Library Progress International
Vol.44 No. 3, Jul-Dec 2024: P. 23558-23570

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article

Available online at www.bpasjournals.com

Prediction of Human Introspection in Facial Expressions through Emotional Intelligence and Machine Learning Techniques

[1]Dr.Maruthaveni, [2]A,Aysha

[1]Reasch supervisor & Associate Professor, Department of Computer Science, Dr.SNS.Rajalakshmi College of Arts and Science, Coimbatore,Tamil Nadu, dhanuvene1@gmail.com
[2]Research Scholar & Associate Professor, Department of Computer Science, J.B.A.S College for Women(Autonomous), Teynampet,Chennai-18, syedaysha@gmail.com

How to cite this article: Maruthaveni, A,Aysha (2024) Prediction of Human Introspection in Facial Expressions through Emotional Intelligence and Machine Learning Techniques. *Library Progress International,* 44(3), 23558-23570

Abstract
Human introspection in facial expression is the process through which people analyze and analyze their own emotions and feelings through the observation and interpretation of their facial expressions. The basic concept behind this idea is that people can learn about their thoughts and feelings by observing the expressions on their faces, which can act as a window into their emotional state. This study analyses the methods by which machine learning for facial expression interpretation is influenced by emotional intelligence. Its goal is to increase our comprehension of how emotional intelligence contributes to precise and thorough facial expression analysis. Our proposed work investigates the efficacy of a Convolutional ReLU Bidirectional Long-Term Neural Network (CRBLNN) for facial expression detection, utilizing an image dataset sourced from Kaggle. Building upon the foundation of machine learning and emotional intelligence, our approach employs a comprehensive multi-step strategy, integrating pre-processing, feature extraction, and feature selection techniques. We present a new approach to facial expression recognition that considers emotional intelligence. Initially, we apply histogram equalization as a pre-processing step to improve the overall quality and contrast of the images. Next, we extract features using Local Binary Pattern (LBP) to identify important patterns and textures in the facial expressions. Finally, we select the most informative features out of the enriched feature set using Select K-best feature selection. Our proposed CRBLNN architecture combines bidirectional Long-Short Term Memory (LSTM) units and convolutional layers with Rectified Linear Unit (ReLU) activation. The model can learn both temporal and spatial dependencies from the facial expression data due to this strong framework that the design offers. Beyond the technical aspects, we also explore the domain of emotional intelligence, recognizing its significance in the interpretation of facial expressions. The model's performance is thoroughly examined through the use of common metrics like F1-score, accuracy, precision, and recall. The results demonstrate how well the model can recognize and interpret facial expressions, highlighting the potential for machine learning and emotional intelligence to work together to improve facial expression recognition systems.
Keywords: *Image dataset, Histogram equalization, Local Binary Pattern (LBP), Select K-best, Convolutional ReLU Bidirectional Long-Term Neural Network (CRBLNN), Performance Analysis*

1. Introduction

The easiest and most effective method of recognizing emotions comes through facial expressions. For humans, faces are an integral component of daily existence [1]. People gesture as they encourage one another. Every day, they engage in face-to-face communication, whether it is through computers or in reality [2]. Monitoring a person's facial or vocal interactions could reveal their emotional state. Understanding people's emotions is essential for social engagement in public. A considerate human emotion requires a grasp of facial expressions. We can interpret human thought processes by recognizing emotions [3]. Humans engage effectively through emotion and a common language, according to assessments [4]. Facial expression recognition (FER) is widely used in human-computer interaction for a variety of purposes, including real-time portable expression recognition, detecting driving exhaustion, and identifying behavioral changes in offenders [5].The urgency of comprehending what facial movements can reveal about an individual's emotions arises from the widespread belief that such information has been established [6, 7].The term emotional intelligence (EI) refers to a collection of abilities that let people accurately assess and express their own emotions as well as those of others. EI is used to examine how people behave in social situations. EI is complexly intertwined, as human reflection in facial expressions indicates [8]. A brief grin or raised eyebrow is subtle cues that reveal the complexity of us as individuals. Recognizing these signs promotes feelings, a sense of connection, and a deep respect for the nonverbal communication that is contained within facial expressions. EI is a broad theory that explains the relationship between various emotional domains along with the way these domains could be responsible for individual responses that are neurological as well as emotional [9]. Sympathetic conscious emotions usually entail recognizing the kind of emotion that is being experienced (such as rage or sorrow); at a minimum, it entails understanding or learning about the relevant emotion and its characteristics. The unique first-personal method by which can be learn about, or at least understand, the conscious perceptions is called introspection [10]. Numerous researches have been conducted on emotions, and several models of emotion have been offered in the research; sadly, it is too difficult to determine that model was most accurate. Emotions have a significant role in how the human mind works. However, there has been a misapprehension and misinterpretation of the part that emotions play in our thoughts, behaviors, and actions. The goal of this research is to use EI to get an extensive understanding of emotional reactions as it examines and analyzes human introspection in facial expressions.

Key contributions

- The expression identification performance could be improved and provided the CRBLNN architecture can capture both spatial and temporal connections in face expressions.
- HE is employed as a pre-processing step to increase the general clarity and standard of images illustrating facial emotions, signalling that image enhancement is being explored to enhance the input to the model.
- Feature extraction using LBP demonstrates an emphasis on finding significant patterns and textures in facial expressions for more reliable identification.
- Selecting relevant features is crucial for achieving the greatest possible model performance, as shown by the introduction of Select K-best feature selection, which helps identify the most informative features.

The remaining sections of this work are organized as follows: In the second part discuss the related work. In third part, examines the suggested method. The findings are analyzed in the fourth part. The fifth part describes the conclusion.

2. Related works

The study [11] proposed a conditional generative adversarial network-based method to mitigate the intra-class variations by concurrently learning the generative and discriminative representations and individually regulating the face expressions. One of the most important tasks in emotion detection was FER. FER challenge, however, demonstrates a large difference between human and machine performance. The article [12] investigated novel adoration-based perceptual architecture for cooperation human-robot interaction (HRI), in which automation was anticipated to identify the emotional states of humans; it promotes an inherent connection between the person and the automated object. They provide a closed-loop technique that grades HRIs based on measured emotions. The measure could be as a way of rewarding the helper to modify their behavior adaptively. Deep neural networks use audio and image inputs to identify the emotional states of individuals. The study [13] examined the emotional states of 12 men volunteers in reaction to image- and video-content stimulation, as measured by electroencephalography (EEG) signals functional Near-Infrared Spectroscopy (fNIRS), and spontaneous facial expressions. The approach they proposed was to assess fNIRS and EEG data simultaneously to determine affective states. The findings of the experiment indicate that the observed emotional valence and spontaneous facial affective responses were strongly correlated.

The paper [14] focused on the 6 most commonly used emotion groups by researching customer's emotions: frustration, annoyance, anxiety, joy, sorrow, and delight. They begin by scrutinizing the prevalent belief known as the common perspective, delving into specific instances. Subsequently, they delve into an analysis of the scientific evidence that contradicts this assumption. A wealth of empirical evidence to support the conventional belief that humans sometimes display more emotion that would be predicted by chance, such as smiling when pleased, frowning when sad, scowling when furious, and other such behaviours. The research [15] provided a new kind of immediately affective interface in which the user participates in the conceptualization of their emotional state. The goal of (Mirror Ritual) which draws inspiration from Barrett's Theory of Constructed Emotion is to increase the user's comprehension of ideas related to emotions while eventually encouraging introspection and management of emotions. The user panel generates verse continuously by using categorized emotions that are gathered through FER. The article [16] discussed the construction and assessment of an emotion-aware chatbot that uses empathy to carry out experience sampling. We assess it in a week-long human-subject study including 39 people. Based on their findings, extroverts had a far stronger preference than introverts for the emotion-aware robot. Participants who interacted with the sympathetic bot also reported a larger proportion of good mood evaluations. In the end, they provide principles for creating emotion-aware robots that could be used in health settings.

The research [17] investigated the automatic identification of nonverbal actions from learners, such as head and gaze movements, emotional expressions, and hand-over-facial behaviors. The suggested computer vision-based approach to behavior monitoring makes use of an inexpensive camera and could be seamlessly incorporated with contemporary coaching tools. During a 40-minute class period that includes reading and problem-solving activities, they examine these habits in more detail over time. The research [18] examined the outcomes of the video-based analysis to the more traditional assessments, which included photoplethysmography (PPG) heart rate and emotion levels using symbols. They provided samples of grain and 3 different soybean sauces, with and without brand identification, to forty individuals who had varied levels of soybean sauce experience to find facial expressions. The findings demonstrated that particular flavours, rather than branding and familiarity, had more effect on pleasure and excitement. The goal of the research [19] discovered a method for identifying those individuals who are experiencing a challenging time in their own lives. Applied machine learning (ML) and artificial intelligence (AI)could determine a person's sentiment by their facial expressions recorded continuously, and then recommend to them some actions or ideas that could aid them achieve control beyond their emotions when feeling dissatisfied, scared, or otherwise to fix the situation in some level, assists because it wasn't always apparent that someone who was experiencing a difficult time could speak up about their feelings to people around them. The research [20] examined the emotion recognition deficits as well as processing, cooperation, and emotional expression difficulties in recognizing and describing these children primarily. It behaves that by exploring recent health and scientific understanding of psycho-emotional and behavioral development, and also EI, in children diagnosed with coexisting attention deficit hyperactivity disorder and autism spectrum disorder.

3. Methodology

This study analyzes the methods by which ML for facial expression interpretation is influenced by emotional intelligence. The work investigates the efficacy of a Convolutional ReLU Bidirectional Long-Term Neural Network (CRBLNN) for facial expression detection, utilizing an image dataset sourced from Kaggle. Initially, we apply histogram equalization as a pre-processing step to improve the overall quality and contrast of the images. Next, we extract features using Local Binary Pattern (LBP) to identify important patterns and textures in the facial expressions. Finally, we select the most informative features out of the enriched feature set using Select K-best feature selection. Our proposed CRBLNN architecture combines bidirectional long-term memory (LSTM) units and convolutional layers with Rectified Linear Unit (ReLU) activation. Figure 1 depicts the proposed flow of the research.

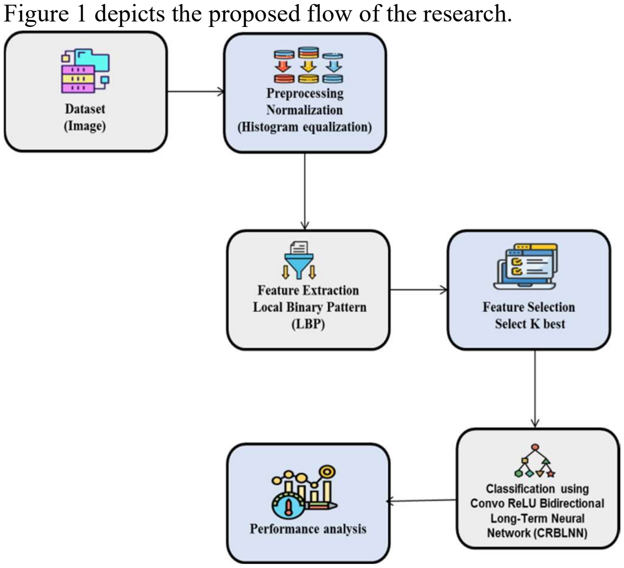

Figure 1: Proposed methodology

i. Dataset

For investigating the efficacy of a CRBLNN for facial expression detection, this study utilized an image dataset sourced from Kaggle. The dataset is composed of expressions such as surprise, sadness, neutrality, happiness, fear, disgust, contempt, and anger. https://www.kaggle.com/datasets/sudarshanvaidya/random-images-for-face-emotion-recognition

ii.Data Pre-processing using Histogram Equalization (HE)

To improve the overall appearance and contrast of the images the Histogram Equalization (HE) pre-processing phase. One method to increase the contrast in pictures is to use HE to make the images more contrast.

$$O(q_l) = \frac{m_l}{M} l = 0, \dots, K-1 \quad (1)$$

Where$m_l$the number of pixels with a gray level isq_l is the$l$ ^th^gray level. Another way to compute the Cumulative Distribution Function (CDF) is as follows in Equation (1-2):

$$D(q_l) = \sum_{j=0}^{j=l} O(q_j)(2)$$

$$l = 0, \dots, K-1, \;\; o \le D(q_l) \le 1$$

Gray level T_l is appropriated to gray level q_l viaHE. Thus, we have in Equation (3):

$$T_l = (K-1) \times D(-q_l)(3)$$

The variations in the gray level $T_l$can be calculated using the standard histogram equalization method as shown in Equation (4):

$$\Delta T_l = (K-1) \times O(q_l) \quad (4)$$

The input image at gray level $q_l$is directly related to the distance between T_l and T_l +1. Figure 2 depicts the visualization of HE.

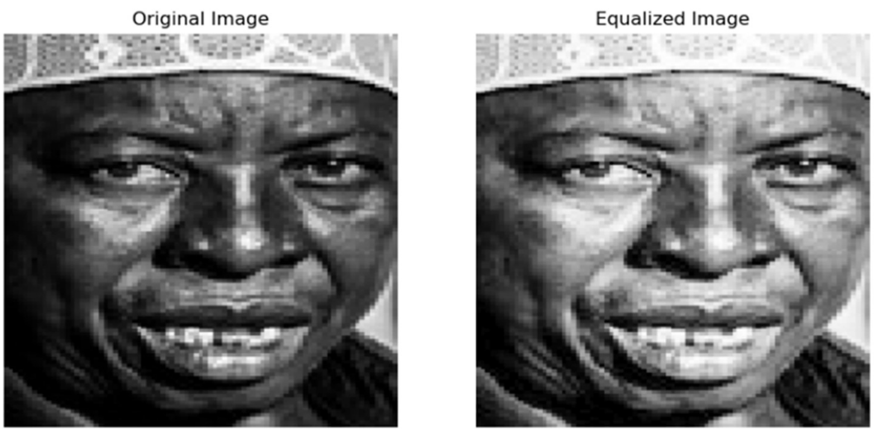

Figure 2: Visualization of HE

The process of histogram equalization involves shifting intensity values to improve an image's contrast. During the preparation stage, could apply histogram equalization to the image data. It enhances detail visibility in images with different lighting situations, which facilitates the capturing of pertinent patterns by feature extraction techniques later on.

ii. Feature extraction using Local Binary Pattern (LBP)

To recognize significant patterns and textures in facial expressions, extract features using Local Binary Pattern (LBP). Images can be characterized using textures by applying the LBP technique. A central pixel's intensity is compared to its nearby pixels to determine the image's local structure and patterns. The resulting binary patterns then describe the texture of the image. Neighborhoods of three-by-three rectangles comprised the first LBP definition as shown in Equation (5-7). The values of the eight surrounding pixels were compared to the core pixel to produce a binary pattern.

$$LBP_{p,R(Q_c,A_c) = \sum_{p=0}^{p-1} O(i_p - i_c)^{2^p}} \quad (5)$$

$$o(Q) = \begin{cases} 0, Q < 0 \\ 1, Q \ge 0 \end{cases} \quad (6)$$

$$Q = \begin{pmatrix} i_0 \; i \; i_2 i_2 \\ i \; i_c i_c i_3 \\ i_6 i_5 i_5 i_4 \end{pmatrix} \quad (7)$$

The central pixel's gray values are its radius and coordinates, and the symbols represent the neighboring pixels' gray values (Q_d, A_d),Q,i_p, and i_c, respectively.After the LBP patterns for each pixel in the image are identified, the histogram occurrence distribution for the entire image is subsequently employed as a feature descriptor in Equation(8).

$$I(m) = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} e(LBP(x,y), m), m \in [0, L], \quad (8)$$

Where $f(O, A) = \begin{cases} 1, \text{if } O = A \\ 0, \text{otherwise} \end{cases}$

The maximum value of the LBP code is displayed when the input image has a size of $G * H$and S.Figure 3 represents the LBP.

```
:  1  features

: array([[    0,   207,     0, ...,  6755,   817, 18227],
        [ 2536,  4404,  2948, ...,  5214,  3848, 11780],
        [ 5064,  4694,  2890, ...,  2934,  4050, 16613],
        ...,
        [    0,     0,     0, ...,  1675,   154, 31840],
        [    0,     0,     0, ...,  2974,   259, 26912],
        [    0,     0,     0, ...,  2164,   290, 29628]], dtype=int64)

:  1  labels

: array(['anger', 'anger', 'anger', ..., 'surprise', 'surprise', 'surprise'],
        dtype='<U10')
```

Figure 3: Representations of LBP

LBP is appropriate for texture-based feature extraction in images since it is computationally efficient and resistant to monotonic grayscale variations.

iii. Feature Selection using KBest

KBest models were used to rate the chosen features. The quantity of characteristics based on the K highest score is called KBest. To reduce the amount of chosen features that will be used for classification in this study, KBest is employed. Information Gain value features were arranged from greatest value (KBest). Based on the mutual information between the features and labels, the Select KBest algorithm was able to choose the feature with the highest relevance for the sample data set. The importance of each feature is estimated using two parameters, k, and the score_func, as input variables. K is the number of qualities that must be maintained. The following is a description of regression in Equation (9):

$$q_j = \frac{(w_j - \bar{w})^S (z - \bar{z})}{std(w_j) std(z)} \tag{9}$$

wherew_j and z are the input and output values, $\bar{w}$ and $\bar{z}$ are the corresponding averages, and std is the data's standard deviation. The test correlation coefficientq_j, indicates the significance of the input and output data sets.Figure 4 depicts the representations of KBest.

```
: array([[  207,  9425,   182,  6755,   817, 18227],
        [ 4404,  6087,  6668,  5214,  3848, 11780],
        [ 4694,  4197,  5429,  2934,  4050, 16613],
        ...,
        [    0,  2207,     0,  1675,   154, 31840],
        [    0,  3822,     0,  2974,   259, 26912],
        [    0,  2981,     0,  2164,   290, 29628]], dtype=int64)
```

Figure 4: Depiction of Kbest

Using statistical testing, the top k features are chosen using the SelectKBest feature selection technique. It does this by choosing the most useful features for classification, hence reducing the dimensionality of the feature space. Benefits include a lower chance of overfitting, enhanced interpretability of the model, and the potential to expedite training by utilizing only a subset of the most pertinent characteristics.

iv Prediction of Human Introspection using Convolutional ReLU Bidirectional Long-Term Neural Network (CRBLNN)

CRBLNN is a hybrid model that blends Bi-LSTM and CNN. The hybrid model, which combines the best features of both architectures, is appropriate for sequential and spatial information-related tasks, like sequential data picture categorization.

a. Convolutional ReLU

An input layer, a convolutional layer, a pooling layer, a linked layer, and an output layer constitute a CNN. Figure 5 shows the architecture of Convolutional ReLU.

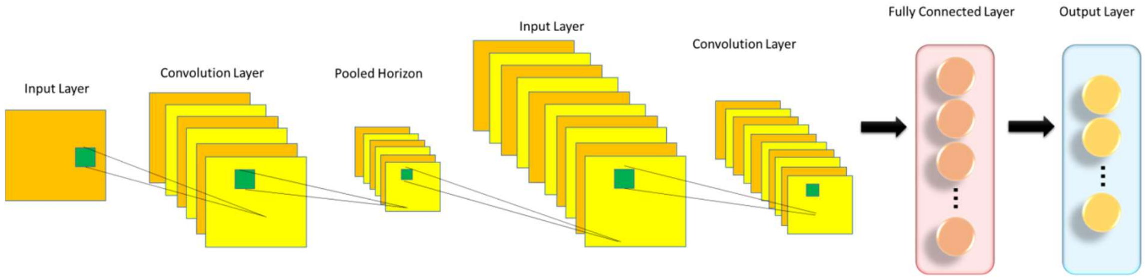

Figure 5: Convolutional ReLU architecture

The convolutional layer uses convolutional kernels, which can be described as follows, to execute convolution operations on the input features as shown in Equation (10):

$$w_i^k = e\left(\sum_{j\in N_I} w_j^{k-1} \times l_{ji}^k + a_i^k\right) \quad (10)$$

$e$is the activation function, $l_{ji}^k$is the input for the$k$-th layer, a_i^k is the bias variable for the i-th layer, and $N_I$is the convolutional kernel for the input feature vector. The i-th feature map output is represented by w_i^k in the equation.The output feature vector is constructed by the activation function. The ReLU function was employed as the activation function for this model to prevent the gradient from vanishing. The following is a representation of it in Equation (11):

$$R_{eLU}(w) = \begin{cases} w, w > 0 \\ 0, w \le 0 \end{cases} \quad (11)$$

The output data of the convolution operation process is represented by$w$in the equation.The predictive model's computational performance can be increased and its size can be decreased by using the pooling layer. Max pooling is the pooling operation selected for this model, and it can be shown as follows in Equation (12).

$$o_j^{k+1}(i) = \max_{(i-1)X+1\le iX}\{r_j^k(s)\} \quad (12)$$

The values corresponding to the neurons in the l + 1-th layer are represented by $O_j^{k+1}(i)$and X, respectively. The s-th neuron of the s-th feature vector in the k layer is represented by $r_j^k(s)$and$s \in [(i-1)XZ+1, iX]$, in the equation.

The completely connected layer will map the learnt features into the label space, as illustrated below in Equation (13):

$$e(w) = X \times w + a \quad (13)$$

The input is denoted by w, the weight matrix of the fully connected layer is represented by X, and the bias vector is indicated by a in the equation.

b.Bidirectional Long Short-Term Memory Network

Since bi-LSTM consists of two layers such as both forward and backward context information. Sequence labeling, machine translation, sentiment analysis, language modeling, time series prediction, and other fields have all made extensive use of it. The problems with vanishing and exploding gradients that can arise in typical RNN training are addressed by LSTM networks. They improve the capacity to manage data from long-term time series. The architecture of Bi-LSTM is shown in Figure 6.

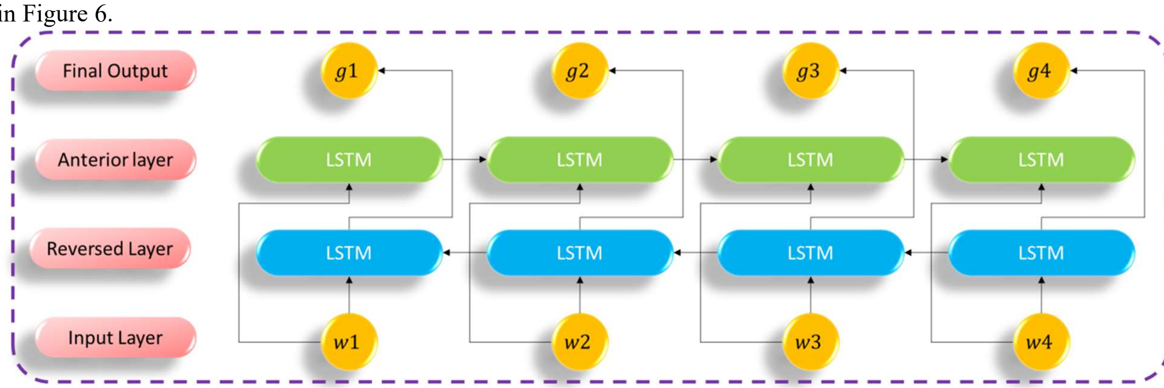

Figure 6: Bi-LSTM Architecture

The LSTM's mathematical calculation formula is displayed below in Equation (14-16).The value of an Oblivion Doore_s is

$$e_s = \sigma(X_e \cdot [g_{s-1}, w_s] + a_e) \quad (14)$$

Input door valuej_s is

$$j_s = \sigma(X_j \cdot [g_{s-1}, w_s] + a_j) \tag{15}$$

Output value$p_s$is

$$p_s = \sigma(X_p \cdot [g_{s-1}, w_s] + a_p) \tag{16}$$

The current state of the cell d_s is in Equation (17-19):

$$d_s = e_s \circ d_{s-1} + j_s \circ \tilde{d}_s \tag{17}$$

The state of the input cell at this moment$\tilde{d}_s[0,1]$ is

$$\tilde{d}_s = \tanh(X_d \ \backslash b7ullet[g_{s-1}, w_s] + a_d) \tag{18}$$

The final output$p_s$is

$$g_s = p_s \circ \tanh(d_s) \tag{19}$$

The CRBLNN prediction model was created by fusing the benefits of BiLSTM and CNN.

c.Convolutional ReLU Bidirectional Long-Term Neural Network (CRBLNN)

CRBLNN is a hybrid model that blends Bi-LSTM and CNN. Whereas Bi-LSTMs excel at identifying sequential patterns in data, CNNs are good at extracting spatial features. Bi-LSTM captures temporal dependencies, whereas CNNs capture spatial hierarchies and patterns. Figure 7 depicts the CRBLNN architecture.

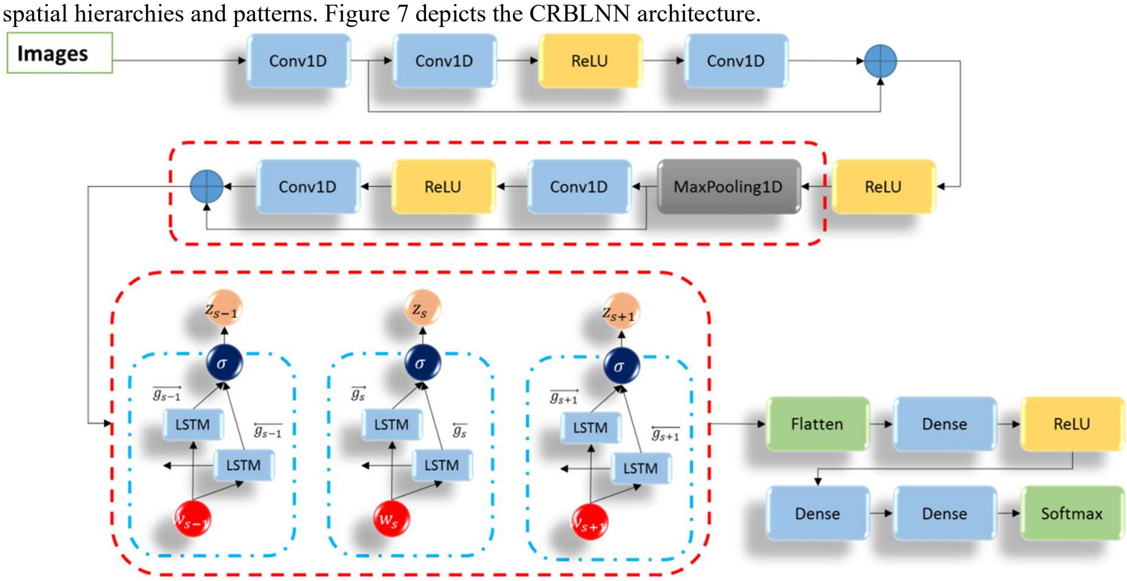

Figure 7: CRBLNN Architecture

4. Results

The suggested system is implemented using MATLAB. Images are used as input to identify and diagnose a person's emotions as well as to recognize and identify their emotions. The CRBLNN framework is assessed for the effective Prediction classification of Human Introspection in Facial Expressions. The effectiveness of the predictive model is calculated by metrics such as Accuracy, Precision, Recall, F1-score, Sensitivity, Specificity, Peak Signal-to-Noise Ratio (PSNR), and Receiver Operating Characteristics (ROC) curves.

i.Accuracy

A model's overall prediction accuracy is determined by dividing the number of correctly anticipated events by the total number of occurrences. When it comes to predictive modeling, accuracy refers to the proportion of precise predictions a model produces out of all of its predictions. The accuracy Equation (20) is :

$$\text{Accuracy} = \frac{TP+T}{\text{TP + FN + TN + FP}} \tag{20}$$

It is calculated as the ratio of all predictions made to all correctly predicted projections. The suggested approach is contrasted with multiple-branch cross-connected Convolutional Neural Networks (MBCC-CNN) to evaluate the model's accuracy. Table 1 and Figure 8 show the results of accuracy.

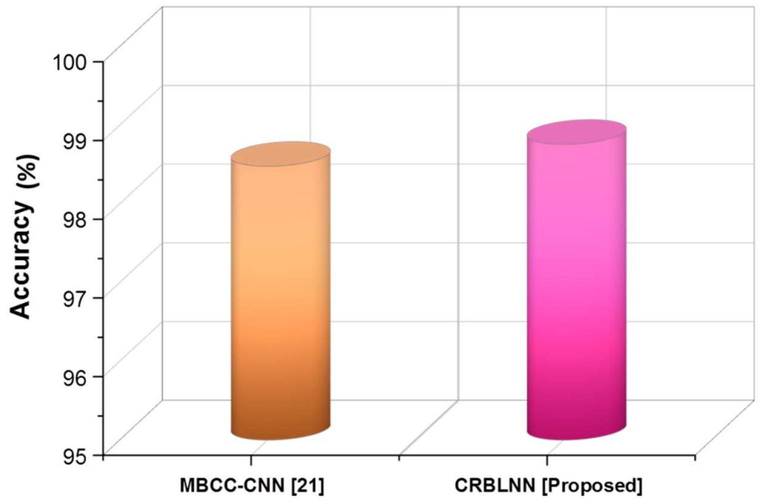

Figure 8: Comparison of Accuracy

Table 1: Numerical outcomes of Accuracy

Methods	**Accuracy (%)**
MBCC-CNN [21]	98.48
CRBLNN [Proposed]	98.76

It shows that the suggested method's accuracy of 98.76% outperforms that of the MBCC-CNN (98.48%). The proposed technique is evaluated by comparing its Precision, Recall, and F1-score to those of the CNN [22].

ii.Precision

A metric called precision is used to evaluate the effectiveness of a classification model, especially in binary classification problems. Precision can be computed mathematically as in Equation (21):

$$\text{Precision}=\frac{TP}{\text{TP + FP}} \tag{21}$$

It determines the proportion of correctly identified positive class instances that are actual positive forecasts among all the cases the model expects to be positive.

iii.Recall rate

It is also known as the "true positive rate," this metric evaluates a model's ability to reliably discern positive instances from all actual positive occurrences within the dataset. Recall is computed mathematically as in Equation (22):

$$\text{Recall} =\frac{TP}{\text{TP + FN}} \tag{22}$$

iv.F1- score

In binary classification tasks, the F1 score is a statistic that is frequently used to combine recall and precision in a single score. The model's performance is represented by a single number that is the harmonic mean of precision and recall. The following formula is used to get the F1 score in Equation (23):

$$\text{F1 score} = \frac{2TP}{\text{2TP + +FP+ FN}} \tag{23}$$

The F1 score is particularly useful when there is an imbalance in the number of positive and negative cases, or when both false positives and false negatives are significant issues. Table 2 and Figure 9 show the results.

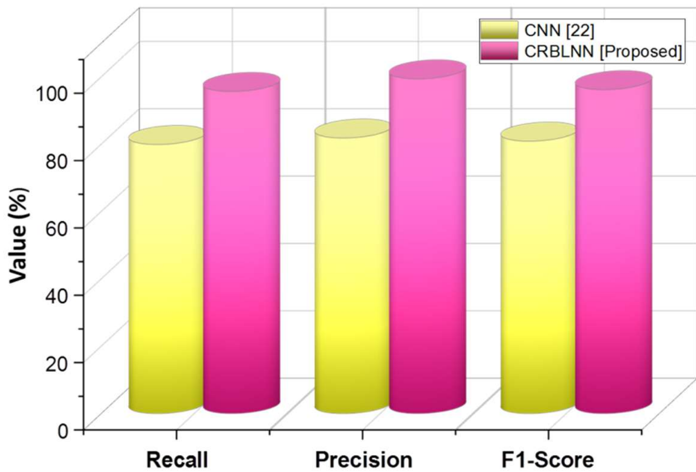

Figure 9: Comparison of Recall, Precision, F1-score

Table 2: Numerical outcomes of Recall, Precision, F1-score

Metrics	CNN[22]	CRBLNN [Proposed]
Recall (%)	79.93	95.68
Precision (%)	81.85	99.43
F1 Score (%)	80.92	96.21

It demonstrates that the suggested approach had a 95.68% recall rate as opposed to CNN's 79.93%. It also depicted the higher precision and F1-score of 99.43% and 96.21% compared to the 81.85 and 80.92% of CNN [22].

v.Confusion matrix

A different perspective on how well a model classifies data is provided by comparing the projected category with the actual category, which is the main purpose of the confusion matrix. A confusion matrix is a table used to evaluate the effectiveness of a classification model. It provides an in-depth analysis of the discrepancies between the model's predictions and the observed outcomes. Its capacity to help identify possible confusion points in the model across classes is the "confusion" matrix. The model's confusion matrix is shown in Figure 10.

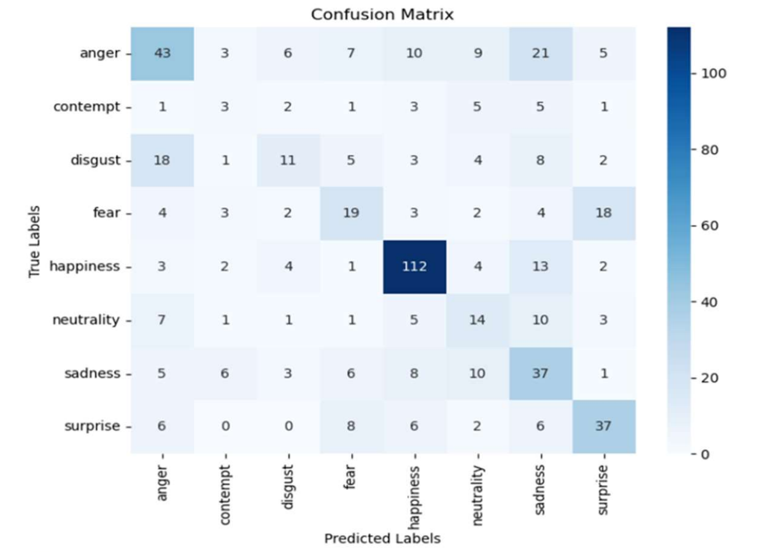

Figure 10: Confusion matrix

A deeper comprehension of the model's performance is offered by the confusion matrix, which highlights the model's accuracy in classifying instances of each class and points out any particular areas where it could be incorrect.

vi.AUC and ROC curves

The efficiency of a classification problem at various limit levels is evaluated using the area under the curve - receiver operating characteristics (AUC-ROC) curve. A probability curve is called a ROC, while an AUC is a measure of separation. It displays the level of class differentiation in the model. The following Equation (24) can be used to find the area encompassed within a curve. AUC and ROC curve comparisons are shown in Figure 11.

$$AUC = \int_1^0 ROC\, d(FPR) \qquad (24)$$

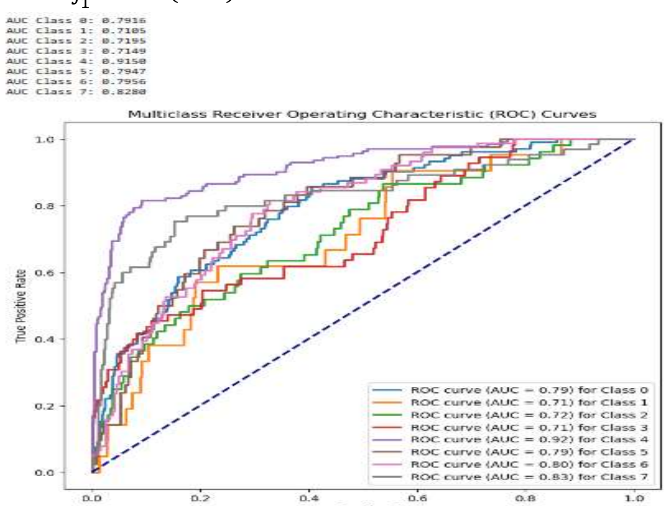

Figure 11: Comparisons of AUC and ROC curves

The model performed better, as demonstrated by the findings, which show greater true positive rates and lower false positive rates across various thresholds.

vii.Specificity and Sensitivity

The specificity parameter specifies how a diagnostic equipment or algorithm detects the absence of a certain disease, such as diseases or visual disorders. Specificity in human introspection of facial expressions entails an extensive investigation of the numerous complications expressed by different face characteristics. It goes beyond basic observations to investigate the minute specifics of emotions exhibited by particular muscle movements, micro-expressions, and the interaction of various facial features. By focusing on these precise signs, people get a better awareness of their own and others' emotional states, resulting in more accurate communication and empathy as shown in Equation (25).

$$Specificity = \frac{TN}{TN+FP} \quad (25)$$

The specificity of the suggested system is displayed in Table 3. The CRBLNN achieved (97.87%) comparatively, with CNN-KNN completed (93.93%). It proves that the CRBLNN's specificity has higher rates than another existing method.

Table 3: Numerical outcomes of specificity

Methods	**Specificity (%)**
CNN-KNN [23]	93.93
CRBLNN [Proposed]	97.87

Sensitivity and human reflection in facial expressions are interconnected components of EI that assist people to converse in difficult social situations. Facial expressions are a strong form of communication that can convey feelings that cannot be expressed directly through discourse. Individuals who are sensitive to these subtle signs are healthier and able to comprehend and sympathize with others, encouraging stronger binds and more effective communication as shown in Equation (26).

$$Sensitivity = \frac{TP}{FN+TP} \quad (26)$$

Table 4 demonstrates the recommended system's sensitivity. CRBLNN achieved (93.88%), whereas CNN-KNN completed (75.26%). It demonstrates that the CRBLNN's sensitivity is greater than other existing approach. Figure 12 depicts the outcomes.

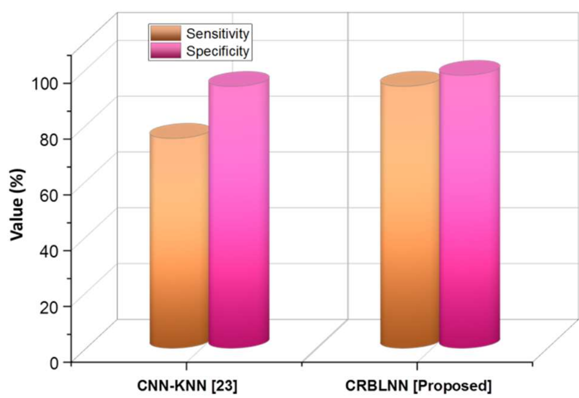

Figure 12: Comparison of sensitivity and specificity

Table 4: Numerical outcomes of sensitivity

Methods	**Sensitivity (%)**
CNN-KNN [23]	75.26
CRBLNN [Proposed]	93.88

viii.PSNR Value

The Peak Signal-to-Noise Ratio is known as PSNR. It's a statistic for estimating the quality of a picture or video signal. To assess the quality of the compressed picture or video in comparison to the original, PSNR is commonly employed in image processing and video compression. For assessing the PSNR value the proposed method is compared with Convolutional sparse auto encoder-Convolutional neural network (CSA-CNN) [24]. Figure 13 and Table 5 depict the outcomes of PSNR value.

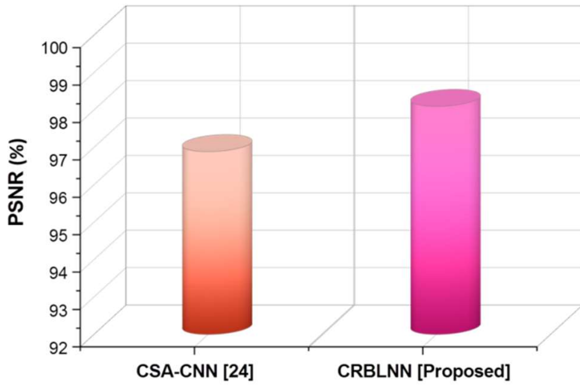

Figure 13: Outcomes of PSNR value

Table 5: Outcomes of PSNR value

Method	CSA-CNN[24]	CRBLNN [Proposed]
PSNR (%)	96.89	98.10

It depicts that the proposed method achieves a higher PSNR value of 98.10% compared to the CSA-CNN [24] of 96.89%.

5. Discussion

This work presented a CRBLNN approach for predicting Human Introspection in Facial Expressions using Emotional Intelligence and Machine Learning Techniques. The subjectivity and variability of human expressions, as well as the scarcity of diverse and labelled datasets for training, are some of the limitations of the various face expression prediction techniques. Some drawback of CNN [22] is its vast amount of labelled training data, which is particularly problematic for applications requiring complex facial expressions. Additionally, interpretability can pose a problem, and if the training data is not correctly regularized, it could result in over fitting. The MBCC-CNN [21] exhibits such types of complicated structures that can have higher computational complexity and need more thorough hyper parameter adjustments. The CNN-KNN [23] has limitations in capturing underlying patterns in high-dimensional spaces and can be operationally costly to use during the testing phase, particularly while working with large datasets. Training CSA-CNN [24] is challenging, and regularization parameter adjustments are likely to be performed carefully. It can be difficult to understand the learned representations in the sparse auto-encoder. This work presents the CRBLNN approach as a means of overcoming these shortcomings. Convolutional ReLU and Bi-LSTM networks are combined to create the CRBLNN, a potent tool for interpreting and forecasting human introspection in facial expressions. It consists of ReLU activation functions that allow for faster convergence during training and alleviates the vanishing gradient problem, while convolutional layers are skilled at extracting spatial information from facial images. Sequential data processing is a good fit for these functions.

6. Conclusion

The prediction of human introspection in facial expressions using machine learning and emotional intelligence techniques is a complex and difficult task that has significant implications for several fields, including psychology, neuroscience, and human-computer interaction. CRBLNN is presented in this paper for data classification and prediction. Deeper comprehension of changing emotional dynamics remains possible through the integration of geographical and temporal data enabled feasible by advancements in deep learning architectures, such as CRBLNNs. When compared to the current methodologies, the suggested approach produced superior results of Accuracy (98.76%), precision (99.43%), Recall (95.68%), F1-score (96.21%), Sensitivity (93.88%), Specificity (97.87%) and PSNR (98.10%).The integrated use of machine learning and emotional intelligence will surely open the door for revolutionary developments in a variety of domains, including human-computer interaction and mental health, as researchers continue to push the envelope of creativity and explore novel approaches.

References

1. Andalibi, N. and Buss, J., 2020, April. The human in emotion recognition on social media: Attitudes, outcomes, risks. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-16).https://doi.org/10.1145/3313831.3376680
2. Lima, M.R., Wairagkar, M., Natarajan, N., Vaitheswaran, S. and Vaidyanathan, R., 2021. Robotic telemedicine for mental health: a multimodal approach to improve human-robot engagement. Frontiers in Robotics and AI, p.37.https://doi.org/10.3389/frobt.2021.618866
3. Rocha, J.M., Rossi, G.N., de Lima Osório, F., Bouso, J.C., de Oliveira Silveira, G., Yonamine, M., Campos, A.C., Bertozi, G., Hallak, J.E.C. and Dos Santos, R.G., 2021. Effects of ayahuasca on the recognition of facial expressions of emotions in naive healthy volunteers: a pilot, proof-of-concept, randomized controlled trial. Journal of Clinical Psychopharmacology, 41(3), pp.267-274.10.1097/JCP.0000000000001396
4. Brand, N., Odom, W. and Barnett, S., 2021, June. A Design Inquiry into Introspective AI: Surfacing Opportunities, Issues, and Paradoxes. In Designing Interactive Systems Conference 2021 (pp. 1603-1618).https://doi.org/10.1145/3461778.3462000
5. Zhdanova, S.Y., Puzyreva, L.O., Mishlanova, S.L., Seredkina, E.V. and Zhdanov, M.A., 2022. Human-robot interaction: perception and reflection. In Science and Global Challenges of the 21st Century-Science and Technology: Proceedings of the International Perm Forum "Science and Global Challenges of the 21st Century" (pp. 791-801). Springer International Publishing.https://doi.org/10.1007/978-3-030-89477-1_73
6. Papoutsi, C., Drigas, A. and Skianis, C., 2021. Virtual and augmented reality for developing emotional intelligence skills. Int. J. Recent Contrib. Eng. Sci. IT (IJES), 9(3), pp.35-53.https://doi.org/10.3991/ijes.v9i3.23939
7. Segrin, C. and Flora, J., 2019. Fostering social and emotional intelligence: What are the best current strategies in parenting? Social and Personality Psychology Compass, 13(3), p.e12439.https://doi.org/10.1111/spc3.12439
8. Megahed, M. and Mohammed, A., 2020. Modeling adaptive E-learning environment using facial expressions and fuzzy logic. Expert Systems with Applications, 157, p.113460.https://doi.org/10.1016/j.eswa.2020.113460
9. Pozo-Rico, T. and Sandoval, I., 2020. Can academic achievement in primary school students be improved through teacher training on emotional intelligence as a key academic competency? Frontiers in Psychology, 10, p.2976.https://doi.org/10.3389/fpsyg.2019.02976
10. Alkozei, A., Smith, R., Demers, L.A., Weber, M., Berryhill, S.M. and Killgore, W.D., 2019. Increases in emotional intelligence after an online training program are associated with better decision-making on the Iowa gambling task. Psychological reports, 122(3), pp.853-879.https://doi.org/10.1177/0033294118771705
11. Deng, J., Pang, G., Zhang, Z., Pang, Z., Yang, H. and Yang, G., 2019. cGAN-based facial expression recognition for human-robot interaction. IEEE Access, 7, pp.9848-9859. https://doi.org/10.1109/ACCESS.2019.2891668
12. Erol, B.A., Majumdar, A., Benavidez, P., Rad, P., Choo, K.K.R. and Jamshidi, M., 2019. Toward artificial emotional intelligence for cooperative social human–machine interaction. IEEE Transactions on Computational Social Systems, 7(1), pp.234-246.https://doi.org/10.1109/TCSS.2019.2922593
13. Sun, Y., Ayaz, H. and Akansu, A.N., 2020. Multimodal affective state assessment using fNIRS+ EEG and spontaneous facial expression. Brain Sciences, 10(2), p.85.https://doi.org/10.3390/brainsci10020085
14. Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M. and Pollak, S.D., 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. Psychological science in the public interest, 20(1), pp.1-68.https://doi.org/10.1177/1529100619832930
15. Rajcic, N. and McCormack, J., 2020, April. Mirror ritual: An affective interface for emotional self-reflection. In Proceedings of the 2020 CHI conference on human factors in computing systems (pp. 1-13).https://doi.org/10.1145/3313831.3376625
16. Ghandeharioun, A., McDuff, D., Czerwinski, M. and Rowan, K., 2019, September. Towards understanding emotional intelligence for behavior change chatbots. In 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 8-14). IEEE.https://doi.org/10.1109/ACII.2019.8925433
17. Behera, A., Matthew, P., Keidel, A., Vangorp, P., Fang, H. and Canning, S., 2020. Associating facial expressions and upper-body gestures with learning tasks for enhancing intelligent tutoring systems. International Journal of Artificial Intelligence in Education, 30, pp.236-270.https://doi.org/10.1007/s40593-020-00195-2

18. deWijk, R.A., Ushiama, S., Ummels, M., Zimmerman, P., Kaneko, D. and Vingerhoeds, M.H., 2021. Reading food experiences from the face: effects of familiarity and branding of soy sauce on facial expressions and video-based RPPG heart rate. Foods, 10(6), p.1345.https://doi.org/10.3390/foods10061345
19. Darapaneni, N., Choubey, R., Salvi, P., Pathak, A., Suryavanshi, S. and Paduri, A.R., 2020, October. Facial expression recognition and recommendations using a deep neural network with transfer learning. In 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0668-0673). IEEE.https://doi.org/10.1109/UEMCON51285.2020.9298082
20. Bakola, L. and Drigas, A., 2020. The technological development process of Emotional Intelligence as a therapeutic recovery is implemented in children with ADHD and ASD comorbidity. https://www.learntechlib.org/p/217998/
21. Shi, C., Tan, C. and Wang, L., 2021. A facial expression recognition method based on a multi-branch cross-connection convolutional neural network. IEEE Access, 9, pp.39255-39274.
22. Wang, X., Ma, Y., Cammon, J., Fang, F., Gao, Y. and Zhang, Y., 2023. Self-Supervised EEG Emotion Recognition Models Based on CNN. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, pp.1952-1962.https://doi.org/10.1109/TNSRE.2023.3263570
23. AbWahab, M.N., Nazir, A., Ren, A.T.Z., Noor, M.H.M., Akbar, M.F. and Mohamed, A.S.A., 2021. Efficientnet-lite and hybrid CNN-KNN implementation for facial expression recognition on Raspberry Pi. IEEE Access, 9, pp.134065-134080.https://doi.org/10.1109/ACCESS.2021.3113337
24. Mohana, M. and Subashini, P., 2023, March. Convolutional Sparse Autoencoder for Emotion Recognition. In The International Conference on Artificial Intelligence and Computer Vision (pp. 3-15). Cham: Springer Nature Switzerland.https://doi.org/10.1007/978-3-031-27762-7_1