International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-11 Issue-5, June 2022

Facial Expression Recognition using Robust Algorithm based on Modern Machine Learning Technique

Shashank M Gowda, H N Suresh

Abstract*: *Recently facial expression recognition has turned out to be an interesting field in research because of more demand for security and the advancement of mobile devices. Due to many serious incidents like terrorists' attack, there arises more concern to develop the security systems mainly in certain places like airports and border crossings where identification and verification are mandatory. On the other hand, these surveillance systems aid to identify the missing person, even though it is based on robust facial expression recognition algorithms and on the developed database for facial expression recognition. However, the human faces are complex and multidimensional which make the facial gesture extraction to be very challenging. Obviously, in high secured applications facial expression recognition (FER) systems are mandatory to avoid incidents. In this paper, the automatic facial expression recognition system is developed based on the machine learning algorithms for classification. This research reveals the identification of FER for the ease of communication. Hybridization of Adaptive Kernel function based Extreme Learning Machine with Chicken Swarm Optimization (HAKELM-CSO) algorithm is introduced for identifying the accurate facial expression among the large database. In this work, an approach is developed by applying the machine learning techniques for the automated classification on the image region. The major purpose of this research work is to overcome the flaws of traditional algorithms and to improve the process of facial expression recognition which could be used in various applications.

Keywords*: *Facial, Machine Learning, HAKELM-CSO, Pre-processing, Weiner Filter.

I. INTRODUCTION

In the vast wide world, and today's paradigm of shifting times, we people are talking about equality in each and every case possible. Though we consider animals to be our fellow creatures of social being, with sympathy towards them, there is one specific thing that varies humans from them, that is sensual emotions and channelizing them in the facial expressions. Human being, the species of "Homo sapiens" is known for the peculiar indexing of face with respect to the surroundings and situations. A wide saying is that "Humans are emotional beings and face is the index of their minds". This being said and provided, the technology outgrowth in our hands and the diverse rapidisation in the area of computer vision and artificial intelligence(AI), especially facial expression recognition is a soaring upgrade to present day systems. With over 55% of non-verbal cues appear from the lack of speaker (Communicator) that influences the receiver (interpreter), that would be highly effective when the recognition happens to be done in that state itself. To overcome this problem and to put up a milestone in the developing fields, recognition of facial expressions remains the best answer for this enigmatic problem in face recognition and detection by utilization of high end image processing algorithm. The basic laid out, 'six universal expressions' that are portrayed widely in the faces, can be identified and the recognition can be done for the faces captured at any time with any provided expression in the face and this makes it easier, efficient and effective to any labyrinthine faces with expressions not just for face detection but also in other technology which simply reveals the process of image in and image out. Over a decade the image processing has been revealed a remarkable progress on several applications like Human Computer Interface (HCI), biometric investigation, content-based coding of images and video surveillance advanced technology ontogenesis.It is accomplished with the use of digital technology to manipulate multi-dimensional signals from simple digital circuits to highly developed parallel computers, in which the image processing techniques plays a vital role in modern digital. HCI aims to put into practice of knowledge from behavioural and social sciences into machines depend on interdisciplinary researches which steps forward the machines to contribute significantly to the human experience. General tendency of HCI systems [1] is to generate an automatic recognition system, which is able to understand the environmental and to interact with the existent system. The face is the basic part of day to day interpersonal communication. Facial expressions symbolize the changes of facial appearance with respect to an individual inner emotional state, social communications or intentions [2]. The facial expression is the best and instant way for humans for conveying the emotions and to express the intentions by natural and nonverbal. When compared to verbalization facial expression is the quicker way to communicate the emotions through facial expressions. For human beings the emotions are irrational and intelligent behaviour which plays a significant role in everyday life. The inner emotion is a strong feeling shows one's situation, mood, or relationships[3]. It is a mental state that does not occur often which is accompanied by physiological changes which is needed to be monitored.

Manuscript received on 20 April 2022.
Revised Manuscript received on 25 April 2022.
Manuscript published on 30 June 2022.
* Correspondence Author
Prof. Shashank M Gowda*, Assistant Professor, Department of Electronics and Communication Engineering, Yenepoya Institute of Technology, Moodbidri (Karnataka), India. Email: shashank@yit.edu.in
Dr. H. N. Suresh, Professor and Head, Department of Electronics & Instrumentation Engineering, Bangalore Institute of Technology, Bangalore (Karnataka), India. Email: hn.suresh@rediffmail.com

It contains information about several states of emotions which aid to understanding the individual behaviors. Humans use the face beside with facial expressions to express their emotional states and the decisions are taken depending upon the person's expressions which includes happy, sad, angry, neutral and surprise shown in Figure.1. The individual activities and hobbies are decided based on the emotions they provoke. Among the nonverbal behaviors such as body movements, postures, gaze, voice, etc. Face is the efficient part to accessible window into the mechanisms which govern our emotional and social lives. It is a complicated task for generating a face recognition system which depends on emotion translation because of complex multidimensional visual model.

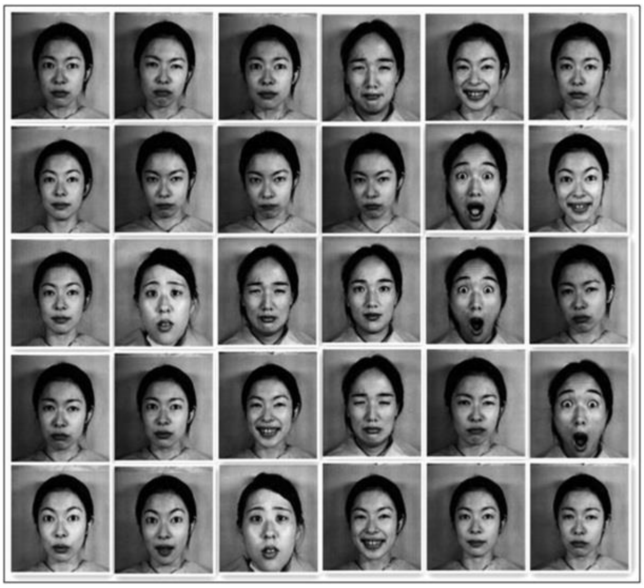

Figure.1. Sample examples of JAFFE data set

A machine that reads passenger passports and IDs (Identification) can process numerous numbers of passports than a person in a similar time. This sort of use spares time, cash and takes out the necessity that a human performs such a tedious task. In this unique situation, the benefit of making machines to read the facial expressions is a tremendous thought. The facial expression reveals the internal characteristics of the expresser (individual). There is a need to introduce a new model for facial expression recognition system. Machine Learning (ML) [4] is a branch of AI that is anxious about planning and designing frameworks that can gain from the given input. It is a branch of computerized analysis that automates analytical model. It is a branch of AI that machines ought to have the capacity to learn and adjust through involvement. Machine Learning (ML) is mainly utilized to develop computer programs that create solutions and understanding. In recent years, numerous studies have been conducted on machine learning algorithms such as Bayesian networks, Decision Trees, Genetic Algorithms, Nearest Neighbors, Neural networks, Principle Component Analysis (PCA), Support vector machines (SVM) and Extreme Learning Machine, each of which has its own set of advantages and disadvantages. The advantages of these algorithms are simplicity, fast training time, and the easy evaluation of parameters which is adjusted for every data set. Additionally, other researchers also get a best result of classifying images, and it can be widely implemented in various domains such as classification of medical images, data mining, cloud computing, image processing and signal processing.

Image classification is to identify the standard form of images which can be automatically recognized, the images into a finite set of classes [5]. The algorithms which are used to classify the images are mainly pre-processing of an image before classification. This process includes extracting relevant features and segmenting the extracted images into sub-components and finally classification. To classify the image datasets, recently machine learning and image processing have a lot of attention in various fields. The medical industry would greatly benefit from an innovative integration of machine learning in image processing, which will help it better analyse complex pictures. Processing images requires a variety of image processing techniques that combine several different types of knowledge. The machine learning technology may be used to tackle image processing issues if the adaptability is increased. A vast number of photos necessitates an ability to deal with enormous data sets, which is a problem for most machine learning algorithms. As a result, successful picture categorization necessitates interaction with image data and a machine learning system.

The primary goal of this study is to overcome the flaws of traditional algorithms and to improve the process of facial expression recognition which could be used in various applications. Section.2. Presents an survey related to facial expression recognition and its process based on image processing technologies. Section.3. explains the proposed scheme as HAKELM and CSO algorithm. The image processing techniques such as pre-processing, gesture extraction, gesture selection and classification based HAKELM are discussed with mathematical description. Finally, the Section.4. Presented the experimental results with the comparison of Indian and Japanese datasets in terms of certain performance metrics to show the effective performance of the proposed method and finally the research conclusion is summarized in Section.5.

II. RELATED WORK

Nowadays facial expression recognition has turned into an incredible concern in the modern scientific research. The most expressive approach to express human emotions is the facial expressions. The doctor can diagnose the patients' condition through their facial expressions. Facial expressions have the most important part in social communication. As a result, numerous researchers are actively involved and have promoted relatively a good number of techniques to attempt on benchmark database images, especially JAFFE. Although the images are used precisely for various applications, they have errors such as processing time and poor contrast. These problems can be resolved using machine learning and optimization techniques. This section discusses about the various researchers' approaches and different techniques that are used for facial expression recognition. The collected literatures are classified and the outcomes are presented below. The most significant way of expressing human emotions is only possible by facial expression. Facial expression is a non-verbal way of communication [6]. The physical movements of muscles underneath the skin of the face are utilized for facial expressions. These expressions are completely natural and occasionally people do not realize that they show their emotions.

Retrieval Number: 100.1/ijeat.E35350611522
DOI: 10.35940/ijeat.E3535.0611522
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Facial expression recognition In our daily lives, it is utilised in a variety of contexts, such as behavioural science, clinical practice, business negotiation, safe driving, remote education, intelligent family robot and analysis of human emotions[7]. The Facial expression recognition process consists of three steps, acquisition of face image, extracting facial gestures and Classification of facial gestures. There are six types of facial expressions such as happy, anger, sad, fear, surprise, and disgust. These expressions can be identified from the facial gestures of eyebrows, eyes, mouth, and nose that are detected from a face image. The gesture segments are extracted by using morphological image processing algorithms and edge detection technique.

The facial expression recognition approach is broadly classified into three parts[8], namely such as global, local and hybrid. In global method, whole face image data to generate the gestures for face recognition. Mainly this approach is used for frontal view which generates the good performance. Local method is better than the global approach because it does not get affected by the imaging factors. This approach uses the gestures of the facial parts like eye region, mouth, and nose. Most of the local approaches are needed to set the facial component's position and to extract them for a gesture vector generation. The hybrid strategy makes use of both the global and local face approaches, and it is equivalent to the human recognition procedure in terms of computing cost, albeit it is more complex. The foremost steps for the facial expression recognition system are extraction and classification. Initially, it is necessary to represent the standard representation form of face images which contain adequate data for the purposes of identification. The second stage is to compare a fresh face image with the query image and categorise it accordingly. Certain facial gestures such as eye blinking, head movement, speech and facial expression are detected through sensors and processed by computer. So, the people can communicate with an Electric Powered wheelchair (EPS). Most of the EPS are handled by joystick manually. Human Machine Interfaces (HMIs) [9] and the head movements are meant to distinguish between the patient's facial emotions. Development of the advanced hands-free HMI is mainly utilized to access assistive tools and robots to enhance the lifetime. Now a day's the image processing techniques such as pre-processing, gesture extraction, gesture selection and classification are widely used in medical diagnosis for efficient classification [10].

Table.1. Various methodologies and Limitations

Sl. No.	Author's Name	Methods	Drawback
1.	Han, JS et.al	Active Appearance Model and Twofold Random Forest Classifier	Classification rate was less and high computation speed
2.	Yacoob et.al	Singular Value Decomposition and PCA	Minimum degree of efficiency and high speed of computation
3.	Mavani et.al	Gabor Texture Features and Local Phase Quantization	Less recognition rate and sensitive to light and position
4.	Valenti et.al	Boosted Deep Belief Network	Not adaptable for large datasets and high computational cost
5.	Datcu et.al	Active Contour-Based Face Detection	High less susceptible to the presence of noise and less robust to large datasets
6.	Mahmud et.al	Principal Component Analysis	Less accuracy and difficult to detect the features
7.	Ferreira et.al	SVM and HMM	Less classification accuracy and difficult for the polynomial mid-level parameters
8.	Tsui et.al	Particle Swarm Optimization	Highly irrelevant, noisy and computationally very expensive
9.	Komala et.al	Gabor wavelets and neural networks.	Inability to detect the variations in pose, illumination

Image processing techniques with respect to medical images are mainly used for the retrieval of images or for the analysis and modification of images. The detailed process of the proposed methodology based image processing techniques is discussed in the further sections. The various methodologies and their limitations are tabulated in Table.1.

III. FER USING HAKELM – CSO ALGORITHM

Image processing techniques are used to identify facial emotions utilising machine learning and optimization approaches. HAKELM-CSO, a suggested face detection algorithm, is depicted in Figure.2. Human expression 256x256 data sets are used to collect the datasets. The wiener filter does pre-processing to minimise and improve the image quality caused by salt and pepper noise that is generated during image capture. Facial detection is GLCM's gesture extraction for a specific area of skin. The redesigned firefly selects the extracted facial skin gestures to save processing time. Face detection and picture expression are classified and recognised using the HAKELM technique, which is tuned by CSO for effective classification

A. Image Acquisition

The 256x256-pixel datasets used in this analysis are all freely available. The skin seems brighter than the backdrop when the picture is converted to a grayscale image. Following this, the algorithm uses the skin region's orientation as a starting point to identify the skin region.

B. Pre-processing

Normally, all images have some amount of unwanted noise due to the machine performance, apparatus and the surrounding environment that lead to serious issues. As a result, pre-processing method is necessary to reduce noise. Prior to computer processing, the precise values may be acquired Low-frequency background noise, levelling the intensity of individual particle pictures, and getting rid of the data images that had been boosted. Wiener filter is adopted mainly to remove the unwanted noise from the given input image. The Wiener filter is designed on two approaches, the spectral characteristics of the original signal and the noise of the given image and the linear time-invariant filter. Discrete Fourier Transform (DFT) is applied to obtain $I(u,v)$. The original image spectrum is evaluated to generate the $I(u,v)$ with the Wiener filter $F(x,y)$. The Wiener filter is estimated as

$$F(x,y) = \frac{H^{*}(x,y)P_s(x,y)}{(|H(x,y)|^2 P_s(x,y) + P_n(x,y))} \quad (1)$$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$H(x,y)$ is the blurring filter and $P_s(x,y)$ $P_n(x,y)$ denotes the power spectrum of the signal and noise. The Wiener filter is performed to reduce the blurring (filtering) part and to smooth the noise. The deconvolution process (high pass filtering) is carried out to eradicate the noise incorporating with a compression process (low pass filtering). Wiener filter estimates the power spectra of the original picture and the noise to produce a smooth image. By dividing each term by $P_s(x,y)$ the result is obtained.

$$F(x,y) = \frac{H^*(x,y)}{\left(|H(x,y)|^2 + \frac{P_n(x,y)}{P_s(x,y)}\right)} \quad (2)$$

In which $\frac{P_n}{P_s}$ can be denoted as the reciprocal of the signal-to-noise ratio and the signal is extremely strong related to the noise $\frac{P_n}{P_s} \sim 0$ and the Wiener filter becomes $H^{-1}(x,y)$ which is denoted as inverse filter. Where the signal is very weak $\frac{P_n}{P_s} \to \alpha, F(x,y) = 0$. But in some cases, if there is increased white noise and no blurring, then the wiener filter can be simplified as

$$F(x,y) = \frac{P_s(x,y)}{P_s(x,y) + \sigma_n^2} \quad (3)$$

σ_n^2 denotes noise variance. In case wiener filters is incapable to rebuild frequency component, then the pre-processed image is obtained from the specific skin region which is detected for gesture extraction.

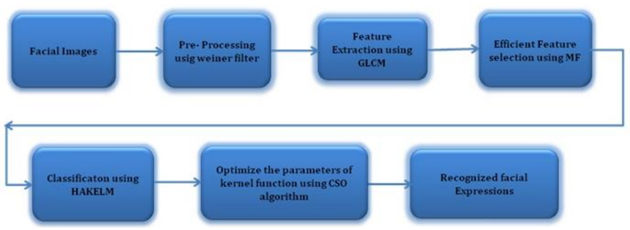

Figure.2. HAKELM-CSO block schematic of FER

C.Skin Detection by QEA

Quantum Evolutionary Algorithm (QEA) is used for skin detection to differentiate the skin region and non-skin region. pixels in the skin region are analogous to the red and blue components (C_r, C_b)of a computer monitor. It is required to segment each pixel either of the skin region or non-skin region. QEA is an evolutionary computation method which is inspired by the ideas of quantum material science as the uncertainty principle and the observer effect. Generally, this method comprises of the search space and speed up convergence to identify the best skin region. The optimal threshold value has been calculated for automatic segmentation. The two kinds of information are approximation image with low frequency data and gradient image with high frequency data. These two-image co-occurrence matrices are calculated by means of enhanced two-dimensional gray entropy[11].

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (4)$$

where $|\alpha|^2$ is the probability of the non-skin region and while $|\beta|^2$ is the probability of the skin region. The skin region is detected by QEA for the face recognition. The detected face skin region is extracted for face recognition to identify the facial expressions.

D.Gesture Extraction

This process is to extract the known gesture which contains some information about the image. Gesture extraction by Gray Level Co-occurrence Matrix is to extract the detected skin region. The extracted skin region gives the detailed information about the gestures which are needed to decrease the process complexity.

GLCM is a second order measure which measures the relationship between neighboring pixels and the spatial relationships based on the extraction of a grayscale image and it is important to differentiate the image texture by estimating to extract the statistical measurements from this matrix, the pixel with defined values and a given spatial link to the picture. It is well-known that the following formula [12] is used to extract GLCM gestures such as Cluster Prominence, Cluster Shade, Inertia and Difference Entropy, which all increase processing time. Other GLCM gestures include Homogeneity, the Angular Second Moment (ASM), Local Homogeneity, the Inverse Difference Moment (IDM), Contrast, Entropy, Correlation and the Sum Average. As a result, the processing time will be reduced by extracting these motions. Let us consider *P* as the normalized GLCM of the input texture image. Some of the gestures estimated equations are given below:

$$Energy = \sum_{u,v} P(u,v)^2 \quad (5)$$

$$Contrast = \sum_{u,v} |u-v|^2 \; P(u,v)^2 \quad (6)$$

$$Homogeneity = \sum_{u,v} \frac{P(u,v)}{1+|u-v|} \quad (7)$$

$$Correlation = \sum_{u,v} \frac{P(u,v)}{1+|u-v|} \quad (8)$$

$$Clustershade = \sum_{u,v=0}^{N-1} P_{uv}(u - M_i + v - M_j)^3 \quad (9)$$

$$Clusterprominence = \sum_{u,v=0}^{N-1} P_{uv}(u - M_i + v - M_j)^4 \quad (10)$$

E. Gesture Selection

The extracted gestures are selected by Modified Firefly (MF) algorithm to select the facial detection skin region. Gesture selection is a process which selects a subset of image gestures of the given original extracted image without any alteration and maintains the physical structure of the original gestures. Gesture selection methods are utilized to search through the subsets of gestures and to identify the best one among the challenging *2N* candidate subsets according to the various estimate functions. Consider n number of flies as M_i, where $i = 1,2,3, \ldots\ldots n$. The random positions of n fireflies are initialized and the position is updated by the equation given below.

$$M_i(t+1) = M_i(t) + D_0 \exp(-\gamma q_{ij}^2)(M_j - M_i) + (rand - 0.5) \quad (11)$$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

where D_0 is represented as the initial attractiveness at $q = 0$ and γ is absorption parameter in the range $(0,1)$ and q_{ij} denotes the distance between i and j and the best firefly position are updated by

$$\text{Mbest}_i(t+1) = \text{Mbest}_i(t) + (\text{rand} - 0.5) \quad (12)$$

where $M_i(t)$ and $\text{Mbest}_i(t)$ are represented as the less bright firefly current position and brightest firefly positions.

The intensity I_i of each firefly x_i is initialized with its dependency A and each set of decision gestures are given as

$$I_i = \gamma_{xi}(A) \quad (13)$$

Every firefly i identify its distance with the other firefly j and decide the increment in the intensity of j with the movement of i towards j, by employing Equation (13) respectively, where $I_j > I_i, J = 1,2,3 \ldots\ldots n, \; j \neq i$.

$$a_{ij} = \min(\gamma_c(A) - \gamma(X_i X_j)(A) \quad (14)$$

$$I_j = \gamma(X_i X_j)(A) - I_{xj} \quad (15)$$

Every firefly i move towards its best mating partner j having least distance with i and the movement brings about the most noteworthy increment in dependency. If any firefly is not ready to find any best matting partner, the intensity of firefly i is consumed by the system and it will be unseen to all other fireflies in the search space [13]. The movement among the flies in this manner brings about the subsets of fireflies with an increase in dependency. After finding the best firefly, the position is updated, which are represented as

$$M_i(t+1) = M_i(t) + D_o\exp(-\gamma q_{ij}^2)(M_j - M_i) + (\text{rand} - 0.5) \quad (16)$$

$$\text{Mbest}_i(t+1) = \text{Mbest}_i(t) + (\text{rand} - 0.5)\text{Mbest}_i(t) \quad (17)$$

After the best gesture selection, these gestures are forwarded as an input to the HAKELM classification process and it trains the sample images based on the gesture inputs. After that, the trained images are stored in the database later, which is tested based on these trained results. The HAKELM classification is performed on the trained images to classify the skin region to predict the facial expression. The Pseudocode for the modified firefly are given below:

Input: The set of all conditional gestures (CF)

Output: The set of decision gestures are defined as A.

Step 1: Objective function $OF = \{M : M \subseteq C, \gamma_f(D) = \gamma_c A$.

Step 2: Initial population of fireflies M_i $(i = 1,2,3, \ldots..n)$ are generated according to each conditional gesture.

Step 3: Light intensity I_i at M_i is determined by using Equation.15

Step 4: The conditional gestures (CF) is equal to C i.e., CF = C

While $\gamma_f(A)! = \gamma_c A$

$CF' = CF$

$CF = []$

$For\ i = 1, CF'\ firelies$

$For\ j - 1 : CF'\ fireflies$

Step 5: The best matting partner *j* for *i* is found by using the following favourable conditions.

- $> I_i)$ is defined the intensity of has been greater than intensity of .
- Euclidean distance has been used to measure the distance among and .
- when *i* move towards *j*, the intensity of *j* has been increased End for *j*.

Step 6: The Firefly *i* has been moved towards *j* and is denoted as

$$M_{ij} I_{ij} = \gamma(M_i, M_j)(A)$$

$$CF = CF \cup M_{ij}$$

End for i

Step7: Each f_{ij} is evaluated in CF for need, which is $\gamma M_{ij}(A) == \gamma_c(A)$ and minimally

End while

Step 8: The best gesture has been selected.

F. Facial Expression Classification

Three layers of sigmoid activation function are used in HAKELM. As a result of HAKELM's one-pass learning process and improved generalisation, it is a quick learning algorithm. Back Propagation has a better performance (BP). Though HAKELM gives efficient generalizations, it is one of the most popular learning techniques that are derived from ELM. So, it is widely used in practical applications. HAKELM is very efficient for large datasets. Here HAKELM is utilized to classify the facial expression recognition.

G. HAKELM based classification of facial expressions

HAKELM is utilized for facial expression recognition to classify the facial expressions[14]. Initially the input hidden neurons are considered as n-gestures which are represented as

$$(\{DF_i, t_i) | DF_i \in R^n, t_i \in R^m, \quad i = 1,2,3, \ldots\ldots, n\})$$

DF is represented as decision gestures, t denotes training time and L denotes output neuron.

$$f_L(DF) = \sum_{i=1}^{L} WV_i hl_i(DF) = h(DF)WV \quad (18)$$

Here the output weight vectors are represented as $wv = [wv_1, wv_2, wv_3, \ldots\ldots\ldots\ldots wv_L]$, and the hidden layer output neuron is represented as $hl_i(DF) = [h_1(DF), h_2(DF), h_2(DF), \ldots\ldots h_L(DF)]$. In which, the output weight with less training time has been reduced by using the below Equation (19).

$$\|HWV - OM\|, \|WV\| \quad (19)$$

H denotes the hidden layer output matrix and OM represents the expected output matrix of samples and WV denotes weight vector.

The above Equation can be rewritten as based on KKT (Karush Kuhn Tucker) conditions

$$WV = H^{OM}(\frac{1}{Co} + HH^T)^{-1} OM \quad (20)$$

Where Co denotes the regulation coefficient.

$$M = HH^T : m_{ij} = hk(CF_i)hl(CF_j) = KM(CF_i, CF_j) \quad (21)$$

Here the $hl(CF)$ is unknown, so the Mercer's conditions is utilized to define the Kernel Matrix (KM). Then the HAKELM output function f (CF) is represented as

$$f(CF) = [k(CF, CF_1), \ldots\ldots, k(CF, CF_n)] \left(\frac{1}{Co} + M\right)^{-1} OM$$

In HAKELM, Wavelet kernel function are utilized to define

$$Wk(p,q) = \cos(d\frac{\|p-q\|}{e})\exp(\frac{\|p-q\|^2}{f}) \quad (23)$$

In which (p,q) are single hidden layered Feed-Forward Neural Networks (FFNN) of hidden neurons kernel function. where d, e and f are adjustable parameters which play a major role for NN.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

These adjustable parameters are mandatory to enhance the HAKELM performance (Huang *et al.* 2012). So, these parameters are optimized to obtain the best values by utilizing Chicken Swarm behavior Optimization (CSO) algorithm. In HAKELM, the kernel functions are utilized to improve the generalization performance, however the parameters are not considered, so the optimization method such as CSO is proposed for parameter optimization. Normally the CSO is developed according to the real chicken behavior which is mainly utilized decrease computing complexity while increasing convergence time as compared to existing swarm techniques.

H. Kernel parameters optimization based CSO

In CSO, the position of every kernel parameter is denoted as $P_{ci,d} \epsilon\ d, e, f$ for every gesture depending on the kernel functions. The following steps for the proposed CSO optimization algorithm is

Step 1: Initially, every chicken position is represented by four variables such as number of mother hen m_n, number of hens h_n, number of chicks c_n and number of roosters r_n . Where n denotes the number of entire chicken swarm's population and the dimension of the search space are denoted as di and time t. The initial position of each chicken swarm particle and their dimension are represented by the given formula

$$P_{ci,d}, c_i \in [1,2, \dots \dots, n],\ d \in [1,2,3, \dots \dots \dots, di]) \quad (24)$$

Step 2: In this step, from the above equations, the roosters are represented as the best and the chicks as worst and the remaining are hens. Thus, the r_n values are the best value of identifying the optimal parameters in which the highest values among all parameter values are denoted as the best values for each kernel parameter. Depending upon the highest position value, the fitness value for each parameter is evaluated and the rooster's position is updated by the given formula

$$P_{ci,d}\ (t+1) = P_{ci,d}\ (t)\left(1 + rand\ (0, \sigma^2)\right), \quad (25)$$

$rand\ (0, \sigma^2)$ is denoted as Gaussian distribution with standard deviation σ and mean 0.

$$\sigma^2 = \begin{cases} 1, & fv_{ci} \geq fv_k,\ k \in [1, n],\ k \neq c_i \\ \exp\left(\frac{fv_k - fv_{ci}}{|f_{ci} + \varepsilon|}\right), & \text{otherwise} \end{cases} \quad (26)$$

Here fv_{ci} denotes the fitness value of particle, fv_k represents the fitness value of k, k represents the rooster's index and ε represents the small constant to avoid zero-division-error.

Step 3: After the rooster searches the food, the hens start to search food. The hens' position is updated by the formula given below

$$P^{t+1}{}_{ci,d} = P^t{}_{ci,d} + C_1\, rand\left(P^t{}_{r1,d} - P^t{}_{ci,d}\right) + C_2\, rand\left(P^t{}_{r2,d} - P^t{}_{ci,d}\right) \quad (27)$$

$$C_1 = \exp((fv_{ci} - fv_{r1})/(|fv_{ci}| + \varepsilon)) \quad (28)$$

$$C_1 = \exp(fv_{r2} - fv_{ci}) \quad (29)$$

From the Equation (27) represents a uniform random number over $(0,1)$, r_1 denotes the index of the, r_n and r_2 represents chicken index which is considered as hen and it is randomly selected from the chicken swarm $r_1 \neq r_2$ in which the best position is updated.

Step 4: In this step, the kernel parameters optimization is considered as the objective function in which the best r_n chickens are related to the ones with r_n is denoted as the best with the highest fitness values.

$$P^{t+1}{}_{ci,d} = P^t{}_{ci,d} + F\left(P^t{}_{m_{n,d}} - P^t{}_{c_{i,d}}\right) \quad (30)$$

Where $P^t{}_{m_{n,d}}$ denotes the mother hen position and the coefficient value represents the $F, F \in [0,2]$ which specifies that the chick should follow its mother to hunt for food. Based on the fitness evaluation, the optimal kernel parameters are updated in kernel function and the generalization performance is improved to classify the facial expression. The algorithm for CSO optimization is given below:

Input: A population of n chickens (i.e kernel parameters) is initialized.

Output: Optimized kernel parameter values

n chickens fitness values are evaluated at $t = 0$

While (t < Max_Generation (G))

If (t %G == 0)

- The chickens' fitness values are ranked and established in a hierarchal order in the swarm.
- The swarm is divided into different groups and it determines the relationship among the chicks and the mother hens.

End if

For i = 1,...n

Check If i == rooster update its solution or fitness value using equation.

End if

Check If i == hen update its solution or fitness value using equation.

End if

Check If i == chick update its solution or fitness value using equation.

End if

The new solution is evaluated

Check if the new solution is better than its previous one and the new one is Updated.

End for

End while

The final best optimal solution has been displayed.

The KELM based adaptive kernel parameters with CSO optimization is developed and denoted as AKELM to classify the facial expressions and the steps for CSO are given below:

Step 1: Initialize the chicken population.

Step 2: Update the positions depending on the kernel parameters.

Step 3: Evaluate the fitness value based on r_n, m_n, h_n, c_n.

Step 4: Each particle position is updated.

Step 5: Update the fitness value.

Step 6: The process continues until iteration is satisfied.

Step 7: Updated kernel parameters are used for training.

Finally, based on the training images, the query image is tested, and the expressions are classified as neutral, happiness, sadness, surprise, anger, disgust and fear.

IV. RESULTS AND DISCUSSION

The MATLAB Tool simulates the performance of HAKELM with CSO in a windows environment. This method's primary objective is to classify diverse facial expressions in order to accurately identify the patient's expressions.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-11 Issue-5, June 2022

The performance of HAKELM is compared to that of established techniques such as the SVM [15] and PCA. The trained image is compared to the query image, with facial emotions rated as normal, happy, sad, surprised, furious, and disgusted.

A.Datasets Description

The evaluation of HAKELM performance is done on the datasets of facial expression which is collected from the Indian and Japanese facial expressions. The facial expression of the individual is captured from the high resolution camera and then compared with the query images. Indian Face Database – Indians' facial expressions are collected which constitute 6 types of expressions.

JAFFE Database – This collection comprises 3 types of face expressions and 213 photographs of ten Japanese female models posing for the camera. Each person has between two and four photos for each expression[16]. Each picture topic contains six distinct expression photographs, totalling 60 in this database. When compared to Indian face, Japanese face is very complicated to identify the facial expressions. To evaluate the performance of HAKELM, both Indian and Japanese faces are considered for accurate classification.

B.Process of Training and Classification of Facial Expressions

From Indian and JAFFE databases, all images are evaluated by measuring the parameters of the datasets by utilizing HAKELM. During the classification training, each dataset contains 6 expressions for each Indian and JAFEE image. Initially the input image is pre-processed to reduce the noise and the skin region is extracted by employing gesture extraction. The extracted skin region is selected by gesture selection which is given as the input to the HAKELM to classify the facial expressions.

Figure.3. Indian training pictures

The HAKELM processes the input image depending on kernel function to identify the best parameter values to optimize by CSO. The kernel function optimizes the parameter values to improve the training process. Followed by the training values, the tested images are classified by the machine learning algorithm HAKELM into six different expressions such as normal, happy, sad, surprise, angry and disgust[17,18]. In this method, totally 300 images are taken for evaluation in which 200 images are allotted for training and 100 images for testing. Figure.3. and 4 shows the samples of Indian and JAFEE datasets with different type of expressions.

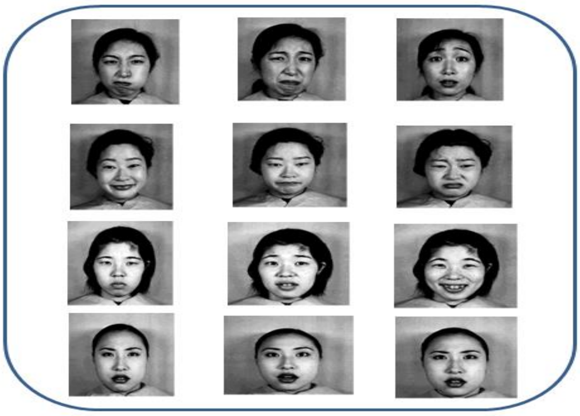

Figure.4. JAFFE Training pictures

The step by step process of proposed pre-processing, gesture extraction, segmentation, detection and HAKELM classification is demonstrated in Figure.5. and Figure.6. The classification result shows that the facial expression results as normal, happy and angry.

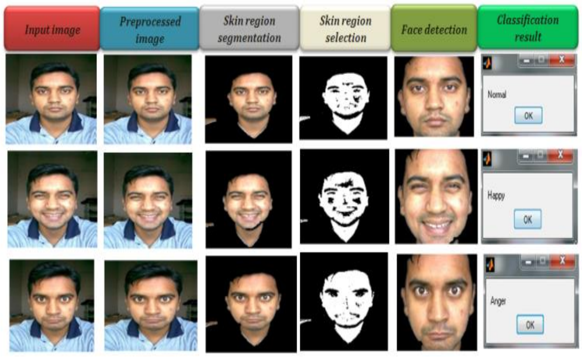

Figure.5. Step by step process of evaluation using Indian Face Database

Figure.6. Process of evaluation using JAFFE database

C.Performance Metrics

The accuracy, sensitivity, and specificity of HAKELM are measured using the confusion matrix, and the results are used to assess its overall performance. The results of these measurements are compared to $tr_p = true\ positive\ rate$, $tr_n = true\ negative\ rate$, $fa_p = false\ positive\ rate$, $fa_n = false\ negative\ rate$

The sensitivity and specificity are estimated by the formula given below

$$sen = \left(\frac{tr_p}{tr_p + fa_n}\right) \quad (31)$$

$$spec = \left(\frac{tr_n}{tr_n + fa_p}\right) \quad (32)$$

ROC (Receiver Operating Characteristic) analysis is used to determine the accuracy of facial expression classification.. The x and y axes of this ROC are shown as:
$x = falsepositiverate\ (1 - spec)\ and\ \ y = truepositiverate\ (sen)$

Table.1. Compares performance indicators including accuracy, sensitivity, and specificity for the HAKELM with the current SVM and PCA. Because of high true positive rate, the HAKELM attains the efficient classification of facial expressions recognition with accuracy of 95.84%, sensitivity of 90.12% and specificity of 96.12%.

Table.1. A comparative study of performance

Performance metrics	Accuracy (%)	Sensitivity (%)	Specificity (%)
HAKELM	95.84	90.12	96.12
PCA	94.45	91.1	86.4
SVM	83.5	82.1	84.21

Table.2. Results of all classifiers' ROC analysis

1-spec	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
HAKELM	0.25	0.31	0.39	0.48	0.62	0.73	0.8	0.96	1.01
PCA	0.17	0.21	0.25	0.33	0.45	0.7	0.7	0.78	0.94
SVM	0.11	0.15	0.15	0.22	0.3	0.31	0.43	0.6	0.67

Figure .7. compares the HAKELM to the current SVM, PCA in terms of accuracy, sensitivity, and specificity.

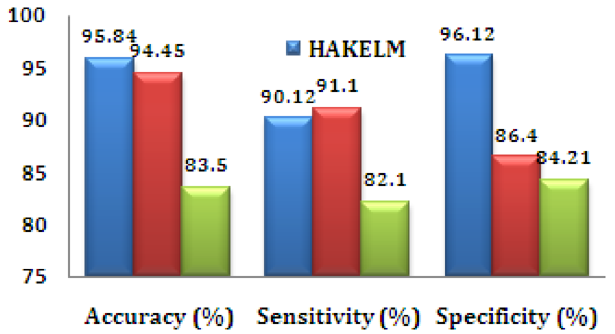

Figure.7. Analysis of the performance of different classifiers

The graph clearly illustrates that the proposed HAKELM method outperforms the other algorithms.

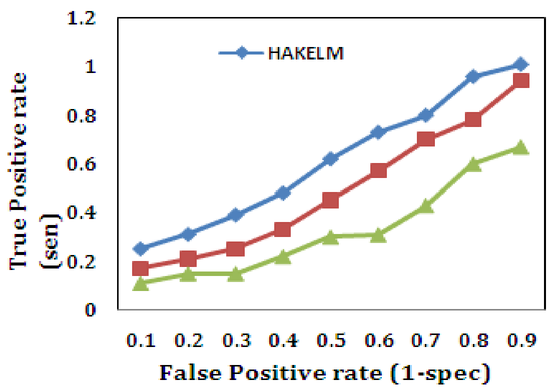

Figure.8. Performance analysis for ROC for various classifiers

Table.2. shows the evaluation of the ROC performance for facial expression recognition by using HAKELM and comparing with SVM and PCA. The efficient pre-processing, gesture extraction and gesture selection with optimization techniques show the best values of specificity 96.12%.

D.Performance Comparison of Processing Time

Comparing the processing time to PCA and SVM is shown in Table.3. The suggested system saves time because of shorter training time and more efficient pre-processing[19]. As shown in Figure.9, the proposed HAKELM and the current PCA and SVM are compared in terms of their processing times. To categorise face expressions from huge datasets, the HAKELM is clearly faster than other methods.

Table.3. Comparison of all classifiers' processing time

No. of Images	HAKELM(s)	PCA(s)	SVM(s)
10	8.012	10.745	16.417
20	16.14	18.147	22.417
30	21.524	24.254	28.154
40	24.47	27.147	33.177
50	29.147	32.2574	39.5412

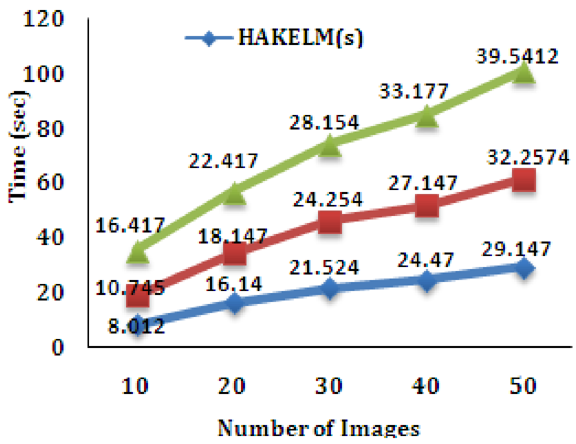

Figure.9. Time required to analyse data for multiple classifiers

E. Comparison of Precision, Recall, and F-measure performance

HAKELM's accuracy, recall, and F-measure are all shown numerically in Table.4. With a lower negative rate and a higher positive prediction rate, the suggested HAKELM achieves accuracy levels of 95.1% in precision, 88.14% in recall, and 87.45% in F-measure performance, respectively.

Table.4. Results of the numerical evaluation of precision, recall, and F-measure for all classifiers

Performance Metrics	HAKELM	PCA	SVM
Precision (%)	95.1	94.35	80.41
Recall (%)	88.14	86.32	78.12
F-measure (%)	87.45	85.14	76.14

Figure.10. Compares the proposed HAKELM with the current PCA and SVM in terms of accuracy, recall, and F-measure performance. HAKELM outperforms previous algorithms in terms of accuracy, recall and F-measure outcomes.

International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-11 Issue-5, June 2022

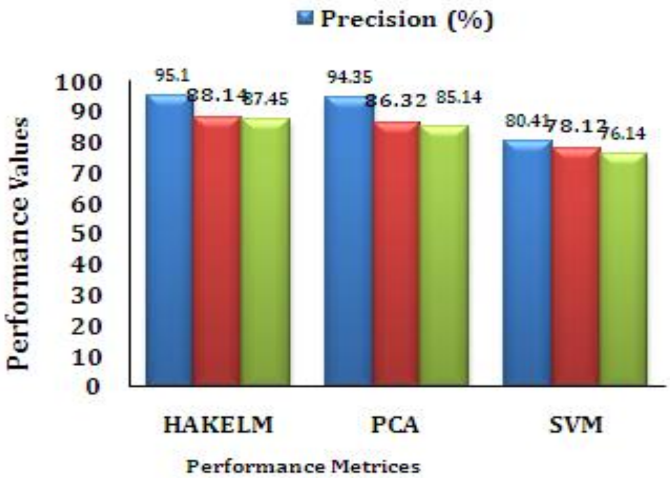

Figure.10. Numerical evaluation of precision, recall, and F-measure for all classifiers

Table.5, 6 and 7 show the evaluation of performance parameters like sensitivity, specificity, precision, recall, F-Measure, accuracy and processing time. Each parameter is evaluated for the combined datasets for both Indian and JAFFE of total 100 images to evaluate the proposed performance [20]. From these tables, it is clear that the proposed HAKELM attains the best results when compared to the other existing classifications, due to the efficient gesture selection and fast training process with the optimized kernel parameter's value. From the above discussion, Hybrid Kernel Extreme Learning Machine (HAKELM) with the optimized Chicken Swarm Optimization (CSO) shows the efficient result for the classification of facial expressions. It is very effective in predicting the patient's illness by using the automatic identification of facial expression which is achieved by employing machine learning and optimization algorithms.

Table.5. Numerical evaluation results of various parameters

Input images-Indian + JAFFE	tr_p	tr_n	fa_p	fa_n	sen (%)	spec (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	6	2	0	2	90.12	98.05	96.14	91.3	95.05	88.12	27.87
10+10	11	5	1	3	90.11	96.1	95.47	91.3	94.98	88.09	28.14
15+15	18	7	1	3	90.12	94.21	95.14	92.47	95.12	88.1	29.04
20+20	27	9	1	4	90.1	93.14	94.78	94	94.97	88	29.21
Total=100	62	23	3	12	90.12	96.12	95.84	87.45	95.1	88.14	29.147

Table.6. Numerical evaluation results of various parameters for existing PCA

Input images-Indian + JAFFE	tr_p	tr_n	fa_p	fa_n	sen (%)	spec (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	3	4	0	3	91.04	86.06	94.52	88.12	94.89	86.45	32.01
10+10	7	8	0	5	91.2	86.24	94.48	87.57	94.54	86.34	32.15
15+15	12	10	1	7	91.1	86.32	94.42	87.45	94.32	86.21	32.22
20+20	18	12	1	9	91.14	86.45	94.44	87.34	94.12	86.12	32.45
Total=100	40	34	2	24	91.1	86.4	94.45	87.45	94.35	86.32	32.25

Table.7. Numerical evaluation results of various parameters for existing SVM

Input images-Indian + JAFFE	tr_p	tr_n	fa_p	fa_n	sen (%)	spec (%)	Accuracy (%)	F-measure (%)	Precision (%)	Recall (%)	Processing time (s)
5+5	1	4	1	4	82.43	84.45	84.14	76.4	80.54	78.24	38.54
10+10	4	10	0	6	82.21	84.31	84.1	76.24	80.34	78.16	39.12
15+15	9	13	0	8	82.11	84.15	83.87	76.1	80.28	78.1	39.54
20+20	14	15	1	10	82	84.1	83.78	76	80.32	78.04	39.78
Total=100	28	42	2	28	82.1	84.21	83.5	76.14	80.41	78.12	39.541

Retrieval Number: 100.1/ijeat.E35350611522
DOI: 10.35940/ijeat.E3535.0611522
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

V. CONCLUSION

Based on machine learning and optimization methods, face expressions may be recognized. Based on specific image processing approaches, the HAKELM algorithm is presented to recognize the human face expression using a hybrid adaptive kernel-based Extreme Learning Machine (HAKELM). The noise in the original image is removed using the wiener filter during pre-processing. QEA is employed to detect the skin region. Then, GLCM gestures are utilized to extract the skin region and the extracted skin region is selected by MF algorithm. Once HAKELM has been classified, the CSO technique is utilised to optimise kernel parameters functions for each expression. Experimental findings demonstrate that the HAKELM scheme obtained a high level of precision (95.1%), sensitivity (90.12%), specificity (96.12%), recall (88.14%), and F-measure (87.55%) when compared to the current algorithms such as PCA and SVM. This is a significant improvement. Various optimization strategies may be used to increase the accuracy of classification to improve facial expression recognition on real-time datasets to further improve this process.

REFERENCES

1. Han, JS, Bien, ZZ, Kim, DJ, Lee, HE & Kim, JS 2018, 'Human machine interface for wheelchair control with EMG and its evaluation', In Engineering in Medicine and Biology Society, Proceedings of the 25th Annual International Conference of the IEEE, vol. 2, pp. 1602-1605.
2. Yacoob, Y & Davis, LS 1996, 'Recognizing human facial expressions from long image sequences using optical flow', IEEE Transactions on pattern analysis and machine intelligence, vol. 18, no. 6, pp. 636-642.
3. Mavani, V, Raman, S & Miyapuram, KP 2019, 'Facial Expression Recognition using Visual Saliency and Deep Learning', ArXiv preprint arXiv:1708.08016.
4. Valenti, R, Sebe, N, Gevers, T & Cohen, I 2018, 'Machine learning techniques for face analysis', In Machine Learning Techniques for Multimedia, Springer, pp. 159-187.
5. Datcu, D & Rothkrantz, LJM 2020, 'Machine learning techniques for face analysis', In Proceedings of EUROMEDIA Conference. Elmansori, MM & Omar, K 2011, 'An enhanced face detection method using skin color and back-propagation neural network', European Journal of Scientific Research, vol. 55, no. 1, pp. 80-86.
6. Mahmud, F & Al Mamun, M 2020, 'Facial Expression Recognition System Using Extreme Learning Machine', International Journal of Scientific & Engineering Research, vol. 8, issue 3, ISSN 2229-5518.
7. Ferreira, RL, Silva, WC, Celeste, TF, Bastos Filho & Sarcinelli Filho, M 2017, 'Human–machine interface based on muscular and brain signals applied to a robotic wheelchair,' Journal of Physics:Conference Series, vol. 90, pp. 012-094.
8. He, B, Xu, D, Nian, R, van Heeswijk, M, Yu, Q, Miche, Y & Lendasse, A 2019, 'Fast face recognition via sparse coding and extreme learning machine', Cognitive Computation, vol. 6, no. 2, pp. 264-277.
9. Tsui, CSL, Jia, P,Gan, JQ ,Hu, H & Yuan, K 2019, 'EMG-based hands-free wheelchair control with EOG attention shift detection', In Robotics and Biomimetic, IEEE International Conference, pp. 1266-1271.
10. Komala Lakshmi, J 2019, 'Face Recognition Using Image Processing Techniques', A Survey. International Journal of Engineering and Computer Science, vol. 3, no. 12. pp. 222-232.
11. Wang, XH, Liu, A & Zhang, SQ 2018, 'New facial expression recognition based on FSVM and KNN', Optik-International Journal for Light and Electron Optics, vol. 126, no. 21, pp. 3132-3134.
12. Albregtsen, F 2018, 'Statistical texture measures computed from gray level co-occurrence matrices', Image processing laboratory, Department of informatics, university of OSLO, pp. 5-13.
13. Carcagnì, P, Coco, M, Leo, M & Distante, C 2019, 'Facial expression recognition and histograms of oriented gradients: a comprehensive study', Springer Plus, vol. 4, no. 1, pp. 12-20.
14. Huang, GB, Zhou, H, Ding, X & Zhang, R 2018, 'Extreme learning machine for regression and multiclass classification,' IEEE Transactions on Systems, Man, and Cybernetics, vol. 42, no. 2, pp. 13–9.
15. Li Xia 2018, 'Facial Expression Recognition Based on SVM,' 7th International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 256-259, 25-26.
16. Valenti, R, Sebe, N, Gevers, T & Cohen, I 2018, 'Machine learning techniques for face analysis', In Machine Learning Techniques for Multimedia, Springer, pp. 159-187.
17. Tsui, CSL, Jia, P,Gan, JQ ,Hu, H & Yuan, K 2017, 'EMG-based hands-free wheelchair control with EOG attention shift detection', In Robotics and Biomimetic, IEEE International Conference, pp. 1266-1271.
18. Siddiqi, MH, Ali, R, Khan, AM, Kim, ES, Kim, GJ & Lee, S 2016, 'Facial expression recognition using active contour-based face detection', Multimedia Systems, vol. 21, no. 6, pp. 541-555.
19. Shoorehdeli, MA, Teshnehlab, M & Moghaddam, HA 2018, 'Feature Subset Selection for face detection using genetic algorithms and particle swarm optimization', In Networking, Sensing and Control, Proceedings of the IEEE International Conference, pp. 686-690.
20. Shan, C, Gong, S & McOwan, PW 2019, 'Facial expression recognition based on local binary patterns: a comprehensive study, Image and Vision Computing, vol. 27, no. 6, pp. 803–816.

AUTHORS PROFILE

Prof. Shashank. M. Gowda, is presently working as Assistant Professor in Department of Electronics and Communication Engineering at Yenepoya Institute of Technology, Moodbidri, Karnataka. He has over 7 years of teaching experience. He completed his B. E. degree from Dr. M. V. Shetty Institute of Technology, Moodbidri in 2012, and M. Tech. Degree from BNM Institute of Technology, Bangalore in 2014. He is pursuing his Doctoral Degree at VTU, Belagavi under the guidance of Dr. H. N. Suresh, Head of the Department Electronics and Instrumentation Engineering, BIT, Bangalore.He is a Lifetime member of ISTE. He has published several articles in various journals. His research areas of interest are embedded systems, digital image processing, and emotion recognition.

Dr. H.N. Suresh, received his BE (E&C) from P.E.S College of Engineering, Mysore University, Karnataka, India, in the year 1989 and completed his M.Tech (Bio Medical Instrumentation) from SJCE Mysore affiliated to University of Mysore., in the year of 1996 and since then he is actively involved in teaching and research and has Twenty six years of experience in teaching. He obtained his PhD (ECE) from Anna university of Technology. He worked at various capacities in affiliated University Engineering Colleges. For Visveswaraya Technical University and Bangalore University he worked as a Chairman for Board of Examiners, and member of Board of Studies etc. At present he is working as Professor and Head, Department Electronics and Instrumentation Engineering in Bangalore Institute of Technology, Bengaluru Affiliated to Visveswaraya Technical University. He has good exposure in the field of signal processing, Wavelet Transforms, Neural Networks, Pattern recognition, Bio Medical Signal Processing, Networking and Adaptive Neural network systems. He has published more than 30 research papers in the refereed international journals and presented contributed research papers in refereed international and national conferences. He is a member of IEEE, Bio Medical Society of India, ISTE, IMAPS & Fellow member of IETE.