*, 2751–2776. [[CrossRef](#)]
6. Ahamed, F.; Farid, F.; Suleiman, B.; Jan, Z.; Wahsheh, L.A.; Shahrestani, S. An Intelligent Multimodal Biometric Authentication Model for Personalised Healthcare Services. *Future Internet* **2022**, *14*, 222. [[CrossRef](#)]
7. Ketab, S.S.; Clarke, N.L.; Dowland, P.S. A Robust E-Invigilation System Employing Multimodal Biometric Authentication. *Int. J. Inf. Educ. Technol.* **2017**, *7*, 796–802. [[CrossRef](#)]
8. Kandasamy, M. Multimodal Biometric Crypto System for Human Authentication Using Ear and Palm Print. *Pattern Anal. Appl.* **2022**, *25*, 1015–1024. [[CrossRef](#)]
9. Thenuwara, S.S.; Premachandra, C.; Kawanaka, H. A Multi-Agent Based Enhancement for Multimodal Biometric System at Border Control. *Array* **2022**, *14*, 100171. [[CrossRef](#)]
10. Vensila, C.; Wesley, A.B. Multimodal Biometric Authentication Using Watermarking Technique. In *Security, Privacy and Data Analytics*; Springer: Singapore, 2022; pp. 79–91.
11. Elavarasi, G.; Vanitha, M. Multimodal Biometric Authentication by Slap Swarm-Based Score Level Fusion. In *Proceedings of Data Analytics and Management*; Springer: Singapore, 2021; pp. 831–842.
12. Ren, H.; Sun, L.; Guo, J.; Han, C. A Dataset and Benchmark for Multimodal Biometric Recognition Based on Fingerprint and Finger Vein. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 2030–2043. [[CrossRef](#)]
13. Channegowda, A.B.; Prakash, H.N. Image Fusion by Discrete Wavelet Transform for Multimodal Biometric Recognition. *IAES Int. J. Artif. Intell. (IJ-AI)* **2022**, *11*, 229. [[CrossRef](#)]

14. Tyagi, S.; Chawla, B.; Jain, R.; Srivastava, S. Multimodal biometric system using deep learning based on face and finger vein fusion. *J. Intell. Fuzzy Syst.* **2022**, *42*, 943–955. [[CrossRef](#)]
15. Sarangi, P.P.; Nayak, D.R.; Panda, M.; Majhi, B. A feature-level fusion based improved multimodal biometric recognition system using ear and profile face. *J. Ambient. Intell. Humaniz. Comput.* **2022**, *13*, 1867–1898. [[CrossRef](#)]
16. Medjahed, C.; Rahmoun, A.; Charrier, C.; Mezzoudj, F. A deep learning-based multimodal biometric system using score fusion. *IAES Int. J. Artif. Intell.* **2022**, *11*, 65. [[CrossRef](#)]
17. Heidari, H.; Chalechale, A. Biometric Authentication Using a Deep Learning Approach Based on Different Level Fusion of Finger Knuckle Print and Fingernail. *Expert Syst. Appl.* **2022**, *191*, 116278. [[CrossRef](#)]
18. El-Rahiem, B.A.; El-Samie, F.E.A.; Amin, M. Multimodal Biometric Authentication Based on Deep Fusion of Electrocardiogram (ECG) and Finger Vein. *Multimed. Syst.* **2022**, *28*, 1325–1337. [[CrossRef](#)]
19. Valsaraj, A.; Madala, I.; Garg, N.; Patil, M.; Baths, V. Motor Imagery Based Multimodal Biometric User Authentication System Using EEG. In Proceedings of the International Conference on Cyberworlds (CW), Caen, France, 29 September–1 October 2020; pp. 272–279.
20. Cherifi, F.; Amroun, K.; Omar, M. Robust Multimodal Biometric Authentication on IoT Device through Ear Shape and Arm Gesture. *Multimed. Tools Appl.* **2021**, *80*, 14807–14827. [[CrossRef](#)]
21. Gavisiddappa, G.; Mahadevappa, S.; Patil, C. Multimodal Biometric Authentication System Using Modified ReliefF Feature Selection and Multi Support Vector Machine. *Int. J. Intell. Eng. Syst.* **2020**, *13*, 1–12. [[CrossRef](#)]
22. Vitek, M.; Rot, P.; Štruc, V.; Peer, P. A Comprehensive Investigation into Sclera Biometrics: A Novel Dataset and Performance Study. *Neural Comput. Appl.* **2020**, *32*, 17941–17955. [[CrossRef](#)]
23. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, 215–220. [[CrossRef](#)] [[PubMed](#)]
24. Liu, H.; Liu, T.; Zhang, Z.; Sangaiah, A.K.; Yang, B.; Li, Y. ARHPE: Asymmetric Relation-Aware Representation Learning for Head Pose Estimation in Industrial Human–Computer Interaction. *IEEE Trans. Ind. Inform.* **2022**, *18*, 7107–7117. [[CrossRef](#)]
25. Liu, H.; Liu, T.; Chen, Y.; Zhang, Z.; Li, Y.-F. EHPE: Skeleton Cues-Based Gaussian Coordinate Encoding for Efficient Human Pose Estimation. *IEEE Trans. Multimed.* **2022**, 1–12. [[CrossRef](#)]
26. Liu, H.; Fang, S.; Zhang, Z.; Li, D.; Lin, K.; Wang, J. MFDNet: Collaborative Poses Perception and Matrix Fisher Distribution for Head Pose Estimation. *IEEE Trans. Multimed.* **2021**, *24*, 2449–2460. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.