An Enhanced Face Detection System using A Novel FIS-CDNN Classifier

Santhosh S.[1], Dr. S. V. Rajashekararadhya[2]

Research Scholar, Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, Karnataka - 572201, India[1]
Professor and Former Principal, Department of Electronics and Communication Engineering, Kalpataru Institute of Technology, Tiptur, Karnataka - 572201, India[2]

Abstract—A computer application that can detect, track, identify or verify human faces as of an image or video capture utilizing a digital camera is Face Recognition (FR). Few challenges like low-resolution images, aging, uncontrolled pose, illumination changes, and poor lighting conditions are not tackled even though huge advancement has been created in the Face Detection and Recognition (FDR) domain. Utilizing the Modified Tiny Face Detection (MTFD) and Fuzzy Interference System - Convolutional Deep Neural Network (FIS-CDNN) classifier, a Face Recognition System (FRS) was proposed here to tackle all complications. Primarily, Gamma correction - Based Histogram Equalization (GBHE) technique is utilized to augment the image's input in the pre-processing phase. The MTFD was employed to detect the face. Following that, the features are extracted. The Improved Chehra (IC) landmark extraction method was employed to retrieve the landmark features. And finally, the Geometric Features (GFs) are extracted. Later, the Gaussian - centered Spider Monkey Optimization (GSMO) Algorithm was employed to choose the vital features. To recognizing the face, the chosen features are fed into the FIS-CDNN classifier. When analogized to the prevailing models, it is concluded via the experiential outcomes that higher accuracy was attained by the proposed method.

Keywords—*Gamma correction - based histogram equalization (GBHE) technique; modified tiny face detection method (MTFD); Improved Chehra (IC) landmark extraction method; Gaussian-based Spider Monkey Optimization (GSMO) Algorithm; fuzzy interference system-convolutional deep neural network (FIS-CDNN) classifier*

I. INTRODUCTION

Since the face contains numerous important details, it is a vital part of the human body for recognizing each individual [1]. This is what led to Human FR being a significant subject in the artificial intelligence field. Owing to their respective pixel data computation, the query face image is harmonized with the template face image in the facial database for the FR, which is the process of recognizing persons [2]. Without using any further equipment, humans have the natural ability to recognize faces in a scene [3]. Additionally, a complex computer vision problem with a variety of aspects is the automatic FR [4]. In every practical aspect of our lives, it has been recognized as a developing field owing to the wide utilization of this technology. The areas where various applications correlated to this field are currently used widely such as defense, surveillance, banking, media field, social media platforms, fraud identification, e-commerce, trade, diagnostics, space

science, research, genomics, bioinformatics, cyber security, internet of things and education [5]. Face Detection, face alignment, face matching and face representation are the four main steps of FRS. Face matching seeks to create a classifier to recognize distinct faces while extracting useful features to differentiate face images as of distinct people is the aim of face representation. The facial landmarks are obtained by the face alignment. Face Detection aims to return the face's position, if they exist.

Some specific conditions still make extremely accurate identification results impossible despite recent advancements in FR. This is typically the case when there is a lack of high-quality data, such as when face images are taken in uncontrolled environments (FR in the wild), where there are numerous variations that can affect facial images, including resolution, background, expression, pose, illumination and occlusions [6]. A critical issue viewed here is the feature's robustness retrieved from face appearance descriptors to tackle these problems. In recent years, a lot of attention was received by deep learning approaches to FR. A turning point was reached when neural networks for facial Feature Extraction (FE) were introduced [7].

Numerous prominent techniques for FE have been developed up to this point, like Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP) etc. [8]. When these features are compared to the tested photographs, the person can be recognized or not. Unfortunately, as there are so many difficult and variable parameters in FD and identification as well as poor evaluation techniques, choosing the optimal facial recognition algorithm is a very difficult procedure [9]. Additionally, redundant features were engendered and struggled to handle high dimensions [10].

A. Problem Definition

Numerous approaches to resolving the FRS issue can be found in prevailing research procedures. Still, there is space for improvement. The following list includes the issues with the prevailing approaches:

- The prevailing approach focuses solely on brightening the images with the issue of locating a face in a hazy image. Thus, it has an impact on the recognition rate.

- Another difficulty for the FR system is the ageing process, which reflects the modifications in the face appearance / texture over a time.

- All the issues like age, expression, illumination, position and occlusion invariants were not addressed by most prevailing works. High recognition performance was attained by the prevailing presented pose invariant FR. However, when age, light, resolution and occlusion were all constant, it was unable to identify faces.

B. Limitations of Existing Works

Mohd Najib Mohd Salleh, Noureen Talpur and Kashif Hussain [27] exhibited Adaptive Neuro-Fuzzy Inference System (ANFIS) is an efficient estimation model not only among neuro-fuzzy systems but also various other machine learning techniques. ANFIS suffers from limitations that halt applications in problems with large inputs; such as, curse of dimensionality and computational expense. Different limitations such as curse of dimensionality, interpretability of rules, and parameter training are the major hurdles that need to be overcome for the implementation in problems with larger number of inputs. This is the reason, ANFIS is often integrated with additional techniques for input selection, rule reduction and parameter tuning, which again increases the complexity of the designed model. Various structural and parameters optimization techniques have been proposed in literature; however, there is enough room of improvement in ANFIS architecture so that applications in larger problems can be achieved easily.

A novel method of FR was proposed in this paper to tackle all the above-mentioned issues. The paper is framed as: the associated work regarding the proposed FR is analyzed in Section II. The concise discussion about FR is displayed in Section III. The proposed system's performance is analyzed in Section IV. The paper is concluded in Section V.

II. LITERATURE SURVEY

Shokoufeh Mousavi et al.[11] presented a Distinctive Landmark - centered FR (DLFR) system to diminish the issues. Grounded on the number of key points acquired as of a modified SIFT together with the most idiosyncratic face's landmark region, novel features were generated. The weights for the features were optimized by a slightly altered genetic algorithm. A Support Vector Machine (SVM) classifier was utilized to classify the weighted features. The experiential outcomes exhibited higher efficacy than the prevailing papers. However, for SVM, a large number of data was not appropriate.

Xue Lv et al. [12] introduced a deep learning algorithm for the application of an FR model. For the face image's stimulation transformation along with crop pre-processing, an Optimized Multi-task Cascaded Convolutional Network (OMTCNN) algorithm was utilized. To diminish the FR's computational complexity in the embedded system, a lightweight FR algorithm grounded on Convolutional Neural Network (CNN) was developed, which was denoted as Lightweight CNN (LCNN). Finally, to build the multi-core embedded FRS, OMTCNN and LCNN were combined. The outcomes demonstrated that better performance was exhibited by OMTCNN in determining face identity. However, owing to the utilization of diverse datasets for training and testing, the developed system's accuracy was missed slightly.

Fanny Spagnolo et al. [13] developed a hardware architecture grounded on a real-time cascade classifier along with a resource-limited model to perform object detection. To accomplish the FD task along with integrating with an entire heterogeneous embedded system, architecture was tailored grounded on a Xilinx Zynq-7000 FPGA-centered System-on-Chip. The experiential outcomes exhibit that a high frame rate was significantly attained by the accelerator than other benchmark competitors however occupied much fewer resources.

Serign Modou Bah and Fang Ming [14] suggested a model tackling a few issues hampering FR accuracy to elevate the LBP codes by merging the LBP with augmented image processing schemes like Contrast Adjustment, Bilateral Filter, Histogram Equalization (HE), Image Blending and so on. Thereby, the overall FRS accuracy was elevated. The experiential outcomes exhibit that it was an accurate, reliable and robust model for FRS. Long histograms were engendered by this LBP that result in retarding the developed system's recognition speed.

Rameswari et al.[15] developed an automated access control system utilizing FR. The HOG was utilized to retrieve the features. The output was engendered by analogizing the face function that contained a classifier named SVM to classify the face encoding. To make the system more effective, Radio Frequency Identification (RFID) sensor together with an Infra Red (IR) sensor was implemented. The authority to the campus management was offered by engendering a webpage. The prevailing models were outperformed by the experiential outcomes. But, the HOG computation speed was tardy.

Bouchra Nassih et al. [16] demonstrated an efficient 3D FR approach named GD-FM+RF grounded on Riemannian geometry's Geodesic Distance (GD) along with Random Forest (RF). The GD between the points of 3D faces' specified pairs' was computed by employing the Fast Marching (FM) algorithm. To analyze the class separability, the Principal Component Analysis (PCA) algorithm utilized the retrieved features offered by the geodesic facial curves. Later, the input to the RF classifier was these features. Finally, the recognition rate was elevated along with attained most promising outcomes analogized to the benchmark models. However, while utilizing PCA, the information might be lost.

Yassin Kortli et al. [17] exhibited an implementation grounded on the Advanced RISC Machines Cortex-A9 processor utilizing the OpenCV library. Grounded on a hybrid ARM-Field Programmable Gate Array (ARM-FPGA) platform, a co-design was developed. Here, the FPGA processed the traditional hardware Visible Light communication (VLC) architecture. Optimized hardware VLC architecture was engendered to respect the real-time constraints. It was processed subsequently to comprehend the second Hardware / Software (Hw/Sw) co-design. The experiential outcome exhibited that a high speed was provided by the optimized hardware VLC architecture when analogized

to the conventional non-optimized framework. The developed system's architecture was more complex.

Masoud Muhammed Hassan et al. [18] exhibited an FR approach grounded on the fusion of Gabor-centered FE, Fast Independent Component Analysis (FastICA) and Linear Discriminant Analysis (LDA). To make the uniform face images, it was converted to grayscale and resized. Later, the Gabor, FastICA and LDA methods were utilized to retrieve the facial features as of the aligned face. To recognize the individual's identity, the nearest distance classifier was utilized. Thus the developed model's efficacy was illustrated by the outcomes, especially under the Cosine distance measure. However, more time consumption along with the high computational cost was the model's major drawback owing to the utilization of '3' sophisticated FE schemes.

Lizzie D'cruz and J. Harirajkumar [19] suggested one-shot learning on the Siamese CNN (SCNN) for FR. Here, the employee attendance was entered in the My Structured Query Language (MySQL) database. For uploading the database, Acute Physiology together with Chronic Health Evaluation (APACHE) Web Server was the application programmable interface employed. Utilizing Send Mail Transfer Protocol (SMTP), get well soon messages were sent to absent employees via Gmail. Better performance was observed by the system with accurate FR along with testing '12' images successfully. However, more memory was requisite by the system to collect the data.

Hao Yang and Xiaofeng Han [20] designed an FR attendance system grounded on real-time video processing. A control group together with an experiential group was comprised in this experiment. Traditional fingerprint check-in was utilized by the control group and real-time video processing was utilized in the experiential group. Two universities were chosen as of a province and a similar number of students were chosen for the experiment, which was then collected, counted and analyzed in the experiment, application space. This may develop FR's prospects along with final issues in the actual check-in. The video FRS' effective performance was exhibited by the experiential data. Owing to the extraction of redundant features, the system was still challenging to offer accurate recognition.

Chaoyou Fu et al. [21] developed integrated large-scale visible data's huge identity information into the joint distribution. Additionally, to ensure their identity consistency, a pairwise identity-conserving loss was imposed on the engendered paired heterogeneous images. As of the noises, immense new distinct paired heterogeneous images with a similar identity could be engendered. To train the HFR network, identity consistency along with identity diversity properties permitted the engendered images via a contrastive learning mechanism. Thereby it yielded both domain-invariant as well as discriminative embedding features. The positive pairs and negative pairs regarded here were the engendered paired heterogeneous images and the images acquired as of diverse samplings. On '7' challenging databases that belong to five Heterogeneous FR (HFR) tasks, encompassing Sketch-Photo, Profile-Frontal Photo, Thermal-Visible and ID-Came,

superior performance was attained by the model. Nevertheless, it was more expensive.

Alireza Sepas-Moghaddam et al. [22] presented a CNN grounded a deep Face Expression Recognition (FER) solution, named CapsField incorporated with an extra capsule network that employed dynamic routing to study hierarchical associations between capsules. The spatial features as of the facial images were extracted by CapsField, which learned the angular part-whole relations for a chosen set of 2D sub-aperture images offered from each Light Field (LF) image. Owing to its capacity in learning inter-view along with intra-view relations available in an LF image, the outcomes exhibited that when analogized with benchmark LF-centered FER, a superior performance. However, the system was susceptible to over fitting.

Weipeng Hu et al. [23] introduced an algorithm to explore the cross-modal image's potential domain-invariant Neutral Face (NF) representations by utilizing Dual Face Alignment Learning (DFAL). Feature-level Face Alignment (FFA), Image-level Face Alignment (IFA), as well as Cross-domain compact Representation (CdR) were the '3' effective components included in the model. To encode features for both VIS NF images as well as non- NF images, Teacher-Encoder CNNs (TeEn-CNNs) together with Student-Encoder CNNs (StEn-CNNs) were designed initially. By executing the feature level alignment betweet non- NF and VIS NF the FFA was brought in to learn NF representations. To restore face images, StDe-CNN was engendered to decode features secondarily. By imposing image-level alignment, an NF image was reconstructed by IFA's design. The DFAL approach's effectiveness on '3' challenging Near-Infra Red and VISual (NIR-VIS) databases was illustrated by the experimental outcomes. However, the position and orientation of input were not decoded by CNN.

Jae Young Choi and Bum Shik Lee [24] demonstrated a "Gabor Deep CNN (GDCNN) ensemble" method for FR applications. Distinct along with multiple Gabor face representations were effectively applied as inputs during the GDCNN's training and testing phases. GDCNN ensemble construction as well as GDCNN ensemble combination were the '2' parts comprised in the generated GDCNN ensemble. The experiential outcomes demonstrated that significantly a better FR performance was attained by the approach than the conventional approaches that were grounded only on grayscale or color face images as input illustration. However, more time was acquired by the system owing to the deep learning characteristics of Neural Networks (NN).

Hua Yang et al. [25] designed a Weighted Feature Histogram (WFH) method of Multi-Scale Local Patches (MSLP) utilizing Multi-Bit Local Binary Descriptors (MBLBDs) for FR. Using an MSLP Generation (MSLPG) method, the local patches were extracted to attain multi-scale information. To extract MBLBDs, the MBLBD learning (MBLBDL) method was engendered to diminish the binary descriptors' quantization information loss. To project Pixel Difference Vectors (PDVs) into the MBLBDs in each local patch, a learning mapping matrix along with multi-bit coding rules were engaged in MBLBDL. To discover a set of robust

weights for each patch, the Robust Weight Learning (RWL) method was employed to incorporate the MBLBDs into the final face representation. The outcomes as of experiments performed on face datasets exhibited that benchmark recognition performance was attained by WFH and a Coupled WFH (C-WFH).

Harish Sharma, Garima Hazrati, Jagdish Chand Bansal [28], Spider monkey optimization (SMO) algorithm is a recent addition to the list of swarm intelligence based optimization algorithms. The update equations are based on Euclidean distances among potential solutions. The algorithm has extensively been applied to solve complex optimization problems. SMO is applied to solve optimal capacitor placement and sizing problem in IEEE-14, 30 and 33 test bus systems with the proper allocation of 3 and 5-capacitors. SMO is used for the synthesis of sparse linear arrays. The amplitudes of all the elements and the locations of elements in the extended sparse subarray are optimized by the SMO algorithm to reduce the side lobe levels of the whole array under a set of practical constraints. The SMO has also been used to synthesize the array factor of a linear antenna array and to optimally design an E-shaped patch antenna for wireless applications. SMO is a meta-heuristic technique inspired by the intelligent foraging behavior of spider monkeys. The foraging behavior of spider monkeys is based on the fission-fusion social structure.

Features of this algorithm depend on social organization of a group where a female leader takes decision whether to split or combine. The leader of the entire group is named here as the global leader while the leaders of the small groups are named as local leaders. With reference to the SMO algorithm, the phenomenon of food scarcity is defined by no improvement in the solution. Since SMO is a swarm intelligence based algorithm, each small group should have a minimum number of monkeys. Therefore, at any time if a further fission creating at least one group with less than the minimum number of monkeys, we define it as the time for fusion. In SMO algorithm, a Spider Monkey (SM) represents a potential solution. SMO consists of six phases: Local Leader phase, Global Leader phase, Local Leader Learning Phase, Global Leader Learning phase, Local Leader Decision phase and Global Leader Decision phase.

III. PROPOSED FACIAL RECOGNITION SYSTEM

For the efficient recognition of the human face, a novel FIS-CDNN Classifier is proposed in this paper. The face is detected initially in this system together with extracting the features and landmarks. Subsequently, the features are chosen. Finally, the face is classified utilizing the FIS-CDNN classifier. Fig. 1 illustrates the proposed model's block diagram.
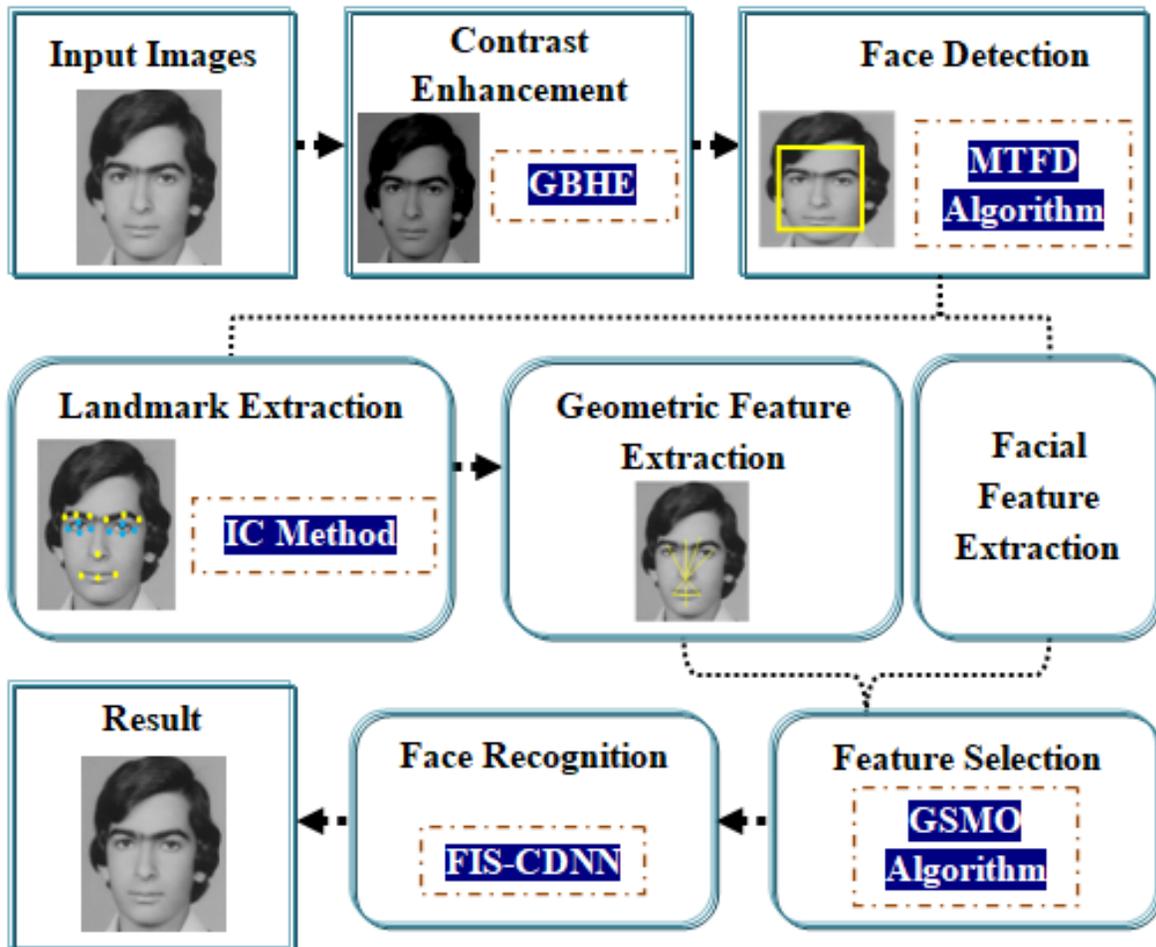
Fig. 1. Block diagram of the proposed methodology

A. Contrast Enhancement

Since the image's quality is low, the face Image (I) is taken as an input and pre-processed in the FRS. By the contrast enhancement technique of Gamma correction-centered Based Histogram Equalization (GBHE), the lower quality is elevated. To augment contrast, HE is a methodology for fine-tuning image intensities. The comparative frequency of various gray levels occurring in the image was represented by the histogram of an image. However, the histogram may bear the intensity saturation issue during the contrast enhancement. So, to control the intensity value, Gamma correction was utilized in this methodology to solve the intensity saturation issue.

A nonlinear operation utilized to encode and decode luminance in the image is Gamma, which is expressed as:

$$G_I = ZI^{\xi} \quad (1)$$

Where, the positive constants are illustrated as Z and I, the gamma correction's output is denoted as G_I, and the gamma value is represented as ξ. Here, the input image's pixel is mapped into the associated pixel of a prepared output image. It is defined as Histogram, which can be expressed as:

$$I_{pre} = \frac{P_{i(G_I)}^{no.oc}}{n(P)} \quad (2)$$

Where, the number of occurrences of a particular pixel with intensity in G_I is denoted as $P_{i(G_I)}^{no.oc}$, the pre-processed image is notated as I_{pre}, the number of pixels is symbolized as $n(P)$ and the image's range of gray level is $[0, L-1]$.

B. Face Detection

The most prominent step in the proposed methodology is face detection. Utilizing the MTFD, the face is extracted by eliminating the unwanted parts from I_{pre}. The Tiny Face Detector is utilized to perform facial extraction on images by implementing the ResNet101 algorithm. In the scale-invariant stage, cosine similarity is utilized rather than Euclidean distance. If '2' similar documents are far distant by the Euclidean distance (owing to the document's size), there may be still a probability of oriented each other as the cosine similarity is having benefits. This is the advantage of cosine similarity.

The input image is shifted into ResNet101 initially. For detecting the face utilizing a pre-trained Artificial Neural Network (ANN), the features are extracted along with amassed in numpy arrays. It expressed the extracted face features as:

$$Q_n = \{Q_1, Q_2, Q_3, ...Q_N\} \quad (3)$$

Where, the number of extracted features is represented as Q_n. By image scale and resolution, the facial extraction process is affected. Between the input image and extracted feature, the similarity score A is computed as:

$$A = \frac{Q_n * I_{pre}}{\|Q_n\|\|I_{pre}\|} \quad (4)$$

The extracted features for FD are valid if the score is identical for extracted features and the input image. The face is not detected otherwise. F denotes the detected face.

C. Feature Extraction

The informative data regarding the face is obtained with the aid of FE. The key features like SIFT, HOGs, Speeded-Up Robust Features (SURF), Binary Robust Independent Elementary Features (BRIEF), LBP, Gabor Feature, Shape, and edge are retrieved as of the face detected region F. The extracted features E_n are expressed as:

$$E_n = \{E_1, E_2, E_3, ...E_N\} \quad (5)$$

1) SURF: It is a fast along with robust algorithm comprising of '4' main parts that are Integral image generation, Fast-Hessian detector, Descriptor orientation assignment and Descriptor generation. By utilizing the integral image for extracting interest points, it is centered on a Hessian-matrix approximation. By applying an approximate Gaussian and derivative scale-space representation, it operates under the entry of integral images. E_1 implies the SURF and the integral image F_{Σ} can be signified as:

$$F_{\Sigma} = \sum_{x=0}^{x \leq p} \sum_{y=0}^{y \leq q} F(x, y) \quad (6)$$

Where, the input image after integral image representation is modeled as $F(x, y)$. For smoothening and space scale representation, a Gaussian filter is also applied. In the pyramid, the scale representation is implemented. For the extraction of the interesting point E_1, the box filter's size doubled the sampling intervals owing to the usage.

2) SIFT: Robustness to illumination, scale and transformation was shown by SIFT. Grounded on specific steps like extremes of scale space detection, determining the feature point's position, feature points direction along with engendering feature descriptors, the features are extracted. To smooth the images, diverse scales of the Gaussian function were utilized by SIFT algorithm. Stable features are chosen along with regarded as a candidate as of the smoothened image. Further, by utilizing the Harris operator, the corner points around the candidate are computed as:

$$E_2 = g(F, S_1, S_2) * F(x, y) \quad (7)$$

Where, the output extracted features are noted as E_2, the Gaussian function is given as $g(F, S_1, S_2)$, the Integral

Length Scale (ILS) is demonstrated as S_1, the scale for the corner point is denoted as S_2, the input image with pixel coordinates are symbolized as $F(x, y)$. If the threshold is lower than the corner points, the corner points are outstanding to represent the object.

3) LBP: A model of retrieving textual features F as of the face detected image F is the LBP, which is a non-parametric operator describing an image's local spatial structure in binary information. Classification, detection and reorganization are the three steps performed by this. The face image's centre pixel is observed as a threshold value along with analogized with the neighbouring pixels. It is denoted as '1' if the neighbouring pixel's value is superior to the threshold value. Otherwise, it is mentioned as 0. An 8-bit value was obtained by merging all these outcomes. Then, the local texture feature's (E_3) LBP value was obtained as:

$$E_3 = \sum_{i=0}^{u-1} \mathrm{T}(P_{(i)\in F} - P_{(c)\in F})2^i \tag{8}$$

$$\mathrm{T}\left(P_{(i)\in F} - P_{(c)\in F}\right) = \begin{cases} 1 & P_{(i)\in F} \ge P_{(c)\in F} \\ 0 & P_{(i)\in F} < P_{(c)\in F} \end{cases} \tag{9}$$

Where, the central pixel value is represented as E_3, the neighboring pixel value is proffered as $P_{(i)\in F}$, the threshold function is modeled as T, the number of pixels in the neighbor of $P_{(c)\in F}$ is notated as u and the centre pixel is signified as $P_{(c)\in F}$.

4) HOG: The gradient structure features E_4 is captured by utilizing HOG. Descriptor blocks, block normalization, Orientation binning as well as Gradient computation are the '4' steps involved here. For each pixel, the face image's gradient is computed. The pixel value's magnitude and direction in each block are computed utilizing the gradient. As the image's few parts might show a slight vanish, block normalization is implemented to oversight the image's brightness. Normalization is employed in specific parts to diminish that. The mathematical representation is exhibited as:

$$E_4 = \frac{v}{\|v\|^2 + \sigma} \tag{10}$$

Where, the constant value to avert deviation is σ, the non-normalized vector is signified as v. Conclusively as of the normalization operation, the HOG features are derived along with obtaining the features E_4.

5) BRIEF: A general-purpose feature point descriptor amalgamated with arbitrary detectors is named BRIEF, which can be notated as E_5. Detection, descriptor and descriptor matching were '3'

steps incorporated. By smoothing the image utilizing a Gaussian filter, the BRIEF commenced averting the descriptor from being sensitive to high frequency. Then, the random pair of pixels was chosen by the BRIEF around the targeted region. Comparisons may occur between utilizing these pixels. The value of 1 is assigned to the corresponding bit if the initial pixel is found to be brighter compared to the second pixel. Or else it is found 0.

$$E_5 = \begin{cases} 1 & F(a) \le F(b) \\ 0 & F(a) \ge F(a) \end{cases} \tag{11}$$

Where, the two selected points are notated as a and b.

6) Shape: The face image's geometric properties were referred to by the shape feature that is utilized to describe the image content. Grounded on the shape boundary information or boundary plus interior elements where the area, parameter and circularity are deemed as key features, the shape features are extracted. It could be computed as:

$$E_6 = \sum_{x, y} F(x', y') \tag{12}$$

Where, the output of the face's shape feature is represented as E_7, the input face image with the extreme pixels is notated as $F(x', y')$.

7) Edge: The input image's edges E_7 is identified by the Canny algorithm. By utilizing noise removal, differentiation, non-maxima suppression, double thresholding and edge tracking. The edge features are acquired as of the image. These images are smoothened initially and gradients ensure the image's edges are taken. By computing the magnitude together with gradients of directions, the points which are not at the maximal are suppressed in the non-maxima suppression. By employing thresholds namely upper threshold and lower threshold in double thresholding, the strong along with weak edges are identified. The edge tracking is aided in determining the final edges. '2' thresholds namely the upper threshold (h_{high}) and lower threshold (L_{low}) are taken in the tracking. The edges E_8 are classified as strong edge (S_{edge}), weak edge (W_{edge}) and non-edge (U_{edge}) along with deciding the concerning gradients under three conditions that are:

$$E_7 = \begin{cases} S_{edge} & if\ F(t, v) > L_{low} \\ W_{edge} & if\ L_{low} < F(t, v) < h_{high} \\ U_{edge} & if\ L_{low} > F(t, v) \end{cases} \tag{13}$$

Where, the gradient in the gradient direction is proffered as $F(t, v)$.

D. Landmark Extraction

Utilizing the Improved Chehra (IC) landmark Extraction method, the face's landmark is extracted as of F. By training a cascade of regressors, this discriminative facial deformable model was obtained. For extracting the Face appearance, the Chehra method utilized PCA in normal. A drawback here is the original features turn into Principal Components (PC) after implementing PCA on the dataset. The linear combination of original features is the PC. As original features, PCs are not readable and interpretable; also it took more time for executing multiclass datasets. So, LDA is employed in the Face appearance stage rather than this.

For facial region extraction, 49 facial landmark points were computed by the Chehra model on the whole facial image. The face image's 10 regions computed following the landmark detection are the forehead region, right eye region, right face region, lower nose region, left eye region, entire eye region, lip region, left face region and upper nose region. For facial appearance, FE has been executed by utilizing LDA after obtaining the blocks.

LDA is a feature diminution technique grounded on supervised linear transformation. It does so in a way that elevates the ratio between the intra-class scatter $\wp_{\text{int }ra}$ and the inter-class scatter $\wp_{\text{int }er}$ when projecting a high-dimensional feature vector onto a low-dimensional space. For multiclass l, the scatters are determined as follows:

$$\wp_{\text{int }ra} \in F == \sum_{i=1}^{l} \sum_{\Omega \in F} \left(X_\Omega - \mu_i\right)\left(X_\Omega - \mu_i\right)^T \tag{14}$$

$$\wp_{\text{int }er} \in F = \sum_{i=1}^{l} \left(\mu_i - \mu\right)\left(\mu_i - \mu\right)^T \tag{15}$$

Where, the average landmark point in i^{th} class is represented as μ_i, the Ωth instance in i^{th} class is indicated as X_Ω. The best projection is then computed along with maximizing the ratio between the projected samples' intra-class scatter matrix and inter-class scatter matrix. It is defined as:

$$X^{best} = \arg\max(X) \frac{\left|X^T \wp_{\text{int }er} X\right|}{\left|X^T \wp_{\text{int }ra} X\right|} \tag{16}$$

Where, the best projection point is denoted as X^{best}, the sample point is denoted as X. For further process, the projects are taken as landmark features.

E. Geometric Feature

The face's GF is extracted as of the X^{best} in this phase. The utilization of the distance between the facial landmark points as features is allowed owing to the facial geometry's invariance to the illumination. Hence, the feature utilized here

is the distance between landmarks. The GFs like height, width, length of chin, eyebrows, nose, mouth etc. are extracted. The extracted GF is expressed as:

$$G_n \in X^{best} = \left\{G_1, G_2, G_3, \ldots\ldots\ldots G_N\right\} \tag{17}$$

Where, the number of extracted GFs is signified as G_n. The extracted facial together with GFs is amalgamated and the output can be expressed as:

$$\lambda_n = G_n \cup E_n \tag{18}$$

Where, the combined features are represented as λ.

F. Feature Selection

For elevating the proposed facial detection system's learning accuracy, the features are chosen as of λ_n utilizing the GSMO Algorithm after FE. A meta-heuristic method named SMO is motivated by the clever foraging behavior of spider monkeys. The Fission-Fusion Social Structure (FFSS) serves as the base for the foraging habits of spider monkeys. The group's social organization is in such a way that the female leader makes a decision on whether to split or combine the group. In cases of food scarcity, a female creates mutable smaller groups as well as typically serves as the swarm's Global Leader (GL). Availability of food as of a specific territory is the factor upon which the group sizes rely. A direct proportionality is identified between the size and food availability. The GL deemed here is the entire group's leader and the smaller group's heads are named the Local Leaders (LLs). The variation between the current position and positions engendered at random is averaged out to update the leader position. For a specific issue, a random position is engendered within a predetermined range. As a result, the research proposed the alteration to quicken convergence and boost reliability. The updating stage uses the Gaussian Mutation Technique rather than the Convergence rate.

The number of monkeys deemed here were the extracted features. With the uniform distribution, each Spider Monkey (SM) is initialized as:

$$E_i \in \lambda_n = E_{\min(j)} + \Omega(0,1) \times \left(E_{\max(j)} - E_{\min(j)}\right) \tag{19}$$

Where, E_i represents i^{th} SM, the lower and upper limits in j^{th} dimension are indicated as $E_{\min(j)}$, and $E_{\max(j)}$, the random number distributed uniformly in the range of $[0,1]$ is signified as Ω.

The following '6' steps are followed by an algorithm for foraging after initialization. The steps are explained as follows,

a) Local leader phase (LLP): Utilizing the past behaviours of both the LL and the local group members, the SM in LLP shifts its present location. Only when the new location's fitness value is greater than the prior site, the SM's location is updated with the new location. Using the Gaussian

mutation technique E_{ij}^*, the ith SM of the dth local group's location update equation is:

$$E_{ij}^* = E_{ij} + \Omega(G) \qquad (20)$$

Where, the jth dimension of ith SM is proffered as E_{ij}, the random Gaussian number is symbolized as $\Omega(G)$.

b) Global leader phase (GLP): Similar to the LLP, but diversely, the swarm updates throughout this phase. Only one randomly chosen solution's dimension is updated here, unlike the LLP. Depending on its probability, which spider monkey will have a probability of being updated is identified.

$$E_{ij}^* = E_{ij} + \Omega(0,1) \times \left(\Re_j - E_{ij}\right) + \Omega(-1,1) \times \left(E_{Rj} - E_{ij}\right)$$
$$(21)$$

Where, the GL position in j^{th} dimension is notated as $\Re_j$. Grounded on the probability that is derived utilizing their fitness, which is computed centered on the classifier's accuracy, the SMs location is updated. The probability of ith SM Φ is calculated as:

$$\Phi = \frac{9}{10} \times \frac{f_i}{f^{max}} + \frac{1}{10} \qquad (22)$$

Where, the fitness of ith SM is illustrated as f_i, the maximum fitness is depicted as f^{max}. The SM's fitness for the freshly engendered position is then analogized to the old position, and a better position is adopted.

c) Global leader learning phase (GLLP): Here, an SM is updated as the swarm's GL whose position has the maximum fitness value of all the members. The global limit count is elevated by 1 if the position of GL is not updated, representing the number of iterations during which the GLs position was not updated.

d) Local leader learning phase (LLLP): At this point, the group's greedy selection is utilized to update the LL's position. The LL's upgraded position is then coordinated with the previous one. The Local Limit Count (LLC) is then elevated by 1 if the LL position has not been updated.

e) Local leader decision phase (LLDP): In this phase, either by random initialization or by means of collective information as of GLs experience utilizing the equation below, the position of every member in that group is updated if any LL does not become restructured to a specific verge known as LL Limit.

$$E_{ij}^* = E_{ij} + \Omega(0,1) \times \left(\Re_j - E_{ij}\right) + \Omega(-1,1) \times \left(E_{ij} - \aleph_{dj}\right)$$
$$(23)$$

f) Global leader decision phase (GLDP): LLD and GLD are equivalent. The GL splits the swarm into a lower maximum number of groups if it does not restructure to a specific boundary known as the GL limit. At that point, the LL was chosen by commencing the LL Learning procedure. Additionally, the GL will unite the entire group if the maximal numbers of groups are engendered but the GLs position is not updated. Until enough food is obtained to feed every SM, the procedure is unremitting. Several features are chosen in this method of opting for leaders, which are expressed as:

$$E_n^{sel} = \left\{E_1^{sel}, E_2^{sel}, E_3^{sel}, \ldots\ldots E_N^{sel}\right\} \qquad (24)$$
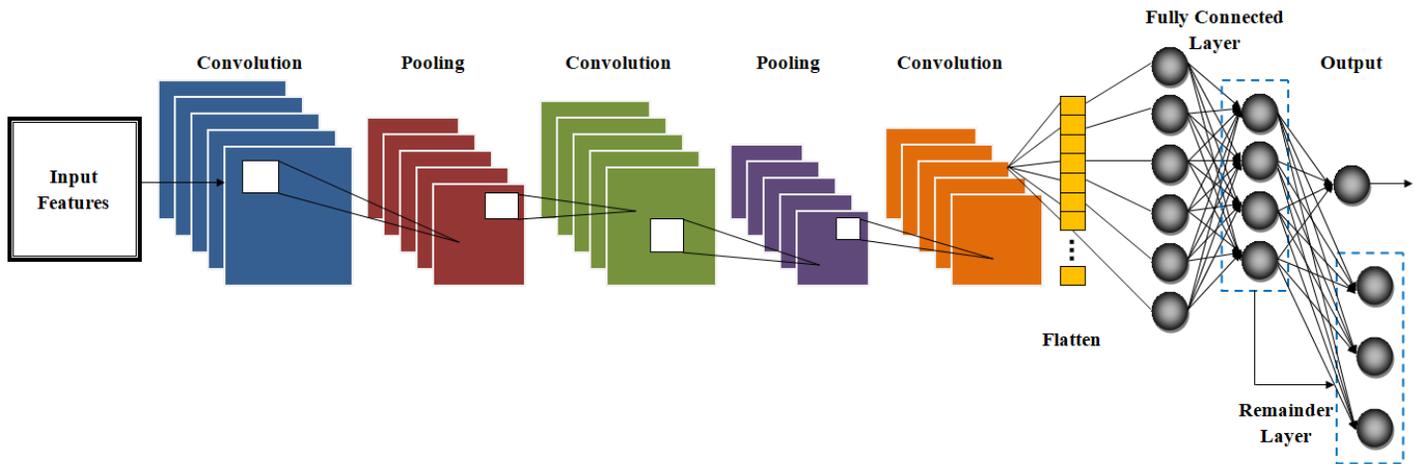
Fig. 2. Architecture of proposed FIS-CDNN

Where, the number of selected features is denoted as E_n^{sel}. Likewise, utilizing a similar algorithm, the features G_n are chosen. Followed by, it is utilized for choosing features E_n along with obtaining the output G_n^{sel}.

G. Face Recognition

Here, for detecting the human's face, the chosen features E_n^{sel} are fed as input to the FIS-CDNN classifier. To process the image's Red, Green and Blue elements at the same time, a 3D NN was utilized by a Convolutional Deep Neural Network (CDNN). Compared to traditional feed-forward NN, the

number of artificial neurons requisite to process an image is diminished considerably. Normally, for the reason of weight value selection, the classifier suffers from the accuracy issue. The remainder layer comprising the previous training images behavior is deemed in this proposed methodology to solve this issue. The previous training weight values are amassed in the remaining layer. Therefore, the weight selection is very much aided by the remainder layer along with elevating the accuracy. Using the Fuzzy Interference system (FIS), the weight value is initialized in the proposed classifier. Fig. 2 exhibits the proposed system's architecture.

a) Weight initialization: By utilizing the five layers, the weights are initialized by the FIS system. The fuzzification process is executed by the first layer, which contains the adaptive nodes. The fuzzified output Ψ_i is:

$$\Psi_i = \nabla_1\left(\Omega_{\tau_k}\right) \quad (25)$$

$$\Psi_i = \nabla_2\left(\Omega_{\tau_j}\right) \quad (26)$$

Where, the input nodes are notated as ∇_1 and ∇_2, the weight value acquired as of the previous iteration is symbolized as τ_k and τ_j, the membership function is represented as Ω, which indicates how much given nodes satisfy the quantifier.

The output signals acquired as of the preceding layers are multiplied in the second layer along with the second layer's ϖ_i output is:

$$\varpi_i = \nabla_1\left(\Omega_{\tau_k}\right) * \nabla_2\left(\Omega_{\tau_j}\right) \quad (27)$$

For normalizing the second layer's output, normalization is executed in the third layer and is expressed as:

$$\left(\varpi_i\right)^* = \frac{\varpi_i}{\varpi_1 + \varpi_2} \quad where, i = 1,2 \quad (28)$$

Where, the normalized features are specified as $\left(\varpi_i\right)^*$. The fuzzy rule's consequent part is executed by the fourth layer that contains the adaptive nodes. Node function has a form:

$$\left(\bar{\varpi}_i\right)^* = \left(\varpi_i\right)^*\left(\varepsilon_i\nabla_1 + \vartheta_i\nabla_2 + \ell_i\right) \quad (29)$$

Where, the linear adaptive parameters are symbolized as ε_i, ϑ_i and ℓ_i, defuzzification is proffered as $\left(\bar{\varpi}_i\right)^*$.

Finally, the proposed system's total weight value W is computed in the final layer as:

$$W = \sum\left(\varpi_i\right)^*\left(\varepsilon_i\nabla_1 + \vartheta_i\nabla_2 + \ell_i\right) \quad (30)$$

b) Convolution layer: A specialized sort of linear operation employed for FE is convolution. Here, a kernel is a small array of numbers applied across the input. At each tensor's location, an element-wise product between kernel's each element and the input array is computed. It is again summed to acquire the output value in the output array's corresponding position. By applying multiple kernels, this procedure is recurring to generate an arbitrary number of feature maps representing the input array's diverse characteristics. The given input data's convolution is mathematically expressed as:

$$c = \sum_d \sum_d \left(E_n^{sel}\right)(\alpha - d, \beta - d) * W(d,d) \quad (31)$$

Where, the convolution operation's output is signified as c, the weight value in the dimension size $d \times d$ is represented as $w(d,d)$, the input matrix's dimension size is α and β. Each convolution layer is specifically pursued by a non-linear Activation Function (AF) implemented element-wise to the preceding layer's output to permit the network to learn non-linear decision boundaries. ReLu is the activation function utilized in this work. The activation function is mathematically expressed as:

$$R = \max(0, c) \quad (32)$$

Where, the ReLu activation function's output is signified as R.

c) Pooling layer: By passing via the AF, the output is obtained, which is again passed to the pooling layer to aggregate the information along with diminishing the representation. The pooling operation's (Δ) result is:

$$\Delta = \frac{c - W}{S} + 1 \quad (33)$$

Where, the kernel's strides are notated as S.

d) Fully connected layer (FCL): The final convolution or pooling layer's output feature maps are transferred into a single-dimensional array of numbers, which means typically flattened. By a learnable weight, each input is allied to every output in this layer. The amount of output nodes contained in the final FCL is typically similar to the number of classes. Then, the flattened output is computed as:

$$\Gamma = \Delta - \left(W(d \times d) - 1\right) \quad (34)$$

Where, the FCLs output is indicated as Γ.

e) Softmax layer: To the softmax activation function, the FCL's output is fed directly into this layer. In the output layer, the activation function is utilized along with normalizing the output real values in the range of $[0,1]$ as of the last FCL to target class probabilities. The softmax function Γ_{soft} is given as:

$$\Gamma_{soft} = \frac{e^{\Gamma_\varphi}}{\sum_{\varphi=1}^{h} \Gamma_\varphi} \quad (35)$$

Where, the FCL output at φ^{th} node is notated as Γ_φ, the total number of output nodes is represented as h. With higher accuracy, the faces are recognized by this process. The proposed FIS-CDNN's pseudo code is:

Input: selected features E_n^{sel}

Output: Recognized face I

Begin
Initialize parameters c, R, Γ, layer b
Compute weight value
$$W = \sum (\varpi_i)^* (\varepsilon_i \nabla_1 + \vartheta_i \nabla_2 + \ell_i)$$
For $b = 1 \, to \, n$

while $b = 1$ (first round of convolution and pooling layer)
Compute convolution operation c
Evaluate activation function
$$R = \max(0, c)$$
Compute pooling operation Γ
End while
End for
Flattening all the layers

Evaluate softmax activation function Γ_{soft}

End begin
Return I

IV. RESULTS AND DISCUSSION

The proposed facial recognition system's performance with FIS-CDNN is analyzed in this section. MATLAB (MATrix LABoratory) is the working platform in which the model is implemented. It is a 4th generation programming language with a multi-paradigm numerical computing environment. For rapid along with simple scientific computations and I/O, it is designed particularly.

A. Database Description

FG-NET, collected from a publically available source, is the dataset used in the proposed framework. The dataset for age assessment along with FR across ages is FG-NET. A total of 1,002 images obtained as of 82 people with an age ranging between 0 to 69 and an age gap of up to 45 years were composed in the dataset.

B. Performance Analysis of Classification Method

Grounded on a few quality measures along with prevailing models like Z- Normalization and Moore Penrose-based Deep CNN (ZMP-DCNN) [26], NN, CNN and Adaptive Neuro-Fuzzy Inference System (ANFIS), the proposed FIS-CDNN method's effectiveness is examined here.

TABLE I. A COMPARATIVE ANALYSIS OF THE PROPOSED MODEL AND EXISTING MODEL WITH RESPECT TO TP, TN, FP AND FN

Classifier	TP	TN	FP	FN
FIS-CDNN	1.853659	161.8537	0.146341	0.146341
ZMP-DCNN	1.52439	161.5244	0.47561	0.47561
CNN	1.487805	161.4878	0.512195	0.512195
NN	1.317073	161.3171	0.682927	0.682927
ANFIS	1.402439	161.4024	0.597561	0.597561

Table I summarises the performance analysis of the proposed FIS-CDNN and the prevailing models Regarding True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). The proposed model accurately recognized the face in the outcome named TP. The mismatched face is predicted by the proposed model in the outcome TN. Hence, the proposed model's TP is 1.853659, which is higher than the prevailing techniques. Similarly, 161.8537 is the proposed FIS-CDNN's TN. The face is recognized incorrectly by the model in an outcome named FP. Here, the proposed model's FP and FN are similar at 0.146341. The face is classified more effectively by the proposed model than the prevailing models, which is obtained as of the entire analysis.
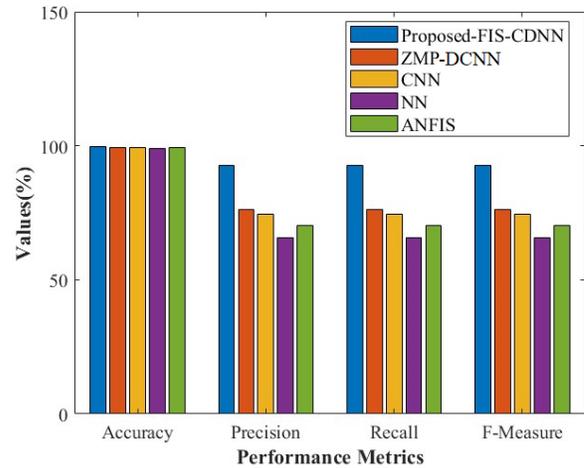
Fig. 3. The performance of FIS-CDNN with the existing methods

The proposed FIS-CDNN and prevailing method's performance analysis are illustrated in Fig. 3, grounded on the Accuracy, Precision, Recall along with F-Measure which is defined as the amalgamation of precision as well as recall. The proposed system's higher performance was exhibited by the proposed method's higher value in every metric. The face was recognized by the proposed method with 99.82% Accuracy. It is higher than the prevailing methods and provided better performance. Likewise, 92.68% is the proposed FIS-CDNN's Precision and Recall value. An F-Measure of 16.47% higher than the existing ZMP-DCNN was exhibited by the proposed one.
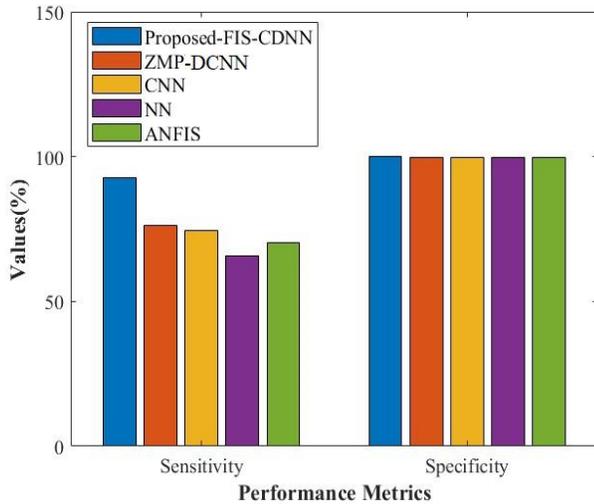
Fig. 4.   The performance of FIS-CDNN with the existing methods

The proposed FIS-CDNN and the prevailing method's performance analysis are depicted in Fig. 4, grounded on Sensitivity and Specificity. The degree of similarity between views of an object that evokes an exact recognition response as of a given subject is Specificity. The capability to classify a test between the True and False positive performance correctly is named Sensitivity. While the ZMP-DCNN attains 99.7%, and CNN attains 99.6%, the proposed model's Specificity is 99.9%, which is superior to the prevailing models. Correspondingly, a higher sensitivity of 18.29% was attained by the proposed scheme than the prevailing CNN.
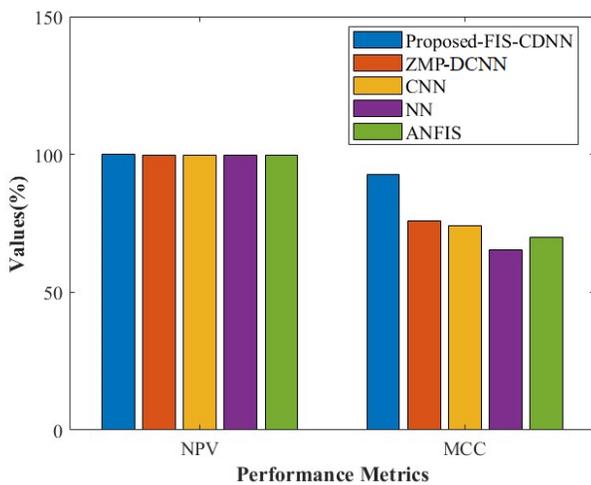
Fig. 5.   The performance of FIS-CDNN in terms of NPV and MCC

The proposed model's performance is demonstrated in Fig. 5, on the basis of Negative Predictive Value (NPV) together with Matthew's Correlation Coefficient (MCC). The proposed model's better performance is proved by higher NPV and MCC. 99.9% is the proposed model's NPV, which is superior to the ZMP-DCNN and all other prevailing models. Likewise, 92.59% is the proposed scheme's MCC.

TABLE II.   THE PERFORMANCE OF FIS-CDNN IN TERMS OF FPR, FNR, FRR, FDR

Classifier	FPR	FNR	FRR	FDR
FIS-CDNN	0.000903	0.073171	0.073171	0.073171
ZMP-DCNN	0.002936	0.237805	0.237805	0.237805
CNN	0.003162	0.256098	0.256098	0.256098
NN	0.004216	0.341463	0.341463	0.341463
ANFIS	0.003689	0.29878	0.29878	0.29878

Table II summarizes the proposed FIS-CDNN and the prevailing model's performance analysis, Concerning False Positive Rate (FPR), False Negative Rate (FNR), FRR and FDR. The proposed scheme's higher performance was exhibited by lower values of FPR, FNR and FRR. The proposed model's FPR is 0.002033 and 0.002259 lower than ZMP-DCNN and NN respectively. But, it is 0.002786 higher than ANFIS. Similarly, the proposed methodology's FNR, FRR and FDR are identical along with exhibiting an augmentation by 0.073171. This exposed that it is lower than the prevailing CNN along with other prevailing models. So, it is concluded from the metrics that better performance was exhibited by the proposed model for facial recognition.
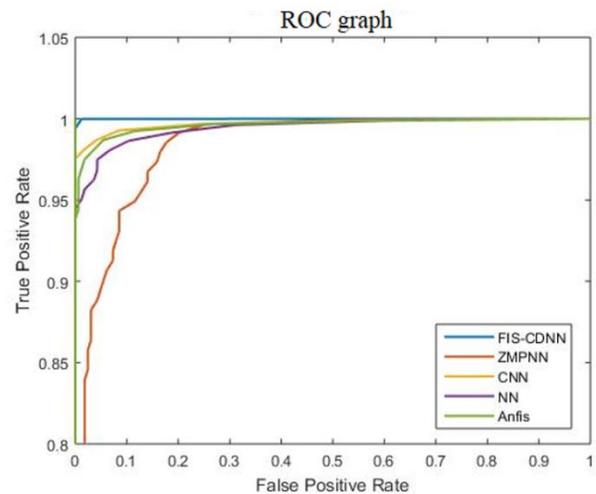
Fig. 6.   ROC curve of proposed FIS-CDNN and existing model

Fig. 6 exhibits the FIS-CDNN and the existing model's ROC curve analysis. The facial recognition's performance or quality was evaluated by an effective model named ROC. The True Positive Rate (TPR) is plotted against the False Positive Rate (FPR) to engender a ROC curve. The ROC curve illustrates that higher sensitivity together with specificity than the prevailing models like ZMP-DCNN, CNN, NN and ANFIS was exhibited by the proposed model.

V.   CONCLUSION

A novel FIS-CDNN classifier was proposed in this work for facial recognition. Pre-processing, FD, facial FE, face landmark, geometric FE, feature selection and classification are all processes involved in the proposed technique. The experimental analysis is performed after all the steps have been completed, and regarding some quality criteria, the proposed

FIS-CDNN's performance is analogized with that of the prevailing approaches. The final outcomes exposed that a 99.82% accuracy rate was attained by the proposed model. Similar outcomes are obtained for all other metrics, including sensitivity, specificity, recall, precision, recall, recall, f-measure, NPV, MCC, FDR, FPR, FNR and FRR by the proposed scheme. The proposed model is therefore found to be significantly more effective than the prevailing approaches for classifying faces grounded on the findings of all metrics. This work will be extended in the future by extracting features as of various head poses employing utilizing advanced models.

REFERENCES

[1] Kolipaka Preethi and swathy vodithala, "Automated smart attendance system using face recognition", Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021), 06-08 May 2021, Madurai, India, 2021.

[2] Naing Min Tun, Alexander I Gavrilov and Pyae Phyo Paing, "Human face recognition using combination of ZCA feature extraction method and deep neural network", International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 25-29 March 2019, Sochi, Russia, 2019.

[3] Pallavaram Venkateswar Lal, Gnaneswarar Rao Nitta and Ande Prasad, "Ensemble of texture and shape descriptors using support vector machine classification for face recognition", Journal of Ambient Intelligence and Humanized Computing, 2019, https://link.springer.com/article/10.1007/s12652-019-01192-7.

[4] Asif Mohammed Arfi, Debasish Bal, Mohammad Anisul Hasan, Naeemul Islam and Yasir Arafat, "Real time human face detection and recognition based on haar features", IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020, Dhaka, Bangladesh, 2020.

[5] Jannat Binta Alam, Momarul Islam Md, Taskeed Jabid and Sabbir Ahmed, "System development using face recognition", International Conference on Automation, Computational and Technology Management (ICACTM), 24-26 April 2019, London, UK, 2019.

[6] Alireza Sepas-Moghaddam, Ali Etemad, Fernando Pereira and Paulo Lobato Correia, "CapsField: Light field-based face and expression recognition in the wild using capsule routing", IEEE Transactions on Image Processing, vol. 30, pp. 2627-2642, 2021.

[7] Belen Rios-Sanchez, David Costa-da-Silva, Natalia Martin-Yuste and Carmen Sanchez-Avila, "Deep learning for facial recognition on single sample per person scenarios with varied capturing conditions", Applied Sciences, vol. 9, no. 24, pp. 1-11, 2019.

[8] Eric-Juwei Cheng, Kuang-Pen Chou, Shantanu Rajora, Bo-Hao Jin, Tanveer M, Chin-Teng Lin, Ku-Young Young, Wen-Chieh Lin and Mukesh Prasad, "Deep sparse representation classifier for facial recognition and detection system", Pattern Recognition Letters, vol. 125, pp. 71-77, 2019.

[9] Farah Deeba, Hira Memon, Fayaz Ali Dharejo, Aftab Ahmed and Abddul Ghaffar, "LBPH-based enhanced real-time face recognition", (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 10, no. 5, pp. 274-280, 2019.

[10] Genevieve M Sapijaszko and Wasfy B Mikhael, "Facial recognition system using mixed transform and multilayer sigmoid neural network classifier", Circuits, Systems and Signal Processing, vol. 39, no. 1, pp. 6142-6161, 2020.

[11] Shokoufeh Mousavi, Mostafa Charmi and Hossein Hassanpoor, "A distinctive landmark-based face recognition system for identical twins by extracting novel weighted features", Computers and Electrical Engineering, vol. 94, pp. 1-17, 2021.

[12] Xue Lv, Mingxia Su and Zekun Wang, "Application of face recognition method under deep learning algorithm in embedded systems", Microprocessors and Microsystems (Pre-proof), 2021, https://doi.org/10.1016/j.micpro.2021.104034.

[13] Fanny Spagnolo, Stefania Perri and Pasquale Corsonello, "Design of a real-time face detection architecture for heterogeneous systems-on-chips", Integration, vol. 74, pp. 1-10, 2020.

[14] Serign Modou Bah and Fang Ming, "An improved face recognition algorithm and its application in attendance management system", Array, vol. 5, pp. 1-7, 2020.

[15] Rameswari R, Naveen Kumar S, Abishek Aananth M and Deepak C, "Automated access control system using face recognition", Materials Today: Proceedings, vol. 45, no. 2, pp. 1251-1256, 2020.

[16] Bouchra Nassih, Aouatif Amine, Mohammed Ngadi, Youssef Azdoud, Driss Naji and Nabil Hmina, "An efficient three-dimensional face recognition system based random forest and geodesic curves", Computational Geometry: Theory and Applications, vol. 97, no. 1, pp. 1-8, 2021.

[17] Yassin Kortli, Maher Jridi, Mehrez Merzougui and Mohamed Atri, "Optical face detection and recognition system on low-end-low-cost Xilinx Zynq SoC", Optik, vol. 217, pp. 1-16, 2020.

[18] Masoud Muhammed Hassan, Haval Ismael Hussein, Adel Sabry Eesa and Ramadhan J. Mstafa, "Face recognition based on gabor feature extraction followed by FastICA and LDA", Computers, Materials & Continua, vol. 68, no. 2, pp. 1637-1659, 2021.

[19] Lizzie D'cruz and Harirajkumar J, "Contactless attendance system using Siamese neural network based face recognition", Materials Today: Proceedings (In Press), 2020, https://doi.org/10.1016/j.matpr.2020.10.462.

[20] Hao Yang and Xiaofeng Han, "Face recognition attendance system based on real-time video processing", IEEE Access, vol. 8, pp. 159143-159150, 2020.

[21] Chaoyou Fu, Xiang Wu, Yibo Hu, Huaibo Huang and Ran He, "DVG-Face: Dual variational generation for heterogeneous face recognition", Journal of Latex Class Files, 2020, https://doi.org/10.48550/arXiv.2009.09399.

[22] Alireza Sepas-Moghaddam, Ali Etemad, Fernando Pereira and Paulo Lobato Correia, "CapsField: Light field-based face and expression recognition in the wild using capsule routing", IEEE Transactions on Image Processing, vol. 30, pp. 2627-2642, 2021.

[23] Weipeng Hu, Wenjun Yan, Haifeng Hu, "Dual face alignment learning network for NIR-VIS face recognition", IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 4, pp. 2411-2424, 2021.

[24] Jae Young Choi and Bum Shik Lee, "Ensemble of deep convolutional neural networks with gabor face representations for face recognition", IEEE Transactions on Image Processing, 2019, http://dx.doi.org/10.1109/TIP.2019.2958404.

[25] Hua Yang, Chenting Gong, Kaiji Huang, Kaiyou Song and Zhouping Yin, "Weighted feature histogram of multi-scale local patch using multi-bit binary descriptor for face recognition", IEEE Transactions on Image Processing, vol. 30, pp. 3858-3871, 2021.

[26] Santhosh S., Dr. S. V. Rajashekararadhya, Efficient Face Recognition System Using Z-Normalization and Moore Penrose-Based Deep Convolutional Neural Network, Advanced Engineering Science, vol. 54, Issue 02, pp. 3177-3202, 2022.

[27] Mohd Najib Mohd Salleh, Noureen Talpur and Kashif Hussain, Adaptive Neuro-Fuzzy Inference System: Overview, Strengths, Limitations, and Solutions, Conference Paper · August 2017, DOI: 10.1007/978-3-319-61845-6_52.

[28] Harish Sharma, Garima Hazrati, Jagdish Chand Bansal, "Spider Monkey Optimization Algorithm", Chapter in Studies in Computational Intelligence · January 2019, DOI: 10.1007/978-3-319-91341-4_4.