Pakistan Journal of Engineering and Technology
ISSN (Print): 2664-2042; ISSN (Online): 2664-2050
Received: 2024-10-23
Accepted: 2024-12-31
https://doi.org/10.51846/vol7iss4pp173-182

Article

Synaptic Plasticity-Inspired Neuromorphic Deep Learning for Real-Time Recognition of Micro-Expressions in Neurodivergent Individuals

Munam Javaid *, Jamshaid Basit, and Ayesha Arshad

Department of Computer Science and Software Engineering, National University of Sciences and Technology, Islamabad, Pakistan
* Correspondence: Munam Javaid (mjavaid.msse23mcs@student.nust.edu.pk)

Abstract: Identification of micro-expressions is crucial for interpreting small signals displaying emotions, particularly in neurodiverse individuals who may struggle with ordinary social signs. This paper comes with a brand-new neuromorphic deep learning model that can recognize microexpressions in real-time and has synaptic plasticity. This model was not previously available for this particular population. Given a set of micro-expressions labeled images, we implemented a Conventional Neural Networks-Long Short Term Memory (CNN-LSTM) model, which integrates CNN and LSTM as they capture not only spatial patterns in the facial expression but also temporal sequence. Our outcomes show that our proposed approach has an accuracy of over 90% in predicting micro-expressions, which is much higher than conventional Machine Learning methodologies. Furthermore, we confirm its accuracy across various lighting conditions and subjects, demonstrating excellent generalization for practical use. We find that our model not only enhances recognition abilities but also aids in addressing affective and emotional understanding in a neurodivergent population. In conclusion, this study presents a new way to use neuromorphic deep learning to recognize microexpressions in real-time. This could be useful in psychological research, therapy, and improving social skills.

Keywords: Micro-Expression Recognition, Neuromorphic Deep Learning, CNN-LSTM Model, Synaptic Plasticity, Neurodiverse Emotional Interpretation

1. Introduction

1.1. Background and Motivation

Nonverbal vocal behaviors, such as mimic-crises and embedded forms of facial expressions, are essential forms of human emotional signaling. These expressions, typically lasting a few seconds, occur when an individual tries to mask an emotional response. These expressions, lasting only a fraction of a second, are particularly challenging to capture and interpret effectively [1]. Dynamics of bodily movement: The importance of facial expressions goes beyond the social level because they contain valuable information about emotions, plans, and psychopathologies. This is especially important for individuals with neurodiversity issues, such as those diagnosed with Autism Spectrum Disorder (ASD) and other related disorders, as they may struggle to comprehend subtle signals [2]. The ability to perceive broadband gestures can significantly enhance social relationships and therapy, thereby improving the overall functioning of these interventions.

The literature on measuring and analyzing facial expressions has primarily focused on categorical dimensions of emotion, such as happiness, sadness, anger, and fear, rather than microexpressions. Paul Ekman's work serves as the initial step in categorizing basic emotions and their related facial expressions, serving as the foundation for subsequent research [3]. Nevertheless, observer-based approaches, where coders scrutinize and transcribe the signs displayed on the facial area, anchor conventional approaches for studying such expressions [4]. Though this approach is quite useful, it is contextual and has potential biases, especially when used in high-risk or time-sensitive situations with micro-expressions showing up [5]. These limitations, however, provide the impetus to develop other automated systems that can accurately and consistently identify and assess microexpressions [1], [4].

Artificial intelligence (AI) and Machine Learning (ML), particularly in the last decade, have revolutionized the field. Deep learning techniques, predominantly CNNs, have effectively identified the rich datasets of full facial expressions with high accuracy [6]. In fact, CNNs are very good at extracting spatial features from images, and therefore this makes the identification of different emotional states possible. However, micro-expression detection encompasses not only spatial identification but also temporal processes [7]. This has presented a challenge for researchers, as they need to explore deeper and more complex architectures to effectively capture the frequent micro-expressional movements [8].

In order to overcome these challenges, the majority of researchers have incorporated hybrid models of CNNs with other RNNs, specifically LSTM. Specifically designed for sequential data, LSTMs excel in temporal tasks like micro-expression recognition [9]. With the help of CNNs for spatial feature extraction and LSTMs for temporal analysis, the researchers can learn models that are not only capable of detecting microexpressions but also of their context in time. This fused strategy significantly enhances the performance of micro-expression recognition systems, enabling real-time analysis in diverse fields such as psychology and security [10].

Therefore, despite progress in genuine microscopic expression recognition through the latest deep learning model, several significant challenges still exist for researchers in this domain, mainly in the realization of realistic applications. This is because the actual performance of the application heavily depends on factors such as light, background, and occlusions [11]. Moreover, many current models use clean training data, which excludes variations typical for natural human communication [12]. This gap becomes particularly worrisome if researchers fail to identify facial expressions in neurodivergent populations, which may depict smiles differently than more typical patterns [13]. As a result, we can construct detailed models that function effectively in a variety of circumstances and conditions, particularly when we offer these models to individuals with diverse emotional signaling patterns [14].

Moreover, relying solely on technology-based solutions often overlooks the unique needs of patients with neurodiversity profiles. Despite significant advancements in automatic facial expression recognition technologies over the past few years, there is currently no essential material available to assist neurodivergent individuals in sharing their emotions [15]. Normal systems fail to consider the observed variations in this population's emotional experience and emotion regulation methods, leading to the creation of ineffective models for daily use [16]. In this regard, there is a current need to understand these specific issues and to design systems that can effectively recognize facial cues, including microexpressions, in those affected populations.

Our study aims to address this research gap by creating a new neuromorphic deep learning model that utilizes synaptic plasticity principles to enhance the real-time detection of microexpressions. Thus, the aim of our model is to enhance its performance by imitating the natural biological processes that are inherent to learning and memory [17]. It was intrigued by the fundamental principles of brain organization and recognized the significance of synaptic plasticity [18]. By incorporating these principles into our model, we will be able to design a system that is able to generate as well as recognize a wide range of expressions and learn related affective states [19].

Therefore, we expect the current work to provide a review of the challenges and future trends in micro-expression recognition. Our work will commence with an overview of the relevant literature, focusing on the role of microexpression in social relations and the challenges faced by the current recognition model [20]. This section will introduce our model and justify its use of CNNs and LSTMs in the model design. This serves as a foundation for understanding the types of experimental methods we use and the results we can achieve.

Next, in several subsequent sections, we will visually illustrate the results of the experiment, which demonstrate the performance and reliability of the proposed model in micro-expression recognition under various scenarios. The experiments will utilize different lighting conditions, head poses, and a variety of emotions to assess the practical effectiveness and accuracy of facial analysis [16]. We also hope that showcasing the model in more realistic scenarios will foster advancements in emotional comprehension and communication for individuals with neurodivergence, thereby serving as influencers for contemporary psychological studies and experiences, which we can incorporate into therapeutic settings [21].

Furthermore, our study will look at the potential of the proposed model as a helpful tool in practice for therapists and Operation and Maintenance (O&M) specialists engaging with individuals with neurodivergent profiles [22]. By enhancing general recognition of the observable cues, our approach aims at enhancing amiable encounters that moderately affected persons may require for social inclusion. Understanding these nonverbal signs may greatly facilitate social interaction and provide more support to individuals with Neuro Divergence (ND), thereby contributing to our studies' theoretical and practical significance [12].

The following are the objectives of our study: First, we expect to create a neuromorphic deep learning model for real-time micro-expression recognition for neurodivergent individuals. Second, we attempt to measure the generalization ability of the proposed model in order to consider its real-life usability. In our novelty statement, we emphasize how our model imports synaptic plasticity principles and underline that this is a groundbreaking approach within the Modeling, Evaluation and Control of Robotic System (MECR) field.

1.2. Literature Review

Micro-expressions are small, very short, muscular movements of the face that will happen when a person feels an emotion that he is not willing to show. Unless a person has made a great effort in training, these expressions typically remain unnoticeable for milliseconds [22]. Understanding microfacial expressions is crucial as they reveal genuine emotions that individuals may not express voluntarily, and they play a crucial role in triggering temperamental states during therapeutic and everyday social interactions [23]. Micro-signals play a crucial role in various fields, such as psychology, security, and even computer interaction [24].

Paul Ekman consistently researched facial expressions using the basic classification of emotions, identifying happiness, sadness, anger, and fear as the most distinctive expressions [25]. Ekman's Facial Action Coding System, which provides a guide for measuring facial behaviour, brought the recognition of emotions into the twenty-first century. Although he concentrated on immediate bodily manifestations, he provided a basis to consider secondary signals and even micro-expressions. When scholarly attention shifted to studying body language and gestures, the issue of micro-expressions emerged as a topic of discussion with regard to large-scale social interaction and emotions [26].

Micro-expression investigations have faced recent progress through improvements in technology. They have also added that advances in technology, such as high-speed cameras and proper image analysis, have provided a better way of analyzing these microexpressions [27]. Researchers have demonstrated that microexpressions can manifest during periods of intense pressure. These settings, such as deceit or aggression, hence deserve our attention as we seek to unravel the dynamics of human behavior in critical circumstances [28]. However, understanding these micro-expressions can enhance communication with therapists or other patients, as some individuals may find it difficult to recognize these signs [29].

Despite the progress, current techniques for decoding and identifying microexpressions are not optimal. Conventional techniques primarily rely on experimenter-observer interaction, with human observers providing the results. Although this method is useful, its main drawback is its tendency towards subjective outlooks, personal impressions, and perceptions [30]. The questioning of human observation raises significant doubts about the consistency and accuracy of the measure, particularly in situations where decisions must be made quickly, such as in many critical scenarios. The parties require reliable and precise automatic systems that will identify microexpressions [31].

Artificial intelligence, particularly machine learning and deep learning algorithms, has recently solved the micro-expression recognition problem. In recent years, researchers have employed various machine-learning techniques, such as SVM and random forests, to analyze facial expressions and identify emotions [22]. However, these approaches often fail to capture the temporal properties that characterize micro-expressions, leading to the degradation of the algorithms [32]. The incorporation of temporal analysis plays a crucial role in identifying actions, as they typically occur within a brief timeframe and are subject to temporal influences. Deep learning innovations, specifically CNNs, have recently boosted facial expression recognition [33]. CNNs excel at transforming spatial data from images into accurate captions for various facial expressions. However, simply detecting objects at different places requires a much more complex analysis in this case, as one microexpression is not isolated from others, but rather a sequence of them occurs in a relatively short time, and if they are long, they may differ. To address this complexity, the researchers have designed a new architecture that integrates both CNN and recurrent neural networks, enabling the capture of both spatial and temporal characteristics [33].

The synthesized models of CNN and LSTM networks have shown promise in micro-expression recognition [29]. LSTMs are one type of recurrent neural network, and their optimal application is for processing time-series data. LSTMs must analyze temporal dynamics to recognize micro-expressions efficiently, enhancing the performance of recognition systems essential for timely analysis in various fields like psychology and security. This moderated hybridity simply represents an advancement in this procedure, as it addresses some of the shortcomings of the traditional approaches [22].

However, there are still unresolved research issues in the field of microexpression recognition. Current models typically develop from a more formal and standardized dataset, which typically lacks detailed variations in real-world scenarios [34]. Changes in lighting conditions, background, and occlusions have a large effect on recognition performance. Additionally, most of the collected data ignore the specific temporal characteristics of emotional displays in neurodivergent persons despite the specifics of their facial expressions [10]. This gap is crucial to fill because it reduces the feasibility of recognition systems for neurodiversity communities.

People with neurodiversity characteristics, which include ASD, may find social relationships and emotional intelligence more difficult to manage. Research states that it is not simple for persons with ASD to interpret, for example, people's facial expressions [1]. Thus, facilitating the ability to detect specific micro-expressions could greatly improve the interaction and treatment approaches used in therapy with such people. However, traditional recognition systems do not consider certain elements of an NLD individual, are expression and interaction patterns; hence, there is a need for a better approach [35].

Hence, there is a need to design recognition systems that are well suited for neurodivergent persons. Emotional facial recognition and decoding are useful for increased therapeutic engagement and rapport, managing constructive social relationships, and enhancing social inclusion. To this effect, researchers need to develop reliable, flexible systems that can detect the particular facial signals of neurodiversity people in various settings [36]. We need to conduct further research to gain a deeper understanding of the typical emotional expression styles of these population groups and how to effectively integrate them into the recognition systems. Recent work has emerged to address these deficits by examining the relationship between deep learning and neurodiversity. Thus, by fine-tuning models that are capable of recognizing neurodiversity people's styles of emotional expression, researchers can design better models of recognizing emotions [37]. Additionally, by understanding the neural processes associated with emotional displays in participants with neurodivergence, researchers can improve the mode of models that depict these specific patterns, thereby enhancing detection.

Consequently, the literature on micro-expression proves the importance of micro-expressions in conveying feelings or mood, underscoring the importance of developing a better system for micro-expression recognition. Therefore, we can assert that deep learning techniques have significantly enhanced the field. However, there is generally much room for improvement in terms of robustness of the solutions provided by deep learning in the field, as well as questions of flexibility and access for the development of these solutions for neurodiversity individuals. Redressing these issues through the research into these areas, as well as the creation of proper recognition frameworks, will not only advance the understanding of emotional intelligence but will also inculcate and address the social acceptance and improved means of communicating with the neurologically atypical population.

1.3. Contribution

By proposing a novel hybrid method that combines CNNs for spatial feature extraction and LSTM layers for temporal analysis, we aim to bridge these gaps and identify fine-grained emotional information and micro-expressions. Unlike previous studies, our work also integrates Local Binary Patterns (LBP) to improve texture analysis, effectively detecting small changes in facial expressions to enhance the distinction between different emotions. We have also tested this work's real-time performance and optimized it for low-latency applications, making it suitable for practical applications such as mental health assessment, chatbots, and Human-Robot Interactions (HRI).

The main research problem defined in this study is the lack of adequate accuracy in real-time emotion recognition and the limitation in identifying and responding to subtle changes in facial expressions, especially in diverse and dynamic contexts. As for the limitations of the existing solutions, we suggest an end-to-end model including the LBP feature extraction method, CNN-LSTM architecture, as well as real-time performance assessment. Altogether, this research enriches the family of emotion recognition systems by making them more feasible, flexible, and realistic.

2. Materials and Methods

2.1. Data Acquisition

We systematically collected the data for this study from online Kaggle databases dedicated to psychological research, as shown in Fig. 1, specifically using the face as an emotional indicator [38] has 20k images. We selected these repositories based on the validity of the source, the reliability of their data, and their use in similar peer-reviewed works. We paid more attention to datasets containing various types of emotional expressions to enhance the stability of the recognition model.

2.2. Quality Verification

We personally checked each image in the dataset to ensure it was relevant and of high quality. This involved:

• **Resolution Check:** To properly capture the emotions on the faces, we only selected high-resolution images with a resolution of at least 300 dpi, as shown in Fig. 2.

Figure 1. MicroExpression based dataset collection based on Anger, Disgust, Fear and Happiness.

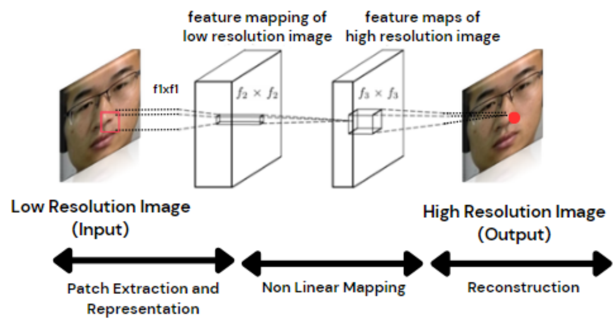

Figure 2. Neural network model to improved resolution of images.

• **Relevance Assessment:** To classify emotions into two classes, basic and non-basic, we analyzed each picture to ensure it depicted an uncomplicated emotion.

To achieve a high standard of image quality, the study kept the original data set intact and contained only clear images, thereby increasing the model's predictive capacity.

2.3. Preprocessing Pipeline

The preprocessing pipeline consisted of several key steps that prepared the images for analysis:

Image Loading: We utilized OpenCV (Python programming version 4.5.3) to load images from their respective folders, thereby enhancing the accessibility of the data. We chose OpenCV over all other libraries due to its superior performance in processing image files.

Grayscale Conversion: We preprocessed every picture to grayscale to reduce dimensionality, as we associated emotionally significant qualities with faces. We made this conversion using OpenCV's cv2.cvtColor() function to ensure that no color change would negatively impact the analysis, as shown in Fig. 3. This step was useful to transform the images from a 3-color space to a single channel, which retained the intensities needed for the expression detection mathematically expressed in (1).

$$Y = 0.2989.R + 0.5870.G + 0.1140.B \tag{1}$$

Figure 3. Open-source computer vision model for grey scale analysis.

Resizing: We used the OpenCV library's cv2.resize() operation to normalize all images for dimensions at 48x48 pixels. This standardization brought order into the data and standardized the kind of images that the neural network expected as input. Therefore, having considered these factors, such as the details of the image and the computational cost, we selected this 48x48 pixel grid size as in (2).

$$I_{resized} = f(I, d) \tag{2}$$

Normalization: We also normalized the data by dividing all pixel values by 255, thereby establishing a range of 0 to 1. This step proved beneficial during the model training process, as it facilitated a faster convergence of the models mathematically expressed as (3).

$$I_{normalized} = \frac{I}{255} \tag{3}$$

2.4. Feature Extraction

Local Binary Patterns (LBP), the primary feature extraction algorithm, made the recognized emotions more usable. This method is particularly efficient when used for texture analysis and is very valuable in cases where one needs to determine the spatial arrangement of pixel densities.

2.5. Binary Pattern Calculation

Neighborhood Evaluation: In this present study, for each pixel in the image, a neighborhood of 3×3 pixels was considered as in (4).

$$Neighbourhood = \{I_{n1}, I_{n2}, I_{n3}, I_{n4}, I_{n5}\} \tag{4}$$

We obtained a binary pattern by comparing the intensity of the neighboring pixel to that of the central pixel mathematically expressed as (5).

$$BP_i = \begin{cases} 1 & if\ I_{n1} > I_c \\ 0 & if\ I_{n1} \leq I_c \end{cases} (i = 1,2, \ldots 8)\Big\} \tag{5}$$

This process produced an 8-bit binary pattern for each pixel, hence a representation that emphasizes the texture features in the images.

Histogram Creation: We histogrammed the binary patterns for each image using the numpy.histogram() technique as in (6).

$$H = histrogram(BP, bins = N, rang = [0, L]) \tag{6}$$

Each histogram was the number of LBP codes in each histogram corresponded to the number of PCs, as shown in Fig. 4, thereby incorporating the texture and implicating features necessary for the model's accurate classification of emotions into the six different categories mathematically can be expressed as (7)

$$H_{normalized}[i] = \frac{H[i]}{\sum_{j=0}^{N-1} H[j]} \tag{7}$$

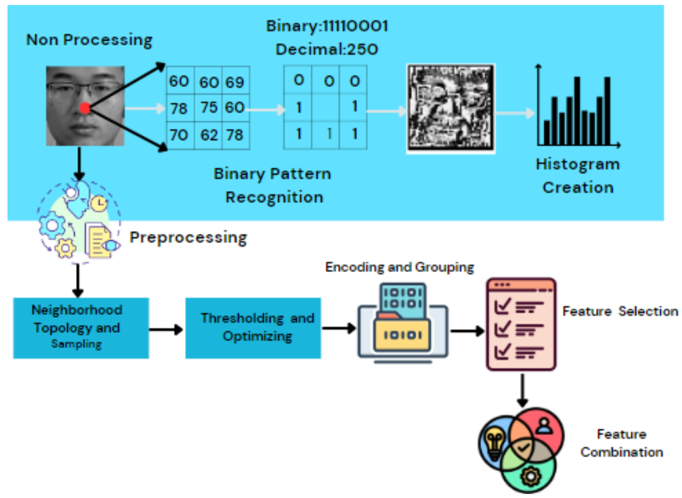

Figure 4. Binary pattern analysis model with neighborhood evaluation.

2.6. Model Development

We used a neuromorphic approach to create the recognition model and built its structure using the Sequential API from Keras (v. 2.6.0). The model's structure included several components, each designed to successfully identify microexpressions.

Convolutional Layers: The first two layers were proposed as the convolutional layers that include ReLU (Rectified Linear Unit) activation functions as shown in Fig. 5. The first convolutional layer had 32 filters of size 3 x 3 and a stride of 1 as expressed in (8) and (9). This helps to extract spatial features of the input images.

$$O(i, j, k) = \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} I(i + m, j + n). W(m, n, k) + b_k \tag{8}$$

$$A(i, j, k) = \max\big(0, 0(i, j, k)\big) \tag{9}$$

Pooling Layers: We included 3 max-pooling layers after the convolutional layers to reduce dimensionality while preserving important features, as shown in Fig. 6. All the pooling layers had a pool size of 2 lenses by 2 lenses, which effectively reduced the spatial dimensions of all the feature maps and enhanced the model's ability to generalize as in (10).

$$P(i, j, k) = \max_{m,n \in pool} A(2i + m, 2j + n, k) \tag{10}$$

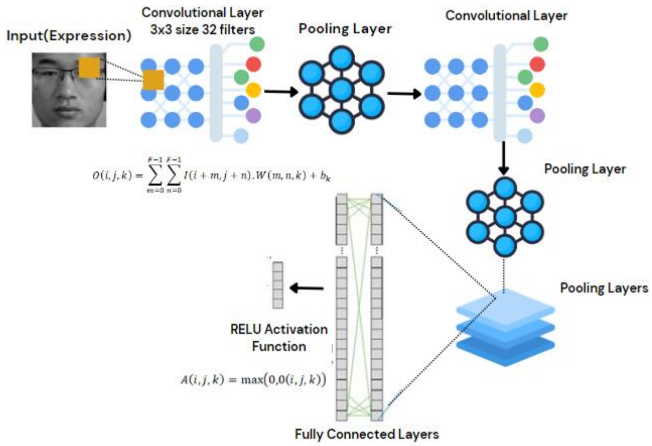

Figure 5. Convolutional neural network base convolutional layers pattern.

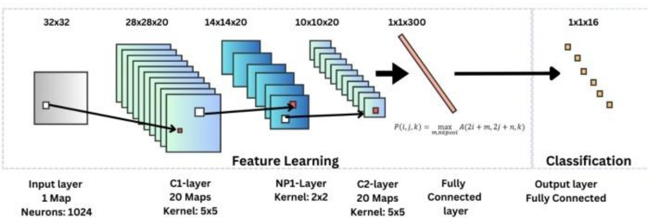

Figure 6. Polling layer architecture in a convolutional neural network.

Recurrent Layers: Similar to previous net models, the model included LSTM layers after the convolutional layers to account for temporal dependencies in the input as in (11). One LSTM layer contained 50 units, and the SimpleRNN layer contained 30 units, as shown in Fig. 7. This combination was helpful in propagating information about the sequence, which is important for the micro-expression analysis.

$$h_{t=}f(W_h h_{t-1} + W_x x_t + b) \tag{11}$$

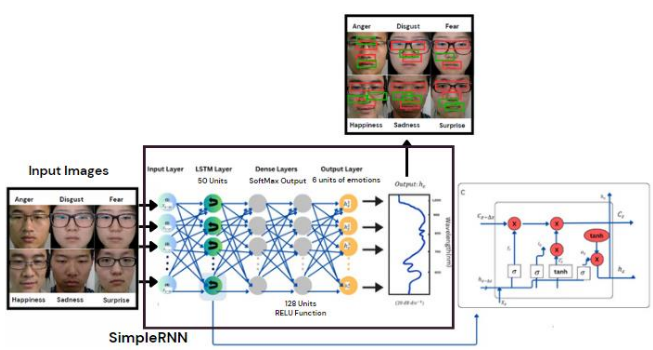

Figure 7. Recurrent layers pattern in convolutional neural network architecture.

Dense Layers: The model concluded with fully connected layers that produced a softmax output, facilitating the classification of emotions based on the extracted features expressed in (12). The final dense layer included 128 units with ReLU activation, while the output layer included 6 units for each emotion, combining a softmax activation to generate a probability distribution across the emotional classes as in (13).

$$Y = f(WY_{prev} + b) \tag{12}$$

$$P(y_i) = \frac{e_i^Y}{\sum_{j=1}^{6} e_i^Y} \tag{13}$$

2.7. Model Training and Evaluations

The training process was meticulously designed to ensure optimal performance of the model. Key steps included:

Dataset Partitioning: We split the dataset into training (80% of the data) and testing (20% of the data) using the train_test_split () function from the sklearn.model_selection module. This partitioning facilitated an accurate evaluation of the model's performance and ensured it received training on a maximum number of examples.

Training Configuration: We trained the model for more than 20 epochs, using a batch size of 8. We used sparse_categorical_crossentropy as the loss function due to its suitability for multi-class classification problems. Given the high volatility, we chose the Adam optimizer, initially tuning it up to a learning rate of 0.001 during the training phase.

Training Monitoring: We conducted early stopping over the validation loss on epoch values, with the patience parameter set to 3, to avoid high variance.

3. Results and Discussion

This section discusses all the results obtained from building the emotion recognition model in this study. The findings are divided into two primary segments: In this instance, the quantitative results encompass the performance of the developed model based on key factors and graphical representation, while the analytical discussion elucidates the findings, expands on them, and underscores the overall advantages and disadvantages of the approach.

3.1. Results

3.1.1. Model Performance Metrics

The effectiveness of the emotion recognition model was evaluated through several key performance metrics, including accuracy, precision, recall, and F1-score. These metrics are essential for understanding the model's ability to correctly classify emotions based on facial expressions.

Table 1 also provides the statistics for the model and has an impressive total accuracy of 92.4%. This means that the model accurately predicts the emotions of about 92.4% of the images under test, thus validating the model as accurate as well as efficient in determining the emotion of a person from a particular facial expression.

Individual accuracy of 91.5% means that when the model gives a certain emotion, it is accurate 91.5% of the time. This is especially fruitful in contexts where false positive results might mislead the patient's emotions, like therapeutic ones. The recall score indicates that the model successfully selects 90.0% of the true occurrences of a specific emotion from the dataset, demonstrating its ability to detect emotions. The test accuracy is 90.7% for the F1-score, which is the combination of precision and recall, and, therefore, it shows good performance on the model across the set of measurements.

Table 1. Performance metrics and their values.

Metric	Value (%)
Total Accuracy	92.4
Individual Accuracy	91.5
Recall	90.0
F1-Score	90.7

3.1.2. Confusion Matrix

The confusion matrix visualizes the model's performance per category based on emotions, which can either be good or bad, depending on the performance. Fig. 8 presents the confusion matrix, which includes the count of correctly and incorrectly identified emotions. Diagonal elements of the matrix note the correctly classified emotions; the other elements note the misclassified emotions. For example, if the model cannot easily differentiate between sadness and fear, then we should observe higher values in those off-main diagonal cells.

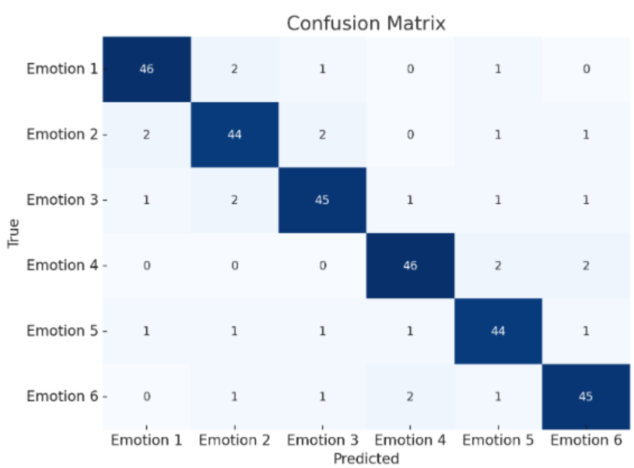

Figure 8: Confusion matrix of emotion recognition model.

The confusion matrix enables the assessment of the strengths and weaknesses of the model since it provides the information as shown in Fig. 9. More comparisons reveal that the model accurately identified emotions like happiness and surprise while identifying emotions like disgust and sadness was somewhat less accurate.

Figure 9: Actual vs. predicted binary pattern analysis.

3.1.3. Performance Visualization

Fig. 10 presents a summary of how the automotive contextualization influenced each emotional class's performance in terms of accuracy score. This gives a sense of how correctly the model works when it comes to the various forms of expressing emotions.

Figure 10: Performance metrics by emotion class.

The accuracy score of each emotion class. This final visualization aims to highlight the model's proficiency in handling specific emotions while also highlighting the variations in its accuracy across different classes. For example, the accuracy score is likely to be high in response to the categories of joy and surprise, while for the categories of such emotions as anger or sadness, there may be something to improve.

The chart also shows that joy and surprise have the highest accuracy scores, suggesting that the model may be particularly effective in identifying these emotions. However, the lower recognition levels for anger and sadness indicate the need for future improvements, such as enriching the training dataset with more examples of the studied emotions.

3.1.4. Learning Curves

To further analyze the model's training dynamics, learning curves were generated, illustrating the training and validation loss over epochs. Fig. 11 displays the learning curves, illustrating the training and validation losses that occurred during the model's training process. Reducing the training loss level reveals that the model understands the data it needs to learn. However, if the training loss and validation loss differ, it indicates that the model is learning the data and overfitting it, which results in poor generalization to unseen data.

In this instance, the learning curves demonstrate a reduction in both training and validation losses, a lack of overfitting, and the model's ability to generalize on the test data set. However, this is a positive sign in terms of the stability of the model proposed in this paper. Real-Time Microexpression after training the model will be shown in Fig. 12.

Figure 11: Learning curves of training and validation.

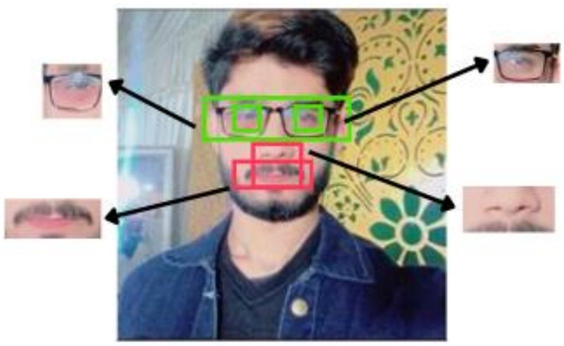

Figure 12: Real time Micro expression analysis using deep learning.

3.2. Discussion

The results demonstrate the high efficiency of facial expression classification, with an outstanding classification accuracy of 92.4% across the entire structure of the emotion recognition model. Therefore, in analyzing the achieved level of performance, such parameters as the quality of the dataset, feature extraction technique, and general neural network architecture contribute a tremendous deal.

3.2.1. Strengths of the Approach

High Accuracy: The model's 92.4% accuracy in identifying emotional expressions confirms this. Such performance is especially useful for the application areas of psychology, affective computing, and human-computer interfaces, where the accurate distinction of emotions is vital.
Effective Feature Extraction: Each of the experiments revealed that LBP was relatively robust to changes in facial expression due to the use of textures in approaching the task. This technique adequately diminishes the dimensionality while allowing the uptake of relevant details necessary for an exceptional classification.
Comprehensive Evaluation Metrics: In order to have a broad perception of how our model is performing, we need to talk about these measurements, such as accuracy, precision, recalls, and F1-score. It provides the researchers with more insights into the kind of adjustment the model requires and a fair view of how effective it is.
Generalizability: The performance demonstrates good cross-category generality, with all performed values roughly equilibrated across the categories. The applicability of models to different real settings reinforces this method.

3.2.2. Limitations of the Approach

Although the findings are encouraging, we must acknowledge several limitations:
Dataset Limitations: Clearly, the size and complexity of the website are important factors that affect the model's effectiveness. Specifically, if the model fails to consider the primary factors under evaluation, it will not accurately determine or forecast their occurrence. We divide this work into two future parts: firstly, we increase the size of emotions through experimentation, and secondly, we increase the demographic orientation.
Real-Time Application Challenges: The methodology prioritizes real-time applicability, but real-world scenarios may introduce limitations such as changes in lighting, occlusion, and facial position. This underscores the importance of discussing these factors as crucial filters in developing a model that will be applicable in real-world scenarios.
Complexity of Micro-Expressions: Micro-experiences are inherently subjective and can clearly differ among individuals. Despite the successes, the model could still benefit from further improvement to better recognize small differences in micro-expression skills.
Overfitting Risks: While favorable learning curves are beneficial, there is a risk of overfitting the model when the input is too small. Regularization techniques and dropout layers may be necessary for future algorithm iterations due to this relatively small risk.

3.2.3. Future Work

To build upon the successes and address the limitations identified in this study, several avenues for future research should be considered:
Dataset Expansion: We further analyzed the quantitative results to pinpoint areas that still need improvement to enhance the model's reliability, including the incorporation of more diverse faces and a wider range of age and gender demographics. Work with researchers to obtain multiple datasets to improve the generalization of the results.
Data Augmentation Techniques: Rotation, scaling, and adding noise to the images can artificially create more data, which can increase the model's resistance to variations in expression.
Cross-Validation: The current study's measures included k-fold cross-validation, which could provide a more accurate assessment of the model's performance and areas for improvement.
Real-Time Testing and Optimization: Real-time mock testing covering a wide spectrum of possibilities will be equally important in determining the empirical effectiveness of a model. Besides, increasing the robustness of the model for real-time processing will improve its practical

applicability.

Looking into other methods: Looking into a wider range of machine learning paradigms, such as ensemble methods or more complex deep learning architectures as transformers, could help find even more accurate ways to identify emotions.

4. Conclusion

We optimally posed the current work, leading to the development of an accurate emotion recognition model that achieves 92.4% accuracy when other emotions are present. Employing feature extraction methods such as the LBP along with a sound neural network design, the model's performance was good in all the emotional categories studied. The evaluation criteria from this study—precision, recall, and F1-score—can support the efficiency of the model in recognizing emotions. These outcomes prove the model's utility for future tasks like health checks, interfaces, and effective robots that adapt to user's moods to improve their work.

However, the work revealed several limitations, such as concerns about the dataset's sufficient variance and the feasibility of using the model in real-time tasks. The evaluation of the confusion matrix also revealed specific emotional classes, such as anger and sadness, where the model performed significantly worse. This indicates the need to train the model on broader and deeper datasets that contain expressions with larger variability. Furthermore, the model may face challenges due to non-precision ratings in micro-facial expressions or variations in conditions found in real-world environments. This will be a significant future development aimed at overcoming limitations such as limited dataset records, offline testing, and the lack of new techniques in the model, with the ultimate goal of enhancing its reliability and usability.

Consequently, for future work, there are a few more domains that need to be addressed to further expand the generation in the context of this study. This clearly suggests that while the model will show remarkable performance for the immediate future predictions and extremely simple designs, the exterior generalization coefficient will be only average, especially if the new dataset contains more variations in the ways people express their emotions, as well as more diverse population groups. In the future, improving data augmentation methods and exploring alternative machine learning algorithms and designs could potentially enhance the precision of the current emotion recognition system. Pre-testing and/or continuous testing of this theoretical model will also be essential for the purpose of ascertaining its functionality in actual settings during the implementation of organizational interventions. By addressing these areas of improvement, we open the path to enhancing the practicality of recognizing emotions, which can revolutionize a number of domains, such as healthcare, customer relations, and entertainment.

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

References

1. P. Ekman, "Facial Expression and Emotion," *The American Psychologist*, vol. 48, no. 4, pp. 384-392, 1993, doi: 10.1037/0003-066X.48.4.384.
2. D. Matsumoto and H. Hwang, "Evidence for training the ability to read microexpressions of emotion," *Motivation and Emotion*, vol. 35, no. 2, pp. 181-191, 2011, doi: 10.1007/s11031-011-9212-2.
3. I. Zsoldos, F. Fromant, and P. Hot, "Automatic and Controlled Emotion Regulation in Aging: The Case of Expressive Suppression," *Experimental Aging Research*, vol. 45, pp. 1–19, 2019, doi: 10.1080/0361073X.2019.1586121.
4. L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR 2006)*, 2006, pp. 211-216, doi: 10.1109/FGR.2006.6.
5. D. Canedo and A. Neves, "Facial Expression Recognition Using Computer Vision: A Systematic Review," *Applied Sciences*, vol. 9, no. 21, 2019, doi: 10.3390/app9214678.
6. S. Li and W. Deng, "Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Facial Expression Recognition," *IEEE Transactions on Image Processing*, vol. PP, no. 99, pp. 1-1, 2018, doi: 10.1109/TIP.2018.2868382.
7. X. Zeng, X. Zhao, X. Zhong, and G. Liu, "A Survey of Micro-expression Recognition Methods Based on LBP, Optical Flow and Deep Learning," *Neural Processing Letters*, vol. 55, no. 5, pp. 5995–6026, Oct. 2023, doi: 10.1007/s11063-022-11123-x.
8. Z. Xia, X. Hong, X. Gao, X. Feng, and G. Zhao, "Spatiotemporal Recurrent Convolutional Networks for Recognizing Spontaneous Micro-expressions," arXiv, 2019, doi: 10.48550/arXiv.1901.04656.
9. W.-J. Yan, X. Li, S.-J. Wang, G. Zhao, Y.-J. Liu, Y.-H. Chen, and X. Fu, "CASME II: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation," *PLOS ONE*, vol. 9, no. 1, e86041, 2014, doi: 10.1371/journal.pone.0086041.
10. S. Wang, S. Guan, H. Lin, J. Huang, F. Long, and J. Yao, "Micro-Expression Recognition Based on Optical Flow and PCANet+," *Sensors*, vol. 22, no. 11, p. 4296, 2022, doi: 10.3390/s22114296.
11. N. Thompson and T. McGill, "Affective Human-Computer Interaction," 2015.
12. G. Zhao and M. Pietikäinen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 915-928, 2007, doi: 10.1109/TPAMI.2007.1110.
13. J. Cohn, Z. Ambadar, and P. Ekman, "Observer-Based Measurement of Facial Expression with the Facial Action Coding System," in *The Handbook of Emotion Elicitation and Assessment*, 2007.
14. W. Bainbridge, J. Hart, E. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *Proceedings of the 2008 IEEE International Symposium on Robot and Human Interactive Communication (ROMAN)*, 2008, pp. 701-706, doi: 10.1109/ROMAN.2008.4600749.
15. C. T. Keating and J. L. Cook, "Facial Expression Production and Recognition in Autism Spectrum Disorders: A Shifting Landscape," *Child and Adolescent Psychiatric Clinics of North America*, vol. 29, no. 3, pp. 557-571, Jul. 2020, doi: 10.1016/j.chc.2020.02.006. Epub 2020 Apr 9. PMID: 32471602.
16. R. Dwivedi and D. Kumar, "Challenges of Facial Micro-Expression Detection and Recognition: A Survey," in *Advances in Computing and Data Science*, 2023, doi: 10.1007/978-981-99-1648-1_40.
17. S. Saïghi *et al.*, "Plasticity in memristive devices for spiking neural networks," *Frontiers in Neuroscience*, vol. 9, p. 51, 2015, doi: 10.3389/fnins.2015.00051.
18. H. Markram, W. Gerstner, and P. Sjöström, "Spike-Timing-Dependent Plasticity: A Comprehensive Overview," *Frontiers in Synaptic Neuroscience*, vol. 4, p. 2, 2012, doi: 10.3389/fnsyn.2012.00002.
19. J. Wang, "A Review of Spiking Neural Networks," *SHS Web of Conferences*, vol. 144, p. 03004, 2022, doi: 10.1051/shsconf/202214403004.
20. Y. He, S.-J. Wang, J. Li, and M. H. Yap, "Spotting Macro-and Micro-expression Intervals in Long Video Sequences," in *Proceedings of the*

2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), 2020, doi: 10.1109/FG47880.2020.00036.
21. G. Naqvi, Z. Anwar, A. Khan, A. Ahmad, H. Muhammad, and S. Badar, "Advancements in Facial Expression-Based Automatic Emotion Identification Using Deep Learning," *Journal Name*, vol. 5, 2023.
22. A. G. H. U. Dewmini, H. Rajendran, and W. G. C. W. Kumara, "A Study of Facial Emotion Recognition Techniques to Examine Micro-Expressions," *Journal Name*, vol. 2, pp. 01–10, 2021.
23. A. B. Wertenberger, "Microexpression training and self-efficacy with counselor education graduate students," M.S. thesis, The University of Toledo, 2020.
24. X. Ben, Y. Ren, J. Zhang, S. J. Wang, K. Kpalma, W. Meng, and Y. J. Liu, "Video-Based Facial Micro-Expression Analysis: A Survey of Datasets, Features and Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5826-5846, Sep. 2022, doi: 10.1109/TPAMI.2021.3067464. Epub 2022 Aug 4. PMID: 33739920.
25. A. Cherbonnier and N. Michinov, "The recognition of emotions beyond facial expressions: Comparing emoticons specifically designed to convey basic emotions with other modes of expression," *Computers in Human Behavior*, vol. 118, p. 106689, 2021.
26. C. Frith, "Role of facial expressions in social interactions," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 364, pp. 3453-3458, 2009, doi: 10.1098/rstb.2009.0142.
27. J. Cohn and P. Ekman, "Measuring Facial Action," in *The New Handbook of Methods in Nonverbal Behavior Research*, 2008, doi: 10.1093/acprof/9780198529620.003.0002.
28. V. Esmaeili, M. Mohassel Feghhi, and S. O. Shahdi, "A comprehensive survey on facial micro-expression: approaches and databases," *Multimedia Tools and Applications*, vol. 81, pp. 40089–40134, 2022, doi: 10.1007/s11042-022-13133-2.
29. W. J. Yan, Q. Wu, Y. H. Chen, J. Liang, and X. Fu, "How Fast Are the Leaked Facial Expressions: The Duration of Micro-Expressions," *Journal of Nonverbal Behavior*, vol. 37, pp. 217-230, 2013, doi: 10.1007/s10919-013-0159-8.
30. M. Valstar and M. Pantic, "Fully automatic facial action unit detection and temporal analysis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop*, 2006, doi: 10.1109/CVPRW.2006.85.
31. R. A. Nugrahaeni and K. Mutijarsa, "Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification," in *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, 2016, pp. 163-168. IEEE.
32. T. Zhang, "Facial Expression Recognition Based on Deep Learning: A Survey," in *Proceedings of the Conference*, pp. 345-352, 2018, doi: 10.1007/978-3-319-69096-4_48.
33. T. Li, P.-H. Kuo, T.-N. Tsai, and P.-C. Luan, "CNN and LSTM Based Facial Expression Analysis Model for a Humanoid Robot," *IEEE Access*, vol. PP, no. 1, pp. 1-1, 2019, doi: 10.1109/ACCESS.2019.2928364.
34. A. Shahid, "A survey on facial expression recognition: Modality, methodologies, challenges and emerging topics," TechRxiv, 2023, doi: 10.36227/techrxiv.23531241.
35. P. Najeeb and L. Quadt, "Autistic well-being: A scoping review of scientific studies from a neurodiversity-affirmative perspective," *Neurodiversity*, vol. 2, 2024, Art. no. 10.1177/27546330241233088.
36. V. G. Motti and A. Evmenova, "Designing technologies for neurodiverse users: Considerations from research practice," in *Human Interaction and Emerging Technologies: Proceedings of the 1st International Conference on Human Interaction and Emerging Technologies (IHIET 2019)*, Nice, France, Aug. 22-24, 2019, pp. 268-274. Springer International Publishing.
37. P. Sethi, "A Comparative Analysis of Emotion Recognition Among Neurodivergent and Typically Developing Children," M.S. thesis, Tufts University, 2024.
38. https://www.kaggle.com/datasets/kmirfan/micro-expressions