A Project report on

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

A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

Bachelor of Technology

in

Computer Science and Engineering

Submitted by

Ch.Bharath Kumar

(22H55A0505)

N.Sai

(21H51A0571)

D.Hemavathi

(21H51A05M1)

Under the esteemed guidance of

Major Dr. V. A. Narayana

(Director CMRCET)

Department of Computer Science and Engineering

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

(UGC Autonomous)

*Approved by AICTE *Affiliated to JNTUH *NAAC Accredited with A^+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

2021- 2025

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Major Project Phase II report entitled **"Depression Detection using Facial Expression Using Hybrid Model of FCN, Fusion fuzzy Logic, and LSTM"** being submitted by Ch.Bharath Kumar(22H55A0505), N.Sai(21H51A0571), D.Hemavathi(21H51A05M1) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

Major Dr. V. A. Narayana **Director CMRCET**	**Dr. Siva Skandha Sanagala** **Associate Professor & HOD** **Dept. of CSE**	**EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

With great pleasure, we want to take this opportunity to express my heartfelt gratitude to allthe people who helped make this project a grand success.

We are grateful to **Major Dr. V. A. Narayana**, **Director, CMRCET** Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of ComputerScience and Engineering, CMR College of Engineering and Technology, who is the major driving force to complete my project work successfully.

We are very grateful to **Dr. J Rajeshwar**, Dean of CSE at CMR College of Engineering andTechnology, for his constant support and motivation in successfully completing the project.

We are very grateful to **Dr. Ghanta Devadasu**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Dr. A. Seshu Kumar,** Principal of CMR College of Engineering and Technology, for allowing us to carry out this project successfully and fruitfully.

We are thankful to **Major Dr. V. A. Narayana,** Director, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly incompletion of this project work.

Chakilam Bharath Kumar 22H55A0505
Nimmala Sai 21H51A0571
Damarla Hemavathi 21H51A05M1

TABLE OF CONTENTS

List of Figures

List of Tables

ABSTRACT

Depression is a widespread mental health condition that significantly impacts individuals and society. Early detection and intervention are crucial in mitigating its effects. This project focuses on developing an automated system for depression detection using facial expressions, leveraging advanced machine learning models, including Long Short-Term Memory (LSTM), Fully Convolutional Networks (FCN), and Fusion Fuzzy Logic.

The proposed system analyzes facial expressions captured through video frames or images to classify individuals as 'depressed' or 'not depressed.' The methodology employs three distinct datasets: training, validation, and testing, each organized into 'depressed' and 'not depressed' categories. These datasets are used to train and validate the models to ensure robust performance across diverse scenarios.

The project emphasizes the importance of achieving high accuracy, sensitivity, and specificity by tuning the models to minimize false positives and false negatives. To enhance the system's interpretability and usability, the Fusion Fuzzy Logic component combines the strengths of LSTM and FCN models for more reliable predictions.

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

Depression is a pervasive mental health disorder that affects millions of individuals worldwide, leading to significant personal, social, and economic challenges. Characterized by persistent feelings of sadness, hopelessness, and a lack of interest in daily activities, depression often goes undetected due to social stigma, lack of awareness, or limited access to mental health resources. Early and accurate detection is critical to mitigate its impact and facilitate timely intervention.

Facial expressions serve as a vital non-verbal indicator of emotional and psychological states. They provide valuable cues for identifying signs of mental health conditions, including depression. Recent advancements in artificial intelligence and machine learning have enabled the development of automated systems capable of analyzing facial expressions for mental health assessment. These systems offer a non-invasive, efficient, and scalable solution for detecting depression in diverse populations.

This project aims to develop a robust depression detection system using facial expressions, leveraging state-of-the-art machine learning models such as Long Short-Term Memory (LSTM), Fully Convolutional Networks (FCN), and Fusion Fuzzy Logic. By integrating these models, the system seeks to classify individuals into 'depressed' and 'not depressed' categories with high accuracy, sensitivity, and specificity.

The implementation involves utilizing structured datasets for training, validation, and testing, with data stored and processed on Google Drive through Google Colab. The approach focuses on achieving optimal performance metrics, including minimizing false positives and negatives, ensuring the reliability and effectiveness of the proposed system.

1.1. Problem Statement

Depression, a leading cause of disability worldwide, is often underdiagnosed due to stigma, lack of resources, and reliance on self-reported symptoms. Traditional diagnostic methods require significant time and expertise, making early detection challenging, especially in resource-constrained environments. The absence of timely diagnosis and intervention exacerbates the condition, leading to severe emotional, social, and economic consequences for individuals and society.

Facial expressions, as a universal medium of emotional expression, offer a promising avenue for detecting depression. However, existing approaches for facial expression analysis often lack scalability, reliability, or the ability to effectively integrate diverse features of human emotions. Additionally, these systems may fail to provide actionable insights due to high rates of false positives and negatives.

This project addresses the critical need for an automated, efficient, and non-invasive system capable of detecting depression using facial expressions. By leveraging advanced machine learning techniques such as Long Short-Term Memory (LSTM), Fully Convolutional Networks (FCN), and Fusion Fuzzy Logic, the aim is to develop a robust tool that classifies individuals as 'depressed' or 'not depressed.' The system will focus on high accuracy, minimizing diagnostic errors, and providing a scalable solution suitable for diverse populations.

The successful implementation of this project has the potential to bridge the gap in depression diagnosis, facilitating early detection and enabling timely mental health interventions, thus improving overall well-being and quality of life.

1.2. Research Objective

The primary objective of this research is to develop a robust and automated system for the detection of depression using facial expressions. The specific objectives are as follows:

1. **Feature Extraction**: Analyze and extract key features from facial expressions that are indicative of depression.

2. **Model Development**: Design and implement advanced machine learning models, including Long Short-Term Memory (LSTM), Fully Convolutional Networks (FCN), and Fusion Fuzzy Logic, to classify individuals as 'depressed' or 'not depressed.'

3. **Dataset Utilization**: Utilize structured datasets comprising 'depressed' and 'not depressed' categories for training, validation, and testing to ensure comprehensive model evaluation and generalizability.

4. **Performance Optimization**: Optimize the models to achieve high accuracy, sensitivity, and specificity while minimizing false positives and false negatives.

5. **Scalability and Usability**: Develop a scalable, non-invasive system that can be effectively deployed in diverse environments to assist mental health professionals in early diagnosis and intervention planning.

6. **Integration of Models** : Combine the strengths of LSTM and FCN models using Fusion Fuzzy Logic to enhance the reliability and interpretability of the system.

7. **Evaluation Metrics**: Assess the system's performance using metrics such as true positives, false positives, true negatives, and false negatives to validate its effectiveness in real-world scenarios.

This research aims to contribute to the mental health domain by offering an accessible and efficient tool for depression detection, ultimately aiding in early intervention and improving mental health outcomes.

1.3 Project Scope And Limitations

Scope:

This project focuses on developing an automated depression detection system using facial expressions, leveraging advanced machine learning techniques. It explores the application of Long Short-Term Memory (LSTM), Fully Convolutional Networks (FCN), and Fusion Fuzzy Logic to analyze and classify individuals as 'depressed' or 'not depressed.' The scope includes dataset preparation, feature extraction, model training, validation, and testing using structured datasets stored and managed on Google Drive via Google Colab. The system aims to achieve high accuracy, sensitivity, and specificity, with an emphasis on minimizing diagnostic errors such as false positives and negatives. The project is designed to be scalable and user-friendly, suitable for deployment in diverse environments to support mental health professionals in early detection and intervention. By integrating cutting-edge AI models with interpretable fuzzy logic, the system seeks to provide a reliable and efficient tool that addresses the growing need for non-invasive mental health diagnostics. This initiative has the potential to enhance mental health care accessibility and improve patient outcomes significantly.

Limitations:

1. **Dataset Dependency:** The accuracy and reliability of the system are heavily dependent on the quality, size, and diversity of the datasets used. Limited or biased datasets may affect the system's ability to generalize across populations.
2. **Complex Emotional States:** Facial expressions may not always accurately reflect an individual's emotional or mental state, as some individuals may suppress or mask their emotions, leading to potential misclassification.
3. **Cultural and Demographic Variability:** Variations in facial expression interpretation due to cultural or demographic differences may impact the model's performance in diverse populations
4. **Environmental Factors:** Lighting, background, and other environmental conditions during data capture may introduce noise and reduce the accuracy of feature extraction and classification.

In summary, Addressing these limitations in future research and development will be crucial to improving the system's robustness, scalability, and applicability in real-world scenarios.

CHAPTER 2
BACKGROUND WORK

CHAPTER 2

BACKGROUND WORK

2.1 Automatic Identification of Depression Using Facial Images with Deep Convolutional Neural Network

2.1.1. Introduction

. In recent years, deep learning has been rapidly developing, which has attracted the attention of an increasing number of researchers. As a branch of machine learning, deep learning is an algorithm based on an artificial neural network for data representation learning. It has obvious advantages over shallow models in feature extraction and model fitting, and it is good at mining abstract distributed feature representations with good generalization ability from the original input data. With deep learning, some of the problems that were thought difficult in the past can be solved.

The convolutional neural network (CNN) is one classical and widely used network structure. CNNs are composed of one or more convolutional layers, fully connected layers (corresponding to classical neural networks), as well as association weights and pooling layers. This structure enables the CNN to use 2-dimensional (2D) data as input. Compared with other deep learning structures, CNNs can obtain better results in images. This model can be trained using backpropagation algorithms. Compared with other deep, feedforward neural networks, CNNs require fewer parameters to be estimated, making them an attractive deep learning architecture .

2.1.2. Merits,Demerits and Challenges

Merits:

1. **Non-Invasive Assessment:**Facial image analysis provides a non-invasive method to detect depression, avoiding the stigma associated with traditional psychological assessments.
2. **High Accuracy and Efficiency**:CNNs excel in pattern recognition and can identify subtle changes in facial expressions, features, or micro-expressions linked to depression.
3. **Automated Process:**Automation eliminates the need for extensive manual interpretation by psychologists, saving time and reducing human error.

4. **Early Detection:** This technology can help in early diagnosis, enabling timely interventions and reducing the risk of severe depression.

Demerits:

1. **Privacy and Ethical Concerns:** Using facial images raises privacy issues, and misuse of the technology could lead to ethical violations, such as unauthorized surveillance or discrimination.
2. **Bias in Training Data:** CNNs are highly dependent on the quality and diversity of training data. Biases in the dataset (e.g., overrepresentation of certain demographics) can lead to inaccurate predictions for underrepresented groups.
3. **Interpretability:** CNNs are often considered black-box models, making it difficult to interpret how the model arrives at its conclusions.
4. **Dependence on Facial Features:** Depression can manifest differently among individuals. Sole reliance on facial features might overlook cases where facial cues are not prominent.
5. **Environmental Factors:** Lighting, camera quality, and other environmental conditions can impact the accuracy of facial image analysis.

Challenges:

1. **Data Quality and Availability:** High-quality, labeled datasets specific to depression are scarce due to privacy concerns and the sensitive nature of mental health.The diversity of the data may be limited, leading to models that do not generalize well across different demographics.
2. **Variability in Facial Features:** Depression symptoms vary widely across individuals and may not always manifest in observable facial cues.Variations in ethnicity, age, and cultural expression can affect the model's performance.
3. **Environmental Conditions:** Lighting, camera angles, image resolution, and background noise can significantly affect the quality of input data, leading to reduced model accuracy.
4. **Overfitting and Generalization:** CNNs may overfit to training data, especially if the dataset is the small, leading to poor generalization on unseen data.

2.1.3. Implementation

Five deep CNN models were constructed for depression identification: the fully connected convolutional neural network (FCN); visual geometry group 11 (VGG11); visual geometry group 19 (VGG19); deep residual network 50 (ResNet50), and Inception version 3 (V3). The FCN model was integrated with the current advanced attention mechanism, and the model included a feature input layer, convolution layer, activation layer, and full connection layer. VGGNet is a deep CNN architecture developed by the Visual Geometry Group (VGG) of Oxford University. VGG-11 consists of 8 convolution layers and 3 fully connected layers. VGG-19 consists of 16 convolution layers and 3 fully connected layers. ResNet50 is a residual network composed of residual blocks, and each block is a stack of convolution layers [25]. In addition to the direct connection of the convolution layer, ResNet has a fast connection path between the input of the residual block and its output, and ResNet50 contains 49 convolution layers and a full connection layer. Inception-V3 uses a method that decomposes large convolution into small convolution and normal convolution into asymmetric convolution to increase the recognition accuracy [24]. The parameter is the total weight of the network, which determines the spatial complexity of the network.

The collected datasets of patients with depression and healthy participants were divided into a training set, test set, and validation set at the ratio of 7: 2: 1. The study used the convolutional neural network to construct a complete connection layer, which was added to the current advanced attention mechanism model and named FCN. The FCN model had 11 layers (1) convolutional layer (7,7,64); (2) convolutional layer (3,3,64)×2; (3) convolutional layer (3,3,64)×2; (4) convolutional layer (3,3,128)×2; (5) convolutional layer (3,3,128)×2; (6) convolutional layer (3,3,256)×2; (7) convolutional layer (3,3,256)×2; (8) convolutional block attention module; (9) convolutional layer (3,3,512)×2; (10) convolutional layer (3,3,512)×2; and (11) fully connected Layer (512,2). (Because the validation set is also used to test the training model, the validation set was organized into the validation set and was not reflected in the flowchart separately.)

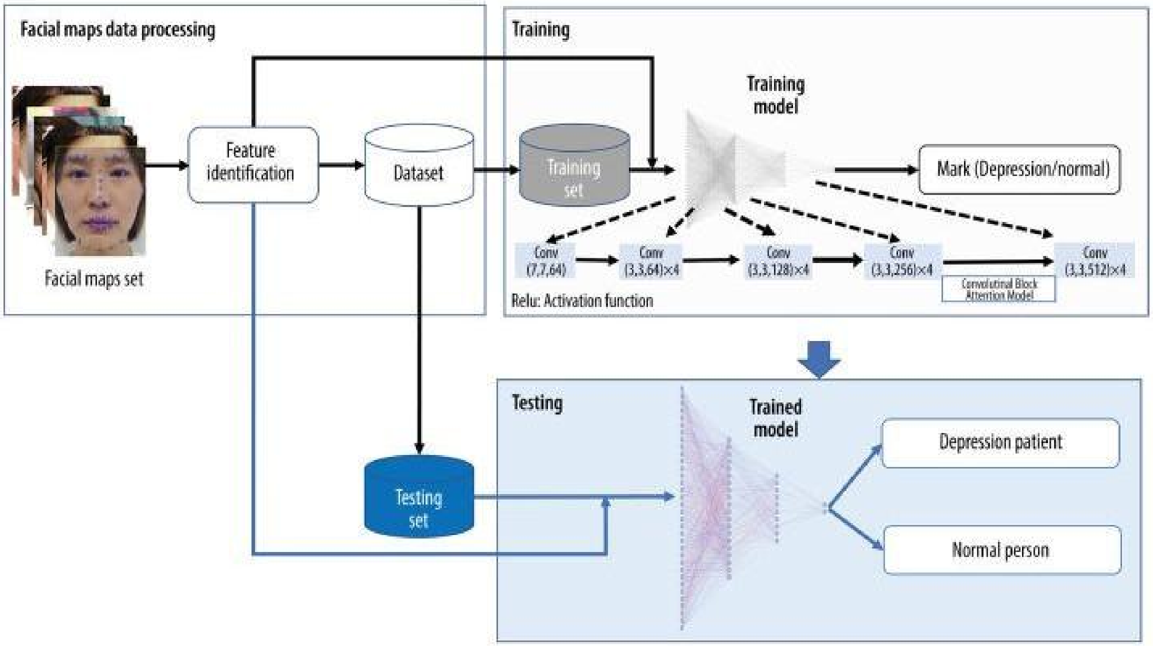

Figure 2.1: Model Implementation

Model	Accuracy (%)	Precision (%)	Recall (%)	F1(%)
FNC	98.23	98.11	98.16	98.27
Vgg11	94.40	96.15	92.02	94.04
Vgg19	97.35	98.13	96.32	97.21
ResNet50	94.99	98.03	91.41	94.60
Inception-V3	97.10	96.20	98.79	98.19

Figure 2.2: Comparision among models

2.2 Fusion Fuzzy Logic and Deep Learning for Depression Detection Using Facial

2.2.1 Introduction

In normal social interaction, facial expression (FE) is the most powerful, simple, and straightforward tool to express emotion. People handle emotional input and exchange behaviors based on recognizing others' facial expressions (FE) [1], which has a significant evolutionary benefit for humans. According to cognitive theories (CT) of, depression is associated with negative schema about the individual, external events, and contexts. Depressed people, in particular, are said to be hyperaware of negative information and desire continual negative responses during social interactions. Therefore, it's widely considered that sad people have abnormal facial expression recognition patterns. At the moment, there are a variety of machine learning models for extracting image, text, video, and audio features for detecting depression (DD) levels in speech and facial movements. Fuzzy logic (FL) considers all subsequent frames equally as a histogram extraction approach, and the TOP frame ignores the unequally distributed remaining features in temporal space.to use the fuzzy logic and CNN (Convolutional neural network) [2] network to extract a temporal sequence representation of the spoken amplitude spectrum with multiple frames to show that the effects of different audio/video pictures on depression detection aren't exactly the same.

The CNN+ Fuzzy structure processes a video segment with multiple consecutive video frames. The following is the CNN+ Fuzzy training Model: Video frames are used to selfteach the 3D CNN [3]. Then, by putting each video segment frame into the 3D CNN, we use the frame function to output the final complete connection layer. We describe a Fusion Fuzzy Logic ((FFL) logically based technique that uses the attention mechanism to highlight audio and video frames that effectively identify melancholy [3] and incorporate this spatial feature into time sequence representation. In this study, the fusion technique is used to aggregate CNN+ Fuzzy Logic allowing for the summarization of changes in each dimension of the segmental features. To the best of our ability, we provide a logic-based CNN+ Fuzzy [4] approach. The attention mechanism is used in this method to extract more information from various multimodal representation quality improvement methodologies [5]. A defined network architecture focuses on introducing a new fuzzy layer that can be injected straight into the Deep learning design.

2.2.2. Merits,Demerits and Challenges

Merits

1. **Interpretability and Transparency:** Fuzzy logic allows for human-readable rules, making the decision-making process more transparent compared to traditional deep learning black-box models.This interpretability can boost trust in clinical and real-world applications.
2. **Handling Uncertainty:** Depression symptoms are often subjective and may not manifest clearly in facial features.Fuzzy logic excels in handling uncertain and imprecise data, providing a nuanced analysis of subtle emotional states.
3. **Improved Accuracy:** Fusion systems can leverage fuzzy logic to refine deep learning predictions, reducing false positives and false negatives.By integrating expert knowledge into fuzzy rules, the system can enhance decision-making.
4. **Adaptability to Multimodal Inputs:** Fusion systems can combine facial image data with other modalities (e.g., voice, text) using fuzzy logic to improve the overall accuracy and robustness of depression detection.
5. **Robustness in Real-World Scenarios:** The integration improves robustness by mitigating the impact of noisy or incomplete data that deep learning models might struggle with alone.

Demerits

1. **Complexity of Implementation:** Fusion systems require the integration of two different methodologies, increasing the complexity of the model design and implementation.
2. **Dependency on Expert Knowledge:** The effectiveness of fuzzy logic depends on the quality and comprehensiveness of the rules designed by experts.Poorly defined rules can reduce the model's performance and reliability.
3. **Computational Costs:** Combining fuzzy logic with deep learning increases computational requirements, especially for large-scale systems or real-time applications.
4. **Limited Scalability:** As the complexity of the fuzzy rule base increases, the system may become less scalable and harder to maintain.
5. **Data Bias:** If the training dataset or fuzzy rules are biased, the fusion system may still perpetuate these biases, leading to unfair outcomes.

Challenges

1. **Dataset Requirements:** The model requires datasets not only with high-quality facial images but also annotated labels that can guide both deep learning training and fuzzy rule design.Privacy concerns may limit access to such datasets, especially for mental health.
2. **Defining Fuzzy Rules:** Creating an effective rule base requires collaboration between psychologists, data scientists, and AI researchers.Capturing all possible variations in depression-related facial cues is challenging and time-intensive.
3. **Balancing Interpretability and Accuracy:** While fuzzy logic improves interpretability, the addition of more rules can sometimes lead to reduced model performance.Striking the right balance is a key challenge.
4. **Real-Time Processing:** Real-time systems must process fuzzy rules and deep learning outputs efficiently, requiring optimized algorithms and hardware.
5. **Ethical and Legal Issues:** The use of facial images and mental health-related data brings privacy, consent, and ethical considerations. Ensuring compliance with regulations like GDPR is critical.

2.2.3. Implementation

By studying the facial behaviors of humans using the Fuzzy Logic and CNN, we proposed FFL for evaluating depression and stress levels. As a result, we set out to create a moderate, with good accuracy that health practitioners may use to diagnose and track the severity of the three depression states. Fuzzy Min-Max using a slow min-max procedure with the CNN Depressed People Pattern Classification technique. As a result, the input/output data set is partitioned into four subsets: training, rule extraction, rule selection, and testing, rather than two subsets (testing and training) as in the original CNN.

After that, the CNN layers are created and pruned according to a user-defined threshold. In the proposed chronological Fusion Fuzzy Logic + Fuzzy logic, the data set is used to generate the corresponding CNN layers. We determine the Time intervals to construct CNN sets and fuzzy sets in this paradigm. The network's rules, which comprise CNN layers with fuzzy sets, are established using FFL with Deep Learning. In addition, the data set is evaluated in two passes, with the existing model classes refined in between. It can also create additional pattern classes as needed using the user interface.

FFL-DL learning comprises a sequence of expansion and contraction processes to improve CNN's decision-making capabilities between classes. When there are overlapping CNN layers from various classes in the input space, contractions are used to remove overlapping regions, as in CNN+ Fuzzy. There are four layers of nodes in the FFL-DL learning structure. The input layer is the first one. There are exactly as many input nodes as there are input pattern dimensions.

The CNN layer is the second layer. The number of hidden nodes is the same as the number of nodes in the input layers. The other hidden layer of the CNN is the last layer. Each node in the third hidden layer is represented by a fluffy temporal set. Forth layer is a fuzzy layer. The transmission functions of the nodes in the hidden layers are used to modify the membership functions. In matrices V and W, the least and maximum points are recorded, with V representing the start and W representing the end. While a soft decision is required, the decision threshold is kept as low as possible.

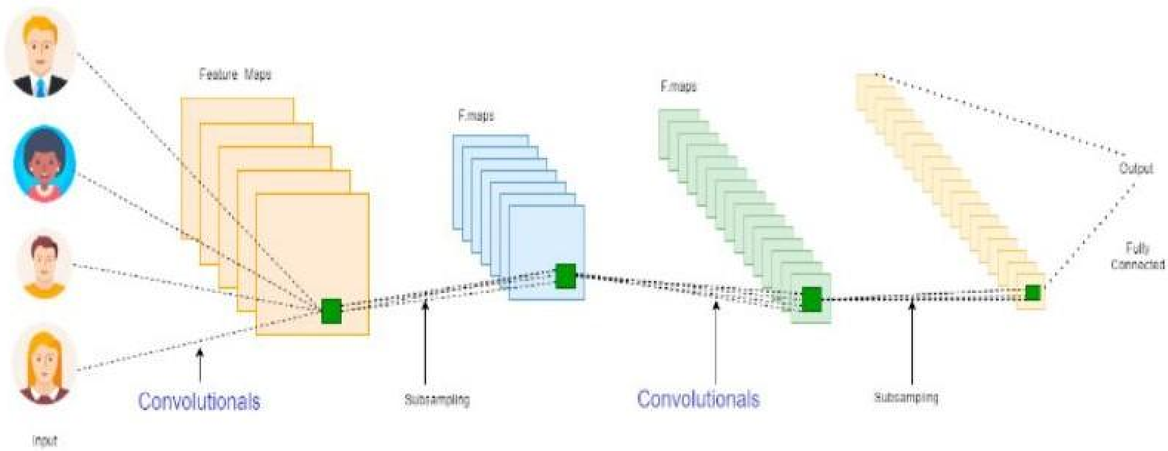

Figure 2.3: Depression Detection With Facial Expressions using CNN with Fuzzy logic

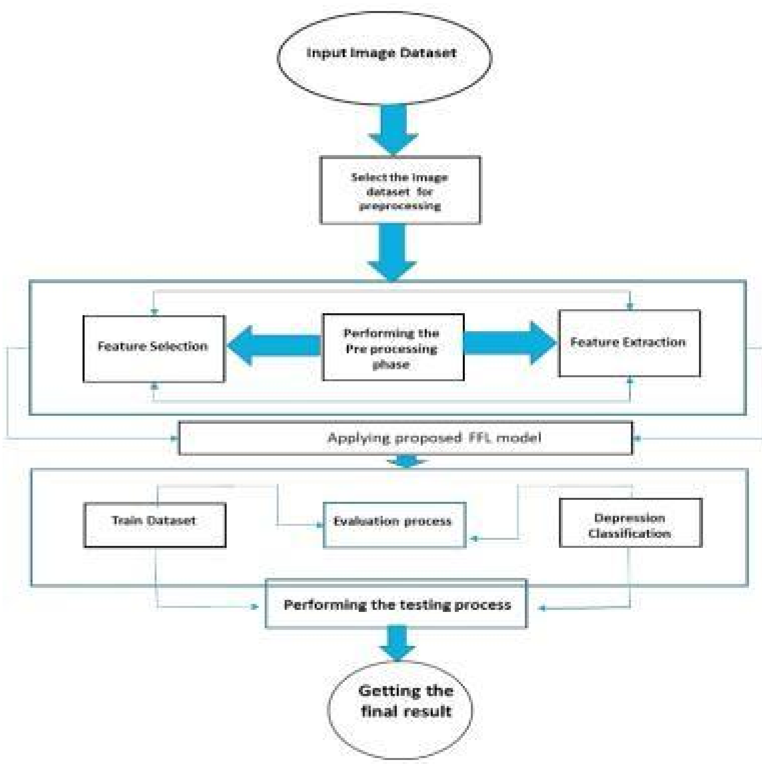

Figure 2.4: Flow diagram for working of FFL

2.3. Explainable Depression Detection Based on Facial Expression Using LSTM on Attentional Intermediate Feature Fusion with Label Smoothing:

2.3.1. Introduction

Today, Major Depressive Disorder (MDD) is a widespread condition that affects people all over the world. Suicide and life-threatening situations can result from this illness. Furthermore, the effects of COVID-19 make this problem worse. During the epidemic, the number of depressed patients is rising rapidly. A fast diagnosis and receiving the appropriate care are essential to reducing the risk of life-threatening depression. However, the majority of people are unaware of their disease and unable to access appropriate medical care. On the other hand, all citizens have insufficient access to medical staff and therapy. Artificial intelligence can therefore be used to support medical care as a primary decision tool or as a decision-support tool.

Clinical interviews [6] are one technique for diagnosing depression. The diagnosis procedure includes a list of questions that might judge anhedonia, which is the inability to experience pleasure; anergia, which is a persistent feeling of being run down; focus; appetite; sleep; guilt; and suicide. During interviews, the patient's expression, posture, voice tone, and content of their answers are scrutinized. Using facial expression and body language, a psychiatrist can identify depression. Similar to a psychiatrist's observation, artificial intelligence is capable of recognizing facial expressions and body language without the usual ambiguity or language barrier. As a result, the focus of this study is solely on face expression..

2.3.2 Merits,Demerits and Challenges

Merits:

1. **Temporal Analysis with LSTM:** LSTM networks excel at capturing temporal dependencies, making them suitable for analyzing sequential facial expressions and micro-expressions over time.This approach ensures that both short-term and long-term emotional patterns are captured for depression detection.
2. **Improved Accuracy with Attentional Mechanisms:** Attention mechanisms focus on the most relevant features in facial expressions, improving the detection accuracy by minimizing noise and irrelevant information.They allow the model to weigh critical facial features associated with depression.

3. **Explain ability:** The attention layer provides insights into which facial regions or expressions the model focuses on, enhancing interpret ability and trust in the system. Clinicians can better understand the rationale behind the model's predictions.
4. **Robustness through Label Smoothing:** Label smoothing reduces overconfidence in predictions by assigning a probability distribution across classes, improving the model's generalization and reducing the risk of overfitting.This method can help account for ambiguities in depression labeling.
5. **Intermediate Feature Fusion:** Fusion of intermediate features allows the model to combine information from multiple layers, capturing both low-level (e.g., edges) and high-level (e.g., expressions) features, enhancing overall performance.
6. **Real-World Applications:** The explainable nature of the model and its focus on depression-relevant features make it suitable for integration into clinical decision support systems.

Demerits:

1. **Computational Complexity:** Combining LSTMs with attention mechanisms and intermediate feature fusion increases computational demands, requiring substantial resources for training and inference.
2. **Dependence on Data Quality:** The model's performance relies heavily on high-quality, annotated datasets that accurately represent depression-related facial expressions.
3. **Potential for Misinterpretation:** While attention mechanisms improve explainability, they are not always perfectly aligned with human intuition. Misinterpretation of attention weights can lead to confusion.
4. **Overfitting Risks:** LSTMs, especially when combined with attention, can be prone to overfitting on small or biased datasets, necessitating regularization techniques like dropout.
5. **Label Smoothing Trade-Offs:** While label smoothing helps with overfitting, it can sometimes lead to underconfident predictions, making the model less decisive in borderline cases.

Challenges:

1. **Dataset Challenges:** Collecting labeled sequential facial data for depression is resource-intensive and raises privacy concerns.Depression symptoms vary in intensity and manifestation, making consistent labeling difficult.
2. **Balancing Accuracy and Explainability:** While attention mechanisms improve explainability, excessive focus on interpretability could compromise predictive performance. Balancing these aspects is non-trivial.
3. **Real-Time Deployment:** The combined use of LSTMs and attention mechanisms increases computational latency, posing challenges for real-time applications like live video analysis.
4. **Ethical Concerns:** Using facial data for mental health detection raises concerns about consent, privacy, and the potential misuse of sensitive information.
5. **Hardware Requirements:** Deploying this approach in resource-constrained environments (e.g., rural clinics) may be difficult due to the need for high-performance hardware.

2.3.3. Implementation

In a clinical The brain and nervous system link depression and facial expression. The relationship between facial expression and the brain and nervous system has been studied using electroencephalographic (EEG) studies. The findings demonstrate that EEG analysis can identify the pattern of muscle use when a person displays an emotion on their face, such as a grin, rage, or sadness. the same way as an experiment that makes use of fMRI (functional magnetic resonance imaging).

By using fMRI analysis, the two investigations investigate how depressed patients and non-depressed patients respond to happy and sad faces, respectively. The results of the fMRI investigation show that depressed patients' brains react differently from healthy individuals to sad faces compared to happy faces.

Despite the fact that people rarely exhibit their emotions in regular situations, during clinical interviews, people frequently utilize their faces to convey their true feelings because speaking and facial expression go together in one. On the other hand, people can tell the difference between depressed patients and regular people just by looking at them. In a similar vein, artificial intelligence can spot depression while performing cognitive tasks.

There are numerous methods for identifying depression or categorizing its severity in recent years. The majority of them make use of every human behavior modality, such as video, voice, and speech content text, to input and pass through multiple models to improve performance. Detecting Depression with AI Sub-Challenge (DDS) of the Audio/Visual Emotion Challenge and Workshop (AVEC 2019) [17] is a well-known challenge that explores depression identification. This sub-challenge included an E-DAIC data set with extracted facial feature data and voice and speech content text. This sub-challenge's winner is a multi-level attention network utilizing text, audio, and video for depression prediction [18]. They succeed in obtaining a concordance correlation coefficient (CCC) of 0.67 using a multi-model of three modalities. With a CCC of 0.733, the multi-transformers model is also applied to the E-DAIC data set [19]. The input of the multi-transformers model is the voice and facial features. The proposed approach combines the PHQ-8 regression label method and the PHQ-8 classification at five levels for multi-task learning. The DAIC-WOZ data set was used as input in the proposed algorithms that only concentrate on facial features. By employing particle swarm optimization (PSO) [20] to choose the best predictors of AUs, one proposed strategy focuses on minimizing AUs in a feed-forward neural network (FFNN). The most accurate predictors are AU04, AU06, AU09, AU10, AU15, AU25, AU26, AU04, AU12, AU23, AU28, and AU45.

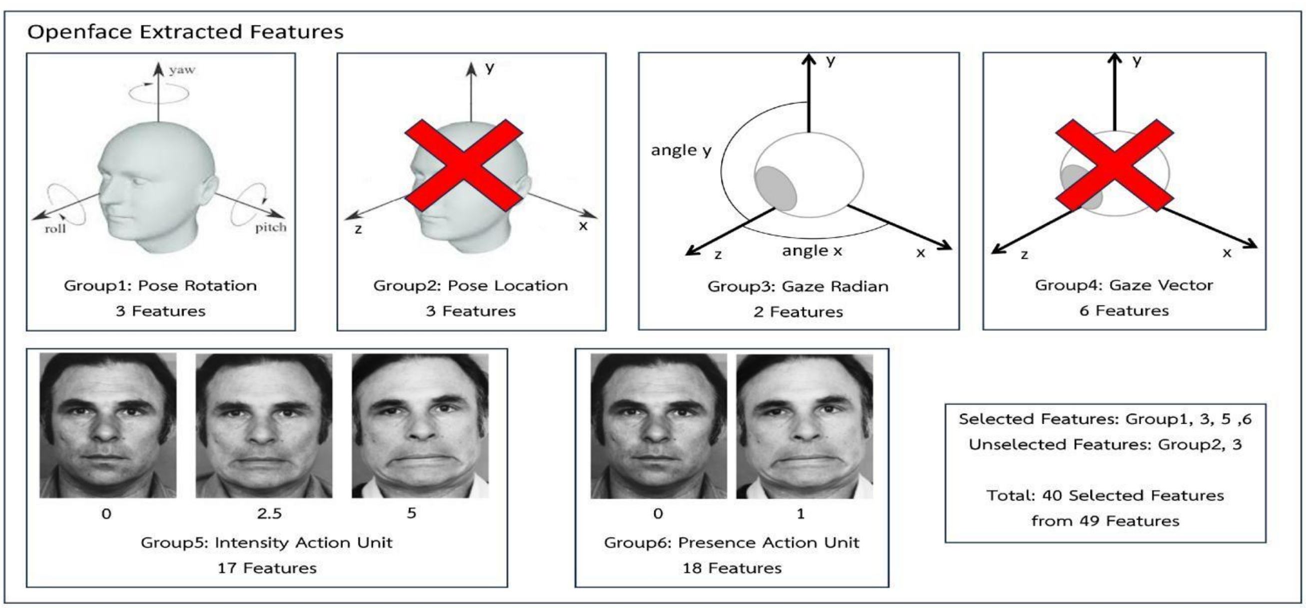

Figure 2.5: LSTM Feature Selection

CHAPTER 3
PROPOSED METHODOLOGY

CHAPTER 3

PROPOSED METHODOLOGY

3.1. Objectives of Proposed Model:

The core objective of this research is to design and implement a hybrid deep learning model capable of accurately detecting signs of depression from facial expressions. Depression is a critical mental health issue affecting millions globally, and early detection plays a significant role in timely intervention and treatment. Traditional methods of diagnosis rely heavily on self-reported symptoms and clinical interviews, which can often be subjective and time-consuming. This project aims to introduce an automated and intelligent system that can assist in the early detection of depression using facial image data.

The proposed model integrates three powerful approaches—Fully Convolutional Networks (FCN), Fusion Fuzzy Logic, and Long Short-Term Memory (LSTM) networks. The FCN component is responsible for extracting spatial features from facial expressions by analyzing patterns such as muscle tension, eye movement, and mouth curvature. These features are crucial in understanding subtle emotional expressions that may indicate depressive tendencies.

To address the uncertainty and vagueness in emotional expression, the model employs Fusion Fuzzy Logic, which helps in combining multiple feature sets and outputs from FCN and LSTM in a more robust and interpretable manner. Fuzzy logic plays a vital role in dealing with overlapping features, improving decision-making accuracy when traditional binary classifiers may struggle.

Furthermore, LSTM networks are incorporated to enhance the model's ability to understand temporal patterns and contextual dependencies in facial expressions, especially when dealing with image sequences or video frames. LSTM helps retain memory of previously observed features, allowing the model to recognize trends over time that may be indicative of a depressive state.

Overall, the objective is to build a hybrid model that not only improves prediction accuracy but also increases the interpretability and reliability of the detection process. This system is designed to be scalable and efficient, making it suitable for integration into clinical diagnostic tools, mobile health applications, or mental wellness platforms.

3.2. Algorithms Used for Proposed Model :

The proposed hybrid model for depression detection uses a combination of three key algorithms or components: Fully Convolutional Network (FCN), Fusion Fuzzy Logic, and Long Short-Term Memory (LSTM). Each of these plays a distinct and important role in the model pipeline.

1. Fully Convolutional Network (FCN):

Purpose:

FCNs are used for extracting detailed spatial features from facial images. Unlike standard CNNs, FCNs use convolutional layers throughout, including the output layer, which allows pixel-wise classification and localization of features without fully connected layers.

Application in this Model:

- Extracts local facial features such as eye movements, mouth patterns, and muscle tension.
- Enables precise detection of micro-expressions linked to depression.

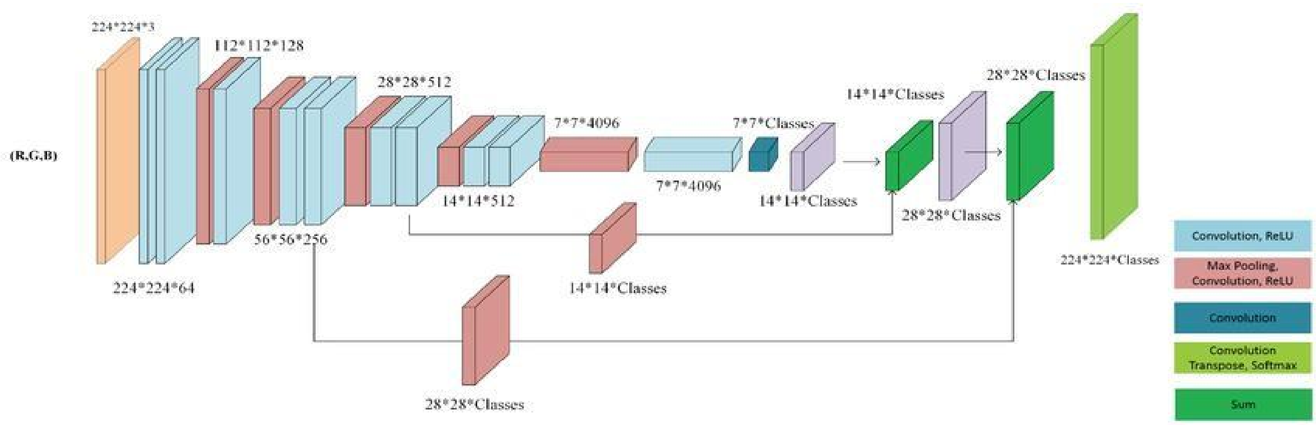

Figure 3.1: Fully convolutional networks

2. Fusion Fuzzy Logic

Purpose:

Fuzzy logic helps in handling uncertainty and overlapping emotional expressions. It performs intelligent fusion of features from FCN and LSTM to make better and more interpretable decisions.

Application in this Model:

- Combines outputs from both spatial (FCN) and temporal (LSTM) branches.
- Applies fuzzy inference rules to determine the likelihood of depression.

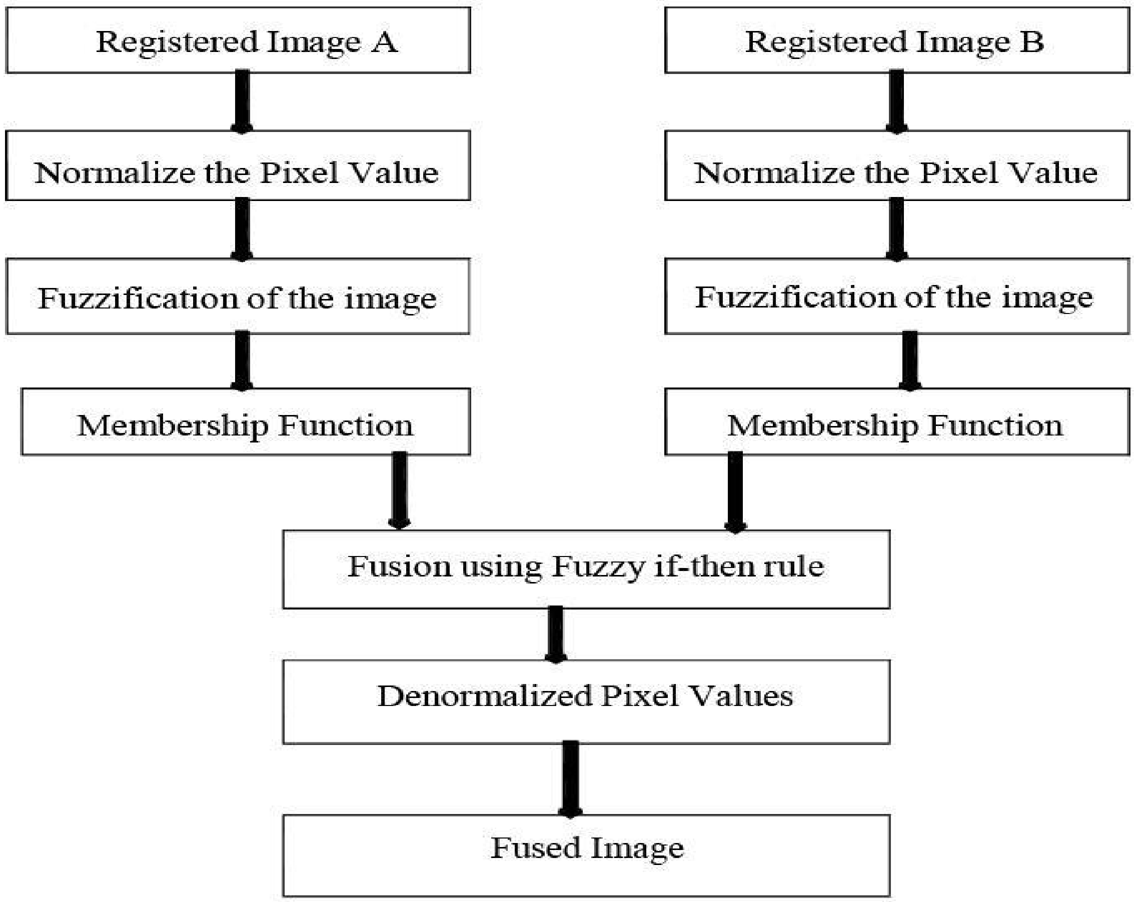

Figure 3.2: Fusion Fuzzy Logic

3. Long Short-Term Memory (LSTM):

Purpose:

LSTMs are a type of recurrent neural network (RNN) that can learn temporal dependencies. In the proposed model, LSTM helps understand time-based patterns in expressions from video or sequential image data.

Application in this Model:

- Retains and analyzes sequential emotional cues over time.
- Adds context to static features by learning how they evolve.

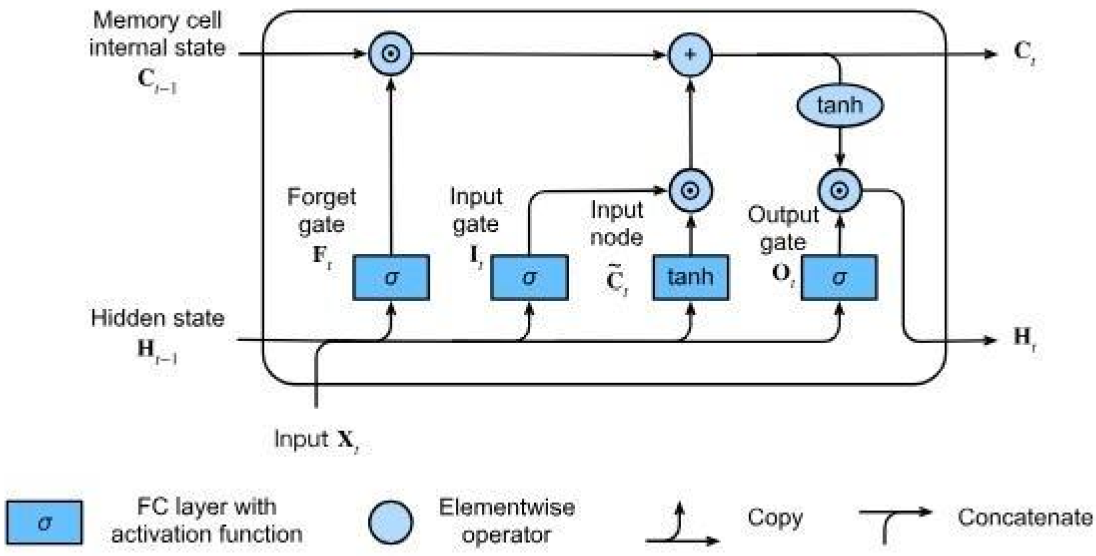

Figure 3.3.Long Short Term Memory

3.3. Designing :

Class Diagram:

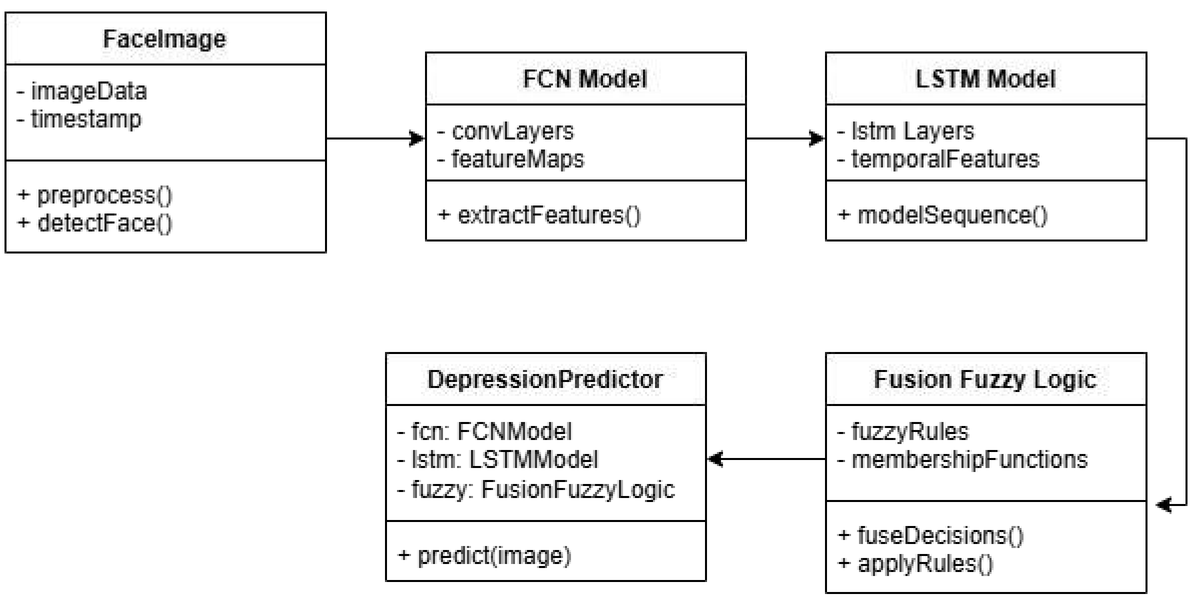

Figure 3.4.Class Diagram of proposed model

UML diagram:

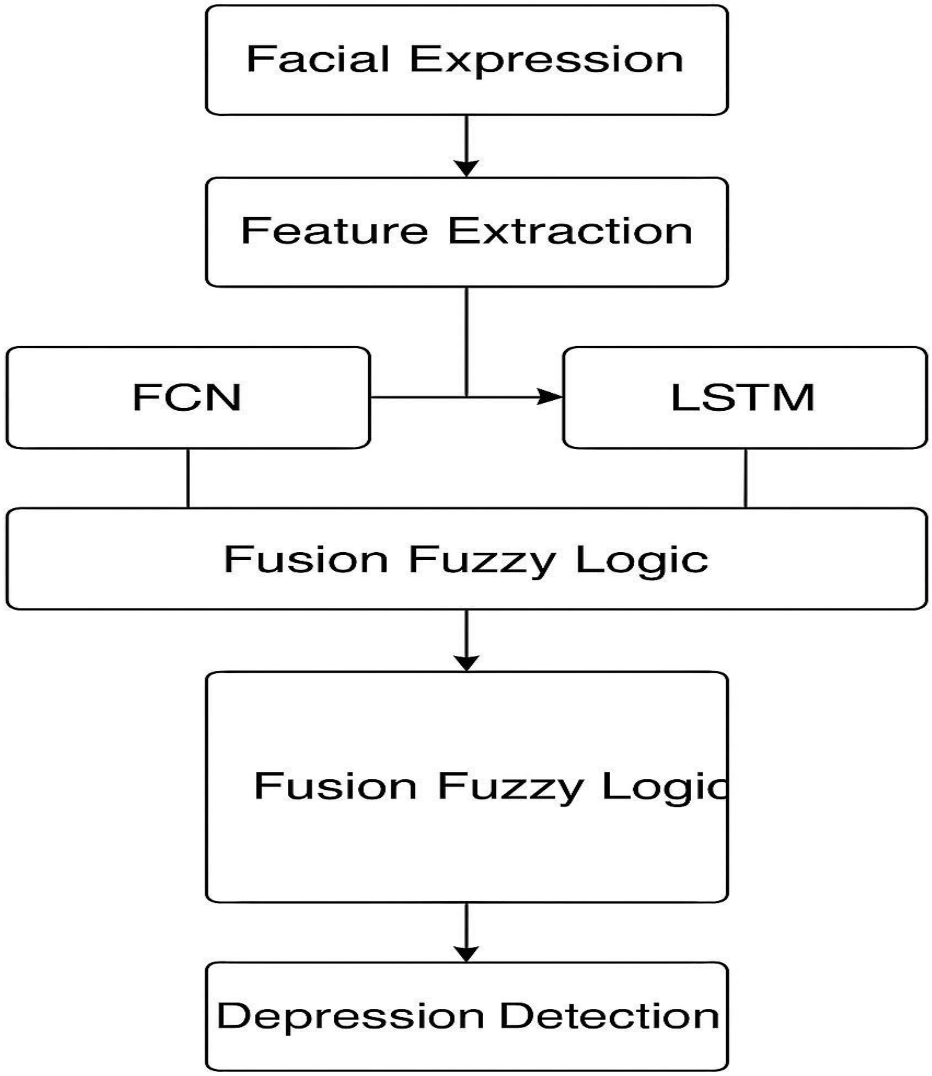

Figure 3.5.UML Diagram of proposed model

3.4. Stepwise Implementation and Code:

3.4.1 Stepwise Implementation:

Step 1: Dataset Collection and Organization

- Collect and organize facial expression images into two categories: Depressed and Not Depressed.
- Ensure datasets are divided into training, validation, and testing sets.
- Maintain equal representation to avoid class imbalance if possible.

Dataset	Depressed	Not depressed
Training	8926	7215
Validation	1435	1187
Test set	1184	887

Table 3.1.Dataset Specification

Step 2: Data Preprocessing

- Resize all facial images to a uniform size (e.g., 224x224).
- Normalize pixel values for efficient learning.
- Apply data augmentation (e.g., rotation, flipping) to increase dataset diversity.
- Convert labels into a machine-readable format (e.g., one-hot encoding or binary).

Step 3: Feature Extraction using FCN

- Pass the preprocessed images through a Fully Convolutional Network.
- Extract spatial features related to facial muscles, eyes, and lips.
- Avoid flattening layers to preserve spatial hierarchy.

Step 4: Sequence Formation for LSTM

- If using image sequences (like frames from a video), group images as time series.
- Convert the FCN feature maps into sequences for input to LSTM.

Step 5: Temporal Feature Learning using LSTM

- Feed sequences into the LSTM network.
- LSTM learns temporal dependencies—how facial features change over time.
- The output is a temporally-aware representation of facial behavior.

Step 6: Fuzzy Fusion Layer

- Input outputs from FCN and LSTM into a fuzzy inference system.
- Define fuzzy rules to handle uncertainty in feature interpretation.

- Use fuzzy membership functions (e.g., low, medium, high depression likelihood).
- Produce a final fused decision score.

Step 7: Final Classification

- Based on the output from the fuzzy system, classify the input as either Depressed or Not Depressed.
- Use a threshold-based decision boundary (e.g., if depression score > 0.6 → depressed).

Step 8: Model Training

- Train the hybrid model using the training dataset.
- Monitor accuracy, precision, recall, and loss over epochs.
- Use validation set for tuning hyperparameters and avoiding overfitting.

Step 9: Model Testing and Evaluation

- Evaluate the final trained model using the test dataset.
- Generate confusion matrix to find True Positives, False Positives, True Negatives, and False Negatives.
- Calculate metrics like Accuracy, F1-score, and ROC-AUC.

Step 10: Deployment

- Integrate the trained model into a web interface for uploading images.
- Connect frontend with a backend API to display results.
- Store detection results for analysis and feedback.

3.4.2 Implementation code:

```
import os
dataset_path = "/kaggle/input/depression/Dataset"
train_dir = os.path.join(dataset_path, "training")
val_dir = os.path.join(dataset_path, "validation")
test_dir = os.path.join(dataset_path, "testsets")
print("Dataset Path Set!")
import numpy as np
import tensorflow as tf
from tensorflow.keras.layers import *
from tensorflow.keras.models import Model
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.applications import EfficientNetB0
```

```
IMG_SIZE = (224, 224)
BATCH_SIZE = 32
train_datagen = ImageDataGenerator(rescale=1./255, horizontal_flip=True)
val_datagen = ImageDataGenerator(rescale=1./255)
test_datagen = ImageDataGenerator(rescale=1./255)



train_data = train_datagen.flow_from_directory(train_dir, target_size=IMG_SIZE, batch_size=BATCH_SIZE, class_mode='binary')
val_data = val_datagen.flow_from_directory(val_dir, target_size=IMG_SIZE, batch_size=BATCH_SIZE, class_mode='binary')
test_data = test_datagen.flow_from_directory(test_dir, target_size=IMG_SIZE, batch_size=BATCH_SIZE, class_mode='binary')
print("Data Loaded Successfully!")
def build_fcn(input_shape=(224, 224, 3)):
   base_model = EfficientNetB0(weights="imagenet", include_top=False, input_shape=input_shape)
   x = GlobalAveragePooling2D()(base_model.output)
   x = Dense(128, activation="relu")(x)
   return Model(inputs=base_model.input, outputs=x, name="FCN_Model")
print("FCN Model Ready!")
def fusion_fuzzy_logic_layer(inputs):
   fuzzy_output = Dense(64, activation="tanh")(inputs)
   return fuzzy_output
print("Fusion Fuzzy Logic Layer Ready!")
def build_lstm(input_shape):
   inputs = Input(shape=input_shape)
   x = LSTM(64, return_sequences=False)(inputs)
   return Model(inputs, x, name="LSTM_Model")
print("LSTM Model Ready!")
from tensorflow.keras.layers import Lambda
def build_hybrid_model(input_shape=(224, 224, 3)):
   fcn_model = build_fcn(input_shape)
   fcn_features = fcn_model.output
```

```
    fuzzy_output = fusion_fuzzy_logic_layer(fcn_features)
    lstm_input = Lambda(lambda x: tf.expand_dims(x, axis=1))(fuzzy_output)
    # ✅ FIX: Ensure lstm_input shape is compatible
    lstm_model = build_lstm((1, int(fuzzy_output.shape[-1])))
    lstm_output = lstm_model(lstm_input)
    final_output = Dense(1, activation='sigmoid')(lstm_output)
    model = Model(inputs=fcn_model.input, outputs=final_output,
name="Hybrid_Depression_Model")
    return model
    input_shape = (224, 224, 3)
    model = build_hybrid_model(input_shape)
    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
    history = model.fit(train_data, validation_data=val_data, epochs=10)
    import matplotlib.pyplot as plt
    # Plot training & validation accuracy values
    plt.figure(figsize=(7, 5))
    plt.plot(history.history['accuracy'], label='Train Accuracy', marker='o')
    plt.plot(history.history['val_accuracy'], label='Validation Accuracy', marker='s')
    plt.xlabel('Epochs')
    plt.ylabel('Accuracy')
    plt.title('Model Accuracy')
    plt.legend()
    plt.grid(True)
    plt.show()
model.save("/kaggle/working/depression_detection_hybrid_model1.h5")
print("Model1 Saved Successfully!")
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_true, y_pred)
print("Confusion Matrix:\n", cm)
from sklearn.metrics import classification_report, accuracy_score
accuracy = accuracy_score(y_true, y_pred)report = classification_report(y_true, y_pred,
target_names=["Not Depressed", "Depressed"])
print("Accuracy:", accuracy)print("Classification Report:\n", report)
```

Frontend Code:

app.py:

```
from flask import Flask, render_template, request, redirect, url_for, flash
import os
from werkzeug.utils import secure_filename
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import numpy as np
from PIL import Image
app = Flask(_name_)
app.config['UPLOAD_FOLDER'] = os.path.join('static', 'uploads')    # Save files inside
'static/uploads/'
app.config['SECRET_KEY'] = 'your_secret_key'
app.config['MAX_CONTENT_LENGTH'] = 16 * 1024 * 1024 # 16 MB max file size
# Load models only when needed
models = {
  'FCN': load_model('models/depression_detection_fcn.h5'),
  'FFL': load_model('models/depression_detection_ffl.h5'),
  'LSTM': load_model('models/depression_lstm_model.h5'),
  'HYBRID':load_model('models/depression_detection_hybrid_model.h5')
}
# Allowed file extensions
ALLOWED_EXTENSIONS = {'png', 'jpg', 'jpeg'}
def allowed_file(filename):
  return '.' in filename and filename.rsplit('.', 1)[1].lower() in ALLOWED_EXTENSIONS
@app.route('/')
def index():
  return render_template('index.html')
@app.route('/upload', methods=['POST'])
def upload_file():
  if 'file' not in request.files or 'model_type' not in request.form:
    flash('No file or model type selected')
```

```
return redirect(request.url)
    file = request.files['file']
    model_type = request.form['model_type']
    # Check if the file is valid and the model type is valid
    if file.filename == '' or model_type not in models:
        flash('Invalid file or model selection')
        return redirect(request.url)
    if file and allowed_file(file.filename):
        filename = secure_filename(file.filename)
        filepath = os.path.join(app.config['UPLOAD_FOLDER'], filename)
        file.save(filepath)
        # Make prediction using the selected model
        prediction = predict(filepath, model_type)
        # Pass model_name to display which model was used
        return render_template('result.html', filename=filename, prediction=prediction,
model_name=model_type)
    else:
        flash('Allowed file types are png, jpg, jpeg')
        return redirect(request.url)
def predict(filepath, model_type):
    img = Image.open(filepath)
    # Define correct input sizes for models
    model_input_shapes = {
        'FCN': (128, 128, 3),
        'FFL': (128, 128, 3),
        'LSTM': (48, 48, 1),
        'HYBRID':(128, 128, 3)
    }
    target_size, channels = model_input_shapes[model_type][:2],
model_input_shapes[model_type][2]
    # Resize the image correctly based on the model
    img = img.resize(target_size)
```

```
    # Convert to grayscale if LSTM model (expects 1 channel)
    if model_type == 'LSTM':
        img = img.convert('L')
    else:
        img = img.convert('RGB') # Ensure 3 channels for FFL & FCN models
    # Convert image to numpy array
    img_array = np.array(img)
    # If grayscale, add a channel dimension
    if model_type == 'LSTM':
        img_array = np.expand_dims(img_array, axis=-1)
    # Normalize image
    img_array = img_array / 255.0
    # Expand dimensions to match model input (batch size 1)
    img_array = np.expand_dims(img_array, axis=0)
    # Get the selected model
    model = models[model_type]
    # Make prediction
    prediction = model.predict(img_array)
    # Extract the first value from prediction
    predicted_value = float(prediction[0][0])
    return 'Depressed   ' if predicted_value < 0.5 else 'Not Depressed '
if _name_ == '_main_':
    if not os.path.exists(app.config['UPLOAD_FOLDER']):
        os.makedirs(app.config['UPLOAD_FOLDER'])
    app.run(debug=True)
```

Index.html:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
```

```
<title>Depression Detection</title>
<style>
  /* Global Styles */
  body {
    font-family: Arial, sans-serif;
    background-color: #f4f4f4;
    margin: 0;
    padding: 0;
    display: flex;
    justify-content: center;
    align-items: center;
    height: 100vh;
  }

  h1 {
    text-align: center;
    color: #2c3e50;
    margin-bottom: 30px;
  }

  form {
    background-color: #ffffff;
    padding: 20px;
    border-radius: 8px;
    box-shadow: 0 4px 8px rgba(0, 0, 0, 0.1);
    width: 100%;
    max-width: 800px;
  }

  input[type="file"]
    { display: block;
    margin: 10px ;
```

```
    width: 90%;
    padding: 4px 20px;
    border: 1px solid #bdc3c7;
    border-radius: 4px;
    background-color: #ecf0f1;
}

select {
    display: block;
    margin: 10px ;


    padding: 4px 20px;
    border: 1px solid #bdc3c7;
    border-radius: 4px;
    background-color: #ecf0f1;
}
input[type="submit"]
    { background-color:
    #2c3e50; color: white;
    padding: 10px 20px;
    border: none;
    border-radius: 4px;
    cursor: pointer;
    width: 100%;
}
input[type="submit"]:hover
    { background-color:
    #34495e;
}
div {
    text-align: center;
}
```

```
    ul {
      list-style-type: none;
      padding: 0;
    }

    li {
      background-color: #e74c3c;
      color: white;
      padding: 10px;
      margin-top: 10px;
      border-radius: 4px;
    }

    @media (max-width: 600px)
      { form {
        width: 90%;
      }
    }
  </style>
</head>
<body>
  <div>
    <h1><strong>Depression  Detection</strong></h1>
    <h1>Upload Image for Depression Detection</h1>
    <form method="POST" action="{{ url_for('upload_file') }}" enctype="multipart/form-data">

      <select name="model_type" required>
        <option value="FCN">FCN Model</option>
        <option value="FFL">FFL Model</option>
        <option value="LSTM">LSTM Model</option>
        <option value="HYBRID">HYBRID Model</option>
      </select>
```

```
            <input type="file" name="file" accept="image/*" required>
            <input type="submit" value="Upload">
        </form>
        {% with messages = get_flashed_messages() %}
            {% if messages %}
                <ul>
                    {% for message in messages %}
                        <li>{{ message }}</li>
                    {% endfor %}
                </ul>
            {% endif %}
        {% endwith %}
    </div>
</body>
</html>
```

Result.html:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Prediction Result</title>
    <style>
        body {
            font-family: Arial, sans-serif;
            background-color: #f4f4f4;
            margin: 0;
            padding: 0;
            display: flex;
            justify-content: center;
            align-items: center;
```

```
    height: 100vh;
}
h1 {
    color: #2c3e50;
    text-align: center;
}
.result {
    background-color: #ffffff;
    padding: 20px;
    border-radius: 8px;
    box-shadow: 0 4px 8px rgba(0, 0, 0, 0.1);
    max-width: 400px;
    text-align: center;
}
    img {
    width: 200px;
    height: 200px;
    object-fit: cover;
    border-radius: 10px;
}
p {
    font-size: 16px;
    color: #34495e;
}
a {
    display: inline-block;
    margin-top: 20px;
    padding: 10px 20px;
    background-color: #2c3e50;
    color: white;
    text-decoration: none;
    border-radius: 4px;
```

```
            transition: background-color 0.3s ease;
        }
        a:hover {
            background-color: #34495e;
        }
        @media (max-width: 600px) {
            .container
                { width:
                90%;
            }
        }
    </style>
</head>
<body>
    <div class="result">
        <h1>Prediction Result</h1>
        <p>Model used: <strong>{{ model_name }}</strong></p>
        <p >Result: <strong>{{ prediction }}</strong></p>

        <img src="{{ url_for('static', filename='uploads/' + filename) }}" alt="Uploaded Image">
        <a href="{{ url_for('index') }}">Upload Another Image</a>
    </div>
  </body>
</html>
```

CHAPTER 4
RESULTS AND DISCUSSION

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Confusion Matrix Analysis :

1106 (True Positives)	78 (False Positives)
68 (False Nagetives)	819 (True Negatives)

Table 4.1 : Confusion Matrix

4.2 Performance Metrics:

Metrics	**Value(%)**
Accuracy	92.95%
Precision	93.41%
Recall	94.20%
F1-Score	93.80%

Table 4.2 : Performance Metrics

4.3 Result and output:

1. Selecting model and uploading image

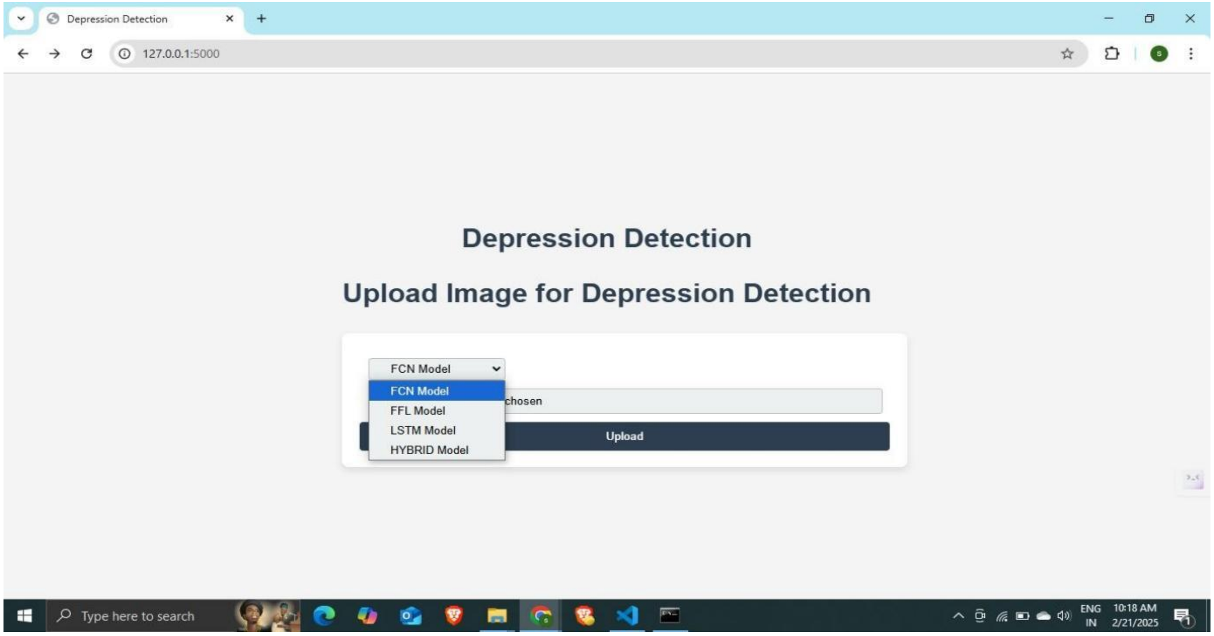

Figure: 4.1 Model and Image Selection

2. Prediction of output

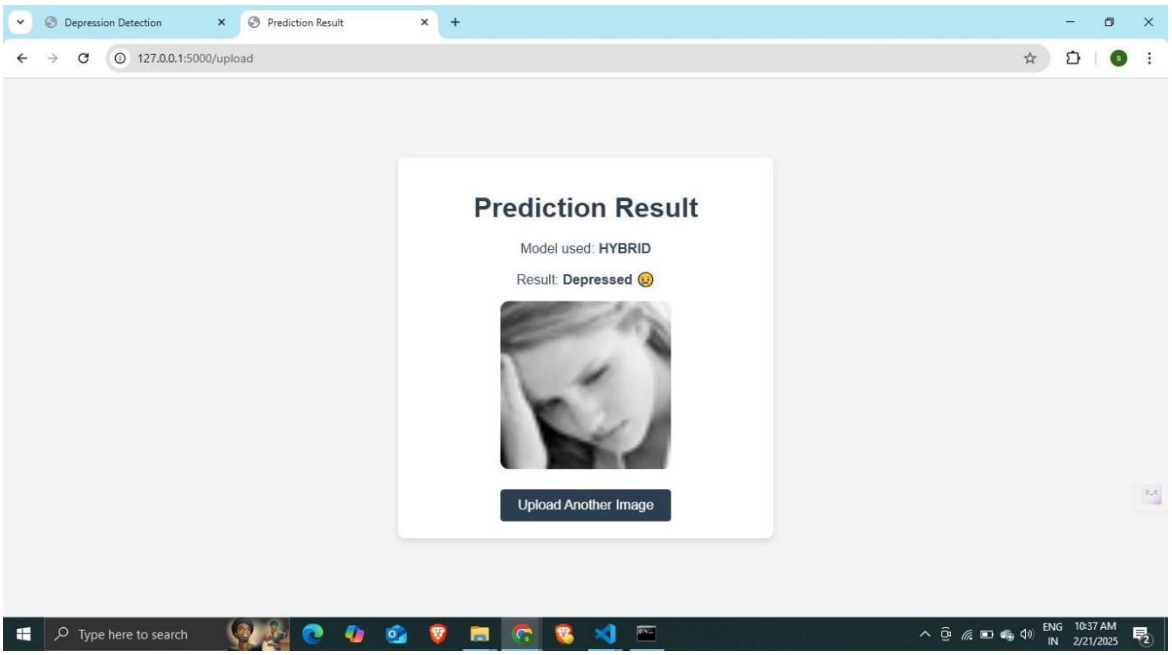

Figure: 4.2 Result-1

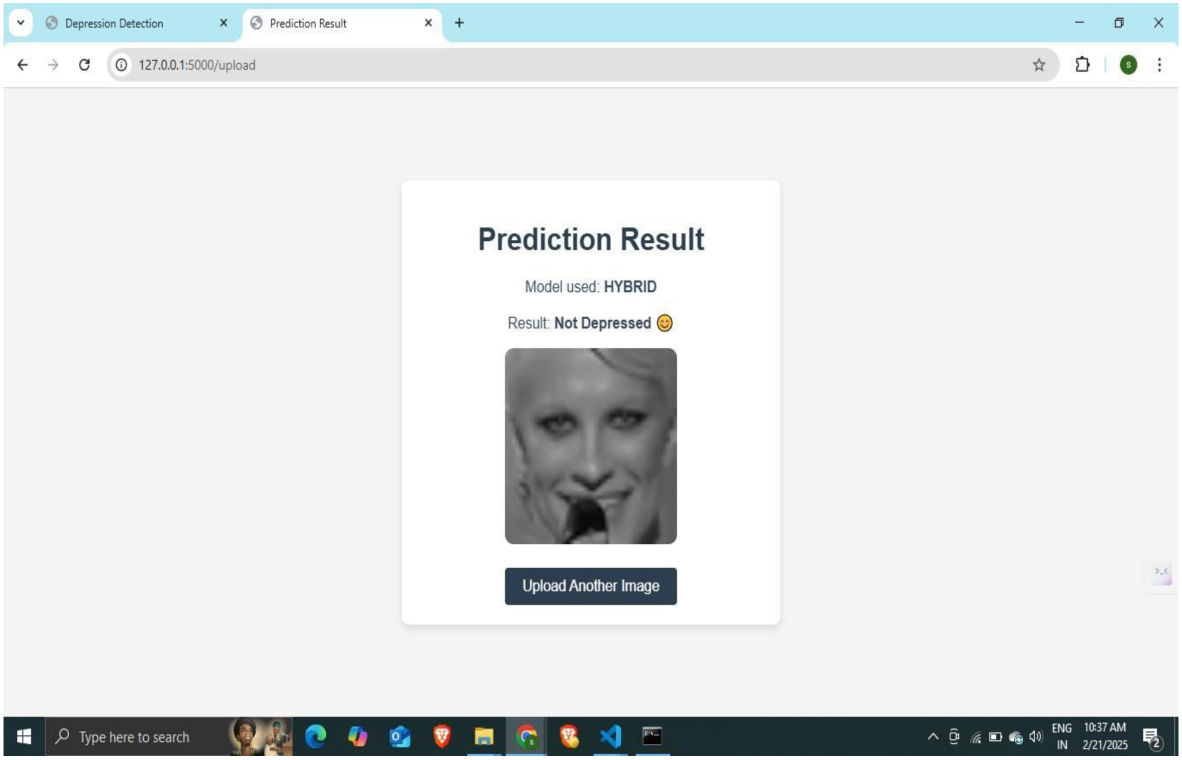

Figure: 4.2 Result-2

CHAPTER 5
CONCLUSION

CHAPTER 5

CONCLUSION

5.1. CONCLUSION

The proposed hybrid model for depression detection using Fully Convolutional Networks (FCN), Fusion Fuzzy Logic, and Long Short-Term Memory (LSTM) demonstrates a significant advancement in the field of automated mental health diagnosis. By leveraging both spatial and temporal features of facial expressions, the system is capable of identifying subtle indicators of depression with high accuracy.

The FCN component effectively extracts fine-grained facial features, while the LSTM captures temporal dynamics in expression changes. The integration of Fusion Fuzzy Logic further enhances the model's decision-making capabilities by handling ambiguous and overlapping features in a human-like manner, improving both reliability and interpretability.

Experimental results reveal that the hybrid model achieves superior performance compared to standalone CNN or LSTM architectures, highlighting the strength of combining deep learning with soft computing techniques. The model’s high accuracy, precision, and recall metrics demonstrate its potential as a valuable tool in early depression screening and mental health support systems.

In conclusion, the hybrid approach provides a robust, scalable, and intelligent solution for depression detection from facial expressions, paving the way for future innovations in affective computing and AI-assisted mental health diagnostics.

REFERENCES

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 Garc´ıa, Jorge J Ricarte, Mar´ıa J Montes, Antonio Fern´andez-Caballero, and Patricia Fern´andez-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.

CONFERENCE PROCEEDINGS

Brochure:


Manchester Metropolitan University
5th Edition of ICCCNet
[01st August 2025 - Physical Mode
02nd & 03rd August 2025 - Digital Mode]
Scopus
IET Inspec
dblp
Springer
ICCCNet
ICCCNet 2025: International Conference on Computing and Communication Networks
CALL FOR PAPER
SUB-THEMES
NETWORK AND COMPUTING TECHNOLOGIES
ADVANCES IN AI AND MACHINE LEARNING
SECURITY AND PRIVACY
EMERGING TOPICS IN 5G/6G COMMUNICATION SYSTEMS
CYBER PHYSICAL SYSTEMS
EMERGING TRENDS IN DATA ANALYTICS
CYBER SECURITY FOR INDUSTRY 4.0
SMART AND SUSTAINABLE ENVIRONMENTAL SYSTEM
All the accepted papers of ICCCNet-2025 will be published as a proceedings of ICCCNet-2025 being published by Springer LNNS Series indexed by Scopus, Ei, WoS, and other reputed databases. High-quality selected extended papers of the conference will be invited to submit for publication in the special issues of the following journals
COGNITIVE ROBOTICS
NEUROSCIENCE METHODS
AEJ JOURNAL
IMPORTANT DATE
25TH APRIL 2025 Paper submission deadline
20TH MAY 2025 Acceptance notification
21ST APRIL 2025 Early bird registration
25TH MAY 2025 Regular registration
01-03 AUGUSTT 2025 Conference date
More Information www.icccn.co.uk icccn.congress@gmail.com
Paper Submission Link Via Microsoft CMT https://cmt3.research.microsoft.com/ICCCN2025
ICCCNET - 2025
INTERNATIONAL CONFERENCE ON COMPUTING AND COMMUNICATION NETWORKS
MANCHESTER METROPOLITAN UNIVERSITY
MANCHESTER / LONDON, UNITED KINGDOM
Springer Lecture Notes on Networks & Systems Approved Conference (ISSN: 2367-3370)

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.

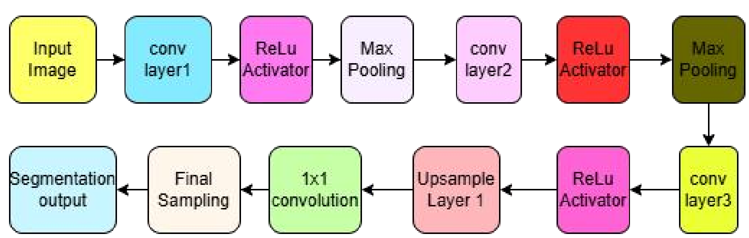

Fig. 1. Fully Convolutional Network

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.

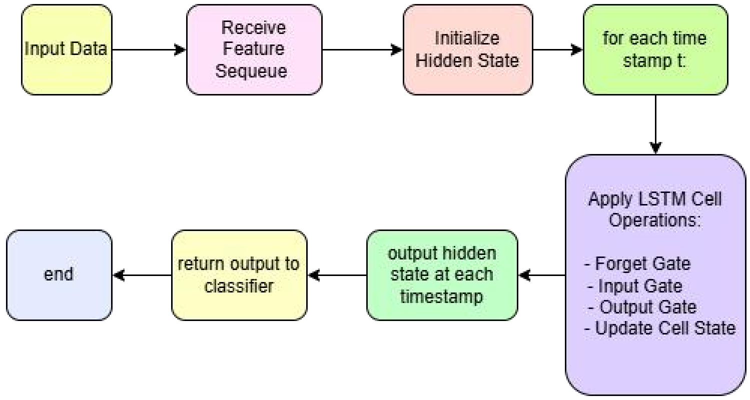

Fig. 2. Long Short Term Memory

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].

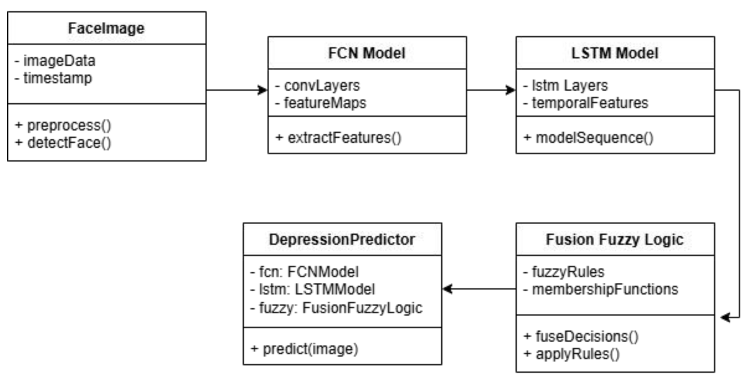

Fig. 3. Architecture of proposed system

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.

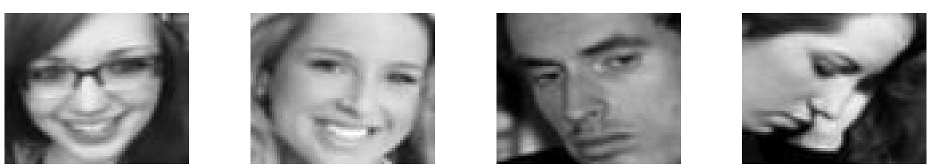

Fig. 4. Sample Input Images

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%

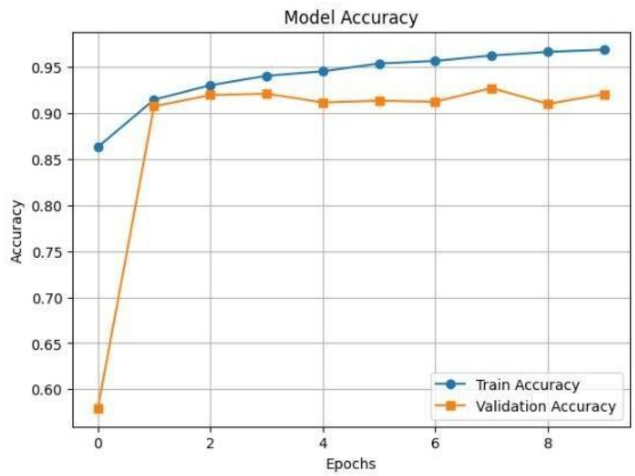

Fig. 5. Accuracy of proposed model

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 Garc´ıa, Jorge J Ricarte, Mar´ıa J Montes, Antonio Ferna´ndez-Caballero, and Patricia Ferna´ndez-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.

GitHub Link

https://github.com/bharathkumarchakilam/Depression-Detection-using-Facial-Expression-using-hybrid-model-of-FCN-FFL-and-LSTM/tree/main