EMOTION RECOGNITION AND SENTIMENT ANALYSIS FOR

RELATIONSHIP IMPROVEMENT.

TMP-2023-24-133

Project Proposal Report

Hewa Gamage Chanika Dilini

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

Department of Information Technology and Faculty of Computing.

Sri Lanka Institute of Information Technology

Sri Lanka

August 2023

EMOTION RECOGNITION AND SENTIMENT ANALYSIS FOR

RELATIONSHIP IMPROVEMENT.

TMP-2023-24-133

Project Proposal Report

Hewa Gamage Chanika Dilini – IT20651442

Supervisor: Dr. Dilshan De Silva.

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

Department of Information Technology and Faculty of Computing.

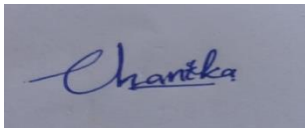
Sri Lanka Institute of Information Technology

Sri Lanka

August 2023

DECLARATION

We 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.

Group Member Name	**Student ID**	**Signature**
H.G.C.Dilini	IT20651442	

Signature of the Supervisor: **Date:** 2023/08/18

Signature of the Co-Supervisor **Date:** 2023/08/18

ABSTRACT

There are different kinds of connections friendships, family relationships, acquaintances, and romantic relationships. These relationships break down for various reasons. Relationship breakdown is a prevalent issue globally. Due to this, it is common today that people suffer from physical and mental suffering. where individuals may grapple with feelings of sadness, anger, loneliness, and confusion. Unable to bear the situation, the youth are motivated to take their life harm. University students and teenagers are more victims of this. Divorce rates range from 40% to 50% in developed countries. Relationship breakdowns in the world are increasing rapidly.

Some people try to seek professional help, which provides a safe space to explore the complexities of their feelings and to overcome these complexities to facilitate the healing process and personal growth. However, not everyone is willing to engage with counselors. There are various reasons for that, cultural stigma can be seen as a one point, a willingness to solve one's problems independently, etc. Some prefer to take the guidance of counselors, while others prefer alternative routes to recovery. Therefore, as an alternative method, a mobile application can be introduced, which is a more accurate and modern way to strengthen relationships. To run this mobile application, need to know what the feeling of the user is, this research component uses an emotion recognition method from the user's video call image processing. The mobile application empowers users to express emotions, track progress, set goals, and build resilience in a supportive environment. This proposed component is a better solution to solve these problems using deep learning techniques and convolutional neural network. Also, we use some technologies and tools python, video analysis by open cv, emotion recognition and build learning models by tensor flow , cloud computing and pre trained models are used for data and calculations. To ensure ethical data usage, the app will prioritize user privacy and consent. Personal data and conversations will be handled securely, and users will have full control over their data sharing preferences.

Keywords: Video call image Processing, Deep learning, Convolutional neural network, Open cv, Tensor flow.

TABLE CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
CNN	Convolutional Neural Network
VGG	Visual Geometry Group
FER	Facial Expression Recognition
CK+	Cohn-Kanade
SDLC	Software Development Life Cycle
AI	Artificial Intelligence
FC	Fully Connected
WBS	Work Breakdown Structure

1. INTRODUCTION

1.1 Background & Literature survey.

Emotion recognition plays an important role in understanding human behavior and interactions. With the increasing use of digital communication platforms, the need to interpret emotions from nonverbal cues has increased dramatically. Advances in image processing techniques, machine learning, and deep learning are opening new possibilities for accurately identifying human emotions through facial expressions, helping to improve relationships and improve the human experience. Facial expressions are the primary means of conveying emotions in interpersonal communication. The relationship between facial expressions and emotions has been studied extensively in psychology and neuroscience[1]. Researchers like Paul Ekman have shown that specific facial muscle movements are associated with specific emotions, forming the basis of automatic emotion recognition systems.

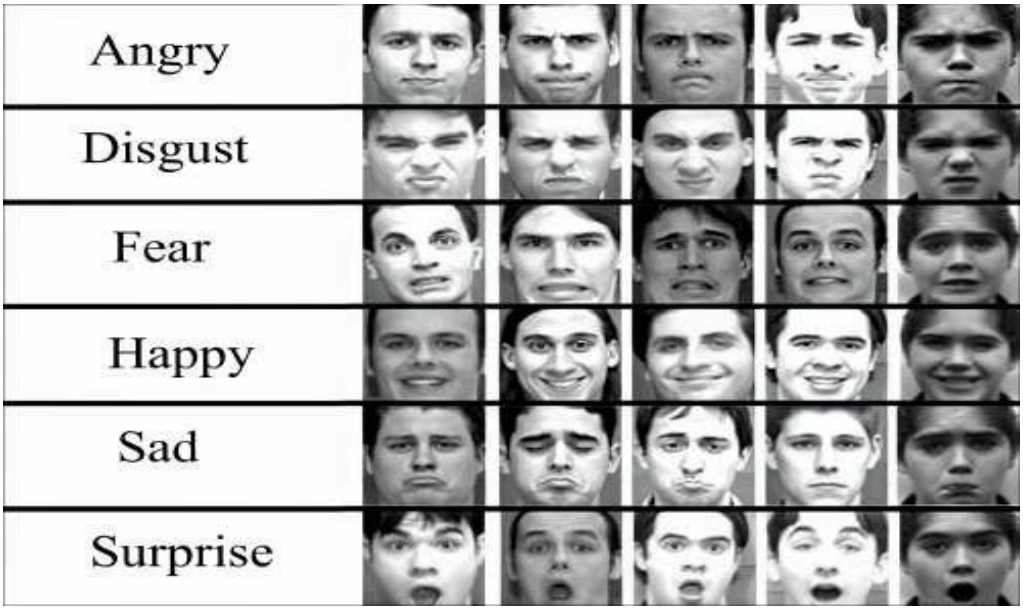


Figure 1. 1 : People expressions different emotions on their faces

Image processing, an important part of this effort, involves the collection, preprocessing, and analysis of image data. Modern computer vision techniques supported by deep learning have been proven to be effective in extracting features related to accurate perspective recognition, motion tracking, and emotion recognition. Open-source libraries like OpenCV have enabled researchers to develop powerful facial recognition and tracking processes. A major challenge in emotion detection is building a model capable of detecting subtle changes in facial expressions. Deep learning methods, especially convolutional neural networks (CNNs)[2],

have achieved considerable success in this area. The researchers trained CNN on a wide range of facial expressions and learned to recognize patterns associated with emotions such as happiness, sadness, anger, and surprise. Pre-trained CNN architectures such as VGG, ResNet, and Inception have been optimized for emotion detection and show promising results.

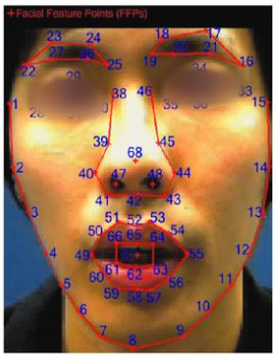
Extracted facial feature points (FFPs)	Facial regions	FAPs Num.	Euclidean distance between FFPs	Comparing FFPs displacement with neutral frame
	Eyebrows	1, 2	$D_{vertical,1}(22, 30)$, $D_{vertical,2}(16, 35)$	$D_{v,1\_Neutral}$-$D_{v,1}$, $D_{v,2\_Neutral}$-$D_{v,2}$
		3, 4	$D_{vertical,3}(25, 30)$, $D_{vertical,4}(19, 35)$	$D_{v,3\_Neutral}$-$D_{v,3}$, $D_{v,4\_Neutral}$-$D_{v,4}$
		5, 6	$D_{vertical,5}(22, 28)$, $D_{vertical,6}(16, 33)$	$D_{v,5\_Neutral}$-$D_{v,5}$, $D_{v,6\_Neutral}$-$D_{v,6}$
		7, 8	$D_{vertical,7}(23, 28)$, $D_{vertical,8}(17, 33)$	$D_{v,7\_Neutral}$-$D_{v,7}$, $D_{v,8\_Neutral}$-$D_{v,8}$
		9, 10	$D_{vertical,9}(25, 28)$, $D_{vertical,10}(19, 33)$	$D_{v,9\_Neutral}$-$D_{v,9}$, $D_{v,10\_Neutral}$-$D_{v,10}$
		11, 12	$D_{vertical,11}(23, 30)$, $D_{vertical,12}(17, 35)$	$D_{v,11\_Neutral}$-$D_{v,11}$, $D_{v,12\_Neutral}$-$D_{v,12}$
		13	$D_{hl,13}(19, 25)$	$D_{h,13\_Neutral}$-$D_{h,13}$
	Eyes	14, 15	$D_{vertical,14}(29, 31)$, $D_{vertical,15}(34, 36)$	$D_{v,14\_Neutral}$-$D_{v,14}$, $D_{v,15\_Neutral}$-$D_{v,15}$
		16, 17	$D_{vertical,16}(28, 49)$, $D_{vertical,17}(33, 55)$	$D_{v,16\_Neutral}$-$D_{v,16}$, $D_{v,17\_Neutral}$-$D_{v,17}$
		18, 19	$D_{horizontal,18}(28, 30)$, $D_{horizontal,19}(33, 35)$	$D_{h,1\_18Neutral}$-$D_{h,18}$, $D_{h,19\_Neutral}$-$D_{h,19}$
	Nose	20, 21	$D_{vertical,20}(52, 68)$, $D_{vertical,21}(58, 68)$	$D_{v,20\_Neutral}$-$D_{v,20}$, $D_{v,21\_Neutral}$-$D_{v,21}$
		22, 23	$D_{vertical,22}(49, 68)$, $D_{vertical,23}(55, 68)$	$D_{v,22\_Neutral}$-$D_{v,22}$, $D_{v,23\_Neutral}$-$D_{v,23}$
	Mouth	24, 25	$D_{vertical,24}(52, 58)$, $D_{horizontal,25}(49, 55)$	$D_{v,24\_Neutral}$-$D_{v,24}$, $D_{h,25\_Neutral}$-$D_{h,25}$
	Facial Contours	26, 27	$D_{horizontal,26}(5, 58)$, $D_{horizontal,27}(11, 58)$	$D_{h,26\_Neutral}$-$D_{h,26}$, $D_{h,27\_Neutral}$-$D_{h,27}$
		28, 29	$D_{horizontal,28}(2, 68)$, $D_{horizontal,29}(14, 68)$	$D_{h,28\_Neutral}$-$D_{h,28}$, $D_{h,29\_Neutral}$-$D_{h,29}$
		30	$D_{vertical,30}(8, 68)$	$D_{v,30\_Neutral}$-$D_{v,30}$

Figure 1. 2 : Extracted facial feature points

Notable datasets such as CK+ (Cohn-Kanade) and FER2013 (Facial Expression Recognition 2013) have enabled progress in this area. These datasets include annotated images that capture a variety of emotions and provide material for training and validating emotion detection models. Transfer learning, which uses a pre-trained model as a base and adapts it to the target task, has become a popular technique due to the limited availability of sensor-specific data. contact, labeled. Researchers are also working on multimodal approaches that combine visual cues with voice and text to improve the accuracy of emotion recognition. This combination of modalities recognizes the complexity of human emotional expression and the potential for more complete information across multiple channels.

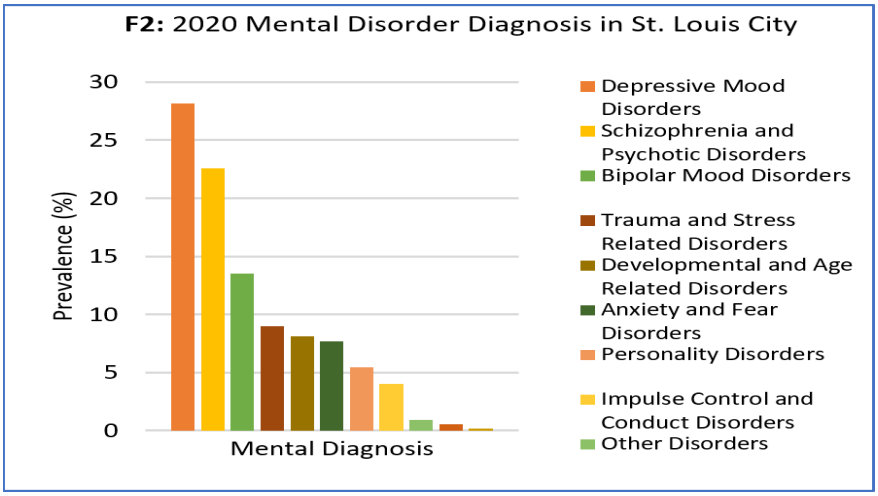


Figure 1. 3 : Survey report on the up to 2020 mental disorder diagnosis

The application of emotion recognition through image processing extends beyond individual interactions. It shows promise in several areas such as healthcare, marketing, and entertainment. In medicine, monitoring a patient's emotional state can help identify mental health problems at an early stage. In marketing, understanding consumer psychology can inform advertising strategy and product design. Additionally, the entertainment industry can use emotion recognition to tailor content to viewers' emotional responses. However, further development of this technology has ethical considerations. Privacy issues arise from the collection and analysis of personal data, especially facial images. Balancing technological innovation and individual rights is essential to ensure responsible use of emotion recognition systems.

In conclusion, image-based emotion recognition for mood analysis is a dynamic field with much potential. The convergence of computer vision, deep learning, and psychology has enabled the development of models capable of deciphering emotions from facial expressions. While ethical challenges and considerations remain, the prospect of enhancing relationships and understanding human emotions through technology remains an exciting avenue for future research. Based on the research background and survey discussed above, the user's sentiment will be identified more accurately and meticulously, and the user will be given the necessary advice. There are very few research reports on appropriate control measures to reduce the stress experienced by users. This is the first attempt of its kind in Sri Lanka and the first research to provide the necessary advice by identifying the user's emotions.

1.2 Research Gap

According to the literature survey conducted above, since only a few studies have been conducted, there is a need to conduct research to improve relationships through emotional recognition and sentiment analysis. Research has been done to identify feelings and emotions through the currently available systems, and those systems do not fully meet the specific requirements for relationship improvement.

"A" research report uses two techniques to simultaneously identify facial expressions in video streams. For this, bilinear pooling B-CNN and fusion feature F CNN are used. This technique is more efficient and more accurate. The method of deep learning, known for its great performance in image classification, becomes essential. The Convolutional Neural Networks (CNN), which have proven their effectiveness in the problems of image recognition and classification. CNN is composed of several layers. Each layer receives information from the preceding layer, interprets this information and sends the result to the next neural layer. Once the system has been trained with enough images, a resolution model is then obtained which is then tested with new images to check its relevance and performance. During the learning and testing phases, each input image passes through a series of convolutional layers with filters (Kernals), pooling layers, fully connected layers (FC) to finally reach the SoftMax function allowing to classify an object with probabilistic value between 0 and 1. This system can quickly and accurately identify human emotions[2]. However, by this system cannot provide several activities such as after relationship breakdown how their relationship improvement and effective management of them. Not only that, but the user also cannot manage their feelings and cannot improve their relationship and interact with the relevant parties. In this regard this research can be misleading.

The research paper "B" presents a system that recognition facial emotion using Convolutional Neural Network (CNN) to classify that. Various gray scale pictures from the dataset and real time videos are taken as input. Then feature extraction is done using the series of convolution and pooling layers of CNN and classification is done using SoftMax layer. To overcome the overfitting problem of the model, various procedures including dropout, cluster standardization and L2 regularization are used. Experiments are performed on Facial expression dataset image folders (fer2013) dataset and achieved better accuracy with their model in predicting the individual emotions[3]. In addition to this, the model is also used to predict the emotion of each image in the real time video data with good performance.

However, this study, like Study "A", is limited to identifying feelings and emotions. Another problem arises here, which is the use of real time images and recordings to identify facial

expressions. It's not realistic because it's not realistic if they can take pictures or record in a situation caused by a breakup or emotional breakdown. Also, once they identify the feeling, they are unable to re-establish the relationship or know what stage they are in. However, the study does not address how to improve their relationship after feelings breakdown or severity classification.

In the published research “C”, a system was developed to analyze emotion using content like this from social networking websites and/or photo-sharing websites like Flickr, Twitter, Tumblr, etc. R6 can offer insights into the general emotion of people[4]. This study attempts to determine five categories such as affection, happiness, cruelty, fear, and sadness an image belongs to. As images play such a significant role in contemporary culture, it is advantageous to understand the various feelings and sentiments that an image may evoke and to quickly assign emotional labels to them, such as happy, affection, sorrow, rage, neutral, etc. A special convolutional neural network and a pretrained convolutional neural network model is used here to detect feelings and analyze feelings of various constant images with a user defined GUI. Sentiment analysis has been used more frequently on CNN in recent years. But there is no information available on how to improve their relationship.

However, the current system lacks a mobile or web application to display find solutions such as relationship goal trackers that can improve relationships after identifying emotions, and as know the status of user's relationships.

Table1.1: Comparison between existing system

Research	Emotion Recognition	Display remedies to control after identifying the emotion.	Progression level of user emotion.	Mobile application, Web application.
Research A	✓	✗	✗	✗
Research B	✗	✗	✗	✗
Research C	✓	✗	✗	✗
Proposed System	✓	✓	✓	✓

1.3 Research Problem

In the world of understanding emotions and improving relationships, there is a big challenge. This challenge aims to find out how people feel when using images and videos, especially during phone live video calls. It's like making computers understand emotions like happiness or sadness on human faces. But here's the tricky part: What works well in the lab doesn't work well in real situations.

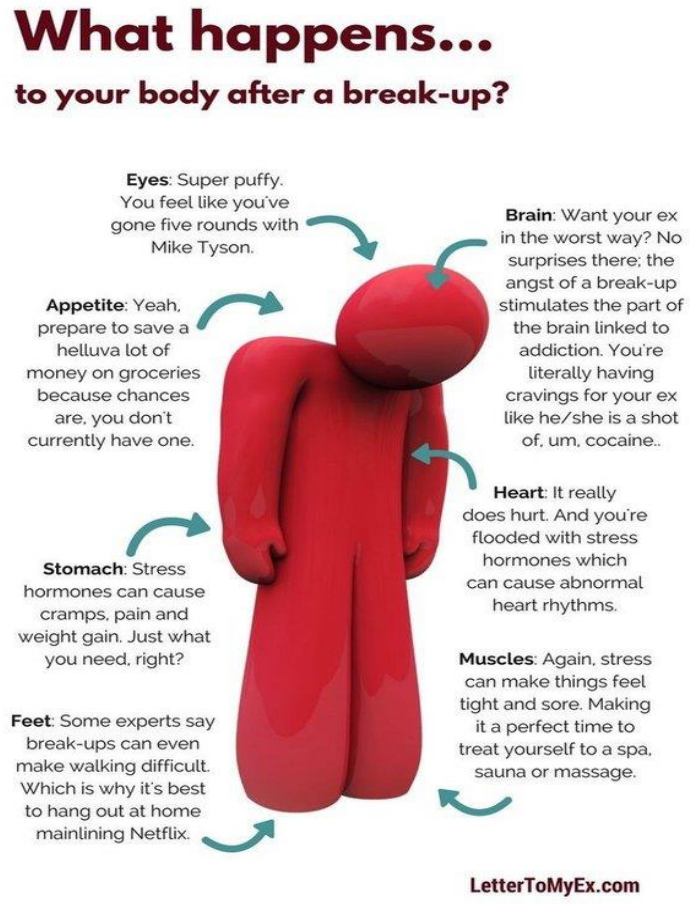


Figure 1.4: Survey of after breakup what happened to your body

Lighting, the way people move their faces, and camera angles make it difficult. So, the problem is how to make these computer programs better recognize emotions in real-time video calls. What's even more difficult is that most of the work is focused on images only. Imagine if we could also use what people say and how they pronounce it to understand their feelings. The combination of all these images, words, and sounds can help computers better understand emotions[4]. Also consider people from different parts of the world. They express their feelings in different ways due to their culture. The question is, how do these computer programs understand cultural sentiment? We want these programs to be smart enough to work for everyone.

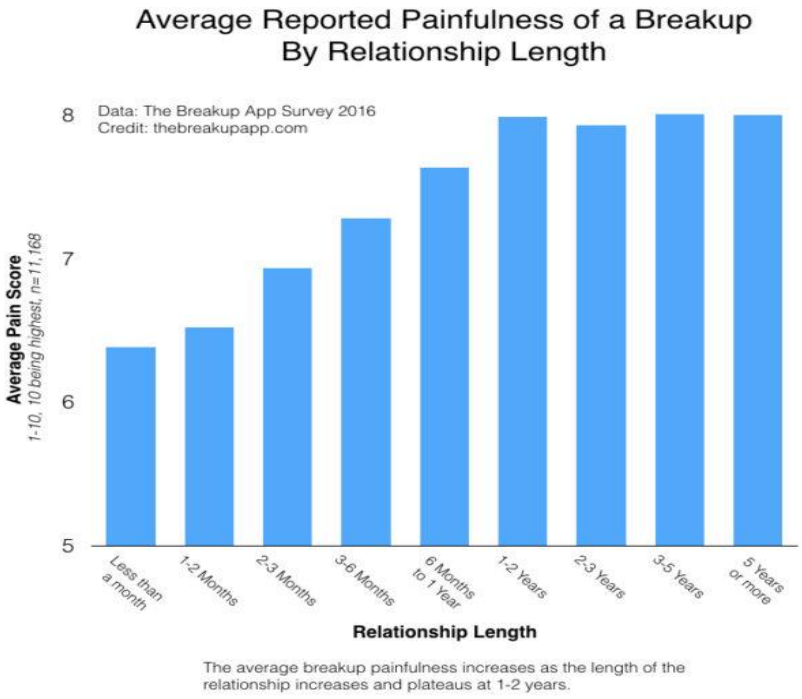


Figure 1.5: Survey report on average painfulness of a breakup

Finally, we must pay attention to people's privacy[5]. Imagine if a computer could tell how, you feel just by looking at you. It sounds unbelievable, but it also means that the computer is watching you closely. So, the challenge is to make computers understand emotions without invading people's privacy or making them feel manipulated. So, the big question is how do you create computer programs that can understand emotions during live video calls on phones? It will help people in a relationship communicate better, even if they don't meet face-to-face. It's like teaching a computer to read emotions like a friend. And that's what we must understand.

2. OBJECTIVES

2.1 Main Objectives

The main objective of this component is to develop a mobile application, an accurate and real-time emotion recognition system using image processing techniques using mobile video calls data. The goal is to enable computers to understand users' emotions during conversations, enhancing relationship communication and offering personalized suggestions for improvement.

2.2 Specific Objectives

To accomplish the goal outlined above, four specific goals must be completed.

Real-Time emotion recognition Model Development: Create an efficient real-time emotion recognition model using image processing techniques. This involves selecting appropriate deep learning models for facial detection and expression recognition and training a Convolutional Neural Network (CNN) for accurate emotion classification.

Multimodal Integration: Investigate the fusion of visual, auditory, and textual cues to enhance emotion recognition accuracy. Develop a system that combines these cues to provide a more holistic understanding of users' emotional states during mobile video calls.

Cross-Cultural Sensitivity: Adapt the model to recognize emotions across diverse cultural backgrounds. Fine-tune the system to account for variations in emotional expressions, ensuring its effectiveness for users from different cultural contexts.

User Interface and Integration using Computer Vision (OpenCV): Design an intuitive user interface that displays real-time emotion recognition results during video calls, using computer vision techniques like OpenCV. Integrate the emotion recognition module into a broader relationship improvement framework to offer personalized suggestions for enhancing communication based on recognized emotions.

Goal Tracking and Emotional Growth: Integrate goal-tracker in the app for users to enhance relationship skills. Real-time emotional analysis offers progress insights and advice, promoting self-awareness and positive changes, fostering emotional growth, and improving relationships[6].

3. METHODOLOGY

The methodology of this component revolves around leveraging video call feed data to accurately identify users' emotions and sentiments, which can contribute to enhancing their relationships. The process can be broken down into several key stages, each serving a crucial purpose in achieving the desired outcomes. Firstly, data collection plays a pivotal role. A diverse dataset of video call sessions, involving individuals engaged in conversations, will be amassed[2]. This dataset should encompass a range of emotions and sentiments that people commonly exhibit during interactions, such as happiness, sadness, anger, and surprise. The dataset will serve as the foundation for training the emotion recognition model.

3.1 System Architecture

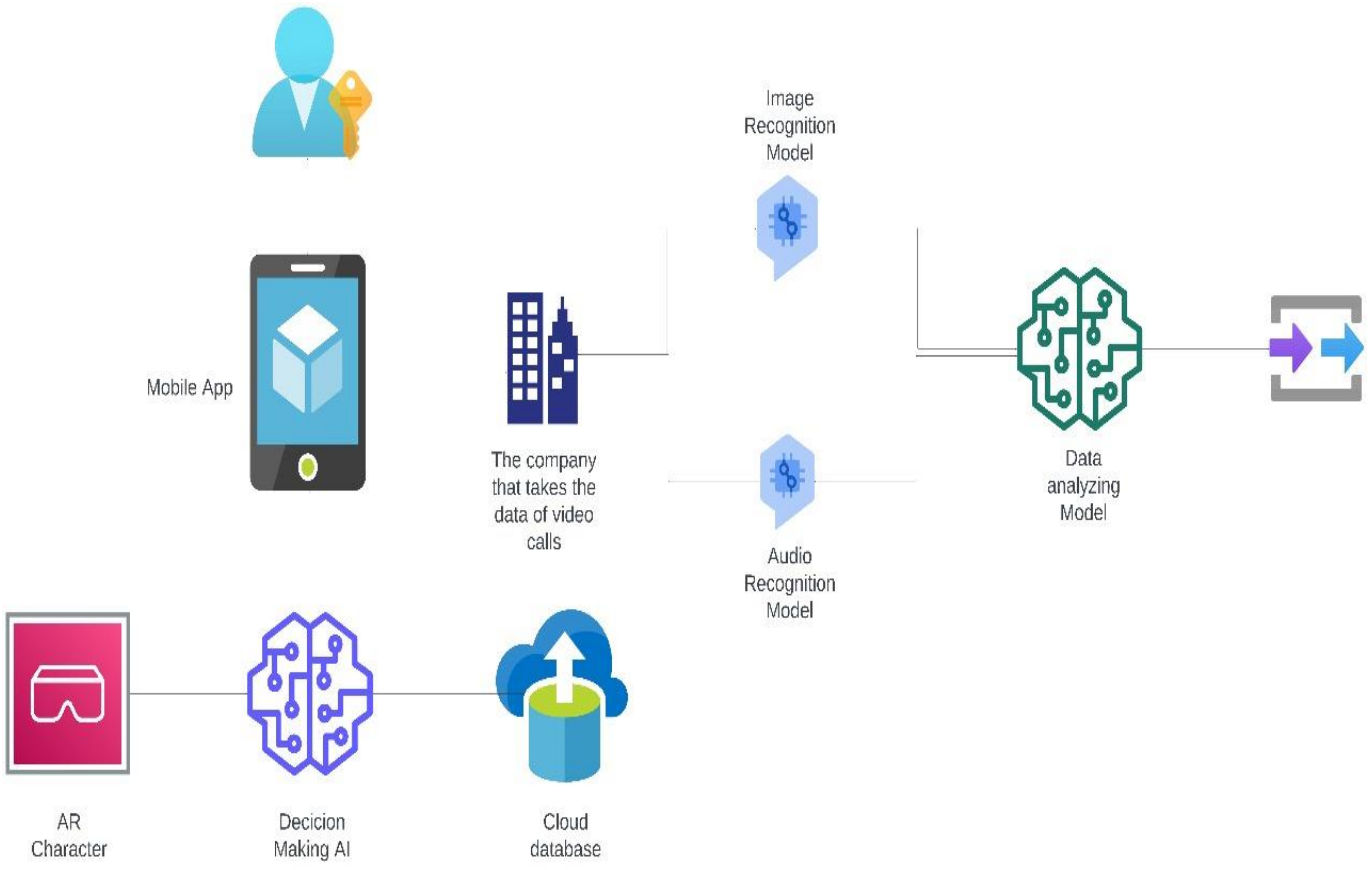


Figure 3.1: Recognize an emotion using image processing system architecture

Once the dataset is assembled, the next step is data preprocessing. Video call feed data is rich and dynamic, capturing various facial expressions, movements, and background elements. The data must be cleaned, aligned, and standardized to ensure consistency across all samples.

This step is crucial for reducing noise and improving the accuracy of subsequent analysis. The subsequent stage involves feature extraction. From the preprocessed data, relevant facial features will be extracted. These features could include aspects like eye movements, eyebrow positioning, and mouth curvature. By focusing on these distinctive cues, the emotion recognition model can better discern the nuanced expressions associated with different emotions.

With the feature extraction complete, the heart of the methodology comes into play: training the emotion recognition model. Machine learning techniques, such as deep learning and convolutional neural networks (CNNs), will be employed. These algorithms learn patterns from the labeled dataset, enabling the model to identify emotions based on the extracted facial features. The model will be trained to accurately categorize emotions, associating certain patterns with specific feelings. After rigorous training, the model's performance will be evaluated using a separate set of data. This validation process ensures that the model generalizes well to new, unseen examples. Adjustments to the model's architecture or parameters can be made to enhance its accuracy and robustness[5].

Once the trained model is deemed satisfactory, it's time for real-time application. During a video call, the model will continuously analyze the facial expressions of the participants. These expressions will be processed through the trained model, which will identify the predominant emotion being displayed. Upon detecting the emotion, the app's interface will display relevant information to the user. For instance, if the model identifies signs of tension or frustration, the app might suggest engaging in active listening or taking a brief pause to cool off. The app could also provide real-time feedback, encouraging users to express themselves more positively or calmly.

Furthermore, the goal tracker feature will be integrated into the app. This component aligns with the broader goal of relationship enhancement. Users will be able to set personalized relationship-related goals, such as practicing active communication or showing appreciation[6]. The app will periodically prompt users to reflect on their progress toward these goals, fostering self-awareness and encouraging positive changes in behavior.

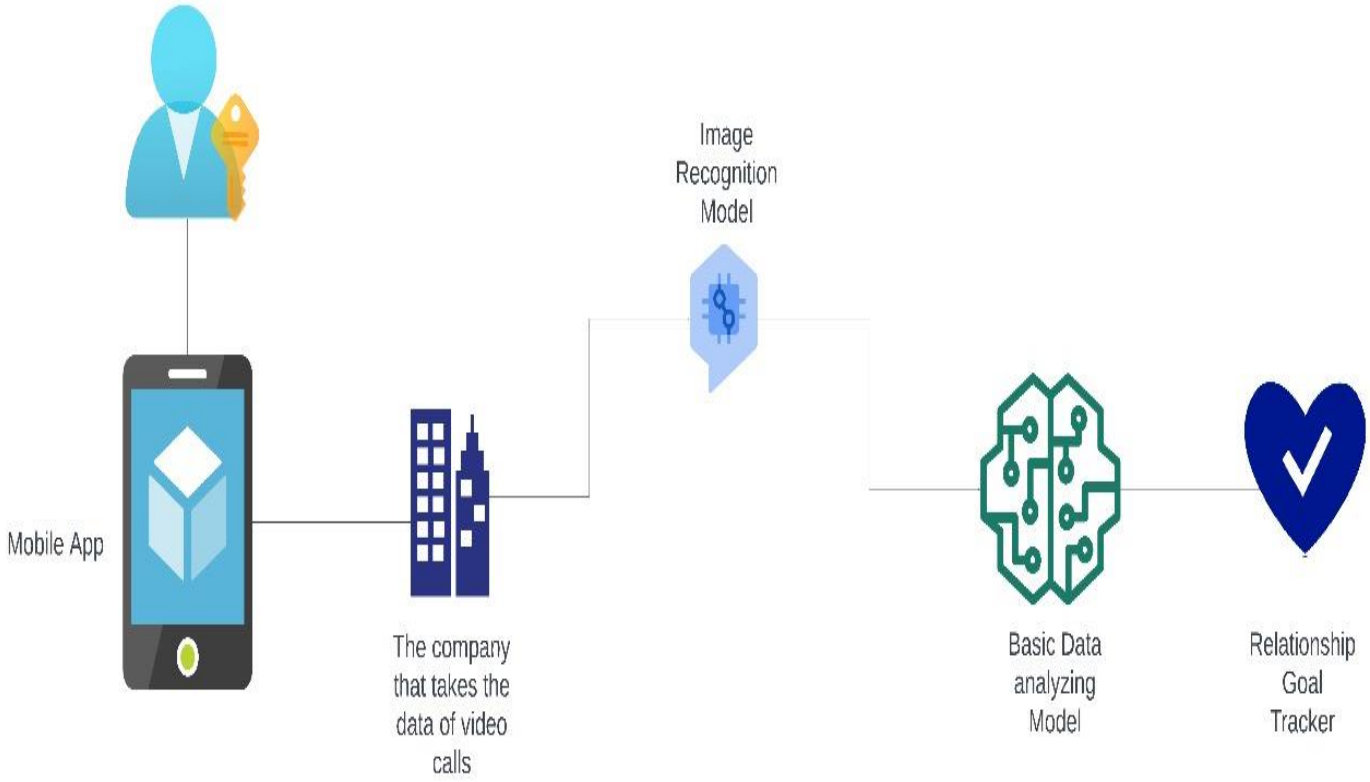


Figure 3.2: Recognize an emotion using image processing system architecture

In conclusion, this methodology capitalizes on video call feed data to facilitate emotion recognition and sentiment analysis. By meticulously collecting, preprocessing, extracting features, training, validating, and applying the emotion recognition model, the app aims to empower users with real-time insights into their emotional dynamics during conversations. The incorporation of the goal tracker feature further bolsters the app's potential to contribute to relationship improvement by promoting self-awareness and constructive communication habits.

Table 3.1: Technologies, techniques, architectures, and algorithms used.

Technologies	React Native, Expo, Keras, OpenCV, Flask, Node.js, ML libraries
Techniques	Deep Learning, Real-time Processing, CrowdSourcing
Architectures	ResNet
Algorithms	CNN, Machine learning algorithms, Facial Landmark Detection

3.1.1 Software Solution

The Software Development Life Cycle (SDLC) serves as a methodical framework for developing software with precision and coherence. In conventional SDLC, developers encounter limitations in revisiting earlier stages when requirements shift, necessitating adherence to a fixed sequence of steps. Yet, in an agile approach, such constraints are eased, as agility emphasizes flexibility and adaptability. Among various agile frameworks, Scrum stands out as a particularly potent one. Functioning as an agile project management methodology, Scrum offers a lightweight yet robust solution for managing intricate and evolving challenges. The illustration in Figure 3.3 vividly portrays the core six processes that constitute the foundation of agile methodology, embodying its fluid and responsive nature.

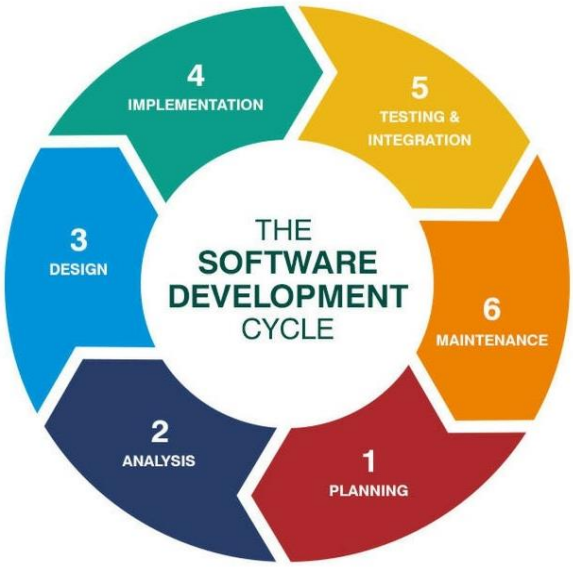


Figure 3.3: Agile methodology(SDLC)

1.Requirement Gathering

This phase involves obtaining video call feed data solely from the Meta company, eliminating the need for direct user input. Collaborating with Meta allows access to anonymized, consented video call streams, ensuring data privacy and regulatory compliance. Gathered data, encompassing facial expressions and vocal cues, forms the essential dataset for emotion recognition. While user involvement isn't required, ethical considerations remain paramount. This approach streamlines data acquisition, establishing a solid foundation for subsequent sentiment analysis.

2.Feasibility Study(Planning):

The feasibility study evaluates the project's viability. Economic feasibility assesses the cost-effectiveness of the solution, including data storage and processing expenses. Schedule feasibility determines the project's timeline, considering the available resources and development time. Technical feasibility examines the compatibility of using CNN (Convolutional Neural Network) and OpenCV technologies for real-time emotion detection. The study ensures that the chosen technologies align with project goals and constraints.

3.Design:

System and software design documents are created in this phase. The system design outlines the architecture, components, and data flow of the emotion recognition system. Software design documents delve into the specifics of implementing CNN and OpenCV for facial expression analysis. These documents serve as blueprints for the development phase, ensuring a structured and organized approach.

4.Implementation(Development):

The implementation phase involves coding and integrating the CNN and OpenCV technologies. The Convolutional Neural Network is trained on a dataset of labeled facial expressions to recognize emotions accurately. OpenCV is utilized for real-time face detection and feature extraction. The integration of these technologies enables the system to process video call feeds, extract facial expressions, and predict emotions.

5.Testing:

Testing is vital to validate the system's functionality and performance. Tracking and monitoring tools are implemented to assess accuracy, reliability, and real-time processing capabilities. The system is rigorously tested using diverse video call scenarios to ensure consistent emotion recognition. Any deviations or errors are identified, addressed, and refined to enhance the system's accuracy and user experience.

3.1.2 Commercialization

Target Audience & Market Space

Target Audience

- Couples and romantic partners
- Families
- Individuals seeking relationship improvement
- Therapists and counselors.

Market Space

- There is no limit in age for users.
- No need to advanced knowledge in technology.

Future Scope

Future scope includes AI advancements for deeper emotion understanding, virtual reality integration, and applications in mental health and education.

4. PROJECT REQUIREMENTS

4.1 Functional Requirements

1. Real-time Emotion Detection: The system should accurately detect and analyze facial expressions and emotional cues in real-time during video calls.

2. Multi-Emotion Recognition: The component should be able to recognize and categorize a range of emotions, including happiness, sadness, anger, surprise, and more.

3. Personalized Goal Setting: Users should be able to set personalized relationship goals and milestones, with the system providing tailored recommendations based on emotional insights.

4. Integration with Video Call Platforms: The component needs seamless integration with popular video call platforms, allowing users to access emotional analysis during their calls.

5. Progress Tracking and Reporting: Users should be able to track their emotional progress over time, receiving regular reports and insights to gauge relationship improvement.

4.2 User Requirements

- Mobile Goal Tracker Access:
  Users should be able to access their personalized goal tracker using the mobile application. This feature will allow users to conveniently track their relationship improvement goals, view progress, and receive updates on their achievements directly from their mobile devices.

- Multi-Lingual User Interface:
  Users can access the application to be available in various languages, ensuring that individuals from diverse linguistic backgrounds can easily navigate and understand the interface.

- Emotion Identification via Mobile:
  Users should have the capability to identify their emotions using the mobile application. This functionality will empower users to analyze their emotional expressions and receive real-time insights into their feelings during video calls or other interactions, enhancing their emotional awareness and communication skills.

4.3 System Requirements

Software requirements delineate the functionalities and resources necessary for a software system to meet user needs effectively. They bridge user expectations and technical implementation, guiding development by defining features, performance, and interactions. By minimizing misunderstandings and aligning stakeholders, requirements ensure the creation of user-centered, reliable software solutions that fulfill business objectives.

- Video Call Data Collection: Obtain video call feed data from the Meta company.

- Emotion Recognition: Convolutional Neural Networks (CNN) for analyzing facial expressions and recognizing emotions.

- Deep Learning Framework: Keras, for a high-level deep learning framework, for training and deploying emotion recognition models.

- Image Processing: OpenCV for real-time face detection and feature extraction from video call data.

- Cross-Platform App Development: Develop a user-friendly application using React Native and Expo, ensuring compatibility with both Android and iOS.

- Backend Server: Use Flask to host and manage the deployed deep learning models for real-time emotion recognition.

- Integration: Establish a Node Server to connect the mobile and web applications with the Flask Server, facilitating real-time communication.

- Crowdsourcing: Implement a mechanism to engage users and gather feedback on the accuracy of emotion recognition results.

4.4 Non-Functional Requirements

1. User Privacy and Data Security: The system must ensure that all collected user data is securely stored, encrypted, and accessible only to authorized personnel, complying with privacy regulations.

2. Performance - The system is expected to demonstrate efficient performance, delivering rapid and precise results to users.

3. Availability - The application's accessibility should extend to users worldwide, accommodating various languages and ensuring uninterrupted availability as required.

4. Scalability and Concurrent Usage: The solution should be able to handle a growing number of simultaneous users without degradation in performance, ensuring a smooth user experience during peak usage times.

4.5 Use Case Diagram

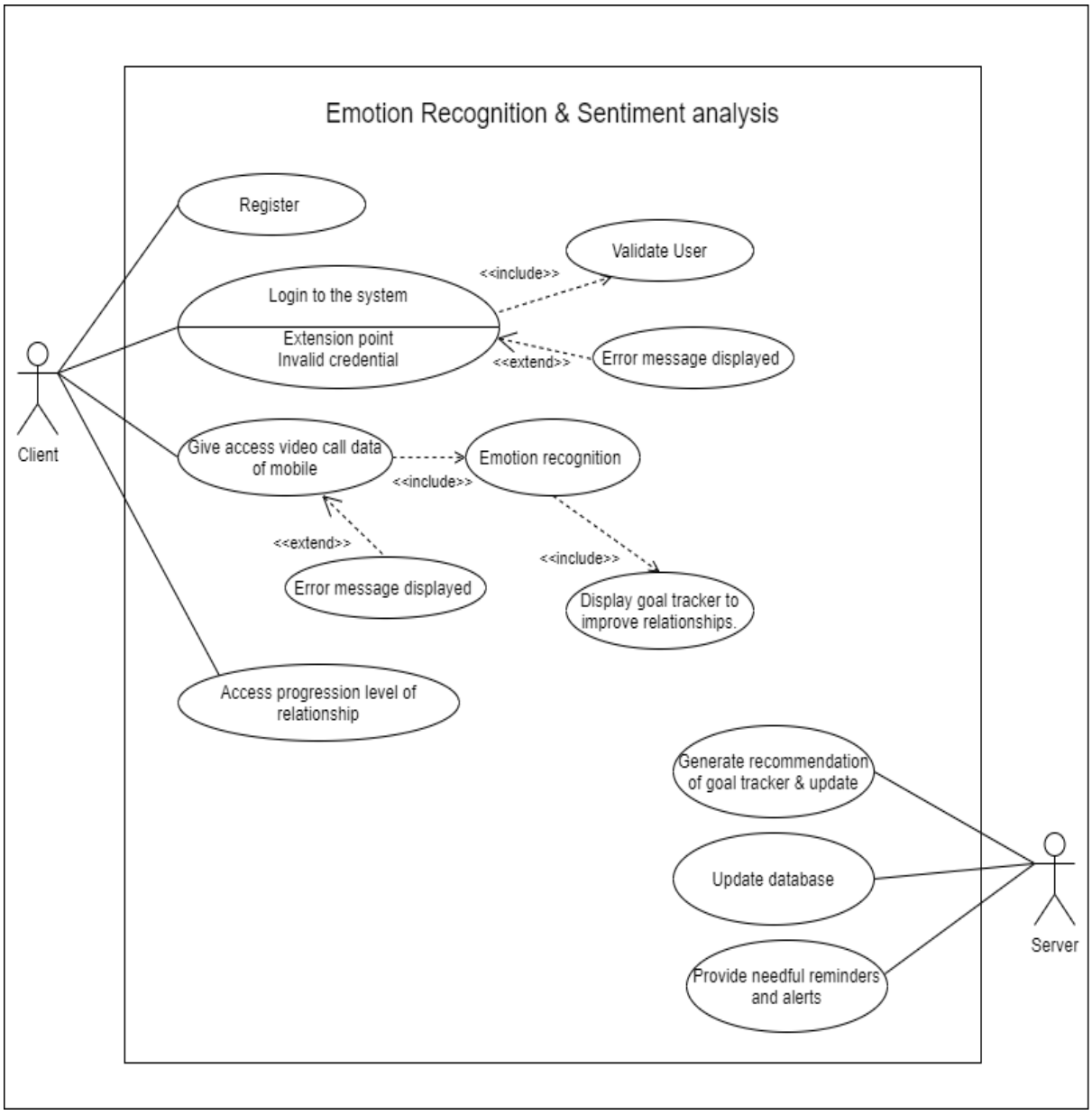


Figure 4.1: Use case diagram

4.6 Tentative Test Scenarios

1. Baseline Emotion Detection:

Scenario: User gives access to video call feed data. The app's recognition model analyzes the expressions of emotion and displays the primary emotion on-screen.

Expected Outcome: The app accurately recognizes and displays the user's emotion, setting the foundation for sentiment analysis.

2. Sentiment Tracking Over Time:

Scenario: Tracking data about user's video calls for a week and identify sentiment. During each call, the app analyzes facial expressions continuously tracks sentiment changes.

Expected Outcome: The app effectively tracks changes in the user's sentiment and presents them graphically, to visualize to the user.

3. Goal-Driven Sentiment Improvement:

Scenario: Set goals to user to improve their relationships. The app maintains a positive expression to user and monitors facial expressions during calls and compares them to the goal. Also, the app provides feedback on goal achievement.

Expected Outcome: The app guides users towards achieving their set goals by analyzing emotions and providing actionable feedback.

5. BUDGET AND BUDGET JUSTFICATION

Table 6. 1 : Budget for the proposed system

Requirement	Cost(Rs.)
Mobile app hosting charge – Play store	5 000.00
Cloud service	6 000.00/month
Internet Charges	5 000.00
Total Cost	**16 000.00**

6. GANTT CHART

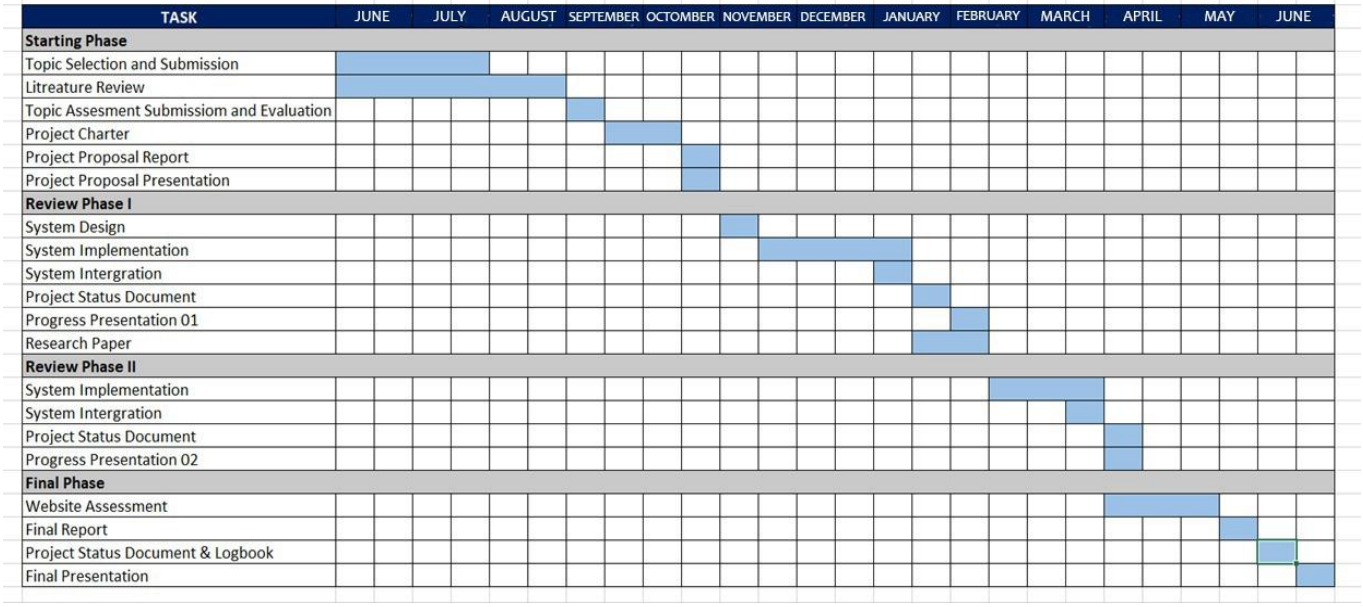


Figure 6.1: Gantt chart

6.1 WORK BREAKDOWN STRUCTURE (WBS)

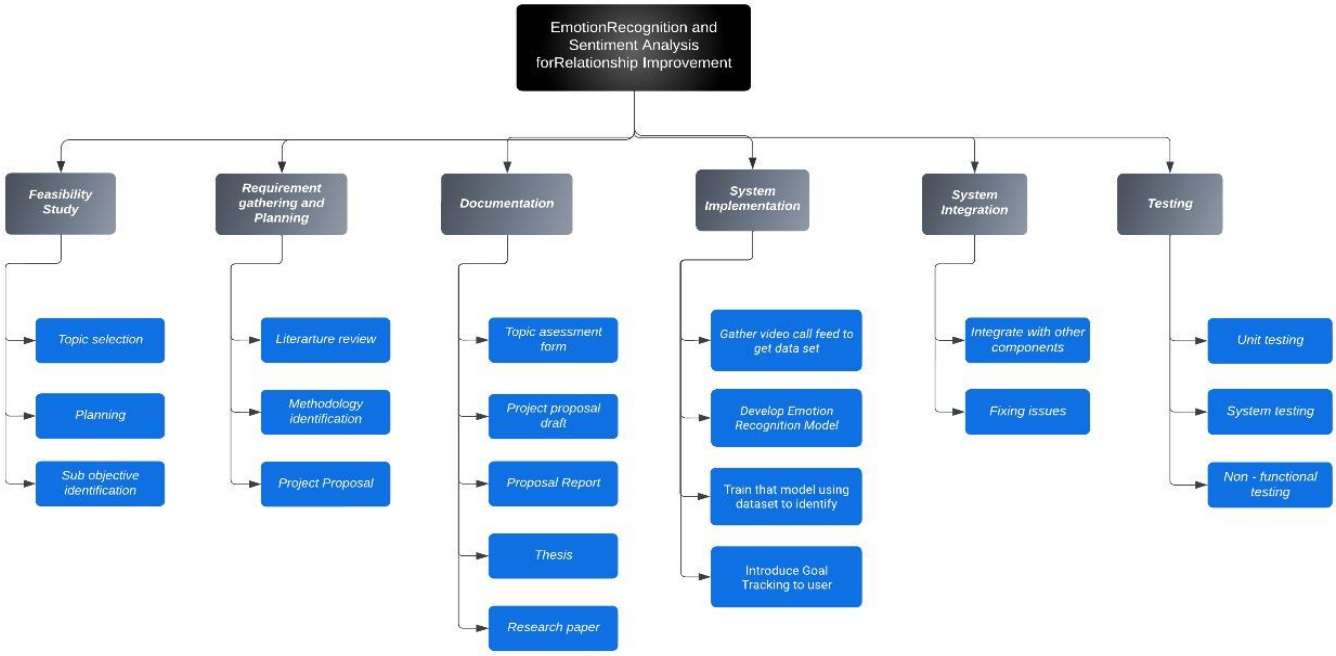


Figure 6.2: Work breakdown chart of Recognize an emotion using image processing component

REFERENCES

[1] Viha Upadhyay, Devangi kotak, "A Review on Different Facial Feature Extraction Methods for Face Emotions Recognition System," in *IEEE*, Coimbatore, India, 2020.

[2] Ramzi GUETARI, Aladine CHETOUANI, Hedi TABIA, Nares KHLIFA, "Real time emotion recognition in video stream, using B-CNN and F-CNN," in *IEEE*, Sousse, Tunisia, 2020.

[3] Swarna Kushibhotla, Chaitanya Voonna, Ram Prasad Penki, Gyan Sai Kumar Pappala, Aruna S, Hima Deepthi Vakayalapati , "Analysis of Facial Emotion Recognition for Image and Video Data using Convolution Neural Networks," in *IEEE*, Coimbatore, India, 2023 January 16.

[4] Yogesh S. Deshmukh, Nikhil s. patankar, Rameshwar Chintamani, Nitin Shelke, "Analysis of Emotion Detection of Images using Sentiment Analysis and Machine Learning Algorithm," in *IEEE*, Tirunelveli, India , 2023 March 14.

[5] Leh Luoh, Chih-Chang Huang, Hsueh-Yen Liu, "Image processing based emotion recognition," in *IEEE*, Taipei, 2010 August 19.

[6] Neetu Faujdar, Shipra Sarawat, Sachin Sharma, "Human Pose Estimation using Artificial Intelligence with Virtual Gym Tracker," in *IEEE*, Mathura , India, 2023 May 04.