BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Oshadi Yashodhika G. B.

(IT18120226)

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka institute of Information Technology
Sri Lanka

October 2021

BEHAVIOR ANALYSIS FOR DEPRESSION DETECTION.

Oshadi Yashodhika G. B.

(IT18120226)

The dissertation was submitted in partial fulfilment of the requirements for the B.Sc. Special Honors degree in Information Technology

Department of Information Technology

Sri Lanka institute of Information Technology
Sri Lanka

October 2021

DECLARATION

I declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
Oshadi Yashodhika G. B.	IT18120226	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date

[Dr. Pradeepa Samarasinghe]

Signature of the supervisor: Date

[Ms. Vijani Piyawardana]

Signature of the supervisor: Date

[Ms. Shalindi Pandithakoralage]

ABSTRACT

A healthy life is founded based on the social and psychological well-being of individuals. It enhances the positive attitude of individuals and improves living standards. However, the depressive condition is one of the leading negative impacts on mental health. Limited access to treatment in low and middle-income countries increases the number of individuals suffering from mental health. The study proposes a mobile application to address the limitation. The study implemented a classifier that analyzes emotion, facial features, and head pose for assessing depressive disorder risk. The analysis employed a deep learning algorithm, a neural network, to train the classifier. In addition, mobile development technology flutter is employed in implementing the proposed application.

Keywords: Depression risk, Neural Network, Emotions, Head Pose

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Assistant Professor Pradeepa Samarasinghe, for her essential assistance in formulating the study objectives and methodology. Your insightful feedback encouraged me to improve the quality of my work. I would also like to express my gratitude to Ms. Shalindi Pandithakoralage, my external supervisor, for her invaluable assistance with the psychological component of my studies. In addition, I am grateful to Ms. Vijani Piyawardana, my co-supervisor, for assisting me in completing the research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AAM	Active Appearance Model
FACS	Manual Facial Action Coding System
AVEC	Audio/Visual Emotion and Depression Recognition Challenge
FFNN	Feed-Forward Neural Network
REST	REpresentational State Transfer
API	Application Programming Interface

1. INTRODUCTION

1.1. Background & Literature Survey

1.1.1. Background

The most effective way to enhance community well-being is to strengthen the emotional well-being of individuals. The enhanced community will create an emotionally and socially improved culture for the next generation. Moreover, healthy life is founded based on the social and psychological well-being of individuals. It enhances the positive attitude of individuals and improves living standards. However, limited attention on the importance of mental wellness is one of the major concerns for the decline of the mental well-being of an individual. A healthy mind and wealthy life guide to longevity along with success. An individual's psychological well-being is impacted by concerning several factors and situations. Depressive condition is one of the leading negative impacts on mental health. The recent reports of the World Health Organization state that depressive disorders have affected over 300 million of the global population [1]. severe level of depression considers in Psychological and pharmacological treatments. However, in low and middle-income countries, often limited or underdeveloped the treatments and support services. Limited access to treatment in these countries increases the number of individuals suffering from mental health [2]. As indicated in [3], underdeveloped countries fronted 77% of global suicide in 2019.

It is necessary to monitor the individuals suffering from depression to provide better treatment. Early depression risk monitoring is crucial as the individuals may obscure their state of mind due to social stigma and fear. Also, the mental health industry in Sri Lanka fronts a key concern of lacking acceptance of depression due to social stigma. Emotions play a vital role in identifying early depressive states in individuals. However, mental health professionals identify symptoms of depression by monitoring the behavior and using standard rating scales. Several studies evaluated depression detection by using different techniques. The studies focused on identifying depression risk using emotions, facial features, and head pose angles.

1.1.2. Literature survey

Facial cues-based automatic depression analysis has gained popularity in recent years. The previous studies employed various Deep Learning approaches to diagnose depression. Different deep learning-based depression detection models have been proposed by considering spatial-temporal facial features [4 - 8], Facial expressions [9 - 11], and face landmarks [1, 12], while few deep learning-based depression prediction models have been proposed by considering spatial temporal facial features [13], Facial expressions [9, 10], and head pose [14, 15].

J. F. Cohn et al. investigated facial and vocal behavior relation in diagnosing depression, which was the first automated facial image analysis and audio signal processing to assess depression. The study used the person-specific Active Appearance Model (AAMs), Manual Facial Action Coding System (FACS), and vocal prosody to detect depression [16]. S. P. Namboodiri and Venkataraman proposed a system to detect depression among college students. This study considered frontal face images of happy, contempt, and disgust face in the video frame to analyze the depression. Viola-Jones face detection algorithm used to extract faces from each image [17].

Few studies have analyzed the link between depression and upper body movement. S. Alghowinem et al. examined the head pose and movement patterns of depressed individuals compared to healthy subjects using clinically validated real-world data. The statical analysis on the head pose and movement patterns showed a significant difference between healthy and depressed individuals [18]. The feasibility of a cross-cultural method to assess depression severity was investigated by S. Alghowinem et al. The study based on temporal aspects of the eye gaze and head position of the participants by using video-recorded clinical interviews (Australian, US, German) from three different datasets (BlackDog, AVEC, Pittsburgh) [15]. J. Joshi et al. study explored the upper body movements and gestures for automatic depression analysis. Upper body movements and intra-facial movements were computed using Space-Time Interest Points (STIP) feature, while head movement analysis was assessed by selecting rigid facial fiducial points [19].

B. Vimaleswaran and G. S. Ratnayake developed a deep residual regression network (DRR DepressionNet), formulated top on a deep-residual regression network (Deep ResNet), to assess depression severity using facial expression [9]. The Dlib toolbox was used to extract facial images. The image-enhancing was considered by applying the Adaptive histogram equalization, YCrCb space enhancement, Histogram equalization, and contrast enhancement. Further, the study applied image rotation as well. The face region was cropped and aligned with the eye position prior to the visual enhancement. The study [13] employed the convolution neural network (CNN) framework, formulated using the 3D CNN and RNN networks, to predict depression levels by accumulating global and local spatial-temporal information from consecutive facial expressions. A feed-forward Neural Network (FFNN) was used in the study [10] to predict depression levels by looking at facial expressions. However, the FFNN model was not tested using a verified dataset in this study. However, the study [10] has not included a performance evaluation. The Active Appearance Model (AAM) was used to extract facial features. Backpropagation was also utilized to train the FFNN model, which involved comparing predicted and expected outputs and changing weights to minimize errors.

The study [4] focused on 3D CNN-based automatic depression identification using local and global face regions to encode spatial-temporal information. It captures Spatio-temporal dynamics in complex movements. The studies used the AVEC 2013 and AVEC 2014 datasets to validate respective observations. Using the Keras framework and the transfer learning approach, the suggested C3D model updated a 3D CNN architecture. Similarly, the study [6] explored a 3D-CNN-based multiscale Spatiotemporal Network (MSN) that accurately explored facial data from video clips linked with depressive behaviors. The research in [5] was based on the CNN-based depression detection technique, which utilizes the Keras framework. The parameters of the convolutional layer were initialized using the VGG Face dataset. The study uses distribution learning to map the association between depression levels and facial images. To examine the distribution of data over depression levels, the researchers used Expectation Loss.

The CNN network was employed in the papers [7] and [8] to detect depression. The study [7] represented a two-stream model composing temporal pooling and stream methods to encode temporal and active facial emotions. Moreover, the analysis formulates an image map by identifying and encoding the Spatio-temporal information in videos. They applied the ResNet-50 model to classify encoded image dependencies. However, the image map can feed to any deep architecture. The two-stream model in [7] comprises a temporal stream and an appearance stream. When comparing the performance of the temporal, appearance, and combined streams, the combined stream outperformed the others. As a result, the model in the analysis referred to the merged stream.

Similarly, the study [8] proposed a Maximization and Differentiation Network (MDN) design based on convolutional networks. To encode the spatial-temporal variability of faces, it uses 3D residual networks. To capture the variation details, the study [8] presented a Maximizing Block and a Difference Block to summarize Spatio-temporal information. The features (Difference and Maximizing Block) were combined using a linear combination in the study [8]. As a result, the feature map dimensions block outputs have to be identical. The feature map was adjusted using the fusion function in this investigation.

1.2. Research Gap

Several studies focused on depression analysis based on facial expression and head pose estimation. The study of [16] focused on facial expression-based depression analysis, while the team in the research [18] explored the head movement-based depression analysis. Similarly, the study of [17] considered facial expressions to analyze the depressive features, while the research team of [18] studied the depression analysis by considering the head movement patterns. The research group in [19] analyzed the depressive features by considering both the facial features and head movement features.

The study of [9] focused on facial expression-based depression analysis by applying the Dlib toolbox, while the team in the research [10] explored facial expression-based depression analysis by employing the AAM approach. Similarly, the study of [13] considered the accumulated global and local spatial-temporal information from consecutive facial expressions to analyze the depressive features.

Similarly, the studies [7, 9] considered spatial-temporal information to capture facial expressions, while the research team of [10] studied depression analysis by encoding temporal and active facial emotions. The research group in [11] analyzed the depressive features by using encoded spatial-temporal variability of faces.

Several studies have been explored facial features using various techniques and technologies in the past few decades. However, none of the studies considered to provide the probability of early depressive features. The summary of the limitations in the previous studies explains in Table 1.1.

Table 1. 1: Comparison between the previous studies

	[2, 4]	[7-12]	[16, 17]	[15, 18]	[19]	Proposed System
Facial expression based depression Analysis	✓	✗	✓	✗	✓	✓
Head pose based Depression Analysis	✗	✓	✗	✓	✓	✓
Facial features based depression Analysis	✓	✗	✓	✗	✓	✓
Probability of early depression features based on facial expression	✗	✗	✗	✗	✗	✓
Probability of early depression features based on head movement	✗	✗	✗	✗	✗	✓

There are mobile applications in the market that recognize the symptoms of depression. New classifications of managing mental health have been explored in various applications.

“MindDoc” is a mobile application that mainly tracks the emotional state based on a set of daily questions. The application can detect patterns by tracking emotional states and identify areas that can improve. Application is generating extensive evaluation regularly, detailing symptoms and providing a summary of your emotional state. However, the “MindDoc” application does not include the technique to detect real-time emotion based on images and videos to analyze the depression symptoms. The analysis merely depends on the user response to the set of questions, which can easily bias from the situation [20 - 22]. “ImoodJournal” is an app that helps individuals to track mood many times per day. A comprehensive summary provides a graph with complete mood history. Although the app keeps tracking the emotion to deliver an extensive mental health status, this is not considered the real-time emotion capture. Direct involvement of the individual in the mood tracking process does not provide a productive analysis since the analysis can bias accordingly [23 - 26].

Mobile App “Sanvello” is a self-care app that helps individuals manage stress, anxiety, and depression. The app includes mood tracking as well as thought and habit tracking. Sanvello progress assessment represents the connection between emotions, activities, and experiences [27 - 29]. Mood journal “Daylio” tracks the mood and related activities and information presented in various graphs as average daily mood and monthly mood chart [30, 31]. UP! is an automated mood diary that provides the warning signs of depression, mania, and hypomania. The app is capable of managing journals of sleep habits, work-life balance, and physical activity. Up! provide the ability to connect with a health provider or trusted person by using a paring code, which enables effortless information sharing. However, the mobile apps do not consider the real-time emotion recognition as well as the head movements pattern that can accurately analyze the early depression symptoms [32, 33].

The existing mobile applications with the proposed system are compared in Table 1.2 as follows.

Table 1. 2: Comparison of existing Mobile Applications for mental health

	Mind Doc [20-22]	Imood Journal [23-26]	Sanvello [27-29]	UP! [32, 33]	Daylio [30,31]	Proposed App
Track emotion state	✓	✓	✓	✓	✓	✓
Identify Emotion	✗	✗	✗	✗	✗	✓
Identify Head Pose angles	✗	✗	✗	✗	✗	✓
Depression analysis based on emotion	✓	✗	✗	✓	✗	✓
Depression analysis based on head pose angles	✗	✗	✗	✗	✗	✓
Analyze depression risk based on emotion, facial expressions and head pose angles	✗	✗	✗	✗	✗	✓
Extensive Evaluation	✓	✓	✓	✓	✓	✓

2. RESEARCH PROBLEM

Treatments would be more effective if depressed individuals have self-awareness of the mental health. Mobile application, which includes emotion-based behavior analysis for depressive disorder, may help people for self-awareness on the mental health state. Generally, the existing mobile applications focused on tracking moods. The analysis might be more biased since it considers daily mood records logged by the end-user. A mobile application that eliminates sort of manipulation by analyzing upper body movement would accurately assess early depression risk. To the best of my knowledge, a similar approach of automated analysis on depressive disorder using unexpected behavior changes that might have a chance of having depression is not yet studied in Sri Lanka.

Shortcomings in the field of depression and the mechanism to diagnosis the symptoms are summarized as follows.

• Individuals do not attend counseling sessions due to social stigma.

• Lack the automated solution to identify early depression risk using facial expression.

• Lack the automated mechanism to identify early depression risk by analyzing head pose angles of a depressed individual.

3. RESEARCH OBJECTIVES

The primary objective of the depressive behavior analysis is to analyze the signs that may lead to being a depressive person. The study focuses on a mobile application to identify behaviors that could be highly anticipated to contribute to depressive disorder.

3.1.1. Main Objective

The main objective of the study focuses on identifying the early depressive behavior risk based on Emotion-based behavior.

3.1.2. Specific Objectives

The Emotion-based behavior analysis for classifying early depressive behavior risk is divided into specific sub-goals as follows.

- Create datasets to retrieve facial emotions of depressed/likely depressed individuals.
    - Datasets will be created of depressed/ likely depressed individuals to retrieve the facial emotions.

- Identify emotions of depressed/likely depressed individuals based on images and videos.
    - Emotions of depressed or likely depressed individuals will be identified using an algorithm.
    - The analysis is based on the images and videos.

- Improving the existing methods to capture subtle emotional changes.
    - Existing methods to capture the subtle emotional changes will be improved when comparing to the existing methods.

- Identify head movements pattern of depressed/likely depressed individuals based on videos.

    - Head movement patterns of depressed or likely depressed individuals will be identified using an algorithm.
    - Videos will be used to analyze the head movement pattern of depressed/ likely depressed individuals.

- Build a classifier to model, abnormal behaviors based on facial emotions and head movements.
    - A classifier will be built to identify abnormal behaviors based on facial emotions and head movements.

- Predict the probability of abnormal behavior towards depression with the designed model.
    - Abnormal behavior towards depression will be predicted using the designed model.
    - The probability of the analysis will be provided in the designed model.

- Integrate the classifier to Mobile Application.
    - Integration of the mobile application and the built model will be done using an Application Programming Interface (API).

Additionally, the sub-goals are achieved with the mobile application.

- Responsible for implementing interfaces.
    - The interfaces of the proposed system are focused on frontend development.

- Mobile Application development for facial movement and head movement analysis.
    - The mobile-based system will develop using flutter, to analyze the facial movement and head movement of depressed/ likely depressed individuals.

4. METHODOLOGY

4.1. Introduction

The following section explores the followed methodology in this study. The subsections are divided into the component overview, commercialization aspects of the product, Tool & technologies, component design & implementation, and testing.

component design & implementation subsection is divided into data acquisition, data analysis, model implementation, and mobile implementation. Frontend test cases and backend test cases are explained in the testing subsection.

4.2. Component Overview

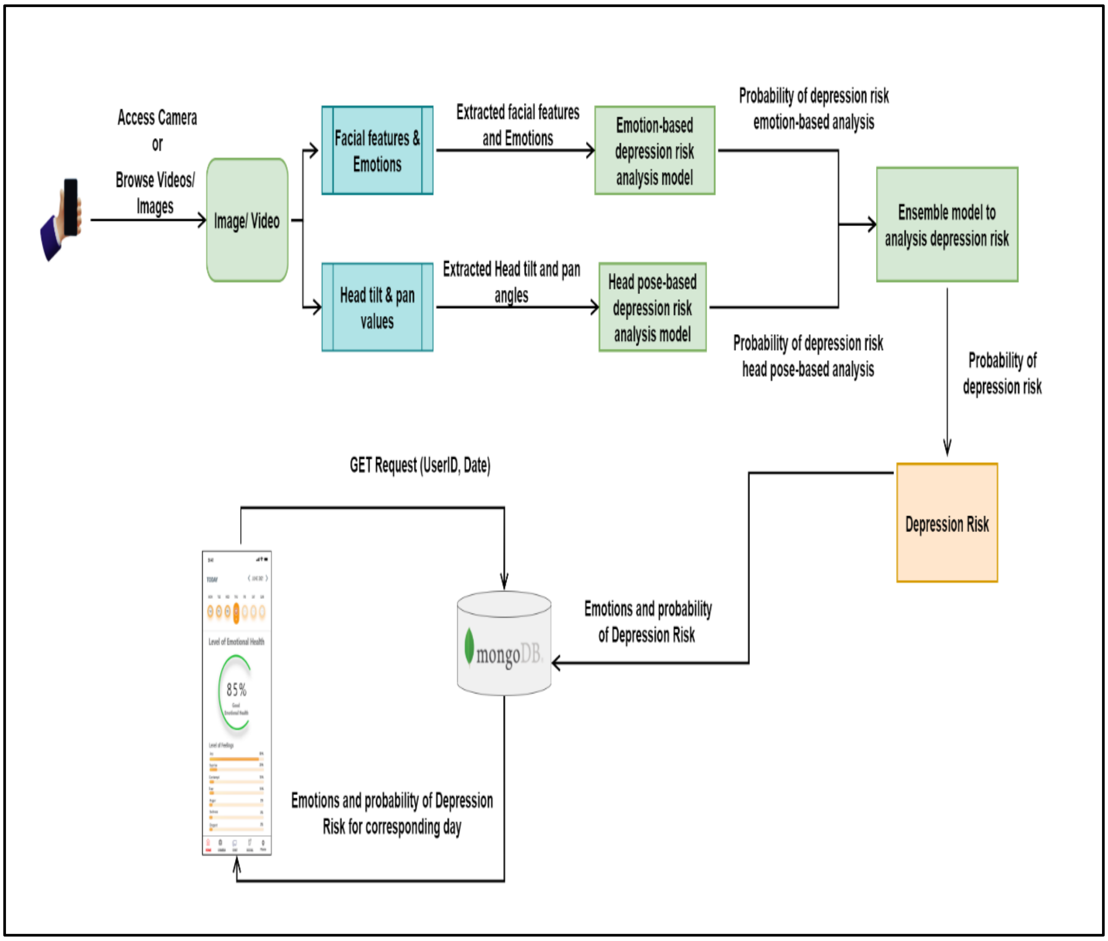


Figure 2. 1: Component overview

The study focused on the emotion and head pose of an individual to identify the depressive risk. As illustrated in Figure 2. 1, images and videos should be input into the system by accessing the camera or browsing the gallery to initiate the procedure. The facial features and emotions will extract from the given image/ video to feed to the classifiers. In addition, the head tilt and pan angles will select from the given image/ video. The extracted emotions and facial features will feed to the emotion-based depression risk analysis model. The head tilt and pan angle will be input to the head pose-based depression risk analysis classifier.

The probability of depressive risk for emotion and head pose will feed to the ensemble classifier. The output probability of the ensemble classifier considers as the probability of depressive risk. The emotions and depression risk values will save in the MongoDB. REST API will connect with the mobile application to explore the data points. The extensive report will appear on the end-user application once the data points have been retrieved and manipulated according to the specification. The summary can be accessed for a particular day using the API endpoints.

4.3. Commercialization Aspects of the Product

Psychologists can use the proposed depression risk analyzer as the initial approach for identifying depressive risk by considering emotions, facial features, and head pose angles. Hence, Psychologists can motivate to use the application by promoting the application through online mental health-related webinars and groups of social media platforms. The psychologists can suggest individuals for using the application daily since the application provides an extensive summary for each day.

Further, the public can have a possibility of self-monitoring the mental health to analyze the depressive risk. Therefore, the public can motivate to use the application by promoting the application through social media platforms. Recently, social media platforms received much attention from the public. The mental health-related groups can be used to introduce the mobile application among the younger generation.

4.4. Tools and Technologies

This section outlines the tools, libraries, and technologies employed in the system implementation. Table 2. 1 explains the tools and technologies used in the model implementation, and Table 2. 2 outlines the applied tools and technologies in the mobile application.

Table 2. 1: Technologies for model implementationn

Technology	**Description**
Python	Object-oriented, high-level, and general-purpose programming language to implement the model.
Anaconda Environment	Anaconda Distribution will manage all the packages required to implement the model. Packages can be installed, upgraded downgraded easily.
OpenCV	The open-source library that uses to acquire the images for the analysis.
TensorFlow	The model will be trained using the open-source library, TensorFlow.
Keras	Open-source software library use to develop the model.
Affectiva API	Extract emotions from given face.
RealHePoNet	Pre-defined architecture to extract head pose for a particular individual.

Table 2. 2: Technologies for mobile implementation

Technology	**Description**
Flutter	Frontend development will be done using the Flutter UI software development kit.
Visual Studio Code	Environment use for mobile app implementation. VS code is a freeware source-code editor.
Image Picker	The package applies to pick images and videos from camera or gallery.
Video Player	The package provides the ability to play videos on the flutter-based mobile applications.
MongoDB	NoSQL database program which includes the document-oriented, cross-platform, source-available features.
Google Cloud	The Google Cloud, a suite of cloud-based computing services, was used to deploy the service.
Postman	An API platform use to test APIs

5. IMPLEMENTATION AND TESTING

5.1. Component Design & Implementation

5.1.1. Data Acquisition

This section outlines the procedure of data acquisition for the emotion-based depressive behavior analysis. The data acquisition phase has initialized with dataset creation. As the initial step, the process started to find the sources for creating the dataset. The study was able to obtain the required data from mental health-specific YouTube channels. Emotion-based depressive behavior analysis has received approval for three YouTube channels in the mental health field to acquire the dataset for videos of Psychiatric Interviews with depressed individuals. Note that YouTube channels for Psychiatric Interviews videos with depressed individuals (Figure 3. 1) were Pika Grape Snack/ University of Nottingham/ Symptom Media. The YouTube channel "Make the Connection" has chosen to acquire a non-depressed dataset. It comprised videos (Figure 3. 2) about True Depression Recovery Stories of US Army Officers.

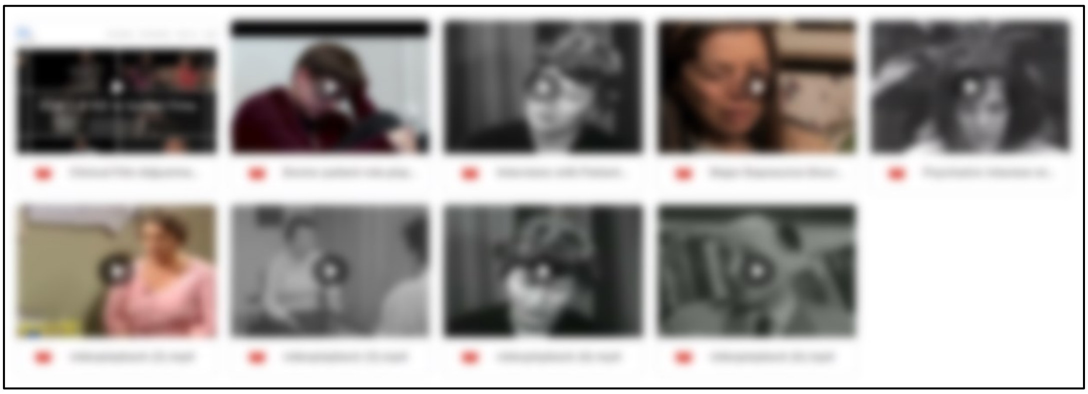


Figure 3. 1: Psychiatric Interviews videos with depressed individuals

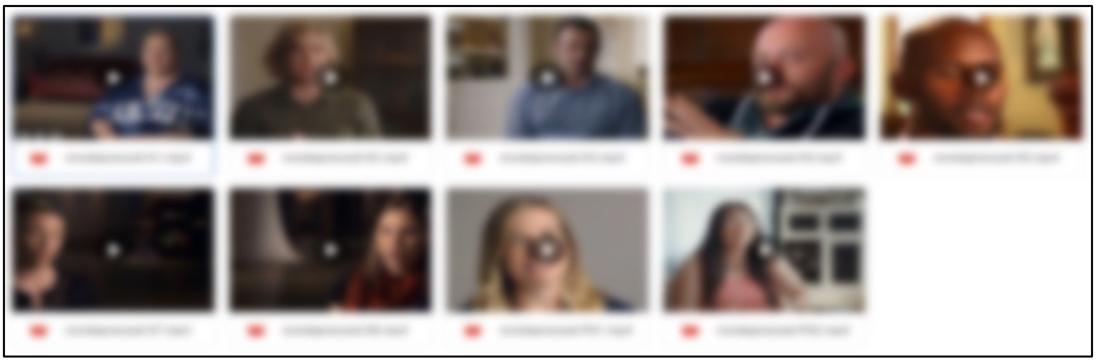


Figure 3. 2: Psychiatric Interviews videos with non-depressed individuals

The videos were separated into video frames as the next step of the data acquisition phase. The 1322 frames from depressed individuals and 1747 frames from non-depressed individuals were chosen and examined in this study. The depressed collection consists of nine female personalities and three male personalities. Moreover, the non-depressed collection comprises seven female Army Officers and eight male Army Officers. The number of individuals considered for the depressed category was fifteen, while the twelve individuals have declared in the non-depression level. In total, 18 videos have been used to build depressed and non-depressed datasets. All of the videos are longer than 1 minute, with the longest-lasting approximately 38 minutes. A professional psychologist approved the movies and divided them into depressive and non-depressive categories.

5.1.2. Data Analysis

The emotion-based analysis was required emotions and facial features as the initial step of implementing the emotion-based risk analysis classifier. Hence, the study employed the pre-trained model Affectiva JavaScript API to extract facial features and emotions. Affectiva is a professional in the development of emotion recognition software [34]. The extracted facial features and emotion values (Figure 3. 3), with the categorical data, were standardized before the training process of the classifier.

joy	sadness	disgust	contempt	anger	fear	surprise	valence	engageme	Timestam	FileName	smile	innerBrow	browRais	browFurr	noseWrin	upperLipF	lipCorner	chinRaise	lipPucker	lipPress	lipSuck	mouthOp	smirk	eyeClosur	attention	lidTighter	jawDrop	dimpler	eyeWiden	cheekRais
0.00183	0.0239	0.42733	0.19267	0.00201	0.00454	0.19272	0	0.07995	0	videoplay	6.99E-13	2.07E-07	7.07E-08	0.00022	0.00065	0.06818	0.00388	0.00206	0.00423	2.98E-08	1.34E-07	0.17007	8.10E-08	0.0612	77.6329	0.22996	0.01125	2.14E-09	5.75E-10	0.00016
0.00183	0.01942	0.42635	0.19267	0.00246	0.00091	0.63814	0	0.12461	0	videoplay	8.23E-14	3.91E-09	6.37E-08	1.83E-05	0.00037	0.02463	0.00034	0.01104	0.84626	3.00E-08	4.68E-07	6.38878	1.32E-08	0.19017	47.842	0.11603	48.0817	2.80E-10	1.55E-10	0.00376
0.00183	0.02388	0.42589	0.19268	0.00207	0.00224	0.32669	0	0.08009	0	videoplay	1.52E-12	9.71E-10	4.67E-08	0.00095	0.00047	0.00026	1.83E-06	0.00021	0.00045	3.99E-06	6.61E-08	0.19266	1.32E-07	8.55544	63.028	0.83779	21.1773	4.35E-09	2.44E-10	0.01182
0.00183	0.00113	0.42941	0.19267	0.04207	0.00017	2.24236	0	35.6572	0	videoplay	5.64E-12	2.71E-10	1.29E-08	2.29E-05	0.00121	0.18306	0.0005	0.01025	1.5564	2.01E-09	0.00205	91.735	1.25E-06	7.2042	67.2094	0.04527	99.0017	3.13E-11	4.27E-11	1.17E-05
0.00183	0.02317	27.8778	0.19267	0.00205	0.00016	2.28515	-42.047	0.08619	0	videoplay	3.14E-11	1.36E-10	5.64E-05	1.61E-09	0.1305	98.8012	2.24E-08	1.48E-08	5.80E-05	1.10E-07	7.46E-08	1.0919	7.67E-07	40.3531	79.4081	0.00131	99.7749	4.45E-10	4.21E-10	0.00263
0.00183	0.01582	30.6249	0.19267	0.003	0.00028	1.51784	-43.159	0.20678	0	videoplay	1.34E-10	9.71E-10	4.76E-06	5.72E-07	0.94998	99.9202	1.72E-06	1.73E-07	1.98E-06	2.53E-07	1.72E-08	12.541	3.42E-07	99.9935	77.0472	0.00216	83.0964	2.79E-09	8.57E-08	0.00145
0.00183	0.00256	27.817	0.19267	0.01853	0.00016	2.27675	-41.612	9.1417	0	videoplay	4.38E-09	1.08E-09	0.0002	5.74E-06	0.70287	97.4738	8.17E-10	4.38E-09	5.78E-06	1.12E-08	3.12E-08	67.1468	1.84E-08	0.07493	71.7767	0.03246	99.6239	9.71E-09	2.92E-06	0.00677
0.00183	0.00093	25.4196	0.19268	0.05177	0.00024	1.73173	-40.401	45.9167	0	videoplay	9.94E-12	1.40E-06	1.85E-06	0.00053	0.21779	95.8405	3.21E-05	0.00027	1.25E-05	9.66E-09	1.91E-05	97.4905	0.00033	0.01035	74.4928	0.36147	88.4583	3.63E-10	7.26E-08	0.0003
0.00183	0.00183	0.53158	0.19268	0.0262	0.0016	0.42094	0	16.462	0	videoplay	1.68E-12	3.56E-07	1.72E-06	0.00018	0.02131	4.85426	0.00013	0.00296	0.00896	6.78E-08	5.11E-05	77.2392	4.42E-05	99.9226	71.754	0.2325	31.3516	2.72E-08	1.06E-07	5.61E-05
0.00181	0.02375	0.77821	0.19332	0.27008	0.00456	0.1904	0	0.13318	0	videoplay	2.68E-10	5.47E-07	1.34E-08	0.17899	4.7331	2.00E-05	0.00024	0.00128	5.25E-05	1.37E-07	1.25E-06	0.75091	1.05E-07	0.01257	97.6861	97.6373	0.06216	2.05E-08	1.06E-09	0.00399
0	0	0	0	0	0	0	0	0	0	videoplay	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18.854	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	videoplay	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7.37792	0	0	0	0	0
0.00183	0.01408	1.21431	0.19267	0.00336	0.00039	1.20822	0	0.28205	0	videoplay	5.70E-10	6.36E-10	0.00927	4.14E-13	1.81847	19.2261	8.05E-10	4.81E-11	0.00012	8.72E-12	9.25E-10	16.0293	3.36E-09	0.00056	87.5052	9.09E-05	73.8221	3.86E-10	5.60E-08	0.01254
0.00183	0.01709	28.6211	0.19267	0.00277	0.0005	1.00252	-39.724	0.22063	0	videoplay	7.99E-11	1.40E-10	0.00725	4.13E-12	4.18034	90.6983	2.38E-07	1.38E-09	7.87E-05	5.89E-10	1.71E-08	10.2205	2.88E-09	0.00265	90.99	0.00033	66.2789	5.59E-10	1.64E-07	0.00027
0.00183	0.02399	0.42586	0.19267	0.00198	0.00454	0.19271	0	0.07925	0	videoplay	7.16E-10	4.31E-09	3.19E-08	5.36E-10	2.94E-05	0.00014	9.14E-07	9.67E-07	0.00035	0.00145	2.08E-06	0.04645	3.58E-05	64.9519	98.4463	5.00E-06	0.00697	1.24E-06	3.72E-10	2.37E-08
0.00183	0.02006	0.4259	0.19268	0.00236	0.00428	0.20131	0	0.11632	0	videoplay	2.77E-08	8.65E-08	1.28E-07	1.92E-08	0.00064	0.00362	5.53E-06	0.01738	0.00362	0.00377	0.00253	5.42356	0.00023	84.7173	96.5895	0.00049	1.75647	2.46E-05	2.12E-09	7.36E-09
0.00183	0.02106	0.42404	0.19269	0.00225	0.00453	0.19307	0	0.10448	0	videoplay	3.91E-09	3.99E-07	2.45E-07	2.94E-07	0.00069	5.53E-05	7.79E-06	0.08044	0.00644	0.00141	0.08028	3.91861	0.00071	97.527	96.5799	0.00206	0.08266	0.014	9.49E-10	1.48E-08
0.00223	0.0019	0.24686	0.10879	0.00066	0.00091	1.4741	2.06259	9.8299	0	nondepre	7.64054	3.76E-07	32.7236	4.66E-06	0.00456	0.17902	0.00072	0.00405	2.99E-08	1.92E-06	5.50E-09	36.2487	0.00548	2.45E-07	98.5374	0.00232	0.10018	3.12E-05	3.54E-07	0.00013
0.04825	0.00127	0.10685	0.04453	0.00147	0.00094	0.31478	14.3633	2.92301	0	nondepre	19.5638	6.63E-06	7.81182	1.02E-05	0.00426	0.80229	0.00045	0.00413	3.55E-09	1.21E-06	9.70E-09	43.1477	0.01012	6.56E-10	98.6889	0.00091	0.154	1.83E-05	4.81E-08	2.11E-05
0.00271	7.83E-05	0.155	0.06732	0.00078	9.08E-06	48.5944	8.8631	98.4952	0	nondepre	14.0932	6.86E-08	59.6721	2.94E-06	0.00158	0.25129	3.06E-05	1.94E-07	6.03E-09	2.75E-08	2.78E-09	98.8668	0.03893	6.06E-11	98.6311	0.00065	98.5703	1.40E-05	1.61E-06	1.54E-05
0.00218	0.00298	0.39386	0.17724	0.01285	0.00026	1.58321	0	5.62311	0	nondepre	1.1387	2.35E-05	0.8125	0.00062	0.02815	0.02751	6.70E-06	8.06E-08	1.26E-07	5.82E-07	1.65E-10	59.7017	0.01627	67.3986	98.4665	0.08695	82.7803	3.35E-05	3.19E-09	0.00062
0.00145	0.00179	0.44716	0.19269	0.01583	0.00181	0.48742	0	18.9202	0	nondepre	0.00052	4.25E-05	6.47665	0.00237	0.20667	0.61849	0.03468	2.29E-05	7.17E-08	3.40E-07	7.38E-08	73.1505	0.00065	91.8408	98.2294	0.00838	21.0586	1.66E-05	4.05E-07	0.00023
0.002	0.00083	4.63231	0.18789	0.05361	0.00016	2.29921	-15.749	50.2775	0	nondepre	0.50526	5.93E-05	0.03739	0.00854	0.16113	53.9108	0.002	5.28E-07	3.38E-08	3.98E-09	1.45E-11	99.9943	0.11796	0.00202	98.487	0.09295	99.9584	1.13E-08	1.05E-08	0.00293
0.0018	0.0237	0.42597	0.1927	0.00192	0.00446	0.19896	0	0.08243	0	nondepre	0.00106	7.68E-05	0.5136	0.00019	0.00336	9.38E-05	0.01783	8.20E-05	2.46E-08	0.08409	8.84E-05	4.95E-05	0.00209	9.45E-10	98.2255	0.00714	0.00299	0.33537	2.00E-08	4.22E-06
0.00051	0.00989	0.43728	0.19272	0.00034	0.00138	1.74048	0	0.98633	0	nondepre	8.38E-05	0.00014	35.3275	3.16E-05	0.00187	0.5806	0.00995	6.25E-05	1.87E-08	8.99E-06	5.93E-08	0.14612	0.00252	2.25E-08	98.18	0.00379	0.34301	3.72E-05	6.77E-07	2.39E-05
0.00187	0.00085	3.17038	0.19076	0.05442	0.00016	2.31515	-10.964	50.2582	0	nondepre	0.15713	1.71E-06	0.14465	0.01363	0.0045	45.0213	0.00086	5.80E-08	6.48E-06	2.70E-08	1.20E-12	99.9999	0.01419	3.78E-07	98.1659	0.03451	99.9883	1.12E-09	7.72E-09	1.29E-07
0.00177	0.00871	2.85383	0.1933	0.00532	0.00428	0.20387	-10.414	0.74352	0	nondepre	5.90E-07	8.53E-06	0.7375	0.06038	0.00105	42.3915	0.95202	0.00077	0.00204	2.23E-05	1.41E-09	30.7456	0.01977	9.73E-09	95.845	0.05197	0.60221	1.73E-07	9.37E-07	9.93E-10
0.00079	0.00074	27.5805	0.19275	0.0101	0.00143	1.08571	-41.994	67.8673	0	nondepre	1.88E-09	3.98E-06	23.2195	0.00901	0.0026	98.7562	0.05181	0.00138	0.00017	6.16E-07	6.67E-10	87.2481	0.00229	3.80E-07	94.8299	0.01948	11.4978	5.49E-10	2.17E-06	8.33E-10
0.00175	0.00358	7.0476	0.19269	0.0121	0.00434	0.20794	-21.495	4.60436	0	nondepre	1.01E-06	2.17E-06	1.20473	0.00112	0.29677	62.5993	0.13635	0.00025	0.00015	2.29E-05	2.63E-12	56.3609	0.0008	97.8085	98.7286	0.02371	0.04793	1.96E-06	1.98E-10	1.51E-08
0.00173	0.00082	1.20682	0.19319	0.05234	0.00018	2.28589	0	52.8728	0	nondepre	0.0182	2.25E-05	1.61528	0.00939	0.18988	22.7004	0.00263	1.45E-06	1.76E-07	1.16E-06	1.92E-10	99.9952	0.03645	0.01266	98.7445	0.69724	95.7783	6.94E-08	8.73E-08	0.0001
0.00182	0.00085	8.79858	0.19573	0.05484	0.00016	2.30882	-24.593	50.2886	0	nondepre	0.00194	0.00014	0.16573	0.00669	0.05832	68.4282	0.00256	1.13E-06	2.46E-07	6.58E-08	2.38E-10	99.9959	0.14875	3.40E-07	98.5229	0.01924	99.8026	3.06E-08	2.44E-09	2.23E-05
0.00182	0.02339	0.42964	0.19261	0.00201	0.00452	0.19411	0	0.08385	0	nondepre	0.00453	0.0001	0.11424	4.52E-05	0.02115	0.15577	0.01149	5.41E-06	1.58E-11	7.82E-06	2.26E-09	0.72236	0.00023	7.55E-07	98.1781	0.03414	0.01253	6.75E-06	7.68E-09	0.00274
0.0015	0.00109	0.48116	0.20281	0.06706	0.00071	0.67666	0	53.0512	0	nondepre	0.00213	8.12E-08	0.00059	2.69532	1.01644	0.46179	0.03055	0.00036	2.56E-05	0.00033	7.12E-10	99.0155	0.00938	0.91358	73.5568	0.035	58.5258	4.20E-05	1.59E-10	4.77E-10
0.00182	0.00977	4.17331	0.19285	0.0049	0.00451	0.19302	-13.009	0.6982	0	nondepre	4.96E-06	1.54E-07	0.03153	0.04806	3.41392	43.5404	0.13175	1.86278	0.09724	0.00377	1.03E-08	27.1822	2.14E-05	9.94729	69.6129	0.0416	0.13675	0.00016	1.43E-11	1.80E-12
0.00183	0.00088	3.86284	0.19277	0.05385	0.00255	0.29635	-13.557	48.5944	0	nondepre	4.33E-07	4.92E-06	0.00691	0.02488	0.04997	49.1739	0.00323	0.00088	8.56E-05	1.04E-06	1.65E-09	99.1627	0.0005	1.49E-05	78.0668	0.00041	17.3207	1.31E-05	2.56E-09	5.38E-14
0.0018	0.0103	0.43344	0.1928	0.00444	0.00443	0.19961	0	0.487	0	nondepre	7.50E-07	0.0008	0.43794	2.09E-06	0.08368	0.20575	0.00443	0.00483	0.00047	6.56E-05	2.54E-07	25.085	0.00609	0.00154	49.7299	4.23E-05	0.32286	2.75E-05	8.74E-12	2.79E-10
0.00219	0.02067	0.39674	0.46813	0.00201	0.00425	0.193	0	0.09487	0	nondepre	0.99615	0.00015	0.02724	0.00043	0.01315	0.0002	0.0442	0.03971	8.50E-10	0.00792	6.85E-06	2.54082	9.08444	8.33E-07	92.2802	0.00383	0.00084	0.24543	4.03E-09	6.73E-05
0.00184	0.02396	0.42536	21.2326	0.00197	0.00452	0.1928	0	0.07954	0	nondepre	0.04916	6.10E-07	0.00879	0.02224	0.01945	0.00102	0.0002	0.00046	1.43E-08	0.01209	7.43E-09	0.02551	46.5159	77.7449	82.2648	0.0004	0.07019	0.07739	4.89E-10	8.62E-09

Figure 3. 3: Extracted facial features and emotion values

Similarly, the head pose-based analysis was required the horizontal (yaw/pan) and vertical (pitch/tilt) angles for implementing the head pose-based risk analysis classifier. Hence, the study employed the pre-trained model RealHePoNet for estimating the Head Pose. RealHePoNet includes a head detector to detect the head in an image/ video and a pose estimator to calculate yaw and pitch values. The RealHePoNet classifier itself holds the best accuracy [35]. The extracted values for title and pan angles (Figure 3. 4) were standardized before the training process of the classifier.

tilt	pan	dep
-15	-13	1
-16	-13	1
-15	-10	1
-12	-11	1
-10	-8	1
-20	-10	1
-18	-10	1
-17	-6	1
-11	-30	1
-12	-32	1
-8	-33	1
9	6	0
8	5	0
4	6	0
1	2	0
2	3	0
5	7	0
6	8	0
6	9	0
6	5	0

Figure 3. 4: Extracted title and pan angles

5.1.3. Model Implementation

As the initial step to implement a classifier for emotion-based early depression risk, the study trained the classifier on machine learning algorithms, Support Vector Machines (SVM), and random forest. The classifiers were implemented based on highly correlated parameters. However, the classifiers achieved nearly 70%-75% test accuracy. The study employed deep learning algorithms to improve the classifier performance by considering the classical machine learning algorithm as the baseline model.

The neural network classifier was defined and implemented by applying the Keras sequential model [36]. The classifier is consists of several dense layers (Figure 3. 5). There were 33 embedded features with the facial features and emotions. Hence, input dimensions of 33 neurons and output dimensions of 100 neurons have consisted of the input layer followed by three fully collected layers and the output layer. The first hidden layer output 50 neurons to the second hidden layer as the input, and 25 neurons were output from the second hidden layer. The third hidden layer utilized the output neurons of the previous layer and output ten neurons to the output layer.

```
patience = 50;
learning_rate = 0.0001
batch_size = 32
epochs = 500
def create_model():
    model = Sequential()
    model.add(Dense(100, input_dim=33))
    model.add(LeakyReLU())
    model.add(Dense(50))
    model.add(LeakyReLU())
    model.add(Dense(25))
    model.add(LeakyReLU())
    model.add(Dense(10))
    model.add(LeakyReLU())
    model.add(Dense(1, activation='sigmoid'))
    return model
```

Figure 3. 5: Define the emotion-based classifier

Similarly, the study trained the head pose-based early depression risk analyzer on machine learning algorithms, Support Vector Machines (SVM), and random forest based on highly correlated parameters. The SVM classifier could achieve 76% test accuracy, while 73% test accuracy has obtained with the random forest classifier. The study employed deep learning algorithms to improve the classifier performance by considering the classical machine learning algorithm as the baseline model.

The head pose-based early depression risk analyzer was implemented by employing the Keras neural network. The implemented classifier comprised an input layer followed by two hidden layers and an output layer (Figure 3. 6). Input dimensions of 11 neurons and output dimensions of 100 neurons have consisted in the input layer. The first hidden layer output 50 neurons to the next hidden layer as the input, and 25 neurons were output from the next hidden layer.

```
model = Sequential()
model.add(Dense(100, input_dim=11))
model.add(LeakyReLU())
model.add(Dense(50))
model.add(LeakyReLU())
model.add(Dense(25))
model.add(LeakyReLU())
model.add(Dense(1, activation='sigmoid'))

model.compile(loss=tf.keras.losses.binary_crossentropy,
              optimizer=tf.keras.optimizers.Adam(lr=learning_rate),
              metrics=['accuracy'])

# fit the keras model on the dataset
model.fit(X_train, y_train, epochs=200, batch_size=2, validation_data=(X_val, y_val), callbacks=[visualizer, earlystops], verbose=1)
```

Figure 3. 6: Define head pose-based classifier

The study trained the ensemble classifier combining the classifier outputs of the emotion-based early depression risk and head pose-based early depression risk. Initially, the study evaluated the ensemble classifier by employing the machine learning algorithm, decision tree. However, the classifier was able to achieve 76% test accuracy. As the next step, the analysis applied deep learning algorithms to improve the classifier performance based on the machine learning algorithm as the baseline model.

The study employed the Keras neural network architecture to define and implement the ensemble classifier. The implemented classifier comprised an input layer followed by two hidden layers and an output layer (Figure 3. 7). Input dimensions of 2 neurons and output dimensions of 100 neurons have consisted in the input layer. The first hidden layer output 50 neurons to the next hidden layer as the input, and 25 neurons were output from the next hidden layer.

```
model = Sequential()
model.add(Dense(100, input_dim=2))
model.add(LeakyReLU())
model.add(Dense(50))
model.add(LeakyReLU())
model.add(Dense(25))
model.add(LeakyReLU())
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])

# fit the keras model on the dataset
model.fit(X_train, y_train, epochs=200, batch_size=2, validation_data=(X_test, y_test), callbacks=[visualizer, earlystops])
```

Figure 3. 7: Define the Ensemble classifier

After each hidden layer, the classifiers activated it with the LeakyReLU activation function. The probability distribution over the defined classes was computed using sigmoid as the activation function in the final dense layer. The study used Adam as the optimizer for the classifiers. Binary cross-entropy was applied as the loss function in each classifier. Moreover, a batch size of 32 and a learning rate of 0.0001 was utilized in defining the emotion-based classifier. The head pose-based classifier and ensemble classifier applied the batch size of 2 and a learning rate of 0.0001 in training the training process.

The classifiers will load and predict the probability of depressive risk. The predicted values will be accessed using the API endpoints. The study employed Express JS and Node JS to implement the RESTful API [37]. The REST API Endpoint for saving the emotions and probability of depression risk is implemented using the POST method (Figure 3. 8).

```
router.post('/files', function (req, res, next)
```

Figure 3. 8: Post method to save emotions and probability of depression risk

Once the images are sent through the API, all the images (Figure 3. 9) will be processed to identify the probability of depressive risk based on emotions and head pose. The videos will be processed to separate frames (Figure 3. 10) before analyzing the depression risk. Once the frames are separated, the frames will be processed to identify the probability of depressive risk based on emotions and head pose.

```
count = req.files.images.length
```

Figure 3. 9: Uploaded image count

```
cv2.imwrite("/uploads/%d.jpg" % count, image)
success, image = vidcap.read()
count += 1
```

Figure 3. 10: process of frame seperation

Further, the endpoint for accessing the data for extensive summary is implemented using the GET method (Figure 3. 11).

```
router.get('/{user_id}/poses', function (req, res)
```

Figure 3. 11: GET method to retrieve emotions and probability of depression risk

The response is formatted as JSON (Figure 3. 12) object to retrieve the data points using the implemented REST API endpoint.

```
axios.get('http://ml-service/api/' + userId + '/poses?count=' + count)
    .then(response => {
        console.log(response.data)
        res.json({
            success: true,
            data: response.data.payload
        })
```

Figure 3. 12: Format of response output

Google cloud platform is used to deploy the emotion-based depression risk analyzer service. Deployed service is accessible through the designed REST API. The service is run on a virtual machine (Figure 3. 13) with 4 virtual CPUs and 32 GB memory.

Machine type
n2-highmem-4 (4 vCPUs, 32 GB memory)

Figure 3. 13: Type of the virtual machine

The OS image (Figure 3. 14) used in the virtual machine is ubuntu with the size of 50 GB.

Boot disk

Name	Image	Size (GB)	Device name	Type	Encryption	Mode	When deletir
ml-instance	ubuntu-2004-focal-v20210927	50	ml-instance	Balanced persistent disk	Google managed	Boot, read/write	Delete disk

Figure 3. 14: Details of the instance

5.1.4. Mobile Implementation

This section outlines the flutter-based mobile implementation procedure followed in the study. The emotion-based depression risk analyzer and head pose depression risk analyzer utilize the videos and images from a mobile phone camera. In addition, the end-user can upload videos and pictures by browsing the files on the mobile. The study employed the flutter package, image picker, to access the phone camera (Figure 3. 15). Further, the analysis obtains the gallery access (Figure 3. 16) to browse files.

```
var file =
    await _picker.pickImage(source: ImageSource.camera);
var file =
    await _picker.pickVideo(source: ImageSource.camera);
```

Figure 3. 15: Access camera

```
var file =
    await _picker.pickVideo(source: ImageSource.gallery);
var file =
    await _picker.pickImage(source: ImageSource.gallery);
```

Figure 3. 16: Access gallery

The user can verify the selected images and videos before uploading the files to the system. The mobile is designed to provide a playing video possibility (Figure 3. 17) before the uploading process. Note that the user can increase the speed of a video to verify the selected video.

```
void initState() {
  super.initState();
  _controller = VideoPlayerController.file(
    File(widget.path),
  )..initialize().then((_) { // VideoPlayerController.file
      chewieController = ChewieController(
        videoPlayerController: _controller,
      );
      setState(() {});
    });
}
```

Figure 3. 17: Implementation of playing video

The API request is triggered once the image or video is uploaded from the mobile. Once the procedure of assessing the depressive risk for a given image/ video, the user interface will display an extensive summary, including the emotions and depressive risk (Figure 3. 18).

```
Text(
  '${(double.parse(data["presentage"].toString()) * 100).toInt()}%',
  style: TextStyle(
    fontSize: 48,
  ), // TextStyle
), // Text
```

```
displayLevel(
  _size,
  "Joy",
  (double.parse(data["joy"].toString()) * 100)
      .toInt(),
),
displayLevel(
  _size,
  "Suprise",
  (double.parse(data["suprise"].toString()) * 100)
      .toInt(),
),
```

Figure 3. 18: Implementation of displaying saved values

The previous summary is available for users. The user can choose a particular day to view the earlier reports (Figure 3. 19) of emotions and depressive risk.

```
@override
void initState() {
  super.initState();
  selectedDate = DateTime.now();
}
```

```
Text(
  DateTime.now().difference(selectedDate).inDays == 0
      ? 'TODAY'
      : DateFormat("E d")
          .format(selectedDate)
          .toString()
          .toUpperCase(),
  style: TextStyle(
    fontSize: 18,
    fontWeight: FontWeight.bold,
    color: Color(0xff334856),
  ), // TextStyle
), // Text
```

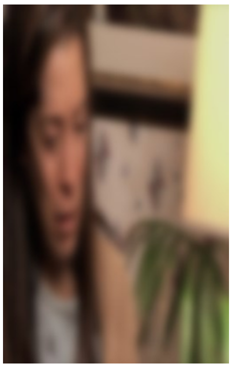
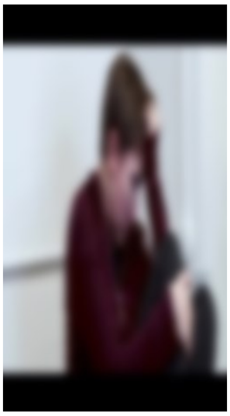
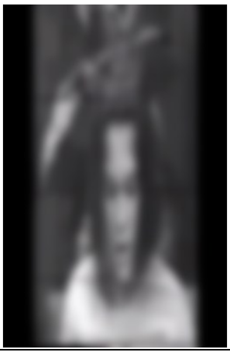
Figure 3. 19: Implementation of accessing date

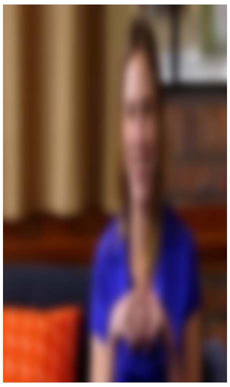
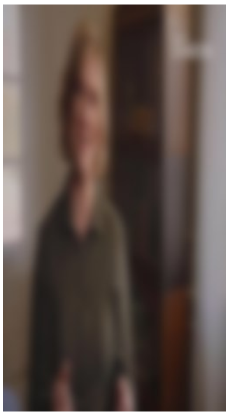
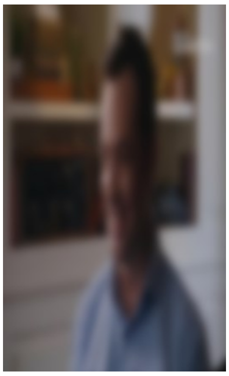
5.2. Testing

5.2.1. Backend test cases

The following Table 3.1 explains the test cases for backend development.

Table 3. 1: Test cases for backend development

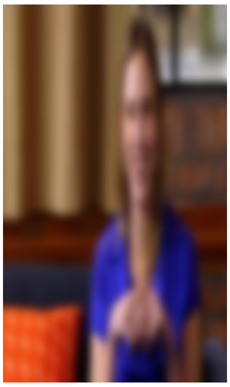
Test case ID	Test scenario	Test steps	Test Data	Expected Result	Actual result	Status
T001	Validate depressive risk of a depressed individual using the emotion-based classifier.	1. Select Image. 2. Input image to emotion-based classifier.		High depression risk	High depression risk (Depression Risk - 88%)	Pass
T002	Validate depressive risk of a depressed individual using the head pose-based classifier.	1. Select Image 2. Input image to head pose-based classifier.		High depression risk	High depression risk (Depression Risk - 66%)	Pass
T003	Validate depressive risk of a depressed individual using the ensemble classifier.	1. Select Image 2. Input image to ensemble classifier.		High depression risk	High depression risk (Depression Risk - 75%)	Pass

T004	Validate depressive risk of a non-depressed individual using the emotion-based classifier.	1. Select Image. 2. Input image to emotion-based classifier.		Low depression risk	Low depression risk (Depression Risk - 0%)	Pass
T005	Validate depressive risk of a non-depressed individual using the head pose-based classifier.	1. Select Image. 2. Input image to emotion-based classifier.		Low depression risk	Low depression risk (Depression Risk - 16%)	Pass
T006	Validate depressive risk of a non-depressed individual using the ensemble classifier.	1. Select Image. 2. Input image to emotion-based classifier.		Low depression risk	Low depression risk (Depression Risk - 10%)	Pass

5.2.2. Frontend test cases

The following Table 3.2 explains the test cases for frontend development.

Table 3. 2: Test cases for frontend development

Test case ID	Test scenario	Test steps	Test Data	Expected result	Actual result	Status
T001	Upload Images/ Videos	1. Access phone camera/ browse files from phone storage. 2. select file or capture image/ record video		Navigate to Extensive summary. Process the image and save data in Mongo DB.	Save data in Mongo DB. Displayed the extensive summry page.	Pass
T002	Check summary for Selected date.	1. Navigate to summary interface. 2. select date	Previous Date – 2021-10-10	Display summary for Selected date. (Depression Risk, Emotions)	Detailed view	Pass
T003	Display current details in the home screen	1. Navigate to home screen	-	Display details in home screen	Detailed summary in home screen.	Pass

6. RESULTS & DISCUSSION

6.1. Results

The following section explores the obtained test results and performance on classifiers. The section is divided into test results and the performance of the emotion-based classifier, head pose-based classifier, and ensemble classifier. Note that the test image for classifiers is included in Appendix E and blurred for privacy concerns.

6.1.1. Emotion-based Classifier

Although the study trained the Random Forest classifier on emotion and facial features, the analysis could achieve 70% test accuracy. The study trained the SVM classifier as well. However, the SVM could obtain 72% test accuracy. Therefore, the study employed a deep learning algorithm, a neural network, to enhance the classifier performance. The trained NN classifier achieved 82% test accuracy (Figure 4. 1) upon emotions and facial features.

```
1177/1177 [==============================] - 0s 40us/sample - loss: 0.3109 - acc: 0.8556
train loss: 0.31091043636671245,
train acc: 0.8555650115013123
471/471 [==============================] - 0s 18us/sample - loss: 0.4494 - acc: 0.7898
val loss: 0.44940553039249087,
val acc: 0.7898089289665222
706/706 [==============================] - 0s 15us/sample - loss: 0.4214 - acc: 0.8229
test loss: 0.4214278268324417,
test acc: 0.8229461908340454
```

Figure 4. 1: Accuracy of the emotion-based classifier

The accuracy curve is indicated in Figure 4. 2, and the loss curve is indicated in Figure 4. 3.

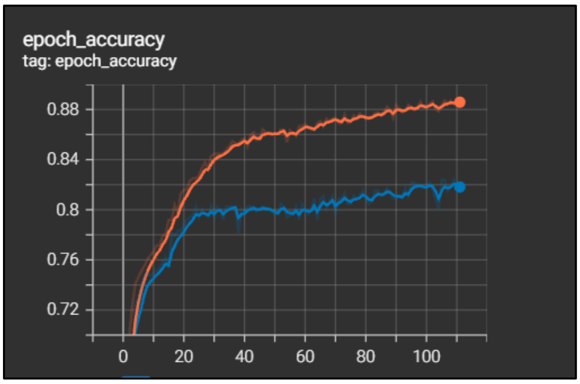


Figure 4. 2: Accuracy curve of the emotion-based classifier

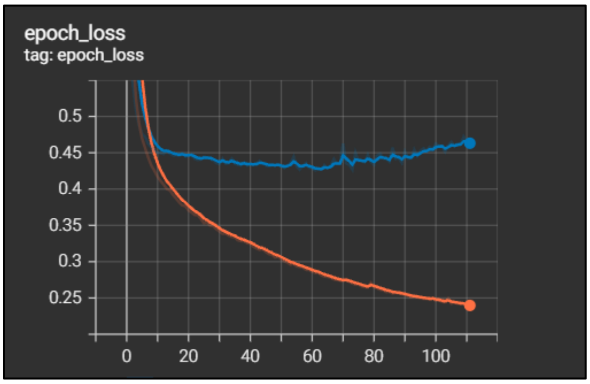


Figure 4. 3: loss curve of the emotion-based classifier

The emotion-based prediction for the image (Appendix E) is included in the following Figure 4. 4. According to the prediction results for this image, the result indicates the person is having no depression risk (0.8%).

	joy	sadness	disgust	contempt	anger	fear	surprise	dep	risk prob (%)
0	99.930618	9.203016e-07	0.000297	0.000107	0.000069	0.000024	0.207015	0	0.008255

Figure 4. 4: Output of the emotion-based classifier

6.1.2. Head pose-based Classifier

Although the study trained the Random Forest classifier on head pose angles, the analysis could achieve 73% test accuracy. The study trained the SVM classifier as well. However, the SVM could obtain 76% test accuracy. The study employed a deep learning algorithm, a neural network, to enhance the classifier performance. The trained NN classifier achieved 77% test accuracy (Figure 4. 5) upon head pose angles.

```
1177/1177 [==============================] - 0s 30us/sample - loss: 0.3971 - acc: 0.8063
train loss: 0.3971469218992235,
train acc: 0.8062871694564819
471/471 [==============================] - 0s 11us/sample - loss: 0.4657 - acc: 0.7962
val loss: 0.4657249729724447,
val acc: 0.7961783409118652
706/706 [==============================] - 0s 10us/sample - loss: 0.4973 - acc: 0.7705
test loss: 0.4973258236818881,
test acc: 0.7705382704734802
```

Figure 4. 5: Accuracy of the head pose-based classifier

T

The confusion matrix summarizes the performances of the classifier [38]. According to the confusion matrix (Figure 4. 6) matrix for the head pose classifier, nearly 320 data points are correctly classified as low risk of depression, and 220 data points are correctly classified as high risk of depression. The analysis considered 403 data points for low-risk depression and 303 data points for high-risk depression in the test phase.

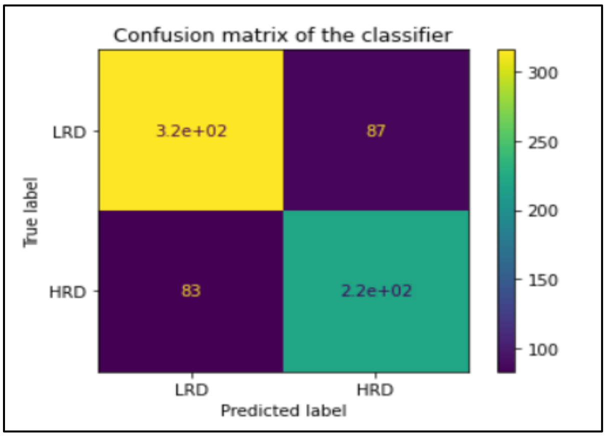


Figure 4. 6: Confusion matrix of head pose-based classifier

The head pose-based prediction for the image (Appendix E) is included in the following Figure 4. 7. According to the prediction results for this image, the result indicates the person is having a low depression risk (1.2%).

	tilt	pan	headpose_dep	risk prob (%)
0	-5.359996	-1.864105	0	0.012613

Figure 4. 7: Output of the head pose-based classifier

6.1.3. Ensemble Classifier

The study trained the Decision Tree classifier by considering the outputs from the emotion-based classifier and head pose-based classifier. However, the Decision Tree classifier could achieve 76% test accuracy. Hence, the study employed a deep learning algorithm, a neural network, to enhance the classifier performance to outperform the emotion-based classifier and head pose-based classifier. The trained NN classifier was able to outperform with 83% test accuracy (Figure 4. 8).

```
650/650 [==============================] - 0s 85us/sample - loss: 0.4174 - acc: 0.8369
test loss: 0.4174381712766794,
test acc: 0.8369230628013611
```

Figure 4. 8: Accuracy of the ensemble-based classifier

Classification reports consider as a metric for performance evaluation [39]. The classifier performance is evaluated using four metrics: precision, recall, F1 score, and support. The classifier based on emotions and head pose has obtained a 0.84 precision value for low-risk (class 0) and high-risk (class 1). As indicated in the classification report (Figure 4.9), the accuracy for the classifier is 0.84.

	precision	recall	f1-score	support
class 0	0.84	0.90	0.87	382
class 1	0.84	0.75	0.79	268
accuracy			0.84	650
macro avg	0.84	0.82	0.83	650
weighted avg	0.84	0.84	0.84	650

Figure 4. 9: Classification report of ensemble classifier

ROC curve implements to acquire a graphical representation of a classifier's performance [40]. It is widely used to analyze the performance of binary classification algorithms. The study evaluated the performance using the ROC curve (Figure 34. 10) since the analysis is a binary classification. The ensemble classifier obtained 0.82 for the ROC curve.

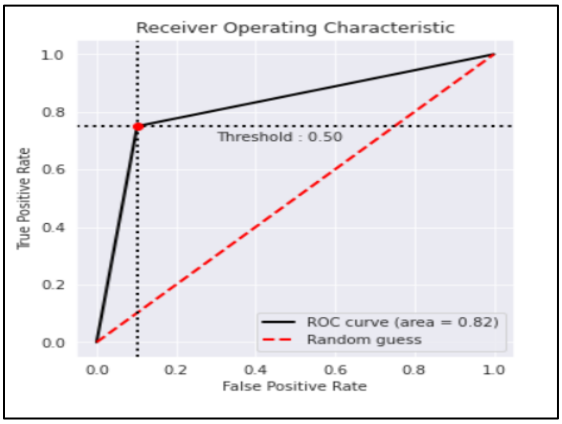


Figure 4. 10: ROC curve of ensemble classifier

6.2. Research Findings

As the initial step of identifying the depressive risk, the study employed random forest and SVM classifiers on emotion and facial features. However, the classifiers could not be able to achieve higher accuracy. The study applied Neural networks considering classical machine learning as the baseline model. Ultimately, the deep learning approach could outperform the random forest and SVM. The trained NN classifier achieved 82% test accuracy upon emotions and facial features.

Similarly, the study employed random forest and SVM classifiers on head pose angles. Even though the random forest classifier could not perform better, the SVM classifier could achieve 76% test accuracy. However, the study focused the performance enhancement. Hence, the analysis applied Neural networks considering classical machine learning as the baseline model. However, the NN classifier could outperform by 1% with the SVM classifier.

The study focused on ensemble classifier as the next step of identifying the depressive risk. The analysis combined the output values of the emotion-based classifier and head pose-based classifier and trained a model to obtain the accurate probability of depression risk. First, the data points were trained on a Decision Tree to find the performance of classical machine learning algorithms. The classifier could achieve 76%, which was similar to the performance of the head pose-based SVM classifier. The analysis applied Neural networks considering classical machine learning as the baseline model. Ultimately, the deep learning approach could outperform the other classification models with 83% test accuracy.

The following summarizes the findings of this study.

- The emotion-based classifier obtained high accuracy (82%) in identifying the depressive risk.
- Depressive risk can be assessed by combining Emotions, facial features, and head pose angles.
- The emotions and head pose combined classifier obtained high accuracy (83%) in identifying the depressive risk.

6.3. Discussion

The study mainly focused on analyzing depressive risk based on emotions, facial features, and head pose. The analysis has obtained approximately 75% test accuracy for head pose and emotion analysis despite training the machine learning algorithms, random forest, and SVM on highly correlated parameters. Hence, as a result, we explored the deep learning approach considering machine learning methods for our baseline model. Note that the classifiers were implemented by employing a neural network.

According to the test results, the emotion-based analyzer could achieve 82% accuracy, while the head pose-based analyzer was able to achieve 77% accuracy. The analysis combined outputs of the classifiers and trained a neural network to enhance the performance. The combined classifier could obtain 83% accuracy. However, the accuracy would have further improved on more data points. The current situation in the country has limited the collection of more data points for the study.

The following summarizes the limitations in this study.

- The classifiers were trained with considerably fewer amount of data points.
- The current situation in Sri Lanka is limited to collect more data points.
- Individuals were denied to share images and videos due to privacy concerns.

7. CONCLUSION

The study aimed to implement a mobile application by analyzing emotion, facial features, and head pose for assessing depressive disorder risk. The analysis trained the emotion-based classifier, which was able to achieve 82% accuracy. Further, the study trained the head pose-based classifier, which was able to achieve 77% accuracy. Finally, the analysis combined outputs of the classifiers and trained a neural network to enhance the performance. The combined classifier could obtain 83% accuracy.

REFERENCES

[1] Y. Wang *et al*., "Automatic Depression Detection via Facial Expressions Using Multiple Instance Learning," *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1933-1936, doi: 10.1109/ISBI45749.2020.9098396.

[2] World Health Organization, "Depression," *World Health Organization*. [Online]. Available: https://www.who.int/health-topics/depression. [Accessed: 10-Oct-2021].

[3] World Health Organization, "Suicide," *World Health Organization*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/suicide. [Accessed: 01-Oct-2021].

[4] W. C. de Melo, E. Granger and A. Hadid, "Combining Global and Local Convolutional 3D Networks for Detecting Depression from Facial Expressions," *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, pp. 1-8, doi: 10.1109/FG.2019.8756568.

[5] W. C. de Melo, E. Granger and A. Hadid, "Depression Detection Based on Deep Distribution Learning," *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, pp. 4544-4548, doi: 10.1109/ICIP.2019.8803467.

[6] W. Carneiro de Melo, E. Granger and A. Hadid, "A Deep Multiscale Spatiotemporal Network for Assessing Depression from Facial Dynamics," in *IEEE Transactions on Affective Computing*, doi: 10.1109/TAFFC.2020.3021755.

[7] W. Carneiro de Melo, E. Granger and M. B. Lopez, "Encoding Temporal Information For Automatic Depression Recognition From Facial Analysis," *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1080-1084, doi: 10.1109/ICASSP40776.2020.9054375.

[8] W. Carneiro de Melo, E. Granger and M. Bordallo Lopez, "MDN: A Deep Maximization-Differentiation Network for Spatio-Temporal Depression Detection," in *IEEE Transactions on Affective Computing*, doi: 10.1109/TAFFC.2021.3072579.

[9] B. Vimaleswaran and G. S. Ratnayake, "E - Therapy Improvement Monitoring Platform for Depression using Facial Emotion Detection of Youth," *2021*

Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 950-955, doi: 10.1109/ICICV50876.2021.9388447.

[10] X. Li, W. Guo and H. Yang, "Depression severity prediction from facial expression based on the DRR_DepressionNet network," *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 2757-2764, doi: 10.1109/BIBM49941.2020.9313597.

[11] A. Mulay, A. Dhekne, R. Wani, S. Kadam, P. Deshpande and P. Deshpande, "Automatic Depression Level Detection Through Visual Input," *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, 2020, pp. 19-22, doi: 10.1109/WorldS450073.2020.9210301.

[12] L. Yang, D. Jiang, W. Han and H. Sahli, "DCNN and DNN based multi-modal depression recognition," *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2017, pp. 484-489, doi: 10.1109/ACII.2017.8273643.

[13] M. Al Jazaery and G. Guo, "Video-Based Depression Level Analysis by Encoding Deep Spatiotemporal Features," in *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 262-268, 1 Jan.-March 2021, doi: 10.1109/TAFFC.2018.2870884.

[14] J. Joshi, "An Automated Framework for Depression Analysis," *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 630-635, doi: 10.1109/ACII.2013.110.

[15] S. Alghowinem, R. Goecke, J. F. Cohn, M. Wagner, G. Parker and M. Breakspear, "Cross-cultural detection of depression from nonverbal behaviour," *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2015, pp. 1-8, doi: 10.1109/FG.2015.7163113.

[16] J. F. Cohn *et al.*, "Detecting depression from facial actions and vocal prosody," *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009, pp. 1-7, doi: 10.1109/ACII.2009.5349358.

[17] S. P. Namboodiri and Venkataraman, “A computer vision based image processing system for depression detection among students for counseling,” *Indones. j. electr. eng. comput. sci.*, vol. 14, no. 1, p. 503, 2019.

[18] S. Alghowinem, R. Goecke, M. Wagner, G. Parkerx and M. Breakspear, "Head Pose and Movement Analysis as an Indicator of Depression," *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 283-288, doi: 10.1109/ACII.2013.53.

[19] J. Joshi, R. Goecke, G. Parker and M. Breakspear, "Can body expressions contribute to automatic depression analysis?," *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, 2013, pp. 1-7, doi: 10.1109/FG.2013.6553796.

[20] "MindDoc: Your mental health companion," *Google.com*. [Online]. Available: https://play.google.com/store/apps/details?id=de.moodpath.android&hl=en&gl=. [Accessed: 11-Oct-2021].

[21] MindDoc: Depression & Anxiety app review," *Onemindpsyberguide.org*, 15-Mar-2018. [Online]. Available: https://onemindpsyberguide.org/apps/minddoc-app-review/. [Accessed: 11-Oct-2021].

[22] "Your mental health companion," *Mymoodpath.com*. [Online]. Available: https://mymoodpath.com/en/. [Accessed: 11-Oct-2021].

[23] "iMoodJournal – Mood Tracking Mobile Application," *Imoodjournal.com*. [Online]. Available: https://www.imoodjournal.com/. [Accessed: 11-Oct-2021].

[24] "iMoodJournal," *Google.com*. [Online]. Available: https://play.google.com/store/apps/details?id=com.inexika.imood&hl=en&gl=US. [Accessed: 11-Oct-2021].

[25] "iMoodJournal," *Mindtools.io*. [Online]. Available: https://mindtools.io/programs/imoodjournal/. [Accessed: 11-Oct-2021].

[26] "IMoodJournal," *Onemindpsyberguide.org*, 07-Jan-2014. [Online]. Available: https://onemindpsyberguide.org/apps/imoodjournal/. [Accessed: 11-Oct-2021].

[27] "Sanvello," *Sanvello.com*. [Online]. Available: https://www.sanvello.com/. [Accessed: 11-Oct-2021].

[28] *Choosingtherapy.com. [Online]. Available: https://www.choosingtherapy.com/sanvello-app-review/. [Accessed: 11-Oct-2021].*

[29] "Progress Assessments," *Sanvello.com*. [Online]. Available: https://www.sanvello.com/product/progress-assessments/. [Accessed: 11-Oct-2021].

[30] "Daylio - journal, diary and mood tracker," *Daylio.net*. [Online]. Available: https://daylio.net/. [Accessed: 11-Oct-2021].

[31] "Daylio," *Mindtools.io*. [Online]. Available: https://mindtools.io/programs/daylio/. [Accessed: 11-Oct-2021].

[32] "UP! - Depression, Bipolar & Borderline Management," *Google.com*. [Online]. Available: https://play.google.com/store/apps/details?id=com.meemo_tec.bip_app&hl=en_. [Accessed: 11-Oct-2021].

[33] "New top 10 mental health apps for 2020," *Buckeyerecoverynetwork.com*, 06-Oct-2019. [Online]. Available: https://buckeyerecoverynetwork.com/10-best-mental-health-apps/. [Accessed: 11-Oct-2021].

[34] L. Kulke, D. Feyerabend, and A. Schacht, "A comparison of the affectiva iMotions facial expression analysis software with EMG for identifying facial expressions of emotion," Front. Psychol., vol. 11, p. 329, 2020.

[35] R. Berral-Soler, F. J. Madrid-Cuevas, R. Muñoz-Salinas, and M. J. Marín-Jiménez, "RealHePoNet: a robust single-stage ConvNet for head pose estimation in the wild," Neural Comput. Appl., vol. 33, no. 13, pp. 7673–7689, 2021.

[36] Keras Team, "The sequential model," Keras.io. [Online]. Available: https://keras.io/guides/sequential_model/

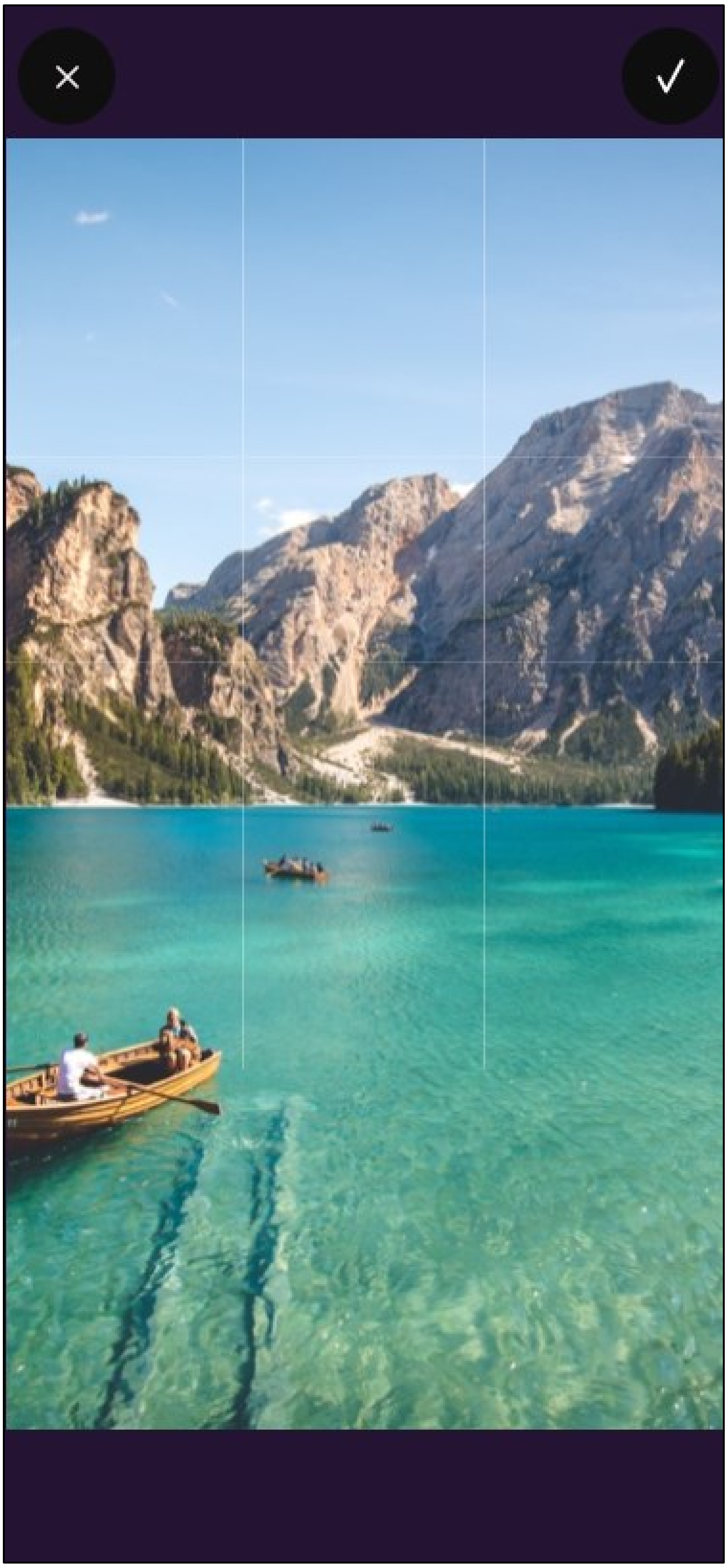
[37] "Building a RESTful API using Node.JS and MongoDB," *Nordicapis.com*, 25-Sep-2019. [Online]. Available: https://nordicapis.com/building-a-restful-api-using-node-js-and-mongodb/. [Accessed: 07-Oct-2021].

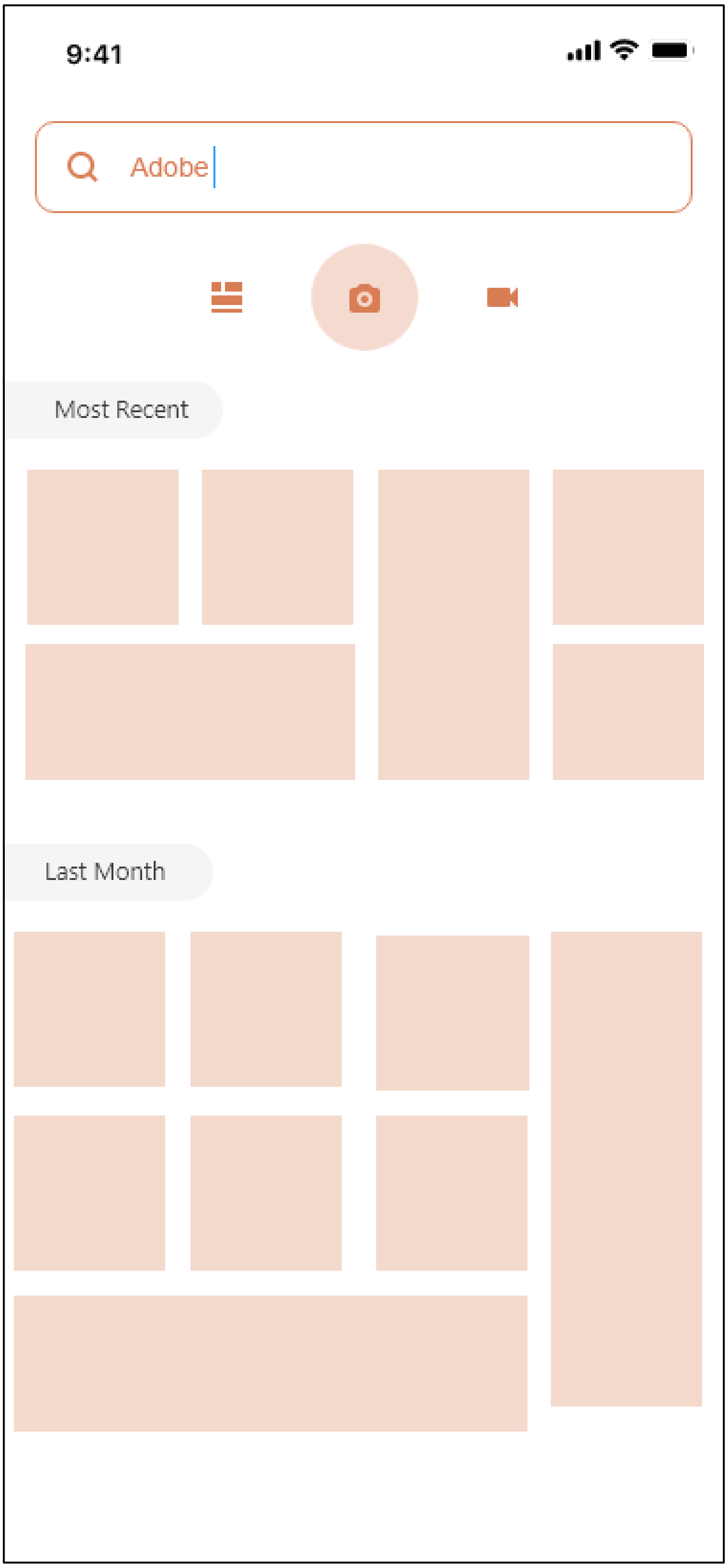
[38] K. Markham, "Simple guide to confusion matrix terminology," *Dataschool.io*, 26-Mar-2014. [Online]. Available: https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/. [Accessed: 08-Oct-2021].

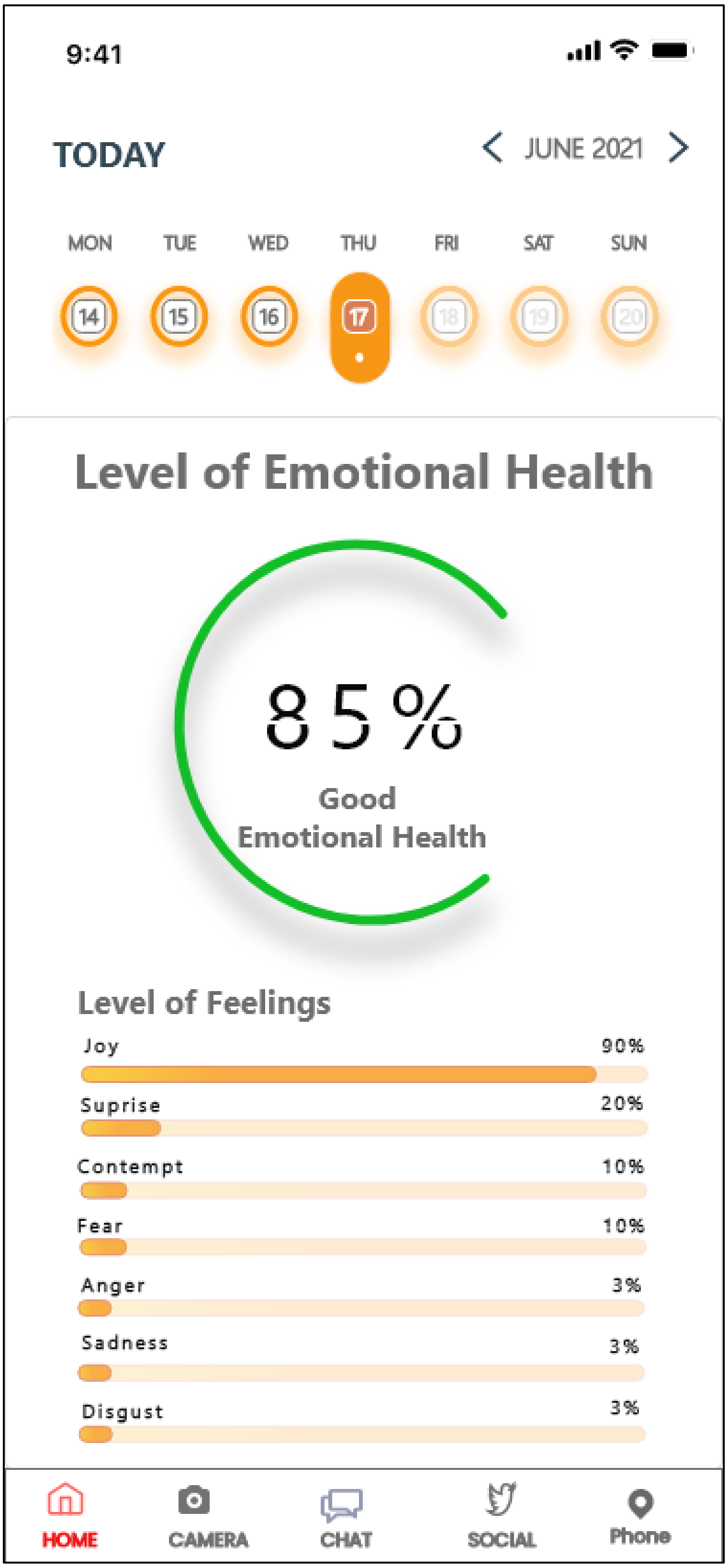
[39] A. Kharwal, “Classification report in machine learning,” *Thecleverprogrammer.com*, 07-Jul-2021. [Online]. Available: https://thecleverprogrammer.com/2021/07/07/classification-report-in-machine-learning/. [Accessed: 08-Oct-2021].

[40] D. Steen, “Understanding the ROC Curve and AUC,” *Towards Data Science*, 13-Sep-2020. [Online]. Available: https://towardsdatascience.com/understanding-the-roc-curve-and-auc-dd4f9a192ecb. [Accessed: 08-Oct-2021].

Appendix A: UI Wireframes




9:41
Adobe
Most Recent
Last Month

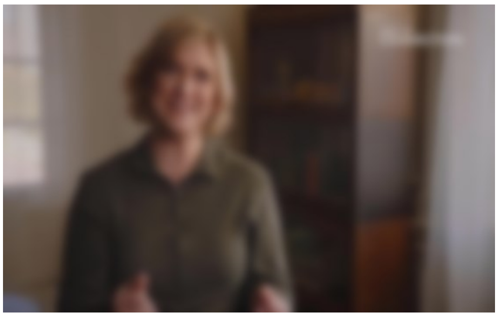


9:41
TODAY
JUNE 2021
MON
TUE
WED
THU
FRI
SAT
SUN
14
15
16
17
18
19
20
Level of Emotional Health
85%
Good
Emotional Health
Level of Feelings
Joy
90%
Suprise
20%
Contempt
10%
Fear
10%
Anger
3%
Sadness
3%
Disgust
3%
HOME
CAMERA
CHAT
SOCIAL
Phone

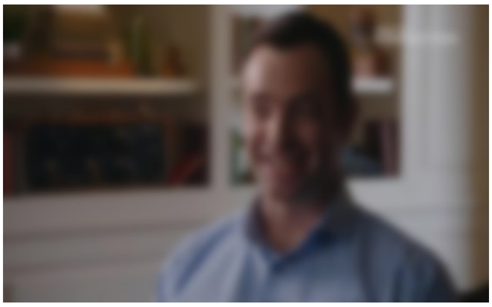
Appendix B: Test cases

Input	Output


```
"emotionPredict": {
    "0": 0.0000949957
},
"headposePredict": {
    "0": 0.1761597991
},
"ensemblePredict": {
    "0": 0.0977558717
}
```

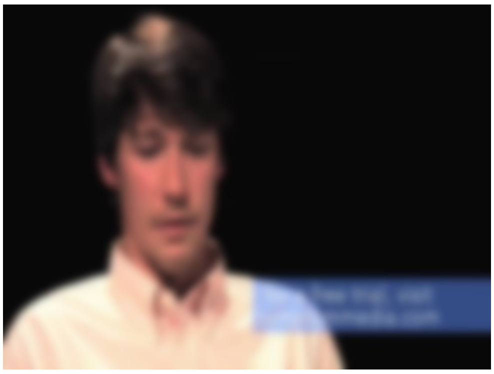




```
"emotionPredict": {
    "0": 0.0000222533
},
"headposePredict": {
    "0": 0.1691427231
},
"ensemblePredict": {
    "0": 0.0918806195
}
```


```
"emotionPredict": {
    "0": 0.0002066397
},
"headposePredict": {
    "0": 0.1799126416
},
"ensemblePredict": {
    "0": 0.1008273214
}
```

"emotionPredict": {
    "0": 0.8826130629
},
"headposePredict": {
    "0": 0.5981444716
},
"ensemblePredict": {
    "0": 0.6345549226
}



"emotionPredict": {
    "0": 0.7643820643
},
"headposePredict": {
    "0": 0.6658686399
},
"ensemblePredict": {
    "0": 0.7306531668
}

"emotionPredict": {
    "0": 0.999350369
},
"headposePredict": {
    "0": 0.6683338881
},
"ensemblePredict": {
    "0": 0.7507427931
}

Appendix C: Save Files in Local Phone Storage

```
class _UploadsScreenState extends State<UploadsScreen>
    with WidgetsBindingObserver {
  List files = [];
  @override
  void initState() {
    super.initState();
    getFilesFromDirectory();
  }

  getFilesFromDirectory() async {
    String dirToBeCreated = "MGuardian";
    Directory directory = await getApplicationDocumentsDirectory();
    var path = Directory('${directory.path}/$dirToBeCreated/');
    setState(() {
      files = path.listSync();
    });
  }

  @override
  void didChangeAppLifecycleState(AppLifecycleState state) {
    getFilesFromDirectory();
  }

  @override
  Widget build(BuildContext context) {
    final _size = MediaQuery.of(context).size;
    return Scaffold(
      body: Column(
        children: [
          Padding(
            padding: const EdgeInsets.all(defaultPadding),
            child: Container(
              padding: EdgeInsets.only(left: defaultPadding, top: 5, bottom: 5),
              decoration: BoxDecoration(
                color: Colors.white,
                borderRadius: BorderRadius.circular(20),
```

Appendix D: API Integration

```
getAnalysis(String id, DateTime date) async {
  var request = await http.get(ANALYSIS_API_ENDPOINT_GET +
      "api/" +
      id +
      "/?date=" +
      DateFormat("yyyy-MM-dd").format(date));
  if (request.statusCode == 200) {
    try {
      return jsonDecode(request.body)['payload'];
    } catch (e) {
      return null;
    }
  } else {
    return null;
  }
}

sendAnalysisRequest(String id, File file) async {
  // open a bytestream
  var stream = new http.ByteStream(file.openRead());
  stream.cast();
  // get file length
  var length = await file.length();

  // string to uri
  var uri = Uri.parse(ANALYSIS_API_ENDPOINT + "files");

  // create multipart request
  var request = new http.MultipartRequest("POST", uri);

  // multipart that takes file
  var multipartFile = new http.MultipartFile('images', stream, length,
      filename: basename(file.path));

  // add file to multipart
  request.files.add(multipartFile);
  request.fields['userId'] = id;
```

Appendix E: Test Image