

Research Article

Analyzing the Role of Emotional Intelligence on the Performance of Small and Medium Enterprises (SMEs) Using AI-Based Convolutional Neural Networks (CNNs)

Suhail H. Serbaya 

Department of Industrial Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Correspondence should be addressed to Suhail H. Serbaya; sserbaya@kau.edu.sa

Received 10 February 2022; Revised 8 March 2022; Accepted 26 April 2022; Published 31 May 2022

Academic Editor: Mukesh Soni

Copyright © 2022 Suhail H. Serbaya. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human emotion detection is necessary for social interaction and plays an important role in our daily lives. Artificial intelligence research is rising, focusing on automated emotion detection. The capability to identify the emotion, which is considered one of the traits of emotional intelligence, is a component of human intelligence. Although the study is limited dependent on facial expressions or voice is flourishing, it is identifying emotions via body movements, a less researched issue. To attain emotional intelligence, this study suggests a deep learning approach. Here initially the video can be converted into image frames after the converted image frames can be preprocessed using the Glitter bandpass butter worth filter and contrast stretch histogram equalization. Then from the enhanced image, the features can be clustered using the hybrid Gaussian BIRCH algorithm. Then the specialized features are retrieved from the body of human gestures using the AdaDelta bacteria foraging optimization algorithm, and the selected features are fed to a supervised Kernel Boosting LENET deep-learning algorithm. The experiment is conducted using Geneva multimodal emotion portrayals (GEMEPs) corpus data set. This data set includes, human body gestures portraying the archetypes of five emotions, such as anger, fear, joy, pride, and sad. In these emotion detection techniques, the suggested Kernel Boosting LENET classifier achieves 98.5% accuracy, 94% precision, 95% sensitivity, and F-Score 93% outperformed better than the other existing classifiers. As a result, emotional acknowledgment may help small and medium enterprises (SMEs) to improve their performance and entrepreneurial orientation. The correlation coefficient of 188 and the significance coefficient of 0.00 show that emotional intelligence and SMEs performance have a significant and positive association.

1. Introduction

The process of anticipating emotive content from low-level data is known as emotional recognition. Physical expressions serve as indicators. These expressions have crucial characteristics that help us to distinguish between various emotions. We have characteristics such as volume and pitch in speech. For example, the posture of the body is one of the aspects of body expressions. Even though heart rate and brain activity cannot be seen, they also communicate emotions. The face is one of the most communicative portions of the human body, serving as one of the primary means of expressing emotions. Emotional recognition via the face is a problem that affective-computing researchers

have studied extensively; this problem is sometimes referred as facial expression recognition [1]. Owing to the widespread interest in technology, facial expression recognition (FER) has a broad variety of applications. It is mostly employed in HCI applications including virtual reality, interactive gaming, robotics, and digital entertainment. It is also employed in surveillance and law enforcement applications, as well as in emotion and behavioral analysis in the medical area (autism, mental problem, and pain evaluation) [2]. Figure 1 indicates different facial expressions, felt and expressed at different situations.

Automated emotion recognition is focused in different researches as in interpersonal conversations; humans often employ nonverbal clues including voice tone, facial

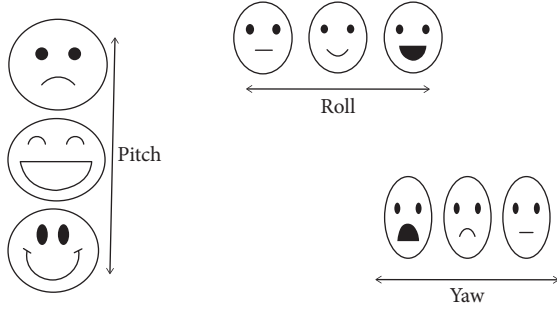


FIGURE 1: Facial expressions.

expressions, and hand gestures to indicate sentiments. Unfortunately, present human-computer interfaces do not fully use these important communication mediums, preventing users from reaping the full advantages of natural contact. If computers could identify users' emotions through their facial expressions and hand movements and respond in a courteous way based on their wants and preferences, human-computer interactions would be considerably enhanced. Aside from human-computer interactions, computerized psychological counseling and treatment, positive role in improving SMEs performance, as well as the identification of criminal and antisocial motivations, have fascinating uses. For the following reasons, identifying human emotions from facial expressions by a computer is a difficult task. To begin with, determining the precise face expression from a fuzzy facial picture is difficult. Second, segmenting a face picture into region of interest (ROI) is challenging, especially when the imaging properties of the parts are not significantly different. Third, unlike humans, robots lack the visual vision needed to translate face expressions into emotions [3].

Deep learning and other machine-learning technologies have produced effective FER solutions to automated emotion recognition. Following that, Group Method of Data Handling (GMDH) and cascade correlations were presented as early forms of deep learning. As previously stated, GMDH-type neural networks have been employed to solve identification difficulties. Evolving neural networks have been used to provide efficient solutions for biometric applications [4].

Deep-learning, in addition to hand-crafted feature engineering, offers an alternate method for formulating acceptable features for the SMEs and job at hand. Deep learning has shown amazing results in a variety of applications in recent years, such as image identification, object identification, and, most recently, voice acoustic modeling. In comparison to hand-crafted characteristics, deep learning can collect a greater stage depiction for task-specific information extracted from annotated information by learning from a maximum count of training examples using a deep architecture. Several academics have looked at the usefulness of deep learning in automatically learning emotional information from signal data in the field of voice emotion detection [5]. Emotion identification from face photos might be simplified using a three-stage technique developed by Donuk et al. [6]. The Fer+ data set is used to

train the convolutional neural network (CNN)-based network in the first step. The feature vector in the fully connected layer of the CNN network trained in the second stage selected using the binary particle swarm optimization technique. Support vector machines are used to categorize certain characteristics. The Fer+ data set has been used to evaluate the proposed system's performance. The exam gave a score of 85.74% correctness. The findings suggest that the combination of BPSO with SVM improves the accuracy and speed of the FER+ data set classification. For voice emotion identification under noisy and clean situations, researchers have explored four kinds of convolutional operations on distinct input characteristics using convolutional neural networks [7]. Every DNN has a deep recurrent subconnection design for the next temporal modeling because emotional behavioral input has been proven to represent temporally shifting mental state and convolutional function is performed locally in time. Analysis of the suggested architectures' information flow is presented in a module-by-module fashion. In particular, we show how effectively the data and other irrelevant data interact along with the transition from one unit to another. On the eINTERFACE corpus, we demonstrate that every deep neural network is performing at the highest level of excellence. The proposed hybrid DNN introduces an adaptive neuro-fuzzy inference method to anticipate a video's mood. Audio waves with a similar emotional experience may be generated using visual signals and a deep short-term memory recurring neural network [8]. Its fuzzy features allow to accurately depicting emotions, while its ability to model data with dynamic time qualities makes it a good fit for data with dynamic temporal properties. This technology is unique in that, it extracts visual emotional elements and then transforms them into audio signals that have comparable emotional qualities for listeners to experience and understand. In both data sets, the presented model can successfully create audio that fits the image generating a comparable emotional response from the viewer, and music produced by method is also picked more frequently. To improve the discriminating strength of features collected from speech and glottal signals, a novel feature augmentation employing the Gaussian mixture model (GMM) was developed. Emotional speech data sets from three distinct sources were used to test the suggested approaches [9]. The numerous sorts of emotions were classified using an extreme learning machine (ELM) and an NN classifier. Several trials were done, and the findings suggest that the proposed approaches considerably improved the recognition of speech emotions compared to previous studies. Convolutional neural networks, sparse autoencoders (SAEs), and deep neural networks are combined in a unique DNN suggested for emotion categorization using EEG devices [10]. For encoding and decoding, features retrieved by the CNN are transferred first to SAE. A DNN is then used to classify the data using characteristics that have less redundancy. The experimental findings reveal that the suggested network outperforms traditional CNN approaches when it comes to recognizing emotional states. It has been proved that the

suggested network is a more efficient and quicker convergent technique than a standard CNN if it is trained independently [10]. Convolutional neural network (CNN) was studied and improved to detect six fundamental emotions, and several preprocessing approaches were compared to illustrate how they affected CNN performance [11]. Face identification, cropping, and noise addition are some of the preprocessing techniques that are compared to each other in terms of data preprocessing. With 85.087% precision, compared to the next preprocessing stage and raw information, face detection was the most important preprocessing step. But when these strategies are combined, CNN can attain a 97% accuracy rate [12]. The facial recognition software may be used to assess the emotional state of students using e-learning throughout the COVID-19 pandemic. This program is capable of deriving emotional states from facial expressions. Using the static frontal facial picture, an emotion type may be determined. Starting with image acquisition, the image is processed using grayscale transformation and contrast stretching, Haar cascade for facial recognition; face model for the mouth and eye location, skin-color categorization approach for image segmentation, and GLCM feature extraction. SVM regression is used to classify emotional types. An LSTM hybrid algorithm based on channel fusion is put up as a possible solution [13]. Eight 3D VR video clips are used to trigger feelings for each emotional state. After an 8-level decomposition using discrete wavelet transforms on the preprocessed data, wavelet and time-domain features are recovered and fed into the hybrid LSTM. With an accuracy of 80.05% and 93.24% for four categories, the hybrid algorithm successfully predicts the eight distinct emotional states (happy, eager, quiet, bored, terrified, worried, sad, and relaxed). It was shown that frequency-domain characteristics on different bands had a better prediction rate compared to time-domain features. Human emotions may be taken into account while configuring cobot's settings [14]. For digitizing and analyzing the emotional state of people, this method makes use of electroencephalography (EEG). After that, the cobot's settings are quickly altered to maintain the human's emotional state in the desired range, which boosts the human's trust and confidence in the cobot. This research study also provides an overview of the latest developments in emotional sensing and identification technology. An ABB YuMicobot equipped with a commercially available EEG headset was used to evaluate this technique, which presented a simple and fast algorithm for recognizing facial emotions while dealing with the issue of interclass pixel mismatch during classification [15]. For example, a min-max metric may be used in the closest neighbor classifier to reduce feature outliers after the use of pixel normalization. Pham et al. have looked at techniques for augmenting data using both traditional and generative adversarial network (GAN) approaches to create hybrid data (HDA) [16]. A deep learning system, known as ADCRNN, is meant to assess the efficacy of HAD approaches by merging deep dilated CNN with an attention method. In addition, the deep learning system uses

3D log Mel-spectrogram (MelSpec) characteristics as inputs. To identify the feelings, we use a loss function that incorporates the softmax loss and the center loss. The experimental findings show that the suggested approaches outperform traditional algorithms on the EmoDB with 86.12% and 87.47% accuracy, correspondingly [17]. A unique classifier that is a combination of a cascaded Gaussian mixture framework and a DNN was suggested for a message and speaker-independent emotion identification system. An Emirati voice resource containing six distinct emotions was used to test this hybrid classifier for the identification of the six different emotions. SVMs and MLPs have been compared, and the sequential GMM-DNN classifier has a performance accuracy of 82.96%, while the other two outperform at 81.33 and 68.77% using SVMs and MLPs, respectively, according to the outcomes of this study. Owing to a dearth of review publications focusing on the three modalities, researchers offered a method to identify related research of emotion identification using face, voice, and textual signals depending on deep learning methods [18]. This study begins with an introduction to commonly recognized models of emotion to understand the term "emotion" itself. Using unimodality, we provide the current state of the art in emotion identification, which includes facial expression, voice, and text-based emotion recognition. Methods for multimodal emotion identification are discussed in depth in this article. As an additional convenience to readers, the definition of key benchmark data sets, their definition, and the effectiveness of the present state of the art in the past few decades are also included. Ultimately, we look at several possible research difficulties and possibilities to help academics better understand how to improve their study on emotion identification. Gou and Suthaharan also offered an approach using the widely known Yale-Faces picture data set to analyze the idea that emotional characteristics are concealed in the frequency range, such that they may be collected using frequency range techniques and masking approaches [19]. Accordingly, researchers use random forest (RF) and ANN classifiers' performance scores as metrics of the collected emotional frequencies' efficacy. Identifying the most essential and discriminative characteristics for each SER was done using the feature selection (FS) method [20]. The categorization of emotions is a complex issue that needs the application of several algorithms. RFs, decision trees, support vector machines, multilayer perceptrons, and KNN are used to identify seven emotions. A total of four publicly available databases are used in all studies. Researchers intend to look at whether there is a perceived barrier between new degree holders and their employers regarding the abilities that are needed to acquire a job, they seek to learn more about the effect of emotional intelligence (EI) on academic achievement [21–25]. In a research, the author suggests a paradigm for analyzing fuzzy system dependability using fuzzy number mathematical operations, where each computer system's stability is represented by a triangular fuzzy number [26]. The strategy relies on simple fuzzy mathematical operations using fuzzy integers rather than

complicated arithmetic and logical intervals. In this work, analytic analysis is undertaken to determine the impact of COVID-19 on the total number of cases, total recovery, and total death reported during the pandemic. This investigation was carried out using a variety of machine-learning algorithms. For this analysis, Indian states were chosen, and all research was conducted using a data set available on the WHO's official website. Finally, the analytical investigation was carried out using the RF and decision tree regression models [27]. The era of computers and digitization has begun. The vastness of cyberspace has transformed the way we do and see things. Most of our activities and future planning now rely on technology. In this new technology-driven and dominated environment, we must rely on machine-learning to optimize our operations and adapt to the new culture. To help the citizen choose the ideal residential location and the police agency fight crime using the data set [28].

The spotted hyena optimizer is a unique metaheuristic algorithm inspired by spotted hyena behaviour. This algorithm is based on spotted hyenas' social relationships and cooperative behaviour. SHO's three basic steps are hunting, circling, and attacking, all mathematically described and implemented. On the basis of 29 well-known standard benchmark functions, eight newly created metaheuristic algorithms are pitted against one another. Emperor penguin optimizer is a new optimization tool dependent on the huddling behaviour of emperor penguins. By constructing the huddle border and calculating the temperature surrounding the huddle, this technique determines the effective mover. Models for all 44 of the most popular benchmark test functions have been created. It is compared to eight cutting-edge optimization methods [29]. The seagull optimization algorithm is a unique bio-inspired method for addressing computationally difficult tasks. This algorithm's major inspiration is a seagull's natural movement and attack tendencies. As part of a specific search area, these behaviours were mathematically expressed and implemented [30]. The Sooty Tern Optimization Method is a bio-inspired method for tackling limited industrial problems. Based on a sooty tern's natural movement and attack patterns, this algorithm was created. Search space exploitation and exploration may be emphasized using these two methods together [31]. For single-objective limited optimization situations, SSA, depending on Hooke's law, is suggested here. In the SSA, search agents are weights connected by springs, each with a force proportional to its length. The text explains the algorithm's mathematics [32]. The results of the proposed method are tested using 25 typical benchmark functions. In addition, the statistical correctness of the proposed technique is examined in this work. It has been shown via research and comparisons that the suggested approach is both more effective and efficient than existing approaches. In the same way, the approach can be applied to the issue of selecting the best features. The binary emperor penguin optimization algorithm is found to be superior, according to the findings [29]. The salp swarm algorithm (ESA) and emperor

penguin are two components of a hybrid bio inspired metaheuristic optimization technique developed in this study to solve optimization issues. Two well-known optimization algorithms, the emperor penguins optimizer and the salp swarm method, both exhibit huddling and swarm tendencies, which this programme replicates [33].

Problems are analyzed rather than unsupervised solutions, previous work on emotion identification has tended to rely on supervised techniques. Furthermore, most research has relied on data sets that only include a few numbers of so-called fundamental emotions, even though daily interactions include a broad range of more complex emotional states. As a result, it is unclear if using a multisolution technique to research a larger number of emotions may lead to new insights into the subject. As a result, the primary issue of previous method is the seeming uniformity of supervised techniques, as well as the traditional usage of very small emotion data sets. The various applications of the proposed algorithm comprises of handwritten digit recognition, detection of handwritten cheques by banks, and traffic signal recognition. The major contributions of this paper include the following:

- (i) This study proposes the development of an EI system using artificial intelligence-based convolutional neural networks
- (ii) The proposed system can identify facial emotions employing AI technique to enhance accessibility
- (iii) The proposed system assists small and medium enterprises (SMEs) in their performance as well as in entrepreneurial orientation

2. Material and Methods

The flow of the recommended technique is explained in this section. The video to picture frame conversion begins with the input video and the geneva multimodal emotion depiction corpus data set. The data set is further preprocessed with a glitter band Butterworth filter and contrast stretch histogram equalization and the features are clustered using the hybrid Gaussian BIRCH algorithm, and the features are grouped and selected using Ada delta bacteria foraging optimization, and it is classified to train and test the features using kernel boosting LENET, and the output is classified, and the performance of the suggested techniques is evaluated. The recommended process is shown in a diagrammatic form in Figure 2.

2.1. Input Video Data Set. The audio toolkit, unlike the video toolkit, only delivers one instance of every file. As a result, a method for achieving information constancy is required. To perform, video examples were combined by document and averaged column-wise, resulting in a total of 1,200 observations in the data set. Finally, training and test subsets were formed by dividing feature values by 5 for normalization reasons. Two major stages were completed in this section. FFmpeg1 was used to extract audio tracks from the database videos. Second, two separate parameter sets were extracted

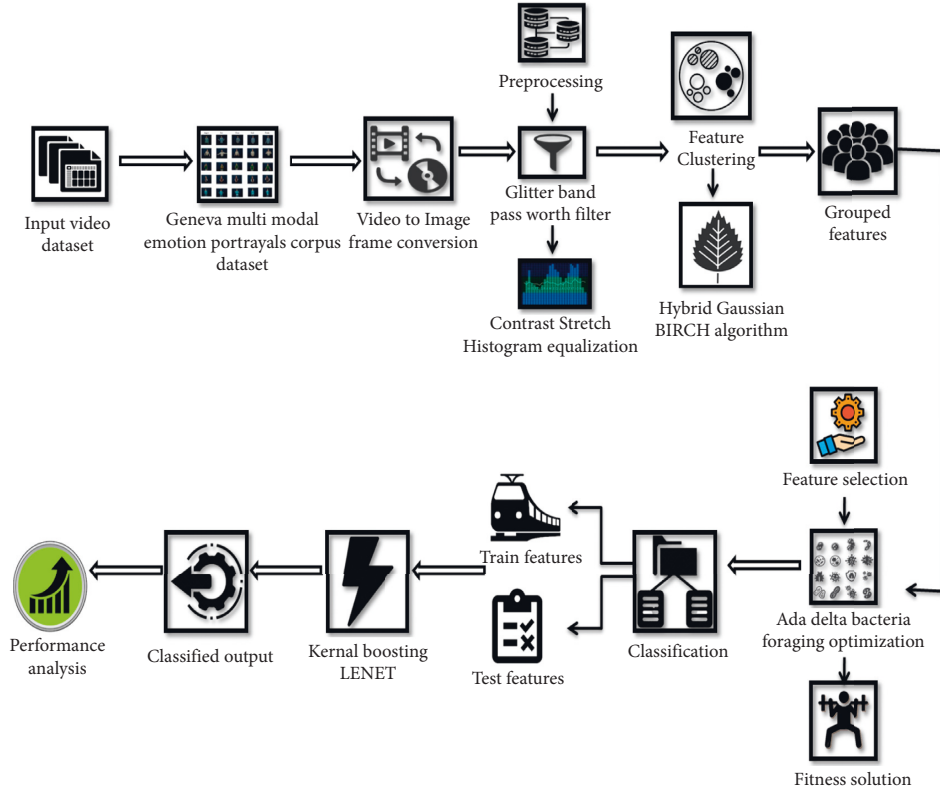


FIGURE 2: Schematic representation of the proposed method.

using openSMILE. It took 15 minutes to complete these two stages. The standard version (GeMAPS) has 62 features, whereas the expanded version (eGeMAPS) has more. More information about the audio features may be found here. One file was destroyed during the video data cleaning operation. As a result, the matching audio track had to be removed from both audio characteristics collection, resulting in a total of 1,259. After the attributes were retrieved and the information was cleaned, both feature collection were stratified into testing and training collection. We next applied a min-max scaler to a training collection first, and then weights were applied to both the training collection and the testing collection. As a result, test cases were updated as if they were in a real-life setting, with fresh input values being altered depending on the prior fixed scale, guaranteeing that new samples are always translated into a decimal from 0 to 1 [34].

2.2. Geneva Multimodal Emotion Portrayals Corpus Data Set. The GEMEP Master collection comprises 1,200 audio and video recordings in which 10 professional actors, guided by a professional director, portray 18 distinct emotional states using two separate pseudolinguistic phonemes patterns or a prolonged vowel “aaa.” The emotions shown are included in the data set. These emotion IDs will be used in some of the charts in the future to make them seem nicer. This data set was selected above others as it includes a broader range of positive and negative emotions than other multimodal databases. As most of the displayed

emotions have 91 documents for each class, while others only have 31 records for each class, the data are unbalanced. Furthermore, performers have not interpreted all of the feelings, and the number of recordings per actor inside each emotion is not proportionately distributed in the majority of situations [35].

2.3. Video to Frame Conversion. The first process is to transform the video clip into a series of picture frames, which will be used to train the model. To get pictures from a movie, we use the cv2 library. The VideoCapture function reads a video file and turns it into an image frame sequence. Each acquired frame will include a two-dimensional array of integers carrying picture data. Pixels make up the pictures, and each pixel is made up of several numeric arrays. Red, green, and blue are the three color channels of colored graphics, each represented by a grid. The intensity of each cell in the grid is represented by a value ranging from 0 to 255. Every 20th frame is sent into our training model to capture a unique emotion each time.

2.4. Preprocessing. The architecture does not include preprocessing. A deep learning model, on the other hand, has the drawback of requiring sophisticated hardware. The application of filters to several numbers of photos with many dimensions puts a significant strain on the CPU. The use of a preprocessing phase may aid in the proper convergence of a deep learning model. The method entails identifying an ROI

in each face picture. The Viola-Jones technique was employed, as well as OpenCV software. After locating the ROI, it is removed from the picture and scaled down to 75×75 pixels. The picture is finally transformed to grayscale and stored in a new database.

2.5. Glitter Bandpass Butterworth Filter. The raw data were then preprocessed to eliminate artefacts such as power-line disturbance, which was eliminated using a 50 Hz bandpass filter, proceeded by a glitter bandpass Butterworth filter to retrieve the essential frequencies in the range of 0.05 to 100 Hz to eliminate noise. In our studies, researchers used a Glitter bandpass Butterworth filter to retrieve data from the signal in four main frequency bands: α (2–8 Hz), β (9–14 Hz), θ (15–31 Hz), and γ bands (31–46 Hz). The data are then transformed into 2D characteristics like PCC, which are used as input for CNN. It is worth emphasizing that the 2D properties encompass not only the frequency of every electrode but also its geographical position. In this work, 1D filtering is used instead of two-dimensional filtering to retain this information.

2.6. Contrast Stretch Histogram Equalization. To begin, the facial picture must be transformed to grayscale format, which consists only of white and black hues. The face picture will next undergo contrast stretching, which is a method of histogram equalization that may enhance contrast by increasing the dynamic range of the histogram. The purpose of histogram equalization is to create a stable histogram. Each pixel is given a new intensity value depending on its previous intensity degree. Histogram equalization begins with histogram generation, which involves collecting and representing the color of picture data in a histogram. Next, using the equation, generate new intensity values for each intensity level (1). When it comes to interacting with computer-dependent learning surroundings, emotions are a critical component of successful learning and fixing the issues. The greatest intensity level that a pixel may achieve is called max, intensity level. The number of pixels with an intensity less than or equal to the output intensity level is expressed in the bracket. Where I is a prior intensity value, and O_i is the current intensity value. The last step is histogram equalization, which involves intensity modification for picture improvement. It uses the picture's cumulative density function to modify the brightness, then flattens the histogram and stretches the contrast of the image to be spread throughout every grey-level intensity degree. At last, the new intensity numbers are substituted for the existing intensity numbers.

$$O_i = \left[\sum_j^i N_j \right] \times \text{Max}_{\text{No.of.pixels}} \cdot \text{Intensity Level}. \quad (1)$$

2.7. Feature Clustering Using Hybrid Gaussian BIRCH Algorithm. Balanced iterative reducing and clustering with hierarchies is developed from hybrid Gaussian BIRCH. It

is the first method in the database field to deal with outlier information point that must be treated as noise and provide a viable solution. In hybrid Gaussian birch clustering, there are two sorts of techniques: probability and distance-dependent techniques. The assumption in chances dependent techniques is that probability dispersion on different qualities is statistically independent of one another. To find clusters, a probability-based tree is constructed. At the granularity of data points, distance-based techniques refer to global or semiglobal procedures. Assume that every information point is provided ahead of time and that they can be inspected periodically. Ignore the notion that not every information point in the data set is equally valuable. None of them can scale linearly in time while maintaining consistent quality. While scanning the data set, the Hybrid Gaussian Birch method creates a dendrogram known as a clustering feature tree. Every clustering choice is performed without scanning all information points in the hybrid Gaussian BIRCH. As each data point is potentially crucial for grouping, the hybrid Gaussian BIRCH derives the fact that the data space is not evenly inhabited. The hybrid Gaussian BIRCH takes maximum use of available memory to extract the finest feasible subgroups, ensuring precision while reducing I/O costs. The cluster of N -dimensional data points of centroid, radius, and diameter which are defined as where $i = 1, 2, \dots, N$.

$$\begin{aligned} \vec{a}_0 &= \frac{\sum_{i=1}^M \vec{a}_i}{M}, \\ S &= \left(\frac{\sum_{i=1}^M (\vec{a}_i - \vec{a}_0)^2}{M} \right)^{1/2}, \\ D &= \left(\frac{\sum_{i=1}^M \sum_{j=1}^M ((\vec{a}_i - \vec{a}_0))^2}{M(M-1)} \right)^{1/2}. \end{aligned} \quad (2)$$

The qualities of a single cluster are determined by R and D . Next, we establish five possible lengths for determining the distance between two groups. The two separate distances $D0$ and $D1$ are defined as given the centroid of two clusters.

$$D0 = \left((\vec{a}_{0_1} - \vec{a}_{0_2})^2 \right)^{1/2}, \quad (3)$$

$$D1 = \left| \vec{a}_{0_1} - \vec{a}_{0_2} \right| = \sum_{i=1}^d \left| \vec{a}_{0_1}^{(i)} - \vec{a}_{0_2}^{(i)} \right|. \quad (4)$$

Mean intergroup distances is $D2$, then mean intracluster distance is $D3$, and also the variance increase distance is $D4$ of the two groups if there are $N1$ d -dimensional statistics in one clusters I and $N2$ data sets in the next group. $D3$ is the D of the connected cluster.

$$\begin{aligned}
D2 &= \left(\frac{\sum_{i=1}^{M_1} \sum_{j=M_1+1}^{M_1+M_2} (\vec{a}_i - \vec{a}_0)^2}{M_1 M_2} \right)^{1/2}, \\
D3 &= \left(\frac{\sum_{i=1}^{M_1+M_2} \sum_{j=1}^{M_1+M_2} (\vec{a}_i - \vec{a}_0)^2}{(M_1 M_2)(M_1 + M_2 - 1)} \right)^{1/2}, \\
D4 &= \sum_{k=1}^{M_1+M_2} \left(\vec{a}_k - \frac{\sum_{i=1}^{M_1+M_2} \vec{a}_k}{M_1 M_2} \right)^2 - \sum_{i=1}^{M_1} \left(\vec{a}_i - \frac{\sum_{i=1}^{M_1} \vec{a}_k}{M_1} \right)^2 - \sum_{j=M_1+1}^{M_1+M_2} \left(\vec{a}_j - \frac{\sum_{i=M_1+1}^{M_1+M_2} \vec{a}_k}{M_2} \right)^2.
\end{aligned} \tag{5}$$

Two clusters are measured using $D0$, $D1$, $D2$, $D3$, and $D4$. The two groups will be utilized to assess whether they are closed. There are four phases in the hybrid Gaussian BIRCH: loading, optional condensing, global clustering, and optional refining. Phase 1's major goal is to scan every piece of information and create a beginning in-memory CF-tree using the available RAM and disc space. Finally, the hybrid Gaussian BIRCH makes no assumptions about the independence of probability distributions for different properties.

2.8. Feature Selection Using AdaDelta Bacteria Foraging Optimization. The filters are used to remove speckle noise in the first stage. The filtered output images and noisy must be given as search space parameters in an AdaDelta BFO technique in the second phase to decrease defects due to fluctuation in filtered and noisy pictures. With constant step size, two significant difficulties influence AdaDelta Bacterial Foraging Optimization.

- (I) If the step size is really low, it will take several generations to find the best solution. It is possible that with less iteration, it will not be able to attain global optimum.
- (II) If the step size is large, the bacterium will quickly achieve the optimum value, but the precision of the optimum value will be minimal. Likewise, in the AdaDelta BFO, the chemotaxis stage establishes a base for local search, the reproduction procedure accelerates convergence, and removal and dispersion aid in avoiding premature convergence. In AdaDelta BFO, mutation by PSO is used instead of removal and dissemination occurrence to acquire adjustable step size, raise velocity, and to prevent premature convergence.

$$(j+1, k) = \theta^j(j+1, k) + *r_1 * C_1(\theta^j(j+1, k) - \theta_{\text{global}}). \tag{6}$$

$\theta^j(j+1, k)$ = Position vector of i th bacteria in j th chemotaxis step and k th reproduction stages.

θ_{global} = good location in the full search area. To attain global optima, the BFpPSO undergoes chemotaxis, swarming, mutation, and reproduction stages.

The BF-pfPSO step-by-step technique is shown as follows.

Initialize specifications S , N_c , N_s , p , N_{re} , N_{ed} , P_{ed} , and $C(i)$, $i = 1, 2, \dots, S$, where, p = dimension of the search area, S = count of bacteria in the density, N_c = count of chemotaxis phases, N_s = count of swimming phases, N_{re} = count of reproduction phases, P_m = mutation probability,

$C(i)$ = step size considered in the random way indicated by the tumble (i, j, k) = Position vector of the i th bacterium, in j th chemotaxis step, in k th reproduction phase and in l th removal and dispersal step.

Stage 1: Reproduction cycle: $k = k \times 1$

Stage 2: Chemotaxis cycle: $j = j \times 1$

- (a) For $i = 1, 2, \dots, S$, take a chemotaxis phase for bacterium i as shown.
- (b) Calculate fitness operation $J(i, j, k, l)$.
- (c) Let $J_{\text{last}} = J(i, j, k)$ to save this value because researchers might search a best cost via a run.
- (d) Tumble: create a random vector $\Delta(i) \in \mathbb{R}^p$ with every element $\Delta_m(i)$ $m = 1, 2, \dots, p$, a random number on $[-1, 1]$.
- (e) Move: Let

$$(j+1, k, l) = \theta^j(j, k, l) + c(j) \frac{\Delta(j)}{\sqrt{\Delta^2(j) \Delta(j)}}. \tag{7}$$

- (f) Calculate $I(i, j+1, k, l)$
- (g) Swim
- (i) Let $m = 0$
- (ii) While $m < N_s$

Let $m = m \times 1$.

If $J(i, j+1, k, l) < J_{\text{last}}$.

Let $J_{\text{last}} = J(i, j+1, k, l)$ and let

$$(j+1, k, l) = \theta^j(j, k, l) + c(j) \frac{\Delta(j)}{\sqrt{\Delta^2(j) \Delta(j)}}. \tag{8}$$

And employ this $\theta_{\text{best}}(i, k)$ and θ_{global} to compute the new $J(j+1, k)$.

Else, let $m = N_s$. This is the end of the while statement.

- (h) Proceed to further bacteria ($i+1$) if $i \neq S$.

Stage 3: Upgrade and If $j < N_c$, go to.

stage 3. In this instance, keep chemotaxis, because the life of bacteria is not over.

Stage 4: Reproductions:

- (a) For the provided k and l , and for every $i = 1, 2, \dots, S$,
Let

$$I_{\text{health}}^j = \sum_{i=1}^{N_c+1} i(j, i, k), \quad (9)$$

be the health of bacterium i . Sort bacteria and chemotaxis parameter $C(i)$ in order of ascending cost J_{health} .

- (b) The $S_r = S/2$ bacteria with the maximum J_{health} values die and other $S_r = S/2$ bacteria with the great values split.

Stage 5: AdaDelta BFO.

For $i = 1, 2, \dots, S$, with probability P_m , modify the bacteria location by pPISO.

Step 6. If $k < N_r$, go to Step 2. Researchers had not attained the indicated count of reproduction phases. Therefore, researchers have to initialize the further generation in the chemotaxis cycle.

Equation (3) mean square error between noisy picture and wiener filter image must be used as cost operation to improve the peak signal-to-noise ratio in the bacterial Foraging technique.

$$\text{Error} = \frac{1}{mn} \sum_{j=1}^{mn} (z(x, y) - p(x, y))^2. \quad (10)$$

A size of both the noisy and the wiener filtered picture is denoted by MN here. Using the PSNR and MAE, the MBFO method's efficacy may be evaluated in terms of its precision (MAE)

$$\text{PSNR} = 10 \log_{10} \frac{2ss^2}{\text{Error}}. \quad (11)$$

Here $f'_{i,j}$ and $f_{i,j}$ denote the pixel value of restored and original image correspondingly.

$$\text{MAE} = \frac{1}{MN} \sum_{j=1}^{M,N} (f'_{i,j} - f_{i,j})^2. \quad (12)$$

2.9. Classification

2.9.1. Train Features. Researchers use a set of in-house corpora for training sets, which totals roughly 300 hours of voice data. The sounds are a combination of single words and small words that encompass the majority of the contextual diversity in the French language.

2.9.2. Test Features. Researchers test the models in a variety of scenarios:

- (i) Commands: short sentences or individual words, including "PerrosGuirec" or "Autrerubrique" (12778 utterances); • Digits (4735 utterances): solitary digits.
- (ii) Numbers: single-word numbers (NN) from 00 to 99, such as "10: dix" or "99: quatrevingtdix-neuf" (7244 utterances).
- (iii) TelNumbers: phone numbers written in the standard French format: 0 80X NN NNNN or NN NNNNNNNN (5664 utterances).

Where X is a number between 0 and 3.

For error accounting in Numbers and TelNumbers, every pair (NN) ranks as a single compound word, identifying 66 soixante.

One mistake ranks as six instead of 70 soixante dix.

2.9.3. Kernel Boosting LENET Classifier. The further enhancement metrics are used in this research is dependent on the kernel boosting LeNet classifier model: maximizing the number of convolution kernels is the initial step. Convolution kernels can be used to understand and interpret many image characteristics. Image attributes can be extracted more effectively by raising the number of convolution kernels. Furthermore, in the LeNet-5 method, the \tanh activation function is substituted with the ReLU activation function, that is given by (13)

$$\text{ReLU}(x) = \max(0, x). \quad (13)$$

The ReLU operation shall transform the negative signal to 0, preventing the negative signal input, and so alleviating the overfitting issue, as shown in equation (13). As a consequence, the ReLU operation can best match the training data, and a huge count of trials indicates that the ReLU operation outperforms the \tanh function in terms of interpretation. Third, to address the issue of a large number of trained specifications in the C5 and F6 levels of the LeNet-5 connections, the C5 and F6 layers are replaced by a convolution layer with an 1×1 convolution kernel counts of 33 and the global mean pooling approach, resulting in a significant reduction in the number of learning parameters. Average global pooling approach produces an $1 \times 1 \times n$ vector output with the same sample size as the average pooling layer. In the fourth step, the subsampling levels are removed. Even though using a subsampling layer shall minimize the number of input specifications and the likelihood of an overfitting issue because the subsampling layer employs an aggregation statistical function, it will eventually dismiss a few of the input picture's attribute values, potentially degrading the model's performance. Fifth, despite Lenet-5, C3 is entirely functional in our design.

3. Results and Discussion

The performance of the suggested classifier using the Kernel Boosting LENET deep learning algorithm in recognizing the human emotions is explained in this section. The simulation results were generated using MATLAB. The performance

metrics utilized to analyze the suggested classifier were accuracy, precision, sensitivity or recall, and F-measure. Emotional intelligence is thought to have a strong and favourable association with SME performance, according to the findings. The correlation coefficient of 0.188 and the significance level of 0.00 suggest a strong and positive relationship among emotional intelligence and the success of small and medium-sized enterprises. As a result, research supports H2's assertion that emotional intelligence has a positive and substantial association with firm success. This indicates that a manager's or owner's ability to recognize, interpret, evaluate, and control his or her own or others'/ groups' emotions will boost a small business's sales turnover and market share, order timely delivery, and product quality. To put it another way, the capacity of SMEs' owners or staff to accept and comprehend sentiments both within and outside the company (customers and suppliers) may lead to better results. The suggested classification model was compared to the existing classifiers namely deep learning neural network-regression activation (DR) [35], Convolutional Neural Network (CNN) [36], and NN-Levenberg Marquardt (NNLM) [37]. Existing techniques are technically difficult and costly. They would categorize images without identifying areas within them, which might be seen as a restriction in terms of interpreting system performance as it relates to various training choices. It would also be very hard to determine if a system has converged and is capable of generalization, or the capacity to categorize data that have not been seen before. Over fitting, exploding gradients, and class imbalances are other key issues when utilizing these existing strategies to trained the model. Table 1 indicates the behavioral analysis of various classifiers.

3.1. Accuracy. The accuracy refers to the emotions that were successfully categorized which are determined by (14)

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}. \quad (14)$$

When true negative is denoted by tn , true positive by tp , false positive by fp , and false negative by fn . Accuracy based comparison of the considered classifiers was given in Figure 3. Table 1 shows that the greatest classification accuracy was achieved by the proposed classifier (98%) which was significantly higher than that of existing algorithms namely DR, CNN, and NNLM. Hence it is confirmed that our proposed Kernel Boosting LENET deep learning algorithm had exceptionally high accuracy for the prediction and recognition of human facial emotions.

3.2. Precision. Precision (P) is the number of human emotions identified that were aberrant within the positive results, which is defined by (15)

$$P = \frac{tp}{tp + fp}. \quad (15)$$

Figure 4 shows that the proposed classifier obtained the highest precision percentage compared to that of DR, CNN,

and NNLM. This in turn denoted that the number of false-positive (FP) emotions recognized was significantly lower in the illustrated classifier.

3.3. Sensitivity. Recall (R) or sensitivity is one of the essential performance criteria in analyzing the effectiveness of the classifier which is determined by (16)

$$R = \frac{tp}{tp + fn}. \quad (16)$$

Figure 5 shows that, when comparing to DR and CNN, and NNLM, the suggested classification method has the highest sensitivity %. This in turn denoted that the number of false-negative (FN) emotions recognized was significantly lower in the illustrated classifier. Lesser number of FP and FN supported the efficiency of our proposed model.

3.4. F-Score. The weighted harmonics mean of P and R , as calculated by equation (17), is the F-score (F1):

$$F1 = \frac{2 * (R * P)}{R + P}. \quad (17)$$

In addition, the highest F1 score percentage was noticed for the proposed approach in comparison with the other classifiers from Figure 6. F1 score is a better metric than accuracy in the estimation of the model's performance. A significantly higher F1 score for the suggested model showed that this algorithm was advantageous in the classification of human emotions. Finally, it is confirmed that our proposed Kernel Boosting LENET deep learning algorithm showed the best classification performance in differentiating and recognizing the human facial expressions compared to existing techniques considered in this research from the outcome analyses.

We found that the suggested approach performed unexpectedly well in real-time classification of facial emotional expression, outperforming current techniques. On the basis of better detection accuracy, the suggested method could ensure the platform's real-time performance while also effectively reducing the platform's complexities. Emotion recognition helps in strengthening work performance and social dealings. Results clearly indicate that emotion recognition and social dealings cannot be separated.

In recent years, AI is providing us with a new scope of development in response to the burgeoning technology trends. AI provides numerous solutions to the difficulties we encounter, as well as increasing the efficiency of solutions that we have solved or working to overcome. However, it is still in the process of being perfected because it still has to deal with the ability to learn and comprehend in order to imitate human intellect. AI work faster than a person through observation and experience, although this is currently limited to only a few domains. Although, AI is highly acknowledged for its contribution to society and influence today, it can be made more human by incorporating Emotional Intelligence into it. AI can be allowed to broaden its realms of knowledge and deliver more advanced answers

TABLE 1: Performance analysis of various classifiers.

| Classifier name | Accuracy (%) | Precision (%) | Sensitivity (%) | F-score (%) |
|---|--------------|---------------|-----------------|-------------|
| Kernel boosting LENET classifier [proposed] | 98.5 | 94 | 95 | 93 |
| DR [35] | 97 | 90 | 92 | 91 |
| CNN [36] | 93 | 79 | 78.5 | 78 |
| NNLM [37] | 88 | 66 | 65 | 64 |

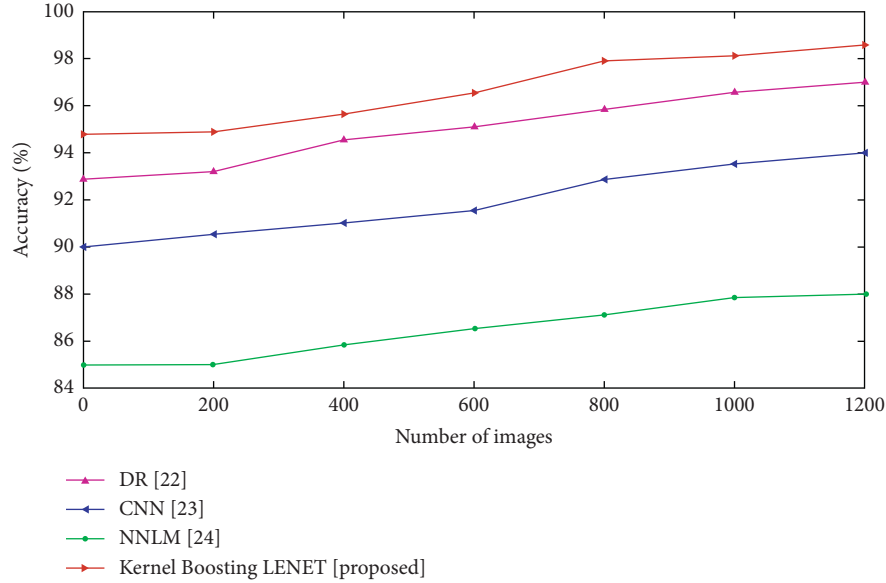


FIGURE 3: Comparative analysis of classifiers based on the accuracy.

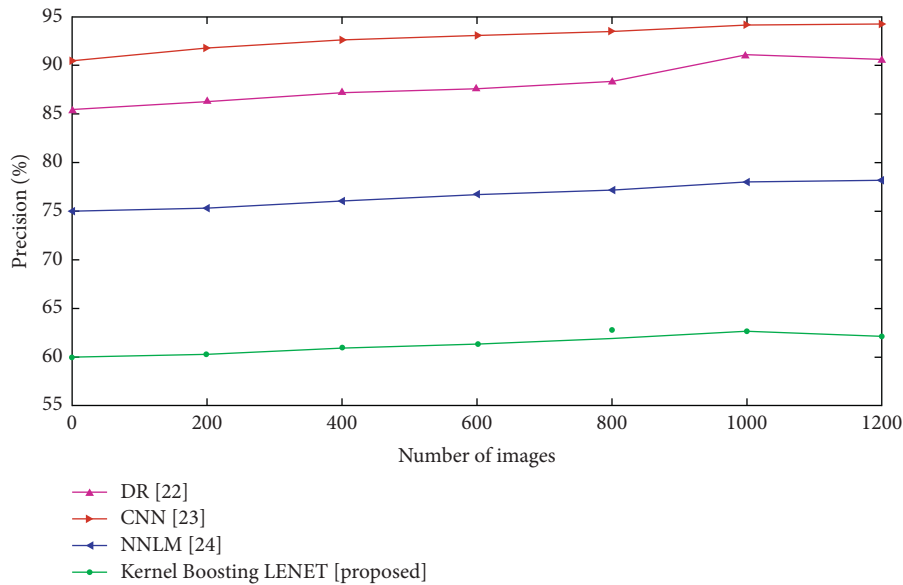


FIGURE 4: Comparative analysis of classifiers based on precision.

to complex situations with the help of Emotional Intelligence. If we can develop Emotional AI, we will be able to bridge the gap between humans and machines. It can help in industries such as medicine, consultation, education, and more, as well as open up new prospects through equal treatment. Emotional intelligence can materialize human-like behavior, resulting in the finest possible human-

machine partnership. It provides reliable analysis and reporting over time, allowing for a better grasp of real-world issues. AI with Emotional Intelligence can help ignite the reach of technology by providing more efficient and precise development procedures. When compared to current AI methods, the application of Emotional AI provides a much more comprehensive understanding of how machines may

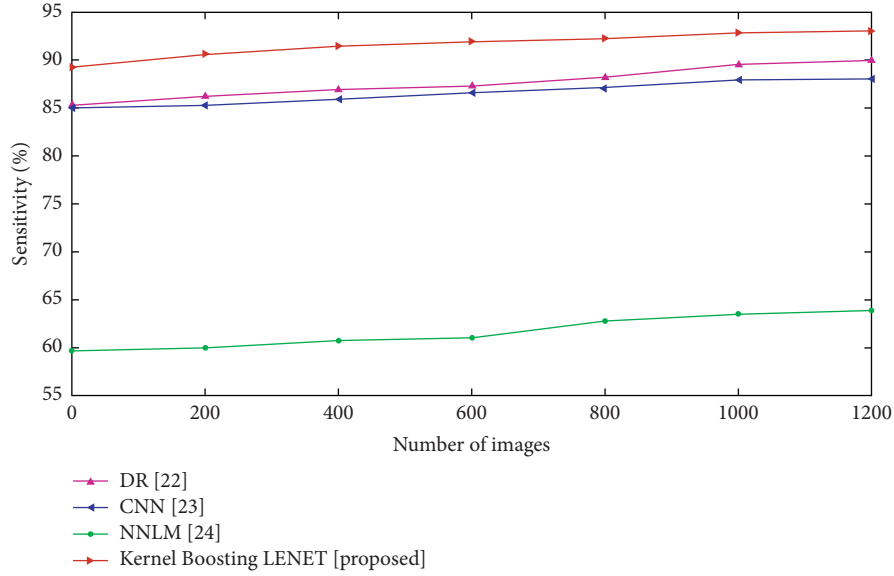


FIGURE 5: Comparative analysis of classifiers based on the sensitivity.

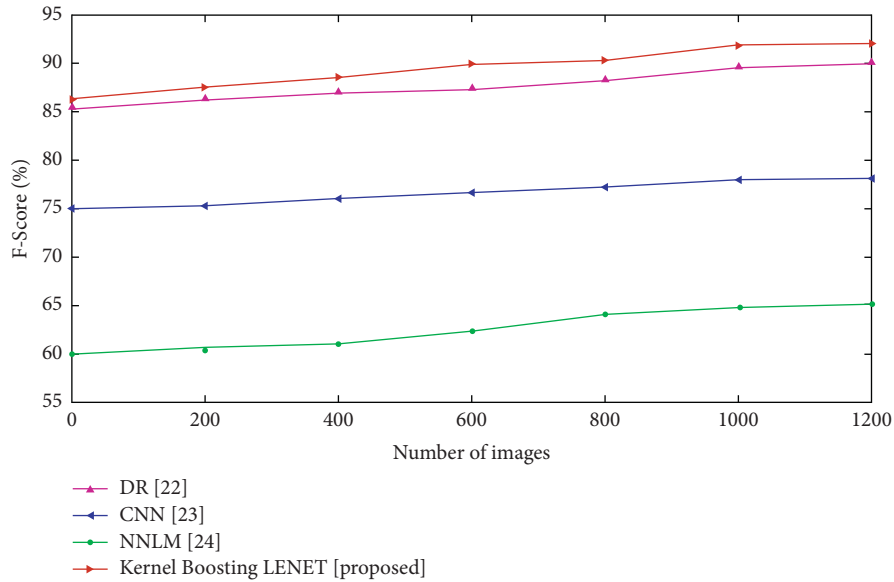


FIGURE 6: Comparative analysis of classifiers based on F-score.

assist humans. Traditional AI relies on logic and efficiency to handle and master problems in a certain field of study, if we incorporate Emotional Intelligence into AI, the technology will be able to expand into new fields of study such as healthcare, education, and consulting. Emotional bias is used in all professions to provide stability for persons who are coping with emotional problems.

AI propels the expansion of an enterprise reach and relationships. It will take the place of humans in repetitive and physical labor jobs in order to boost production. The other element of technology is the development of a quality called Emotional Intelligence, which improves soft skills, communication practice, and analysis. Tone, pitch, eye contact, body language, and facial expression are all detected, analyzed, and processed using AI algorithms. It

also has an impact on other types of verbal and nonverbal communication, assisting in the adoption of the appropriate emotion at the appropriate time with the appropriate individuals. As a result, AI's potential uses will go beyond simply comprehending one's emotions, duties, and obligations. It also tries to broaden the reach of traditional sales and customer service models. The first empirical study carried out by Uma et al. compared how organizations use emotional intelligence training to improve virtual communication and decision-making amid pandemic [21, 39].

A study conducted by Cao et al. demonstrated the application and investigation of AI to recognize emotions. On the basis of using CNN, the premise for emotion recognition using AI is considered. The primary emoji classification

techniques are examined, and the CNN architecture is constructed. The authors used EEG signals from the DEAP data set to identify the subjects' emotions; the data set that represents the emotional classification research. The primary emotional EEG features are determined using principal component analysis, which reduces the dimension of the preprocessed EEG data. The CNN algorithm is then used to test the accuracy of the training and test sample classifications, and the other classification methods are compared to obtain the nerves. The network outperforms classical learning techniques as a robust classifier for brain signals [22, 40].

Mehmet et al. proposed a CNN-based LeNet architecture for facial expression recognition. Initially, they have merged three data sets (JAFPE, KDEF, and their data set) and trained their LeNet architecture for emotion states classification. The results demonstrated an accuracy of 96.43% and validation accuracy of 91.81%. The accuracy has been calculated with seven different emotions through facial expressions [41].

James et al. applied deep neural network architecture for facial expression recognition, leveraging machine-driven design exploration in the design process, and the resulting architecture had architectural characteristics such as high architectural heterogeneity and selective long-range connectivity that were not seen in previous FEC network architectures. The proposed EmotionNet Nano networks showed accuracy comparable to state-of-the-art FEC networks while using much less parameters, according to experimental results using the CK + facial expression benchmark data set. The scientists also showed that the proposed EmotionNet Nano networks could attain real-time inference speeds (e.g., >25 FPS and >70 FPS at 15 and 30 W, respectively) and great energy efficiency (e.g., >1.7 images/sec/watt at 15 W) [42].

The suggested classification model was compared to the existing classifiers namely deep learning neural network-regression activation (DR) [35], Convolutional Neural Network (CNN) [23–27, 36], and NN-Levenberg Marquardt (NNLM) [37]. Existing techniques are technically difficult and costly [28–30, 38]. They would categorize images without identifying areas within them, which might be seen as a restriction in terms of interpreting system performance as it relates to various training choices. It would also be very hard to determine if a system has converged and is capable of generalization, or the capacity to categorize data that has not been seen before [31–33]. Over fitting, exploding gradients, and class imbalances are other key issues when using these existing strategies to trained the model.

4. Conclusion

According to the study, emotional intelligence does have a considerable and favourable association with SMEs performance, social and personal life as well. Facial emotion recognition research would be an interesting area for many years to come, involving many engineers and scientists. In this article, a novel feature selection and feature clustering strategy based on the hybrid Gaussian BIRCH algorithm

with AdaDelta bacterial foraging optimization is used to improve multiclass emotion recognition. Based on the kernels boosting LeNet architecture, this research proposes a low-cost and effective method for real-time categorization of varied moods by facial expression. A novel deep learning technique is also presented, which investigates the distribution of kernel boosting LeNet classifier training parameters. To demonstrate the framework's superiority over existing techniques like DR, CNNs, and NNLM, experiments revealed that the model retains its high accuracy even when a huge number of training data are compressed. As a result, the ability of SMEs' owners or employees to accept and understand attitudes both within and outside the organization (clients and suppliers) may lead to improved outcomes.

Data Availability

On request, the data used to support the conclusions of this research may be obtained from the corresponding author.

Conflicts of Interest

The author declares that he has no conflicts of interest.

References

- [1] F. G. Hernández, R. Z. Cabada, M. L. B. Estrada, and H. R. Rangel, "Recognition of learning-centered emotions using a convolutional neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 34, no. 5, pp. 3325–3336, 2018.
- [2] S. Meriem, A. Moussaoui, and A. Hadid, "Automated facial expression recognition using deep learning techniques: an overview," *International Journal of Informatics and Applied Mathematics*, vol. 3, no. 1, pp. 39–53, 2020.
- [3] A. Chakraborty, A. Konar, U. K. Chakraborty, and A. Chatterjee, "Emotion recognition from facial expressions and its control using fuzzy logic," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 4, pp. 726–743, 2009.
- [4] N. Christou and N. Kanojiya, "Human facial expression recognition with convolution neural networks," in *Proceedings of the Third International Congress on Information and Communication Technology*, pp. 539–545, Springer, Makassar, Indonesia, October 2019.
- [5] C. W. Huang and S. Narayanan, "Characterizing Types of Convolution in Deep Convolutional Recurrent Neural Networks for Robust Speech Emotion Recognition," 2017, <https://arxiv.org/abs/1706.02901>.
- [6] K. Donuk, A. R. I. Ali, M. F. Özdemir, and D. Hanbay, "Deep Feature Selection for Facial Emotion Recognition Based on BPSO and SVM," *PoliteknikDergisi*, no. 1, p. 1, 2021.
- [7] C. W. Huang and S. Narayanan, "Characterizing Types of Convolution in Deep Convolutional Recurrent Neural Networks for Robust Speech Emotion Recognition," 2017, <https://arxiv.org/abs/1706.02901>.
- [8] G. C. Sergio and M. Lee, "Emotional video to audio transformation using deep recurrent neural networks and a neuro-fuzzy system," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8478527, 15 pages, 2020.
- [9] H. Muthusamy, K. Polat, and S. Yaacob, "Improved emotion recognition using Gaussian mixture model and extreme

- learning machine in speech and glottal signals," *Mathematical Problems in Engineering*, vol. 2015, Article ID 394083, 13 pages, 2015.
- [10] J. Liu, G. Wu, Y. Luo et al., "EEG-based emotion classification using a deep neural network and sparse autoencoder," *Frontiers in Systems Neuroscience*, vol. 14, p. 43, 2020.
 - [11] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with preprocessing stage in automatic emotion recognition," *Procedia Computer Science*, vol. 116, pp. 523–529, 2017.
 - [12] N. Sabri, N. H. Musa, N. N. A. Mangshor, S. Ibrahim, and H. H. M. Hamzah, "Student emotion estimation based on facial application in E-learning during COVID-19 pandemic," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 1.4, pp. 576–582, 2020.
 - [13] K. M. Ravikumar and S. Thejaswini, "Electroencephalogram based emotion detection using hybrid LongShort term memory," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 8, pp. 2786–2792, 2020.
 - [14] A. T. Eyam, W. M. Mohammed, and J. L. M. Lastra, "Emotion-driven analysis and control of human-robot interactions in collaborative applications," *Sensors*, vol. 21, no. 14, p. 4626, 2021.
 - [15] O. Krestinskaya and A. P. James, "September. Facial emotion recognition using min-max similarity classifier," in *Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 752–758, IEEE, New York, NY, U.S.A, September 2017.
 - [16] N. T. Pham, D. N. M. Dang, and S. D. Nguyen, "Hybrid Data Augmentation and Deep Attention-Based Dilated Convolutional-Recurrent Neural Networks for Speech Emotion Recognition," 2021, <https://arxiv.org/abs/2109.09026>.
 - [17] I. Shahin, A. B. Nassif, and S. Hamsa, "Emotion recognition using hybrid Gaussian mixture model and deep neural network," *IEEE Access*, vol. 7, Article ID 26787, 2019.
 - [18] T. Zhang and Z. Tan, *Deep Emotion Recognition Using Facial, Speech and Textual Cues: A Survey*, TechRxiv, China, 2021.
 - [19] A. Goulu and S. Suthaharan, "Facial Emotion Characterization and Detection Using Fourier Transform and Machine Learning," 2021, <https://arxiv.org/abs/2112.02729> a.
 - [20] A. Amjad, L. Khan, and H. Chang, "Effect on speech emotion classification of a feature selection approach using a convolutional neural network," *PeerJ Computer Science*, vol. 7, p. e766, 2021.
 - [21] U. Manzoor, A. Rizwan, A. Demirbas, and N. A. S. Hafiz, "Analysis of perception gap between employers and fresh engineering graduates about employability skills: a case study of Pakistan," *International Journal of Engineering Education*, vol. 34, no. 1, pp. 248–255, 2018.
 - [22] H. Alsulami, M. Bashir, A. Rizwan et al., "Impact of emotional intelligence on the academic performance and employability of female engineering students in Saudi Arabia," *International Journal of Engineering Education*, vol. 35, no. 1, pp. 119–125, 2019.
 - [23] A. Rizwan, S. H. Serbaya, M. Saleem, H. Alsulami, D. A. Karras, and Z. Alamgir, "A preliminary analysis of the perception gap between employers and vocational students for career sustainability," *Sustainability*, vol. 13, no. 20, p. 11327, 2021.
 - [24] A. Rizwan, M. S. Alvi, and M. M. Hammouda, "Analysis of factors affecting the satisfaction levels of engineering students," *International Journal of Engineering Education*, vol. 24, no. 4, pp. 811–816, 2008.
 - [25] A. Rizwan, M. Choudhary, M. Jahanzaib, and A. Ammar, "Analysis of factors affecting the stress level of engineering students from remote areas," *International Journal of Engineering Education*, vol. 29, no. 4, pp. 926–932, 2013.
 - [26] R. Kumar and G. Dhiman, "A comparative study of fuzzy optimization through fuzzy number," *International Journal of Modern Research*, vol. 1, no. 1, pp. 1–14, 2021.
 - [27] P. K. Vaishnav, S. Sharma, and P. Sharma, "Analytical review analysis for screening COVID-19 disease," *International Journal of Modern Research*, vol. 1, no. 1, pp. 22–29, 2021.
 - [28] V. K. Gupta, S. K. Shukla, and R. S. Rawat, "Crime tracking system and people's safety in India using machine learning approaches," *International Journal of Modern Research*, vol. 2, no. 1, pp. 1–7, 2022.
 - [29] G. Dhiman and V. Kumar, "Emperor penguin optimizer: a bio-inspired algorithm for engineering problems," *Knowledge-Based Systems*, vol. 159, pp. 20–50, 2018.
 - [30] G. Dhiman and V. Kumar, "Seagull optimization algorithm: theory and its applications for large-scale industrial engineering problems," *Knowledge-Based Systems*, vol. 165, pp. 169–196, 2019.
 - [31] G. Dhiman and A. Kaur, "STOA: a bio-inspired based optimization algorithm for industrial engineering problems," *Engineering Applications of Artificial Intelligence*, vol. 82, pp. 148–174, 2019.
 - [32] M. Dehghani, Z. Montazeri, G. Dhiman et al., "A spring search algorithm applied to engineering optimization problems," *Applied Sciences*, vol. 10, no. 18, p. 6173, 2020.
 - [33] G. Dhiman, "ESA: a hybrid bio-inspired metaheuristic optimization approach for engineering problems," *Engineering with Computers*, vol. 37, no. 1, pp. 323–353, 2021.
 - [34] F. Eyben, K. R. Scherer, J. Sundberg et al., "The geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190–202, 2016.
 - [35] S. D. Suputri and C. Satyanarayana, "An efficient facial emotion recognition system using novel deep learning neural network-regression activation classifier," *Multimedia Tools and Applications*, vol. 80, no. 12, Article ID 17568, 2021.
 - [36] D. Sánchez, P. Melin, and O. Castillo, "Optimization of modular granular neural networks using a firefly algorithm for human recognition," *Engineering Applications of Artificial Intelligence*, vol. 64, pp. 172–186, 2017.
 - [37] J. Shao and Y. Qian, "Three convolutional neural network models for facial expression recognition in the wild," *Neurocomputing*, vol. 355, pp. 82–92, 2019.
 - [38] G. Dhiman and V. Kumar, "Spotted hyena optimizer: a novel bio-inspired based metaheuristic technique for engineering applications," *Advances in Engineering Software*, vol. 114, pp. 48–70, 2017.
 - [39] U. Warriar, A. Shankar, and H. M. Belal, "Examining the role of emotional intelligence as a moderator for virtual communication and decision making effectiveness during the COVID-19 crisis: revisiting task technology fit theory," *Annals of Operations Research*, vol. 13, pp. 1–17, 2021.
 - [40] G. Cao, Y. Ma, X. Meng, Y. Gao, and M. Meng, "Emotion Recognition Based on CNN," in *Proceedings of the 2019*

- Chinese Control Conference (CCC)*, pp. 8627–8630, Guangzhou, China, July 2019.
- [41] M. A. Ozdemir, B. Elagoz, A. Alaybeyoglu, R. Sadighzadeh, and A. Akan, “Real time emotion recognition from facial expressions using CNN architecture,” in *Proceedings of the 2019 Medical Technologies Congress (TIPTEKNO)*, pp. 1–4, Izmir, Turkey, November 2019.
 - [42] J. R. Lee, L. Wang, and A. Wong, “EmotionNet Nano: an efficient deep convolutional neural network design for real-time facial expression recognition,” *Frontiers in Artificial Intelligence*, vol. 3, pp. 1–9, 202.

Copyright © 2022 Suhail H. Serbaya. This is an open access article distributed under the Creative Commons Attribution License (the “License”), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License. <https://creativecommons.org/licenses/by/4.0/>