ARTICLE

An evolutionary and insightful graph neural network based hybrid model for deciphering tenacious stress detection of humans using facial emotion recognition: coinage to affective computing

Megha Bansal* and Vaibhav Vyas

Department of Computer Science, Banasthali Vidyapith, Rajasthan, India

ABSTRACT

In the realm of affective computing, facial emotion recognition plays a pivotal role in ensuring prooper stres detection in humans. Within the domain of facial emotion recognition, accurate detection of emotions from facial expressions holds immense significance. Existing methods often neglect intricate connections between various facial features, potentially leading to challenges in discerning naunced emotional states. In response to this, we introdue approach termed Graph Neural Network-based Facial Emotion Recognition. Our technique harnesses the capabilities of graph neural network (GNNs) to surmount these limitations. The process inititaes with the creation of a graph that encapsulates the similarity between individual facial expression samples. Subsequently, this graph is fed into the model for intricate feature mapping. The outputs generated by this integrated model ingeniously merge the feature insights from adjacent facial expression samples, thereby enhancing the representation of data for downstream emotion recognition tasks. The graph neural network model processed facial expression samples are then directed to an emotion adjacency matrix that discerns various emotional states. The final outcomes are the culmination of results from this classification process. By accurately defining the top samples exhibiting the most pronounced emotional cues, our method efficiency determines the prevailing emotional state. Through comprehensive experimentation involving generation of 99.89% accuracy on publicily available facial expression dataset, our results highlight this algorithms superiority. This approach showcases a substantial increase in recognition accuracy, outperforming the closest competing algorithms by a significant margin. This demonstrates the efficacy and promise of this algorithm in revolutionizing stress detection, potentially ushering in a new era of precise and nuanced stress analysis form facial expressions.

*Corresponding author
Email Address: meghabansal05@gmail.com
Date received: September 4, 2023
Date revised: December 30, 2023
Date accepted: January 15, 2024
DOI: https://doi.org/10.54645/202417SupZAM-51

INTRODUCTION

According to global statistics almost 1.3 billion people around the world faces poor mental health condition and amongst them

KEYWORDS

stress, mental health, machine learning, affective computing

almost 20% ends in suicidal accidents. Studies indicate that uncontrolled emotions, including anger leading to rage and sadness/stress affecting overall mental health, significantly increasing risk for unstable health conditions (Bradvik 2018). It has been found that 94% of suicides and health problems could be prevented through early detection of emotional state of a person. Hence, identifying the emotional state plays a vital role in reducing and preventing anxiety, stress, depression, severe health conditions and suicidal rates, facilitated by assisted stress detection systems. Existing methods for detecting stress can be categorized into three main approaches: Physiological methods, Behavioral methods, and Self-Report methods. Physiological signal-based methods utilize measures like Heart Rate Variability (HRV), or electroencephalography (EEG), or Cortisol Level Analysis to assess stress levels. Behavioral methods focus on analyzing facial expressions, speech analysis or body movements to detect signs of stress. Self-Report method analysis method assess stress by analyzing factors such as questionaries and surveys, diaries or journal entries (Zainudin et al. 2021). However, there are few shortcomings associated with these methods like physiological signal-based methods requires a) Specialized equipment and sensors to measure HRV, EEG and EDA signals, and this complexity and dependency on equipment can limit their practicality and accessibility (Maaoui and Pruski 2018) ; b) These responses can vary significantly between individuals i.e., what may be considered a stress response for one person may not be the same for another, hence this inter- individual variability is challenging in decision making; c) Physiological signals can be influenced by a range of emotions like excitement, fear, or even physical exertion and elicit similar physiological response, making it difficult to differentiate between stress specific patterns from other emotional states (Kushagra Nigam et al. 2021); d) These signals may vary depending upon context and environmental factors such as temperature, noise levels, or presence of others can influence physiological response (Katmah et al. 2021); e) Methods such as cortisol level analysis, may involve invasive or costly procedures (Shahbazi and Byun 2023). Similarly, in Self-Report Method, few major shortcomings are: a) These methods rely on individual's own perceptions and interpretations of their stress levels (Maltman et al. 2023); b) Mainly individual's subjective experiences are captured and this may not provide a complete understanding of the physiological responses associated with stress (Witte et al. 2021); c) Stress perception may vary depending on situational context, mood, or personal factors, leading to inconsistent reporting; d) the accuracy and reliability can be compromised when individuals are asked to recall stress experiences from the past, particularly for long-term or chronic stress assessments; e) these methods often rely on periodic assessments or surveys, which may not capture stress fluctuations in real-time (Theon et al. 2023), f) the diversity of available questionnaires and scale make it challenging to establish a universally applicable measure for stress detection. Behavioral methods also possess their own critics when aided via human: a) these observations can be subjective and prone to interpretation biases (Hajera and Ali 2018), b) behavioral responses to stress can vary significantly between individuals, c) behavioral cues associated with stress can be influenced by the context and environment in which they occur; d) behavioral cues can be present in any psychological changes as well; e) this analysis includes many technical complexities in implementation (Nijhawan et al. 2022). Behavioral methods have their shortcomings majorly when they are curated by human interventions. Instead of that when machine learning is used for stress detection the results are entirely different. Multiple studies (Li and Liu 2020; Mittal et al. 2022; Zhou et al. 2023; Liu et al. 2023) have shown that with intervention of machine learning, stress detection took a new reform by getting a higher prediction accuracy in less and real-life scenario transforming lives of many.

PROPOSED MODEL

Emotion and stress detection have gained significant attention in the field of human-machine systems for intelligent psychological tool. Traditional manual feature extraction methods for stress and emotion recognition have been surpassed by deep-learning based feature extraction methods, which provide superior performances. However, the complex network models pose computational challenges. Existing research has focused on individual detection of either stress or emotional state, but both factors are contributing to a person's health and mutually influence each other (Khosrowabadi 2018). This paper proposes a non-invasive and efficient detection model based on graph neural network (GNN) to identify emotional state of a person by which stress state can be identified, enabling early warning of potential health risks. Unlike traditional neural networks that operate on grid-like data structures such as images or sequential data like text, GNNs can model and learn from complex relational information. GNN is hardly explored fro emotion detection and in our paper, we are proposing the usage of GNN for emotion detection with FER13 dataset.

Graph Neural Networks (GNN) are a class of deep learning models designed to operate on graph-structured data. This algorithm is not exposed much by researchers till now but GNN can effectively capture the relationship and dependencies between elements in a graph, making them suitable for tasks involving structured data like facial emotion recognition (FER) datasets. Unlike traditional convolutional neural network (CNNs) that process regular grid-like data such as images, GNNs can handle irregular and non-Euclidean data represented as graphs. Graphs consists of nodes and edges, where nodes represent entities, and edges represent relationships or connections between nodes. In FER, facial images can be represented as graphs, where individual facial landmarks or pixels are nodes, and the edges between them represent the spatial relationships or connections. GNNs is applied by us to process these graphs and we have also extracted meaningful features that encode the facial expression information. However, early image emotion recognition signals often appear weak and are challenging to distinguish from normal signals. To address this issue, we propose a novel approach called Graph Neural Network based Facial Emotion Recognition. This method consists of various key components which are illustrated in Figure 1; which consists of detection process of this algorithm, showcasing how the different components work together to achieve effective image emotion recognition.

This paper's innovative contributions are three-folds: Firstly, a system is developed for preprocessing the FER13 dataset to create a meshgrid. Further normalization is done feature localization, facial alignment and feature extraction is performed to identify and locate specific points or landmarks on a person's face. These landmarks correspond to distinct facial features, such as eyes, nose, mouth, and eyebrows. Secondly, an adjacency matrix is prepared from the gathered images so that the prediction of edges between any pair of pixels can be done, this 3-layer GNN model with ReLU activation function performs exceptional in feature extraction and provides an accuracy of 99.89% on the FER13 dataset which is exemplary.

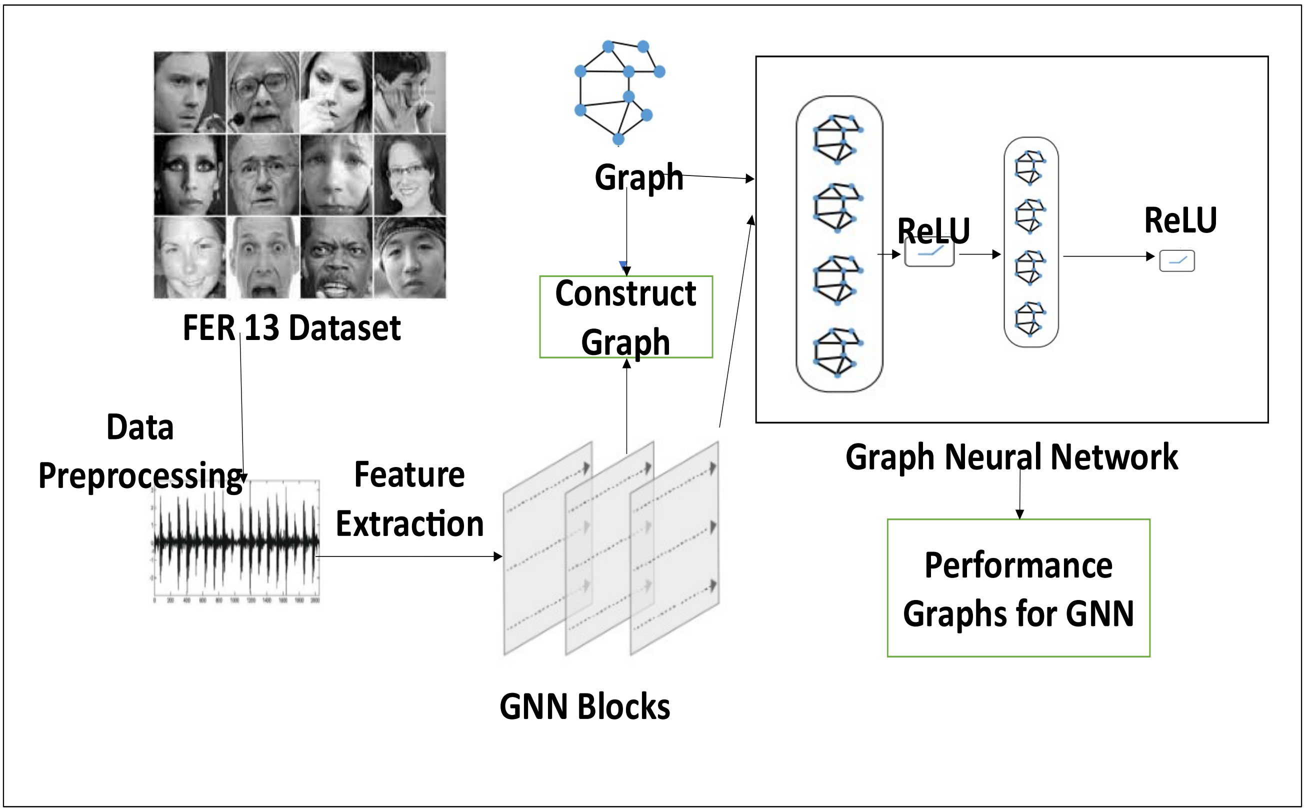

Figure 1: Emotion detection by GNN

MATERIALS AND METHODS

For the facial emotion recognition experiments, we conducted our research on a hardware environment consisting of an Intel® core ™ i5-10500 CPU with a clock speed of 4.50 GHz and 8GBof RAM. The software side of the experiments involved using Jupyter Notebook for implementing the GNN model. The operating system used for the experiment was Windows 10 Professional. GNN-FER consists of six phases which are elaborated below:

- Dataset Preprocessing
- Image-to-Graph Conversion
- Node Features
- Adjacency Matrix
- Graph Visualization
- Node Coloring

Dataset Preprocessing

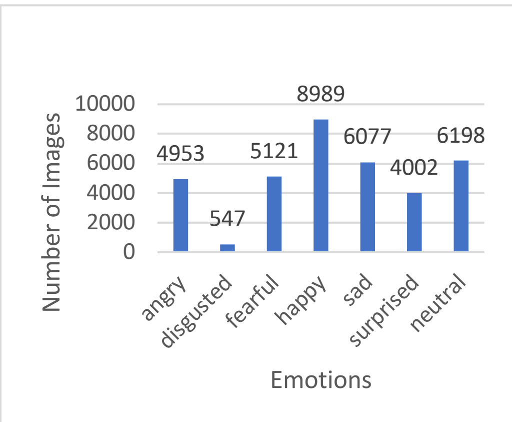

Figure 2: Emotion categorization as per images available in Fer13 dataset

We have utilized publicly available Facial Expression Recognition 2013 (FER13) dataset contains 35,887 grayscale images 48*48 pixels, depicting facial expressions of seven different emotions: anger, disgust, fear, happiness, sadness, surprise and neutral. The first step is to preprocess the image dataset. This includes tasks such as resizing the images to a uniform size, normalizing pixel values, and organizing the images into appropriate 7 classes or categories of emotions. The dataset has a relatively balanced distribution of images across the seven emotion categories, which is shown below in figure 2. Within the FER13 image dataset, there are both occulated images and clear images. To facilitate the emotion detection of images, we transform these images signals into nodes on a graph using data slicing and feature transformation.

Image-to-Graph Conversion

In GNNs, images are converted into graph representations to leverage the power of graph-based learning. One common approach is to treat each image as a node in the graph and define connections (edges) between nodes based on some criteria like facial landmarks here. This conversation allows us to approach image emotion detection as a node classification problem in machine learning. Constructing a graph for FER13 involves creating correlations between subsamples to capture their interconnectivity and similarity. The subsamples processed in the dataset are independent of each other, and traditional deep learning methods treat them separately, overlooking potential correlations. To address this, we propose a “construct graph” method. The construct graph approach comprises three main steps: 1) calculating similarity between subsamples, 2) assigning weights based on similarity, and 3) generating graph.

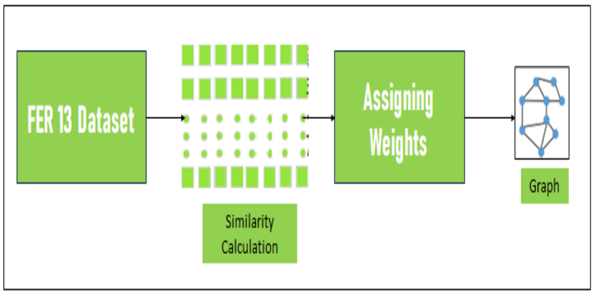

Figure 3: Graph Construct

The flowchart of the "construct graph" method is depicted in Figure 3. Here, first similarity will be evaluated then weights will be assigned to edges which will result in graph as output.

Node Features

Each node in the graph (corresponding to an image) can have associated node features. Here we have extracted features from the images using random horizontal flip. Each node contains intensity value and neighborhood relationship as the edges. For similarity calculation,

$$\text{let } X = \{X1, X2, X3 \ldots., Xm\} \quad (1)$$

represent the processed dataset, where $Xi \leq Xi1, Xi2, Xi3 \ldots..Xid >$ indicates a subsample in X, and d indicates the dimension of X. Xid indicates the value of subsample Xi in the dth dimension. We employ normalised Euclidean distance to measure the similarity between subsamples xi and xj, and it is computed as follows:

$$Dist(xi, xj) = \sqrt{\Sigma(xi - xj)^2}/V \quad (2)$$

Where V is the n-by-n diagonal matrix, and its jth diagonal element is xi^2, where xj is a vector of scaling factor for each dimension. Large values between $Dist(xi, xj)$ indicate higher similarity between two subsamples, capturing their correlation and enabling efficient graph construction for improved Facial Emotion Recognition. Assigning weights involves selecting the top k subsamples with the highest similarity to the xi subsample.

Adjacency Matrix

Once the graph is constructed, an adjacency matrix is created. Figure 4 shows the visualization of the adjacency matrix which is simply an unraveled 2D array. These selected subsamples from the set of neighbors of xi, denoted as Nk(xi).

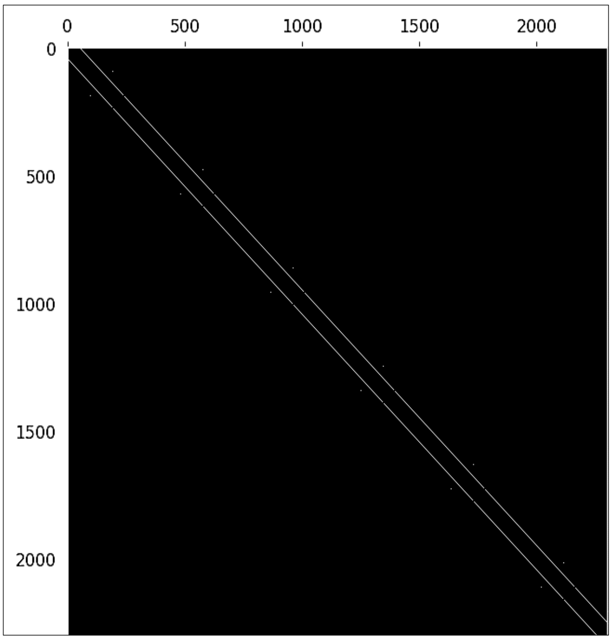

Figure 4: Adjacency Matrix for FER13

Next, we calculate the weights between these k subsamples and xi using the following equation:

$$Weight(xi, xj) = \exp\left(-\frac{Dist(xi,xj)^2}{2\sigma^2}\right) \quad (3)$$

where Dist (xi, xj) is the normalized Euclidean distance between xi and xj, and σ is a hyperparameter that controls the spread of the weights. Once the similarity is calculated and weights are assigned in the first two steps, we can connect the subsamples to each other. The weights between subsamples decides whether a connected edge will exist or not; if the weights linking subsamples exceed 0 then only connection edge exists. By connecting the subsamples, we construct a graph, and this constructed graph is represented by the adjacency matrix A. The diagonal elements of the adjacency matrix A are set to 1, indicating that each subsample is effectively preserved during the subsequent model training process, preventing any loss of essential features. The constructed graph with assigned weights and adjacency matrix A is used to capture the correlations and interconnectivity between subsamples, enhancing the performance of the Facial Emotion Recognition model.

Graph Visualization

To visualize the graph, here we display the nodes and edges in a visually appealing and interpretable way as shown in Figure 5.

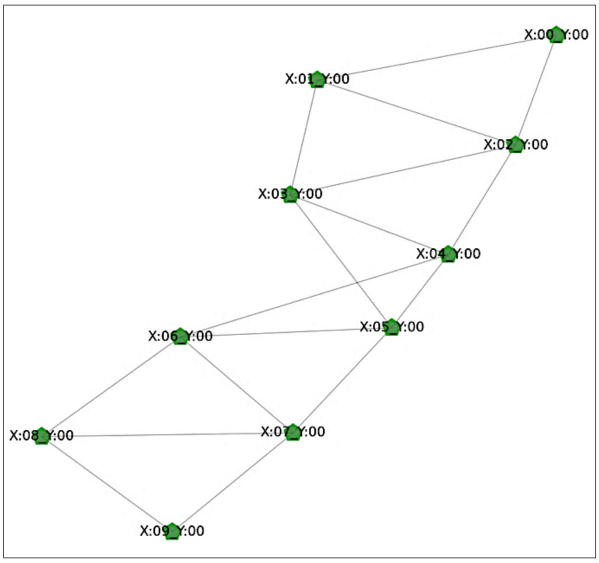

Figure 5: Graph Visualization Topology

Node Coloring

Coloring nodes based on certain attribute or features can help distinguish emotion classes better. For instance, in Fer13 facial images are assigned red color and any other object is colored as green for easy identification as shown in Figure 6 transpose of a sample face.

Figure 6: Node Coloring for facial emotion recognition

Usage of Graph Neural Network-FER 13 (GNN-FER): Revolution for Stress Prediction

In facial emotion recognition conventional neural network architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are designed to work with data in Euclidean space, rendering them unsuitable for handling non-Euclidean data structures like graphs. To address this challenge and capitalize on graph-structured data, we adopt GNN-FER as a powerful framework for deep learning. In our approach, we utilize a recurrent neural structure within the GNN to iteratively propagate neighbor information until reaching a stable equilibrium. In this paper FER13 images were trained over a Graph Neural Network consisting of 3-layer. Firstly, the images are precomputed for adjacency images because graph data structure could be represented either using adjacency matrix or adjacency list. Further, an adjacency matrix is created to predict the edge between any pair of pixels using meshgrid. After that training of GNN model is done. Further ReLU activation function is used over this and further validation is performed. The first layer of our GNN model is the input layer which receives actual image features and pass data to hidden layer that contains 256 neurons then hidden layer is connected with another hidden layer that contains 128 neurons and finally, we have output layer which contains 7 neurons based on our 7 emotion classes.

This process allows us to effectively learn the representation of the target node, enhancing the accuracy and efficiency of facial emotion recognition. By leveraging GNNs for feature extraction, each graph node not only contains its individual information but also incorporates valuable feature information from its neighboring nodes. The forward model of our GNN takes a straightforward form, enabling us to exploit the potential of GNNs and make significant advancements in facial emotion recognition. We implemented the GNN model as follows:

$$Z = f(X, A) = ReLU\ (ReLU\big(X\ AW^{(0)} - b^{(0)}\big)W^{(1)} - b^{(1)} \tag{4}$$

Here, $W \in R^{D*H}$ represents the input to hidden weight matrix for a hidden layer with H feature maps, and $b \in R^{H}$ denotes the input to hidden biases matrix. The GNN weights $W^{(0)}, W^{(1)}$, and biases $b^{(0)}, b^{(1)}$ are trained using gradient descent.

The GNN model operates based on an information propagation mechanism, where each node iteratively exchanges information with neighboring nodes. Through continuous updates, the nodes coverage to a stable state, allowing the smooth flow of information throughout the entire graph. Consequently, each node acquires information not only about itself but also about its neighboring nodes. This process facilitates comprehensive feature extraction and contributes to more accurate FER. The algorithm used is elaborated below:

Input: Given FER dataset X, Adjacency matrix A, Number of iterations n.

1. Initialize $W^{(0)}, W^{(1)}, b^{(0)}, b^{(1)}$.
2. for iteration= 1: n
3. %forward Propagation
4. $Z = \big(X^{*}A^{*}W^{(0)*} - b^{(0)}\big) * (A^{*}W^{(1)*} - b^{(1)})$
5. loss $= \frac{1}{2}(X - Z)^2$
6. %back Propagation
7. Update W and b using batch gradient decent
8. End
9. Return Z

Experiment and Performance Analysis

In this section, we will present a comprehensive demonstration of the effectiveness of the proposed GNN for facial emotion recognition using the FER13 dataset. The evaluation is divided into four parts: a) Introduction to the experimental environment: Here, we will provide detailed information about the experimental setup, including the hardware and software configurations used for training and testing the GNN model. b) Evaluation Method and Comparison Algorithm: In this section, we will define the evaluation metrics used to assess the performance of the GNN in facial emotion recognition, we will also compare the GNN results with those of other relevant facial emotion recognition algorithms to provide a comprehensive analysis. c) Experimental Results: This part will present the final facial emotion recognition results obtained using the GNN algorithm. We will delve into the details of the emotions classified correctly and analyze the model's accuracy on the FER13 dataset. d) Effectiveness Analysis: To verify the effectiveness of the GNN algorithm for facial emotion recognition, we will conduct rigorous evaluations, including cross-validation experiments, to assess the model's generalization capability and robustness. This analysis will demonstrate the reliability and performance of the GNN approach in accurately recognizing facial emotions from the FER13 dataset.

RESULTS AND DISCUSSION

To assess the performance of the GNN based FER model, we employ standard evaluation metrics commonly used in facial emotion recognition tasks. Specifically, we utilize the receiver operating characteristic (ROC) curve and the corresponding area under the curve (AUC) as our primary evaluation metrics. Additionally, we calculate accuracy plot (ACC) and loss plot (LOSS) to further evaluate the model's performance. Higher AUC and ACC and lower LOSS values indicate better recognition performance. These metrics provide valuable insights into the model's ability to correctly classify facial emotion states.

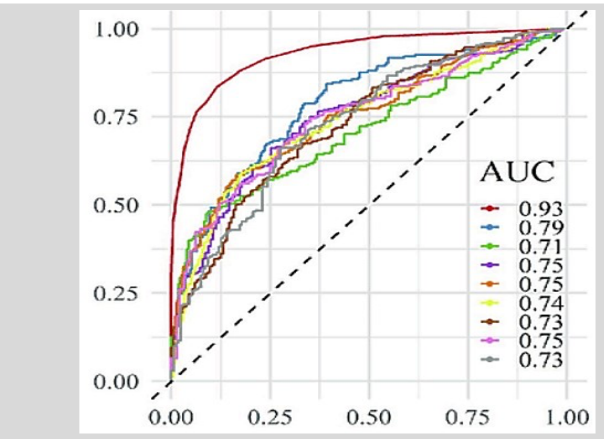

Figure 7: AUC-ROC curve for FER-13

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \qquad (5)$$

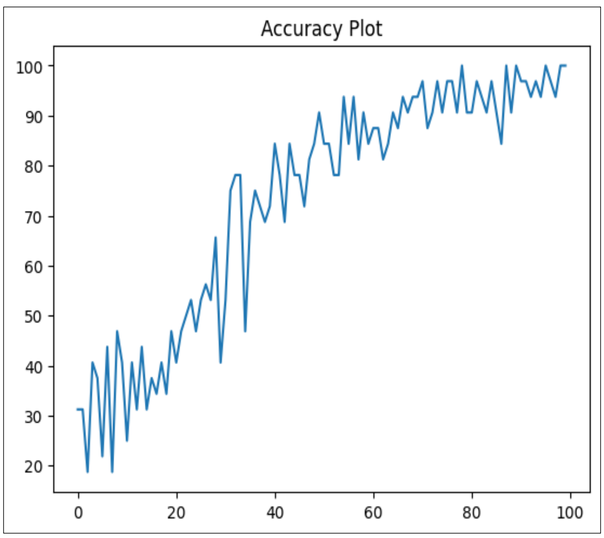

Figure 8: Accuracy Plot for FER-13

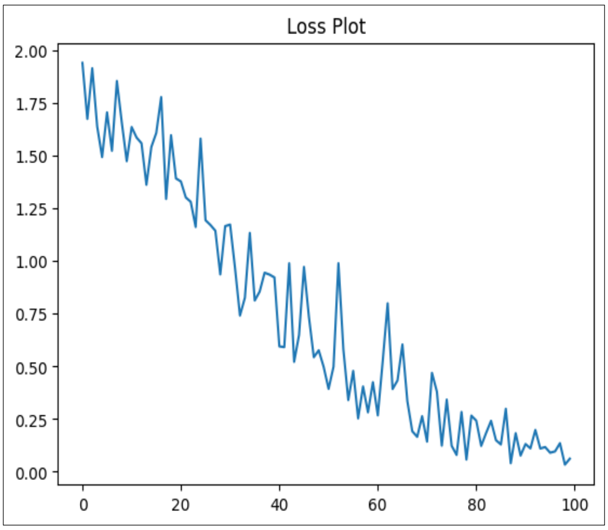

Figure 9: Loss Plot for FER-13

CONCLUSION

In this study we introduced a novel approach for enhancing the precision of facial emotion recognition suing Graph Neural Network (GNN). The GNN exhibits a remarkable capability for intricate feature mapping, allowing it to effectively capture both the inherent features of individual samples and the contextual information from their neighboring samples. The utilization of GNN enables the generation of feature-enriched representations, aiding in a more distinct differentiation between emotion categories or classes. Recognizing the imperative demand for robustness in real world scenarios, we integrate a comprehensive technique to amalgamate the outcomes of facial expression detector. This synthesis approach contributes to the stability and efficiency of our proposed GNN-based facial emotion recognition (GNN-FER) algorithm. Rigorous experiments conducted on the FER13 dataset, which is widely recognized in the field, validate the efficacy of the GNN-FER algorithm in effectively identifying and distinguishing various emotional states within the dataset. The accuracy achieved by GNN-FER is 99.89% which is exemplary in the domain.

To further advance our approach, further investigations will delve into enhancing the detection performance of the GNN-FER model by exploring the potential benefits of deeper network architecture and more intricate feature representations. Moreover, in the future, other facial recognition datasets can also be explored with GNN for better opportunities. This research holds significant promise for contributing to the evolution of facial emotion recognition systems and their practical applicability.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Brådvik L. Suicide Risk and Mental Disorders. International Journal Environmental Research and Public Health 2018; 15(9):1-4. https://doi.org/10.3390/ijerph15092028

Zainudin Z, Hasan S, Shamsuddin SM, Agrawal S. Stress Detection using Machine Learning and Deep Learning. Journal of Physics: Conference Series Asian Conference on

Intelligent Computing and Data Sciences (ACIDS) 2021; 1997:1-8.
https://doi.org.10.1088/1742-6596/1997/1/012019

Maaoui C, Pruski A. Unsupervised stress detection from remote physiological signal. IEEE International Conference on Industrial Technology (ICIT) 2018; 2018: 1538-1543.
https://doi: 10.1109/ICIT.2018.8352409

Nigam K,Godani K, Sharma D, Jain S. An Improved Approach for Stress Detection Using Physiological Signals. EAI Endorsed Transactions on Scalable Information Systems; 8(33):1-8.
https://eudl.eu/pdf/10.4108/eai.14-5-2021.169919

Katmah R, Shargie F, Tariq U, Babiloni F, Mughairbi F, Nashash H. A Review on Mental Stress Assessment Methods Using EEG Signals. Sensors 2021; 21(15):1-26.
https://www.mdpi.com/1424-8220/21/15/5043

Shahbazi Z, Byun CY. Early Life Stress Detection Using Physiological Signals and Machine Learning Pipelines. Biology 2023; 12(1). 90-116.
https://www.mdpi.com/2079-7737/12/1/91

Maltman N, DaWalt SL, Hong J, Baker WM, Kravis BE, Brilliant HM, Mailick M. FMR1 CGG Repeats and Stress Influence Self-Reported Cognitive Functioning in Mothers. American Journal on Intellectual and Develpomental Disabilities 2023;128(1). 1-20.
https://pubmed.ncbi.nlm.nih.gov/36548377/

Witte de M, Kooijamans R, Hermanns M, Hooren VS, Beismans K, Hermsen M, Stams JG, Moonen X. Self-Report Stress Measures to Assess Stress in Adults With Mild Intellectual Disabilities-A Scoping Review. Frontier Psychology 2021; 21:1-23.
https://www.frontiersin.org/articles/10.3389/fpsyg.2021.742566/full

Thoen A, Steyaert J, Alaerts K, Evers K, Tine VD. A Systematic Review of Self-Reported Stress Questionnaires in People on the Autism Spectrum. Review Journal of Autism and Developmental Disorders 2021; 10(2): 295-318.
https://pubmed.ncbi.nlm.nih.gov/34603935/

Hajera S, Ali MM. A Comparative Analysis of Psycholgical Stress Detection Methods. International Journal of Computational Engineering & Management 2018; 21(2):1-8.
https://ijcem.org/papers032018/ijcem_032018_01

Nijhawan T, Attigeri G, Ananthakrishna T. Stress detection using natural language processing and machine learning over social interactions. Journal of Big Data 2022;33(9): 1-24.
https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00575-6#citeas

Li R, Liu Z. Informatics and machine learning methods for health applications. BMC Medical Informatics and Decision Making 2020; 20(11):1-10.
https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01299-4

Mittal S, Mahendra S, Sanap V, Churi P. How can machine learning be used in stress management: A systematic literature review of applications in workplaces and education. International Journal of Information Management Data Insights 2022; 2(2):1-13.
https://www.sciencedirect.com/science/article/pii/S2667096822000532

Zhou S, Wu X, Jiang F, Huang Q, Huang C. Emotion Recognition from Large Scale Video Clips with Cross Attention and Hybrid Feature Weighting Neural Networks. International Journal of Environmental Research and Public Health 2023; 20(2):1-23.
https://www.mdpi.com/1660-4601/20/2/1400

Liu H, Cai H, Lin Q, Zhang X, Li X, Xiao H, FEDA: Fine-grained emotion difference analysis for facial expression recognition. Biomedical Signal Processing and Control 2023;79(2):1-23.
https://doi.org/10.1016/j.bspc.2022.104209.

Khosrowabadi R. Stress and Perception of Emotional Stimuli:Long-term Stress Rewiring the Brain. Basic and Clinical Neuroscience 2018; 9(2):107-120.
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6026093/