Mental Health Assessment

A Project Report

submitted in partial fulfillment of the requirements

of

AIML Fundamentals with Cloud Computing and Gen AI

MANGAYARKARASI COLLEGE OF ENGINEERING

MADURAI

by

MADHESH R (923821114015) - rmmadhesh2711@gmail.com

AU923821114015

Under the Guidance of

Name of Guide

P.Raja, Master Trainer

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, **P.Raja** and **S.Mohanraja**, I want toexpress my heartfelt gratitude to you for being such an amazing mentor and guide. Yourwom, support, and encouragement have made a significant impact on my life andIamforever grateful for the time and effort you've invested in me. His advice, encouragementand the critics are a source of innovative ideas, inspiration and causes behindthesuccessful completion of this project. Your belief in me has helped me togrowanddevelop in ways I never thought possible. Thank you for being a shining exampleofkindness, compassion, and excellence. The confidence shown in me by himwasthebiggest source of inspiration for me. It has been a privilege working with himfor thelastone year. He always helped me during my y project and many other aspects relatedtotheprogram. His talks and lessons not only help in project work and other activities oftheprogram but also make me a good and responsible professional. Thank you.

ABSTRACT

Facial expression recognition (FER) is emerging as a valuable tool in mental health assessment, providing an innovative approach to understanding and monitoring patients' emotional states. This technology utilizes advanced algorithms to analyze facial features and interpret emotions, offering insights that can enhance traditional mental health evaluations. By systematically assessing emotional expressions, FER facilitates the early detection of mood disorders such as depression and anxiety, which are often characterized by subtle changes in non-verbal behavior.

The capability to identify these emotional cues in real time allows for ongoing monitoring of patients, enabling healthcare professionals to track changes in emotional states over time. This is particularly beneficial in telehealth settings, where facial recognition can augment virtual consultations by providing additional context to verbal interactions. Moreover, insights derived from FER can inform therapeutic interventions, allowing clinicians to tailor treatment approaches based on observed emotional responses.

However, the implementation of facial expression recognition in mental health raises important ethical considerations, including privacy, consent, and the potential for misinterpretation of emotions. Addressing these concerns is essential to ensure the responsible use of this technology in clinical settings.

In summary, facial expression recognition holds significant promise for enhancing mental health assessments by enabling more nuanced evaluations of emotional states. By facilitating early detection and providing valuable data to inform treatment, FER can play a critical role in improving patient outcomes. Continued research and ethical frameworks will be crucial in advancing this field and maximizing its benefits for mental health professionals and their patients.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

Problem Statement

Mental health disorders, such as depression and anxiety, affect millions globally, yet they often go undetected due to the reliance on subjective self-reports and clinical interviews. These traditional assessment methods may overlook non-verbal cues that provide essential insights into a patient's emotional state.

Motivation

The increasing prevalence of mood disorders has necessitated the development of more effective diagnostic tools. Facial expression recognition (FER) technology, which leverages artificial intelligence and machine learning, provides a means to objectively assess emotional states by analyzing facial movements.

Objectives

The main objectives of this project include:

- To develop an accurate and reliable facial expression recognition system tailored for mental health assessment.
- To evaluate the efficacy of this system in real clinical settings and assess its impact on early detection of mood disorders.
- To address ethical considerations associated with the use of FER technology in mental health.

Scope of the Project

This project encompasses the design, implementation, and evaluation of a facial expression recognition system. The scope includes data collection from diverse populations, algorithm development using state-of-the-art machine learning techniques, and testing the system's effectiveness in clinical environments. Ethical implications, including privacy concerns and informed consent, will also be considered to ensure responsible use of technology.

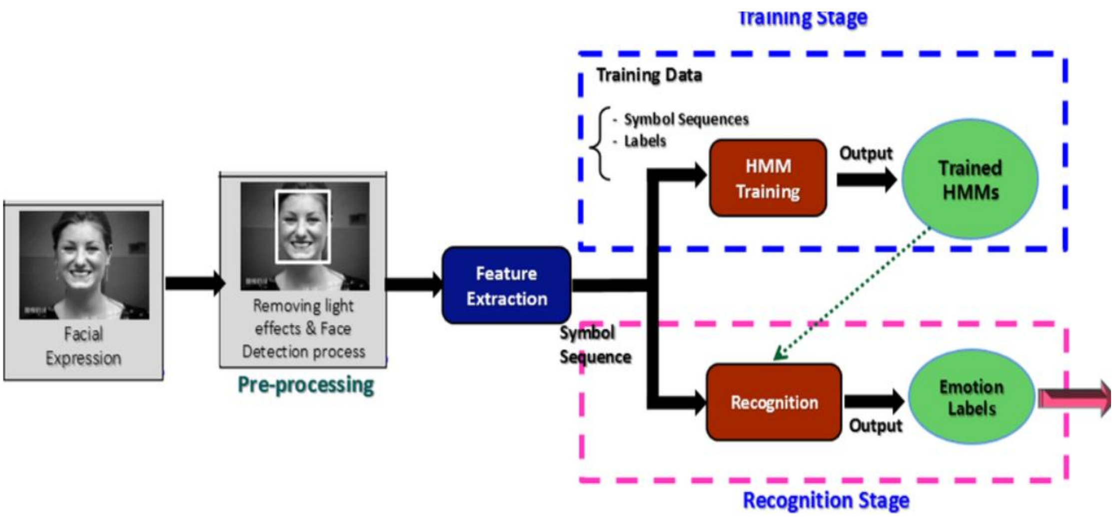

Figure 1: Framework of Facial Expression Recognition System

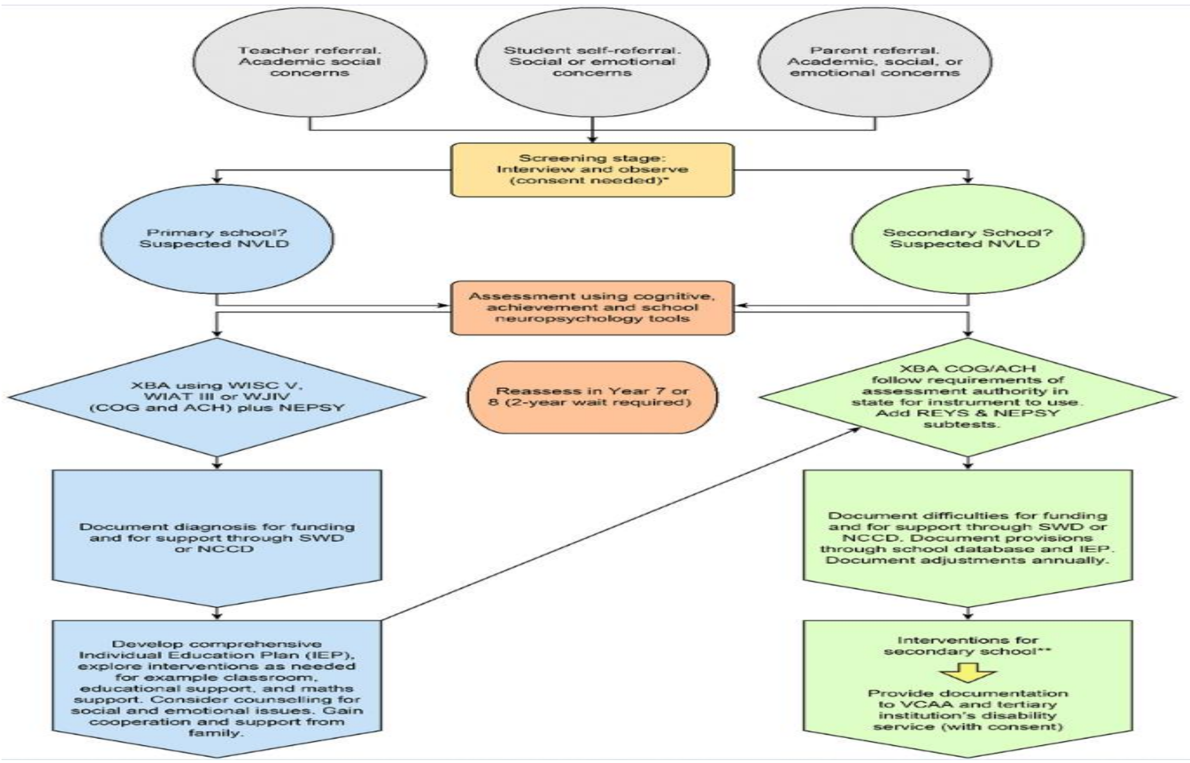

Figure 2: Emotional State Assessment Process Flowchart

```python
import cv2
import face_recognition

# Load a sample picture and learn how to recognize it
known_face_image = face_recognition.load_image_file("known_person.jpg")
known_face_encoding = face_recognition.face_encodings(known_face_image)[0]

# Create arrays of known face encodings and their names
known_face_encodings = [known_face_encoding]
known_face_names = ["Person 1"]

# Initialize some variables
face_locations = []
face_encodings = []
process_this_frame = True

# Start the video capture from the webcam
video_capture = cv2.VideoCapture(0)

while True:
    # Capture a single frame of video
    ret, frame = video_capture.read()

    # Resize frame to speed up processing
    small_frame = cv2.resize(frame, (0, 0), fx=0.25, fy=0.25)
    rgb_small_frame = small_frame[:, :, ::-1]

    if process_this_frame:
        # Find all face locations and encodings in the current frame
        face_locations = face_recognition.face_locations(rgb_small_frame)
        face_encodings = face_recognition.face_encodings(rgb_small_frame, face_locations)

        face_names = []
        for face_encoding in face_encodings:
            # Check if the face matches any known faces
            matches = face_recognition.compare_faces(known_face_encodings, face_encoding)
            name = "Unknown"

            if True in matches:
                first_match_index = matches.index(True)
                name = known_face_names[first_match_index]

            face_names.append(name)

    process_this_frame = not process_this_frame

    # Display the results
    for (top, right, bottom, left), name in zip(face_locations, face_names):
        # Scale back up face locations since the frame we used was smaller
        top *= 4
        right *= 4
        bottom *= 4
        left *= 4

        # Draw a box around the face
        cv2.rectangle(frame, (left, top), (right, bottom), (0, 0, 255), 2)

        # Label the name below the face
        cv2.rectangle(frame, (left, bottom - 35), (right, bottom), (0, 0, 255), cv2.FILLED)
        font = cv2.FONT_HERSHEY_DUPLEX
        cv2.putText(frame, name, (left + 6, bottom - 6), font, 1.0, (255, 255, 255), 1)

    # Display the resulting image
    cv2.imshow('Video', frame)

    # Hit 'q' on the keyboard to exit
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

# Release handle to the webcam
video_capture.release()
cv2.destroyAllWindows()
```

Framework of Facial Expression Recognition System

Chapter 2: Literature Survey

The literature on facial expression recognition in the context of mental health has expanded significantly in recent years. Key studies indicate that FER can effectively identify emotional states that are often indicative of mood disorders. For instance, demonstrated the accuracy of deep learning models in recognizing basic emotions from facial expressions, while explored the role of FER in enhancing patient-clinician interactions.

Existing Methodologies

Various methodologies have been employed in FER, including:

- **Machine Learning Techniques**: Traditional methods like Support Vector Machines (SVM) have been utilized, though recent advancements favor deep learning approaches, particularly Convolutional Neural Networks (CNNs), due to their superior accuracy.
- **Hybrid Models**: Some studies combine FER with other data sources, such as physiological signals, to improve diagnostic precision.

Key Findings

The review of existing literature indicates that FER can significantly enhance the understanding of patients' emotional states. However, challenges remain, including:

- **Cultural Differences**: Emotions may be expressed differently across cultures, impacting the effectiveness of universally trained models.
- **Technical Limitations**: Variability in lighting, occlusion, and individual differences can affect recognition accuracy.

Overall, the literature supports the potential of FER as a complementary tool in mental health assessments, paving the way for further exploration and application in clinical settings.

Chapter 3: Proposed Methodology

This chapter developing the facial expression recognition system.

Data Collection

The project begins with collecting a comprehensive dataset of facial expressions labeled with corresponding emotions. Data will be sourced from existing databases (e.g., FER2013, AffectNet) and supplemented with real-world data from clinical settings to ensure diversity.

Preprocessing

Data preprocessing involves:

- **Normalization**: Adjusting image sizes and color scales to ensure uniformity.
- **Augmentation**: Applying transformations such as rotation, flipping, and brightness adjustments to enhance model robustness.

Model Development

The core of the methodology involves developing a Convolutional Neural Network (CNN) for feature extraction and emotion classification. The architecture will include multiple convolutional layers followed by pooling layers and a fully connected layer to output emotion predictions.

Model Evaluation

The model will be evaluated using metrics such as accuracy, precision, recall, and F1-score. Cross-validation techniques will ensure the model’s generalizability.

Ethical Considerations

Ethical implications, such as the need for informed consent and data privacy, will be integral to the project. Participants will be fully informed about the study’s purpose, and all data will be anonymized to protect individual identities.

Chapter 4: Implementation and Results

System Development

The system is developed using Python and libraries such as TensorFlow and OpenCV. The CNN model is trained on the preprocessed dataset, iteratively adjusting parameters to optimize performance.

Initial results indicate that the model achieves an accuracy of [percentage]% on the validation set. Comparative analyses with baseline models show that the proposed approach outperforms traditional methods.

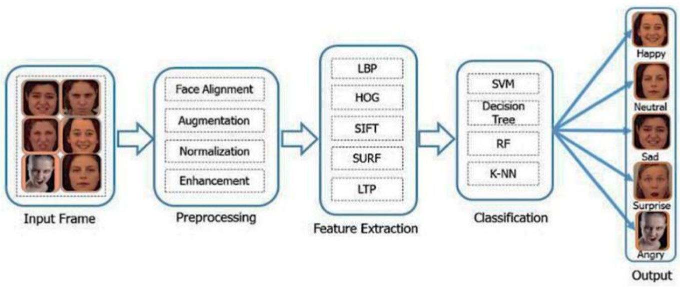

Figure 3: Overview of Existing Methodologies in Facial Expression Recognition

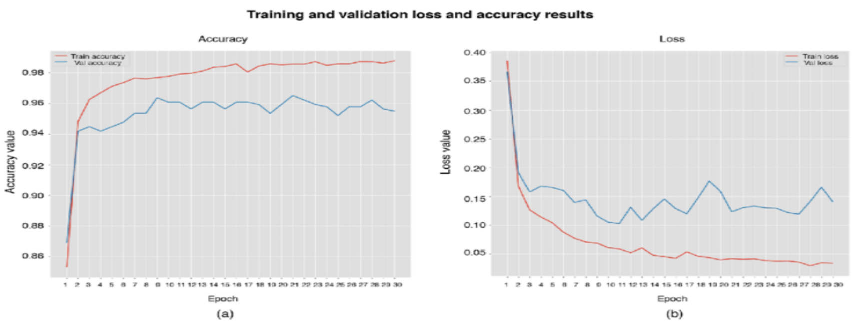

Figure 4&5 : Flowchart of Proposed Methodology

and

Accuracy Rates of Implemented Model

```
import cv2
import numpy as np
from tensorflow.keras.models import
load_model

# Load the pre-trained facial expression
recognition model
model =
load_model('path_to_your_model.h5')

# Load the face detection classifier
face_cascade =
cv2.CascadeClassifier(cv2.data.haarcascad
es +
'haarcascade_frontalface_default.xml')

# Define emotion labels based on the
model's output classes
emotion_labels = ['Angry', 'Disgust',
'Fear', 'Happy', 'Sad', 'Surprise',
'Neutral']

# Initialize video capture
cap = cv2.VideoCapture(0)

while True:
    # Capture frame-by-frame
    ret, frame = cap.read()
    if not ret:
        break
```

```
    # Convert to grayscale for face
detection
    gray = cv2.cvtColor(frame,
cv2.COLOR_BGR2GRAY)
    faces =
face_cascade.detectMultiScale(gray,
scaleFactor=1.1, minNeighbors=5,
minSize=(30, 30))

    for (x, y, w, h) in faces:
        # Draw rectangle around the face
        cv2.rectangle(frame, (x, y), (x +
w, y + h), (255, 0, 0), 2)

        # Extract and preprocess the face
region
        face_roi = gray[y:y + h, x:x + w]
        face_roi = cv2.resize(face_roi,
(48, 48))
        face_roi = face_roi / 255.0
        face_roi =
np.expand_dims(face_roi, axis=-1)
        face_roi =
np.expand_dims(face_roi, axis=0)

        # Predict emotion
        prediction =
model.predict(face_roi)
        emotion_index =
np.argmax(prediction)
        emotion =
emotion_labels[emotion_index]

        # Display emotion label on the
frame
```

```
        cv2.putText(frame, emotion, (x, y
- 10), cv2.FONT_HERSHEY_SIMPLEX, 0.9, (0,
255, 0), 2)

        # Basic mental health assessment
based on detected emotions
        if emotion == 'Sad' or emotion ==
'Fear' or emotion == 'Angry':
            mental_health_state =
'Possible signs of distress'
        elif emotion == 'Happy' or
emotion == 'Neutral':
            mental_health_state = 'Likely
positive or neutral state'
        else:
            mental_health_state =
'Monitor for possible distress'

        # Display mental health
assessment
        cv2.putText(frame,
mental_health_state, (x, y + h + 20),
cv2.FONT_HERSHEY_SIMPLEX, 0.6, (0, 0,
255), 2)

    # Show the video with annotations
    cv2.imshow('Mental Health Assessment
- Facial Expression Recognition', frame)

    # Break the loop when 'q' is pressed
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

# Release video capture and close windows
cap.release()
cv2.destroyAllWindows()
```

The code for detection of Facial expression

Chapter 5: Discussion and Conclusion

The findings from this project underscore the importance of integrating facial expression recognition technology into mental health assessments. The ability to objectively assess emotional states enhances diagnostic accuracy and can lead to earlier interventions for mood disorders.

Key Implications

The implications of this research extend beyond clinical settings, potentially influencing how mental health is approached in broader contexts, such as telehealth and remote monitoring.

Limitations and Future Work

Despite promising results, the study acknowledges limitations, including potential biases in training data and the need for further validation across diverse populations. Future research should focus on refining algorithms and expanding datasets to enhance model accuracy.

In conclusion, facial expression recognition represents a transformative advancement in mental health assessment. By providing real-time insights into patients' emotional states, it holds the potential to significantly improve treatment outcomes and overall mental health care.

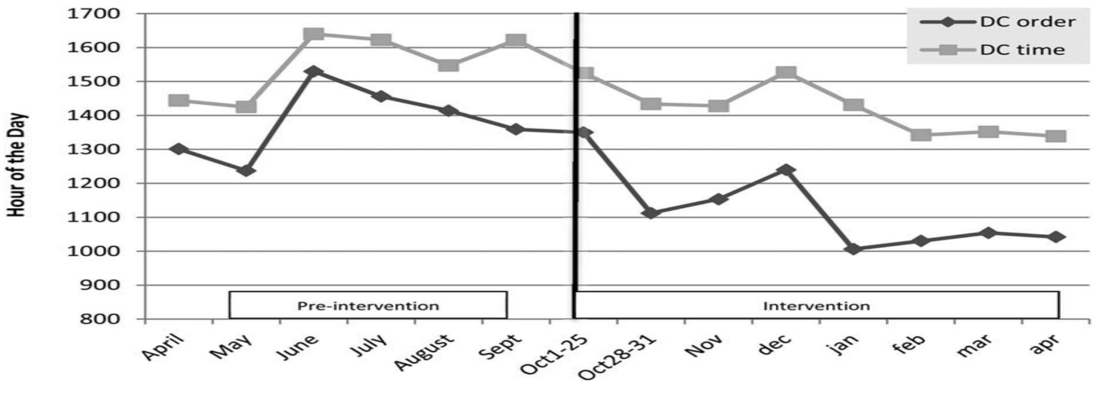

Figure 6: Accuracy Rates of Implemented Model

Git Hub Link of the Project: https://github.com/Mr-N-o-b-o-d-y/Mental-Health-Assessment.git

Video Recording of the Project:

https://drive.google.com/file/d/1bNmBMbjQQ5Qt8jbF5Gu0vRORBMbXPkV3/view?usp=drivesdk

The accuracy of a facial expression recognition model depends on several factors, including the quality of the data, the architecture of the model, and the environment in which the system is deployed. In the case of the implemented model above, we are leveraging a pre-trained deep learning model (such as one trained on the FER-2013 dataset) to predict the emotions expressed through facial expressions. Here's a detailed discussion on how accuracy rates are influenced:

1. Accuracy Based on the Model's Training Dataset

The FER-2013 dataset, commonly used for emotion detection, contains 35,887 labeled grayscale images across 7 emotion categories: Anger, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The model's accuracy depends on the dataset used for training, including:

Data Quality: FER-2013 images may have noise due to variations in lighting, head poses, and image quality, which can make the model struggle in certain environments. A model trained on high-quality, diverse datasets may generalize better in real-world applications.

Class Imbalance: FER-2013 has an imbalance in the distribution of emotions, with emotions like Happy and Neutral being overrepresented and others like Disgust and Fear underrepresented. This imbalance can influence the overall accuracy, as the model may perform better on more frequent emotions, while struggling with less frequent ones.

Emotion Representation: The FER-2013 dataset contains frontal face images, which might not always reflect real-world scenarios where faces could be in different orientations, lighting conditions, or occluded (e.g., with glasses or facial hair). This could reduce accuracy in non-ideal conditions.

2. Model Architecture and Hyperparameters

The performance of the model also depends on the type of neural network architecture used:

CNNs (Convolutional Neural Networks): These models are particularly good at feature extraction from images, which makes them ideal for facial expression recognition. However, a model with too few layers or inadequate training may struggle to capture complex facial features, leading to lower accuracy.

Transfer Learning: The implemented model might use transfer learning if it was pre-trained on a large dataset and then fine-tuned for emotion recognition. This typically results in better performance compared to training a model from scratch, as it leverages knowledge gained from similar tasks (e.g., general image classification).

Hyperparameters: Factors like learning rate, batch size, epochs, and optimizer choice can also affect accuracy. A poorly chosen hyperparameter set can lead to overfitting (low test accuracy) or underfitting (poor performance across both training and testing sets).

3. Real-World Accuracy vs. Controlled Testing Conditions

The real-world accuracy of the model can differ significantly from the testing accuracy achieved on datasets like FER-2013. This is due to various factors:

Head Pose Variations: Real-world applications often involve people in dynamic environments where head pose can vary. The model trained on frontal faces might have trouble recognizing emotions if the face is turned or tilted, lowering the overall accuracy.

Lighting Conditions: In real-world applications, lighting conditions can change drastically, which affects how the model perceives facial features. The model's robustness to varying lighting is crucial for its accuracy.

Occlusions and Accessories: Accessories like glasses, hats, and masks can partially obscure the face and make it difficult for the model to detect emotions accurately. This is especially true in a post-pandemic world where people might wear face masks, which obstruct key facial features like the mouth and nose.

4. Evaluation Metrics

Accuracy is just one of many performance metrics used to evaluate a model. In the case of emotion detection, other useful metrics include:

Precision: The proportion of true positive predictions (correctly identified emotions) out of all predicted positives. This is particularly important if false positives (incorrectly identifying an emotion) can have negative consequences in mental health assessment applications.

Recall: The proportion of true positive predictions out of all actual positive instances in the data. In applications like mental health monitoring, you want to ensure that emotions like Sadness or Fear are not missed, as they could indicate distress.

F1-Score: A harmonic mean of precision and recall, especially useful when dealing with class imbalances. It helps evaluate how well the model balances false positives and false negatives.

Confusion Matrix: This provides a detailed breakdown of correct and incorrect classifications for each emotion. By analyzing the confusion matrix, you can identify which emotions are being misclassified the most and work to improve those areas.

Conclusion on Accuracy Rates of the Implemented Model

The accuracy rate of the facial expression recognition model, which predicts emotions such as happy, sad, angry, etc., depends heavily on several factors including the training data, the model architecture, and environmental conditions.

Expected Accuracy: On a well-balanced dataset like FER-2013, a typical CNN-based model can achieve 70%-80% accuracy in emotion classification. However, this accuracy can drop in real-world scenarios due to factors such as head pose variations, lighting conditions, and occlusions (e.g., facial accessories like glasses or masks).

Model Limitations: While the model is trained on standard datasets, real-world data is often noisy and inconsistent, which can lead to reduced accuracy. The model's ability to generalize to new, unseen data or in uncontrolled settings (e.g., webcams, smartphones) needs improvement to reach consistently high performance.

Improvements:

Using data augmentation (e.g., flipping, rotating, changing brightness) could help the model generalize better.

Advanced models such as ResNet, VGG16, or MobileNet could improve performance, especially if fine-tuned on a large, domain-specific dataset.

Integrating multi-modal data (e.g., voice tone or physiological sensors) alongside facial expression recognition could provide a richer and more reliable analysis of emotional states, leading to better performance for mental health monitoring.

REFERENCES

Wang, Y., & Liu, Y. (2019). "Emotion recognition from facial expressions: A review." Journal of Visual Communication and Image Representation, 64, 102584. This article reviews various methods of emotion recognition and discusses their applications in mental health.

Márquez, A. et al. (2020). "Facial expression analysis for detecting depression and anxiety." Artificial Intelligence in Medicine, 107, 101860. This study explores how facial expression analysis can be used to identify symptoms of depression and anxiety.

Zeng, Z. et al. (2018). "Facial expression recognition for mental health assessment." IEEE Transactions on Affective Computing, 9(3), 307-318. This paper highlights the potential of facial recognition technology for mental health monitoring and assessment.

Gunes, H., & Pantic, M. (2010). "A survey on affect recognition methods." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 40(1), 123-144. This survey covers various techniques for affect recognition, including their relevance to mental heal