Depression Detection from Facial Expression Using Hybrid Model of FCN, Fusion Fuzzy Logic, and LSTM

V.A. Narayana[1], Chakilam Bharath Kumar[2], Nimmala Sai[3], and Damarla Hemavathi[4]

Department of Computer Science and Engineering,
CMR College of Engineering Technology, Telangana, India
vanarayana@cmrcet.org, bharathchakilam2137@gmail.com,
sainimmala2025@gmail.com, damaralahemavathi9@gmail.com

Abstract. Depression is a severe mental condition that affects millions of individuals all over the world. The importance of early diagnosis lies in early intervention and treatment. In this project, a depression detection system using facial expressions through deep learning is proposed. A hybrid model based on Fully Convolutional Networks, Long Short-Term Memory, and Fusion Fuzzy Logic is used to analyze facial features and recognize depressive symptoms. The system processes images using face detection, feature extraction, and classification. Spatial facial features are extracted by Fully Convolutional Networks, temporal dependencies are captured by Long Short-Term Memory, and Fusion Fuzzy Logic enhances decision-making by solving expression variability. The model is trained on labeled images ("Depressed" and "Not Depressed") and tested on Google Colab and Kaggle for optimal performance. The final model is implemented in a web application with image uploading capability for real-time depression detection. The approach provides an efficient, scalable, and accurate means of automated depression diagnosis for the benefit of mental health education and early intervention.

Keywords: Fully Convolutional Networks, Long Short-Term Memory, Fusion Fuzzy Logic, Depression Detection, Face Recognition.

1 Introduction

Depression is an intense and prevalent mental illness that affects millions of individuals worldwide[1], producing emotional distress, cognitive impairment, and even medical disease[2]. Standard diagnostic methods for depression are clinician interviews, self-report questionnaires, and psychiatric rating scales, all of which are subjective, time consuming, and susceptible to biases[3]. As the increasing psychological issues of contemporary, fast-paced life, a need for automated, non-invasive, and efficient early detection mechanisms has been rising[4].

Facial expressions are the most important element of human emotions, and due to ongoing advancements in deep learning, facial expression analysis can be performed for diagnosing mental health disorders[5]. The goal of this project is to design a hybrid deep network model incorporating FCN and LSTM networks along with Fusion Fuzzy Logic for efficiently detecting depression from facial expressions[6]. The model is trained on a dataset of facial images of depressed and non-depressed patients that are labeled to recognize specific characteristics.

The proposed system will be implemented in real-time, and users can upload an image through a web-based interface. The trained model will process the image and mark it as "Depressed" or "Not Depressed," showing a consistent and automated way of detecting depression[7]. By integrating this technology into well-being and healthcare solutions, the project aims to assist in supporting early intervention processes, reducing mental illness stigma and encouraging people to seek professional help in a timely manner.conditions.

2 Literature Survey

Facial expression-based depression detection has become a trending research topic with the advancement of artificial intelligence and deep learning. The traditional method of diagnosing depression relies on clinical assessment, self-report questionnaires, and psychological tests[8]. Standard diagnostic instruments such as the Beck Depression Inventory and Patient Health Questionnaire are used to measure the degree of depression[8]. But these methods are subject to bias and the voluntary cooperation of the patient to report his or her own emotions, and they introduce variability in the diagnosis.

Facial expression analysis has been a promising technique for detecting depression since facial expressions and micro-expressions have strong correlations with psychological and emotional states. Depressed people have been found to typically show lower facial activity, longer neutral facial expressions, and reduced emotional range. Facial Action Coding System has been used in various studies to derive subtle facial movements indicating mental health disorders[9]. Machine learning models such as Support Vector Machines, Random Forest, and K-Nearest Neighbors have been used to deploy depression classification from facial features. Feature extraction methods like Histogram of Oriented Gradients, Local Binary Patterns, and Gabor filters have been used to extract specific patterns in facial expressions. However, these approaches are based on handcrafted feature engineering, which can become restrictive to generalizing to different subjects and datasets.

Deep learning has transformed the performance of depression detection models by eliminating the need for handcrafted feature extraction. Convolutional Neural Networks[9] have been widely used to apply automatic facial feature extraction, and their ability to learn hierarchical patterns from facial images has made them extremely capable for classification. Several studies have integrated CNNs with Recurrent Neural Networks and Long Short-Term Memory networks to explore the temporal dynamics of facial expressions, providing a better picture of emotional states through time.

Hybrid deep learning models[10] combining CNNs and LSTMs have demonstrated improved accuracy by incorporating spatial and temporal variations in the facial expressions. Fusion Fuzzy Logic has also been examined to handle uncertainty in facial expressions, improving the model's ability to differentiate between depressed and non-depressed subjects[11]. A few publicly available datasets, such as the DAIC-WOZ dataset and the AVEC dataset, have been used to test and train deep learning models for depression detection. In-house-labeled datasets with facial images labeled based on depression severity have also been built by researchers.

The integration of deep learning has been coupled with facial expression analysis to reveal new possibilities for automatic and objective identification of depression. Despite the promising research conducted to date, challenges such as dataset bias, variability in facial expressions from individual to individual, and real-time analysis necessities have yet to be addressed. Future studies will attempt to enhance model accuracy and robustness through improved data augmentation techniques, larger and more diverse datasets, and more advanced hybrid models combining multiple deep learning techniques.

3 Methodology

The workflow for depression detection using facial expressions consists of several steps from data preprocessing to model deployment. The hybrid model uses Fully Convolutional Networks for feature extraction, Fusion Fuzzy Logic for handling uncertainty in classification, and Long Short-Term Memory for facial expression temporal analysis. All steps are significant in detecting depression both accurately and robustly.

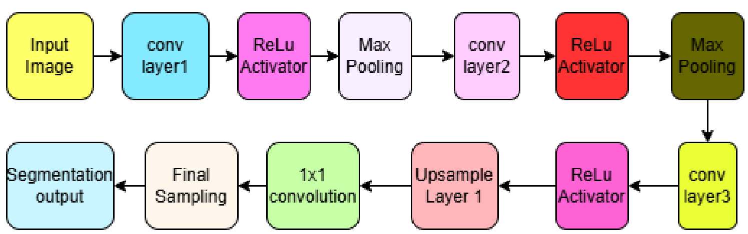

Fig. 1. Fully Convolutional Network

The pipeline begins by preprocessing data: facial images are resized to a uniform size, normalized to settle training, and augmented through rotations, flipping, and contrast adjustment. This enhances dataset variability as well as supports improved generalization to unseen data on the model's part. The images are then processed through an FCN, extracting spatial features encompassing facial points, texture variation, as well as musculature in motion expressing affective states.

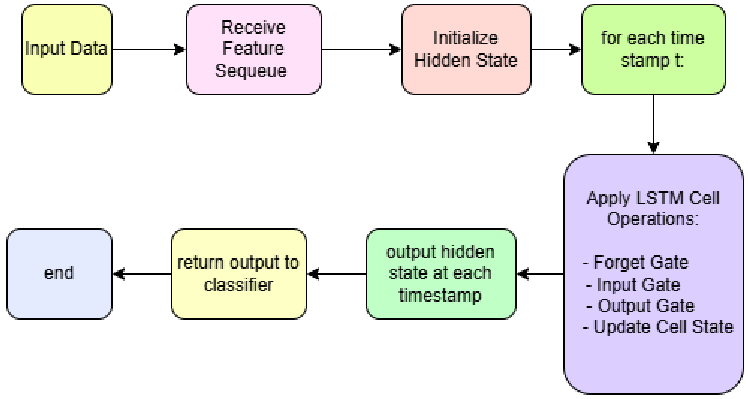

Fig. 2. Long Short Term Memory

Features are then passed on to Fusion Fuzzy Logic[12], which reduces the choices of classification by giving different features confidence values. This module enables the model to be robust to variations in lighting, poses of the head, and partial occlusions, which typically affect facial analysis models. The information is then passed through an LSTM network to investigate sequential patterns of facial expressions in time after feature set reduction. Contrarily to static single-image models, LSTM considers soft expression variations within multiple frames in order to support better detection of depressive states, which are unobservable in an individual image.

The hybrid model is learned using a labeled facial image set separated into "depressed" and "not depressed" groups. Training involves feeding preprocessed images to the FCN to extract features, FFL to process the features further, and LSTM to examine sequential patterns. Backpropagation and gradient descent are used for optimizing the model, with learning rate, batch size, and dropout rate being hyperparameters for improving performance.

3.1 Dataset Specification

The training, validation, and testing dataset of the depression detection model includes facial images grouped into two categories: "depressed" and "not depressed." The dataset is organized into three subsets: training, validation, and test sets to balance the assessment of the model. All three subsets are labeled images of their respective depressive state so that the model can learn and generalize properly.

The distribution of the dataset is as follows:

Training Set: 8926 pics of depressed ones, 7215 pics of non-depressed ones.

Validation Set: 1435 photographs of depressed participants, 1187 photographs of non-depressed participants.

Test Set: 1184 photos of depressed people, 887 photos of non-depressed people.

The dataset is mainly facial expressions taken under varying conditions, such as lighting, head pose, and intensity of emotions. These variations contribute to how the model can be made more resilient to real-world conditions. The dataset is organized such that the model can learn both spatial and temporal characteristics of facial features related to depression, and it is therefore suitable for classification using deep learning.

3.2 Model Architecture

The model begins with an input layer that takes facial pictures, reduces their dimensions to a standard size, and rescales pixel values to improve the stability of training. The preprocessed pictures are then passed through an FCN, which detects spatial features in different regions of the face, including prominent landmarks, movement of muscles, and variations in texture. As compared to traditional CNNs, FCNs do not drop spatial information throughout the process of feature extraction and are hence most appropriate for the use of localizing facial expressions with precise spatial information[13].

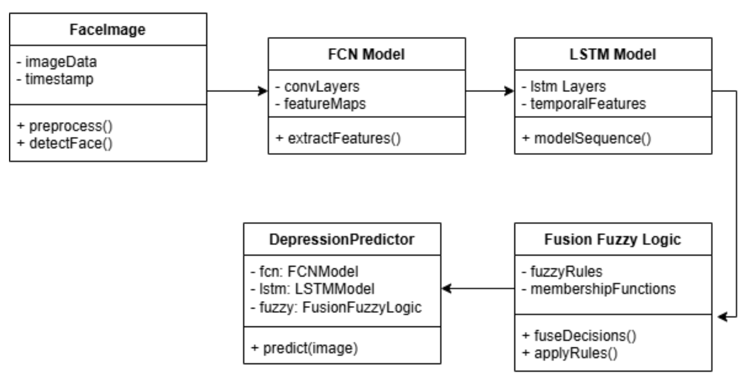

Fig. 3. Architecture of proposed system

The features extracted by the FCN are then fed to the Fusion Fuzzy Logic module[14], which enhances the process of classification by resolving uncertainty in facial expressions[15]. Since depressive states may at times occur consistently within each individual, this component provides confidence levels to identified patterns in order to improve the robustness of classification. It makes the model resistant to variations in lighting, head poses, and expression intensity without being overly sensitive to minor distortions[16].

Finally, the fine-tuned feature set is fed into an LSTM network, which learns sequential dependencies between frames of a video or between minor variations in facial features between static images. LSTM turns out to be useful in detecting depression because depressive facial expressions are typically long and repetitive micro-expressions that are not necessarily detectable from individual images. By considering time-series data, LSTM enhances the ability of the model to detect such patterns and, therefore, make a more informed classification.

The final few layers are the fully connected layers that fuse the features learned and output the final classification result. The final layer uses softmax activation to output the classification as "depressed" or "not depressed." The categorical cross-entropy loss and adaptive gradient-based optimizers like Adam are used to train the model to ensure stable convergence. Batch normalization and dropout layers are also added to prevent overfitting and enhance generalization.

The complete architecture as depicted in fig 1 is formulated to process facial images with maximal information preservation at each stage in a pipeline fashion. By combining FCN for feature learning, Fusion Fuzzy Logic[12] to process uncertainty, and LSTM for processing sequential inputs, the model achieves high accuracy in depression detection from facial expressions. The modularity in the architecture allows for flexibility with space for future developments and data refinement.

4 Results and Discussion

The proposed hybrid deep learning framework, a combination of Fully Convolutional Networks, Fusion Fuzzy Logic, and Long Short-Term Memory, exhibits remarkable performance in depression detection from facial expressions. The model was evaluated on a given test set, and the result indicates strong accuracy and robustness in classification.

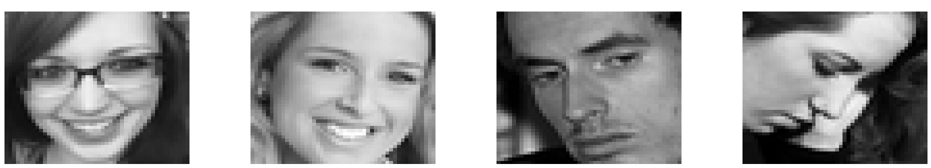

Fig. 4. Sample Input Images

Table 1: Performance Comparison of Models

Model	Accuracy	Precision	Recall	F1
FCN	87%	89%	87%	88%
LSTM	77%	76%	75%	76%
Proposed(Hybrid)	92%	93%	94%	93%

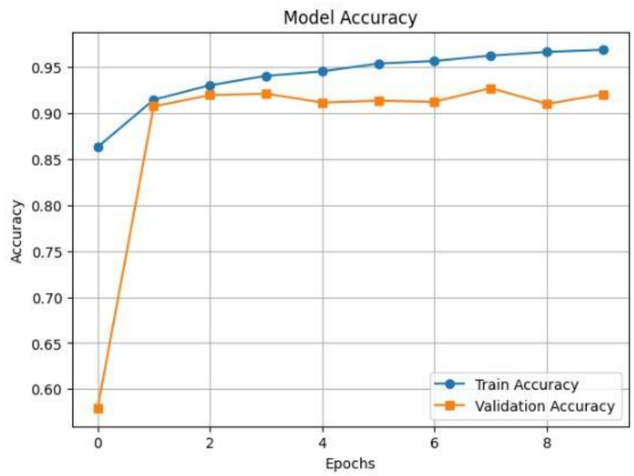

Fig. 5. Accuracy of proposed model

With a total accuracy of 92.95% as shown in Fig.2, the model successfully differentiates between "depressed" and "not depressed" facial expressions. The precision rate of 93.41% indicates that the model accurately classifies depressive cases with few false positives. The recall of 94.20% indicates the model's capacity to identify most depressive cases, reducing false negatives. The F1-score, which is the balance between precision and recall, is 93.80%, affirming the efficacy of the model in depression classification.

5 Conclusion

The hybrid model proposed uses the merits of Fully Convolutional Networks, Long Short-Term Memory, and Fusion Fuzzy Logic in identifying depression from facial expressions. The FCN part is used for extracting informative facial features through the capture of spatial information using convolutional layers, which improves the model's capacity to recognize slight differences in expressions of depressed states. While this, the LSTM module is geared toward examining temporal relations in sequences of facial expressions, enabling the model to perceive how and when the emotions are developed over time. Furthermore, Fusion Fuzzy Logic integration enhances the decision-making ability by mitigating the built-in uncertainty and ambiguity of human feelings, allowing the model to efficiently deal with ambivalent or confounding expressions more accurately. By integrating these methods, the hybrid model obtains a balance between spatial, temporal, and decision-level analysis, resulting in a 92.95% classification accuracy. The high values of precision, recall, and F1-score evident from the experimental outcome also corroborate the confidence with which the model detects depressive facial expressions.

References

1. Xinru Kong, Yan Yao, Cuiying Wang, Yuangeng Wang, Jing Teng, and Xianghua Qi. Automatic identification of depression using facial images with deep convolutional neural network. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 28:e936409–1, 2022.
2. Gajendra Kumar, Tanaya Das, and Kuldeep Singh. Early detection of depression through facial expression recognition and electroencephalogram-based artificial intelligence-assisted graphical user interface. *Neural Computing and Applications*, 36(12):6937–6954, 2024.
3. Yanisa Mahayossanunt, Natawut Nupairoj, Solaphat Hemrungrojn, and Peerapon Vateekul. Explainable depression detection based on facial expression using lstm on attentional intermediate feature fusion with label smoothing. *Sensors*, 23(23):9402, 2023.
4. Marta Monferrer, Arturo S García, Jorge J Ricarte, María J Montes, Antonio Fernández-Caballero, and Patricia Fernández-Sotos. Facial emotion recognition in patients with depression compared to healthy controls when using human avatars. *Scientific reports*, 13(1):6007, 2023.
5. Luna Ansari, Shaoxiong Ji, Qian Chen, and Erik Cambria. Ensemble hybrid learning methods for automated depression detection. *IEEE Transactions on Computational Social Systems*, 10(1):211–219, 2023.
6. Bin Hu, Yongfeng Tao, and Minqiang Yang. Detecting depression based on facial cues elicited by emotional stimuli in video. *Computers in Biology and Medicine*, 165:107457, 2023.
7. Zhuozheng Wang, Yihan Wang, Xixi Zhao, Keyuan Li, Gang Wang, and Lei Feng. A novel depression detection model based on hybrid neural network. In *2024 IEEE 1st International Workshop on Future Intelligent Technologies for Young Researchers (FITYR)*, pages 53–59. IEEE, 2024.
8. Yanfei Wang, Jie Ma, Bibo Hao, Pengwei Hu, Xiaoqian Wang, Jing Mei, and Shaochun Li. Automatic depression detection via facial expressions using multiple instance learning. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1933–1936. IEEE, 2020.
9. Weitong Guo, Hongwu Yang, Zhenyu Liu, Yaping Xu, and Bin Hu. Deep neural networks for depression recognition based on 2d and 3d facial expressions under emotional stimulus tasks. *Frontiers in neuroscience*, 15:609760, 2021.
10. Sanagala S Skandha, Andrew Nicolaides, Suneet K Gupta, Vijaya K Koppula, Luca Saba, Amer M Johri, Manudeep S Kalra, and Jasjit S Suri. A hybrid deep learning paradigm for carotid plaque tissue characterization and its validation in multicenter cohorts using a supercomputer framework. *Computers in biology and medicine*, 141:105131, 2022.
11. Yiwen Meng, William Speier, Michael K Ong, and Corey W Arnold. Bidirectional representation learning from transformers using multimodal electronic health record data to predict depression. *IEEE journal of biomedical and health informatics*, 25(8):3121–3129, 2021.
12. VA Narayana, Madhavi Pingili, K Srujan Raju, and Chilukuri Dileep. An essential hybrid model for developing fuzzy rules in a specific malware identification and

detection system. In *Applications of Mathematics in Science and Technology*, pages 186–191. CRC Press, 2025.
13. VI Osubor and AO Egwali. A neuro fuzzy approach for the diagnosis of postpartum depression disorder. *Iran Journal of Computer Science*, 1(4):217–225, 2018.
14. Meza Silvana, Ricky Akbar, Mia Audina, et al. Development of classification features of mental disorder characteristics using the fuzzy logic mamdani method. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI)*, pages 410–414. IEEE, 2018.
15. Le Yang, Dongmei Jiang, and Hichem Sahli. Integrating deep and shallow models for multi-modal depression analysis—hybrid architectures. *IEEE Transactions on Affective Computing*, 12(1):239–253, 2018.
16. Yan Ding, Xuemei Chen, Qiming Fu, and Shan Zhong. A depression recognition method for college students using deep integrated support vector algorithm. *IEEE access*, 8:75616–75629, 2020.