Shahjalal University of Science and Technology
Department of Computer Science and Engineering

Facial Emotion Recognition Using Mobile Devices: A TensorFlow Lite Approach with ML Kit Integration

ASIF AHMED
Reg. No.: 2018331006
4^{th} year, 2^{nd} Semester

FATIMA NUZHAT
Reg. No.: 2018331106
4^{th} year, 2^{nd} Semester

Department of Computer Science and Engineering

Supervisor
DR. SADIA SULTANA
Associate Professor
Department of Computer Science and Engineering

16th February, 2025

Facial Emotion Recognition Using Mobile Devices: A TensorFlow Lite Approach with ML Kit Integration

A Thesis submitted to the Department of Computer Science and Engineering, Shahjalal University of Science and Technology, in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering.

By

Asif Ahmed
Reg. No.: 2018331006
4^{th} year, 2^{nd} Semester

Fatima Nuzhat
Reg. No.: 2018331106
4^{th} year, 2^{nd} Semester

Department of Computer Science and Engineering

Supervisor
DR. SADIA SULTANA
Associate Professor
Department of Computer Science and Engineering

16th February, 2025

Recommendation Letter from Thesis Supervisor

The thesis entitled *Facial Emotion Recognition Using Mobile Devices: A TensorFlow Lite Approach with ML Kit Integration* submitted by the students

1. Asif Ahmed
2. Fatima Nuzhat

is under my supervision. I, hereby, agree that the thesis can be submitted for examination.

Signature of the Supervisor:

Name of the Supervisor: Dr. Sadia Sultana

Date: 16^{th} February, 2025

Certificate of Acceptance of the Thesis

The thesis entitled *Facial Emotion Recognition Using Mobile Devices: A TensorFlow Lite Approach with ML Kit Integration* submitted by the students

1. Asif Ahmed
2. Fatima Nuzhat

on 16th February, 2025, hereby, accepted as the partial fulfillment of the requirements for the award of their Bachelor Degrees.

Head of the Dept.	Chairman, Exam Committee	Supervisor
Dr. Md Masum	Dr. Md Masum	Dr. Sadia Sultana
Professor	Professor	Associate Professor
Department of Computer Science and Engineering	Department of Computer Science and Engineering	Department of Computer Science and Engineering

Abstract

Facial expression detection is becoming increasingly important in a variety of applications, from improving human-computer interaction to assisting with healthcare diagnoses. With the widespread usage of mobile devices in daily life, there is an increasing demand for real-time emotion detection capabilities built directly into these platforms. This study conducts a thorough investigation of face emotion identification designed exclusively for mobile devices, leveraging the combined capability of TensorFlow Lite and ML Kit integration. It begins with a thorough examination of existing techniques, emphasizing their strengths and limitations. Then, an innovative and simplified technique is provided to improve the efficiency and accuracy of emotion detection on mobile platforms. Through a series of rigorous tests and performance assessments, the proposed integrated system exhibits its ability to achieve high accuracy and real-time responsiveness, notably on Android platforms. Future research efforts are expected to focus on improving emotion identification algorithms, fine-tuning performance for smooth real-time usage, resolving privacy concerns, and guaranteeing seamless cross-platform interoperability. By making substantial advances in this technology for mobile devices, we hope to provide essential insights and practical consequences for academics, practitioners, and developers navigating the dynamic field of mobile-based emotion identification solutions.

Keywords: Tensorflow (TF), Tensorflow Lite, Facial emotion recognition (FER), ML Kit, Android.

Acknowledgements

We would like to thank the Department of Computer Science and Engineering at Shahjalal University of Science and Technology, Sylhet, Bangladesh, for allowing us to conduct this research. We express our heartfelt appreciation to our esteemed supervisor, Dr. Sadia Sultana, for leading, instructing and supporting us with patience and encouragement. Additionally, we would like to convey our sincere gratitude to Dr. Mohammad Khalad Hasan, whose professional direction and knowledge have greatly benefited us. It would not have been possible without their guidance and inspiration to keep moving ahead with our thesis work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Facial Emotion Recognition (FER) is a growing subject in artificial intelligence and computer vision that aims to automate the detection and interpretation of human emotions through facial expressions. This thesis investigates the use of Facial Emotion Recognition on mobile devices, employing the TensorFlow Lite and ML Kit frameworks. Before digging more into the complexities of this technique, it is critical to understand certain essential words relevant to this research.

In this context, mobile devices include smartphones, tablets, and other portable computing devices that include cameras and the processing power to analyze facial expressions in real time. Mobile devices are suitable platforms for delivering FER applications in a variety of scenarios due to their widespread availability and ease.

At its core, FER involves the automated detection and analysis of emotional states depicted in human faces. This process relies on computational algorithms to extract facial features, classify emotions, and interpret subtle cues in facial expressions.

A plethora of research endeavors have delved into the domain of emotion recognition, each contributing unique insights and methodologies. Montagne (2007) [5] devised a paradigm aimed at quantifying the perception of facial emotional expressions, revealing that happiness emerged as the most readily identifiable emotion. Nagalkar (2019) [6] conducted a study focusing on recognizing seven distinct emotions through facial expressions, with the Histogram of Oriented Gradients (HOG) technique yielding superior results compared to the Bag of Features (BOF) approach. Maaoui (2010) [7] proposed an innovative method leveraging physiological signals for emotion recognition, achieving an impressive classification performance of 92%. Furthermore, Tarnowski (2017) [8] employed facial expressions as cues to recognize seven emotional states, successfully employing both a k-Nearest Neighbors (k-NN) classifier and a Multilayer Perceptron (MLP) neural network. These diverse studies collectively underscore the multifaceted potential of various methodologies in the realm of emotion recognition.

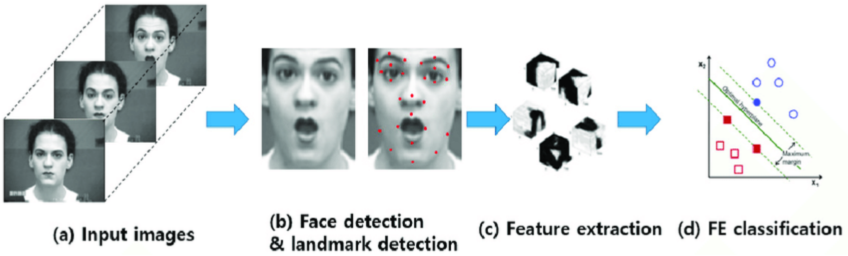

Figure 1.1: Conventional FER method

TensorFlow Lite is a more lightweight version of TensorFlow, Google's open-source machine learning framework. It is designed for mobile and embedded devices, allowing developers to effectively deploy machine learning models on resource-constrained platforms such as mobile phones.

Silberman (2017) [9] released TF-Slim, a lightweight library designed for designing, training, and assessing complicated models in TensorFlow. The significance of TF-Slim stems from its contribution to the creation of TensorFlow Lite, a version of TensorFlow designed for mobile and edge computing devices. Pavithra (2019) [10] noted TensorFlow's critical significance in deep learning, citing its versatility and smooth deployment across several platforms. Abadi (2016) [11] presented an overview of TensorFlow's dataflow model and its performance in actual applications, emphasizing its significance in the creation of TensorFlow Lite. These findings highlight the importance of TensorFlow Lite in enabling the deployment of machine learning models on devices with minimal processing capabilities.

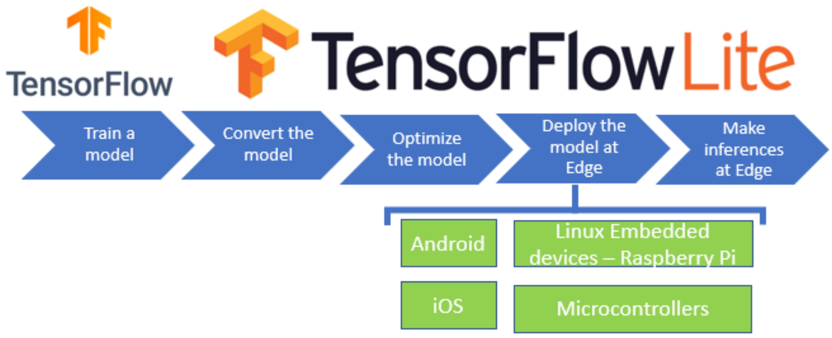

Figure 1.2: Tensorflow Lite integration concept

ML Kit, a suite of machine learning tools developed by Google, encompasses pre-trained models and APIs tailored for various tasks such as image classification, text recognition, and facial identification. Integrating ML Kit into mobile applications simplifies the execution of complex machine learning assignments, including facial expression recognition (FER).

Moreover, ML Kit fosters the development of ML-driven products by involving designers throughout the machine learning lifecycle, as indicated by Sun (2022) [12]. Additionally, the Open MatSci ML Toolkit, particularly the OpenCatalyst Dataset, provides a flexible platform for deploying deep learning models in the materials science domain, as highlighted by Miret (2022) [13]. Alongside, the i-Virtual Lab Kit with MBL system delivers an economical solution for conducting fundamental electronics experiments, featuring real-time data capture and interactive educational features, as discussed by Abdullah (2019) [14]. Furthermore, the Google Cloud Machine Learning Engine (Cloud MLE) furnishes a managed infrastructure for training and managing large-scale ML models, as outlined by Bisong (2019) [15]. Galaxy-ML enhances accessibility to supervised machine learning for biomedical researchers, per Gu (2021) [16], while Dlib-ml offers a C++ toolkit for crafting machine learning applications, as described by King (2009) [17].

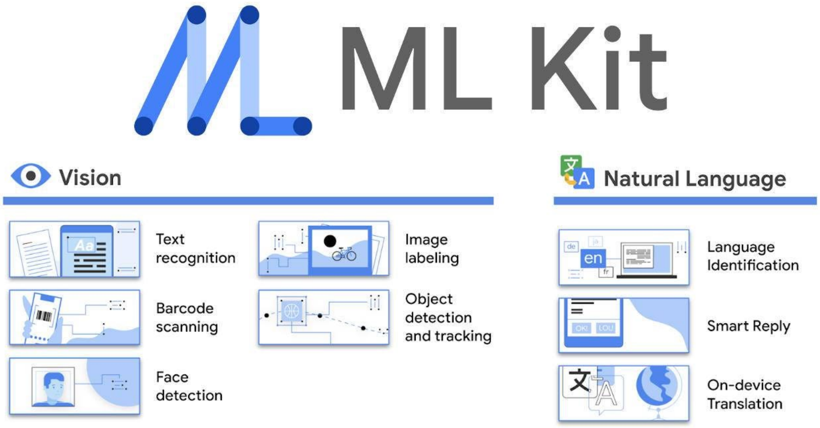

Figure 1.3: Overview of Google ML Kit Services [1]

The thesis presents a unique technique that combines TensorFlow Lite and ML Kit to create a powerful FER system designed exclusively for mobile devices. By combining these two frameworks, the system becomes capable of rapidly processing facial expressions directly on the device, eliminating the need for constant internet access or reliance on cloud-based services.

Using TensorFlow Lite and ML Kit, this thesis investigates the feasibility and efficacy of deploying FER systems on mobile devices. This integration not only improves the performance and efficiency of facial expression recognition, but it also expands the possibilities for implementing FER applications in a wide range of real-world settings. These scenarios cover a wide range of domains, from assistive technology, where precise and timely recognition of facial expressions can greatly benefit people with disabilities, to interactive user interfaces, where seamless integration of FER capabilities can transform user experience and engagement. Thus, the combination of TensorFlow Lite and ML Kit not only paves the door for sophisticated FER technologies on mobile platforms, but also demonstrates mobile devices' potential as powerful and adaptable machine learning tools.

1.1 Motivation

So far, there have been several approaches and systems created to accurately recognize emotions. Emotion recognition techniques have been developed in both the psychology and engineering disciplines. These approaches try to automatically identify and comprehend human emotions using a variety of information such as speech, facial expressions, physiological signs, and motion data. One strategy is to use verbal and nonverbal communication signals to determine signal characteristics, which are then linked to elements of an emotional ontology for decision-making [18]. Another way is to analyze electroencephalogram (EEG) data and extract time-frequency domain features, nonlinear characteristics, and brain network attribute properties. Additionally, speech-based emotion detection can be accomplished by identifying distinctive tones associated with various emotions and assessing the speaker's state based on the occurrence of these tones in phonemes [19]. Furthermore, music emotion recognition focuses on recognizing emotions in music and has been treated from a dimensional viewpoint, with valence and arousal as the fundamental emotion dimensions [20]. Furthermore, the accuracy and reliability of emotion detection models or algorithms utilized in these techniques may vary based on the quality and variety of the training data, as well as the models' generalizability to other populations or cultural settings. It is also crucial to address the ethical implications of emotion recognition systems, such as privacy issues or the possibility of misusing or misinterpreting emotional data.

1.2 Research Objectives

The comprehensive objectives of this study encompass a multifaceted exploration, development, and assessment of facial emotion recognition (FER) systems on mobile devices, with a nuanced emphasis on the integration of TensorFlow Lite (TF Lite) and Google's ML Kit. At the outset, the study aims to conduct an extensive review of existing FER methodologies, algorithms, and frameworks, specifically focusing on their adaptability and efficacy in the mobile environment. This foundational investigation seeks to provide a robust understanding of the landscape of FER techniques, setting the stage for subsequent developments.

In parallel, the integration of TF Lite and ML Kit into the FER framework presents a pivotal opportunity to leverage the capabilities of these cutting-edge technologies. While TF Lite offers lightweight and efficient model execution on mobile devices, ML Kit provides a suite of pre-trained models and APIs for tasks such as facial structure extraction from photos. By evaluating the compatibility, performance, and ease of integration of TF Lite and ML Kit within the FER context, this study aims to elucidate their potential synergies and limitations in real-world applications.

Integral to the FER system's functionality is the extraction of facial structures from photos, a task facilitated by ML Kit's sophisticated algorithms. This additional layer of analysis enriches the FER system's capabilities, enabling a deeper understanding of facial expressions and enhancing emotion recognition accuracy. Through meticulous experimentation and evaluation, the study endeavors to optimize the integration of facial structure extraction into the FER pipeline, ensuring seamless compatibility and performance in diverse scenarios.

Central to the study's objectives is the development of a prototype FER system leveraging TF Lite, ML Kit, and the extracted facial structures. This entails the design and implementation of algorithms for facial feature extraction, emotion classification, and real-time inference, tailored specifically for mobile devices. The system's architecture will be meticulously crafted to optimize resource utilization, minimize computational overhead, and ensure robust performance across a range of mobile platforms.

Furthermore, the study aims to refine and optimize the developed FER system to enhance its efficiency and accuracy. Techniques such as model quantization, compression, and runtime optimization will be explored to maximize the system's performance while minimizing its footprint on mobile devices. Through iterative experimentation and refinement, the study seeks to achieve a balance between computational efficiency and emotion recognition accuracy, ensuring the system's practical viability in real-world applications.

Critical to the success of the FER system is its evaluation against standardized datasets and performance metrics. Rigorous testing will be conducted to assess the system's ability to accurately detect and classify facial emotions in diverse real-world scenarios. By systematically analyzing the system's performance across various metrics, the study aims to provide comprehensive insights into its efficacy and limitations, informing further refinements and optimizations.

Finally, the study will evaluate the practical applications and deployment considerations of the developed FER system. Usability, scalability, privacy implications, and deployment challenges will be meticulously analyzed to determine the system's suitability for real-world deployment across diverse settings. Through a holistic assessment of these factors, the study aims to provide actionable recommendations for researchers, practitioners, and developers seeking to deploy FER solutions on mobile platforms.

In summary, this study embarks on a comprehensive exploration and development journey, aiming to advance the state-of-the-art in FER technology on mobile devices. By integrating TensorFlow Lite, ML Kit, and facial structure extraction capabilities, the study seeks to develop a robust and efficient FER system with practical applications across a range of domains. Through rigorous experimentation, optimization, and evaluation, the study aims to provide valuable insights and recommendations for the broader FER research community, driving forward innovation in this exciting field.

Chapter 2

Background Study

2.1 Emotion Detection

Emotion detection is an important aspect in Affective Computing and has gained significant attention in recent years. Analysis of facial expressions is one of the fundamental techniques for emotion identification. This technique takes use of the fact that facial signals are frequently used to vividly convey human emotions. Notable works by Ekman and Friesen laid the groundwork for understanding universal facial expressions, forming the basis for automated emotion recognition systems [21, 22].

2.1.1 Recent Works

Training emotion detection models relies heavily on machine learning methods. Unsupervised learning, which includes grouping and pattern discovery from unlabeled data, is more common than supervised learning, which involves models learning patterns from labeled data. Researchers such as Neal [23], Burr [24], Anzai [25], and Bishop [26] have contributed to the development of pattern recognition and machine learning methods. Their work includes the introduction of Bayesian approaches, approximate inference algorithms, and the use of graphical models to describe probability distributions.

Hybrid strategies that combine many modalities have also become more popular. These multimodal systems take use of how multiple data sources work well together. According to a study conducted by Wei et al.[27], the accuracy and dependability of detection systems are improved by the integration of face expressions, physiological data, and text analysis.

Here, table 2.1 provides a graphic representation of a recent assessment of emotion detecting methods.

Study	Technique	Classifier	Attribute and Input Format
Lee et al. (2014)	AAM and Fuzzy K-NN	No preprocessing (template Matching Method)	Face MPEG-4
Liu and Yin (2015)	SIFT (Scale Feature Invariant Transform)	SVM	Face (skin temperature and head motion) FLIR thermal camera videos
Georgakis et al. (2016)	DICA (Discriminant Independent Coherent analysis)	Linear Regression	Face Database images (32x32 pixel)
Le Ngo et al. (2016)	EMM	SVM	Face Short movies
Sarma and Bhattacharyya (2016)	ANN	LDA	Face Image
Li and Nan (2011)	BLD, FLD	KPCA	Face Images

Zen et al. (2016)	Novel transfer learning Framework	SVM	Face, Gesture Video
Zhao et al. (2015)	Space Based FER	2-D median filter	Face 2.5 D Facial Data
Zhang et al. (2016)	Discrete Wavelet Transform and Biorthogonal Wavelet Transform	SVM	Face Image
Shojaeilangari et al. (2015)	Extreme sparse learning	Non-Linear classifier	Face Videos
Lee and Ro (2015)	Partial matching	Sparse representation classifier	Face Video
Agarwal and Mukherjee (2017)	Local motion pattern	Motion descripts/SVM	Face Video
Kung et al. (2015)	3D HMM	Histogram, Gaussian	Face Video
Mistry et al. (2016)	Micro GA, Embedded PSO	Ensemble classifier	Face Image
Goyal et al. (2016)	Mixture of experts based fusion model	MFCC and HOF (Histogram of optical animated movie as flow)	Face, Audio Video

Reney and Tripathi (2015)	KNN and multi resolution decomposition	Violo Jones and Mel frequency component	Face, Voice Sound file and image
Shah and Kaushik (2015)	Euclidian distance and neural networks	DCT and canny edge detection	Face Images
Oh et al. (2016)	Higher Order Riesz Transforms	SVM	Face Videos
Qayyum et al. (2017)	Stationary Wavelet Transform (SWT)	DCT	Face Videos

Table 2.1: Various algorithms for facial emotion recognition

Deep learning has completely changed the way emotion detection models work and made it possible to build complex neural network designs that can detect minute emotional nuances [28]. This topic has become more well-known as a result of its numerous uses in a variety of industries, including psychology, marketing, healthcare, and the design of human-computer interfaces.

Study	Response time and Accuracy	Best features	Limitation	Application
Lee et al. (2014)	Dynamic 76%	Beneficial for quick application in phone	Happy, sad and neutral emotions are identified. Difficult to differentiate false dismissals	Iphone and mobile application.
Liu and Yin (2015)	Dynamic 91%	Eight spontaneous expressions are educed	Cannot differentiate real/ fake expression.	To find chronic low back pain.
Georgakis et al. (2016)	Static 95.4%	Subject independent. Expression recognition. Training set may have scarf, sunglass and occlusion	Different accuracy database, AU detection, gives low performance.	Biometric. Face analysis.
Li and Nan (2011)	Static 83.55%	Seven basic emotion classification Inducement intersection of facial expression.	Person dependent. It is hard to differentiate the inducement expression.	Find the inducement of facial expression
Zen et al. (2016)	Static 90%	User independent. Computational time reduced. Increases accuracy. Computational cost is low because of pre trained regression.	Accuracy based on user independent classifier.	Used in smart watch. AU (Action Unit) detection. Pain detection.

Zhao et al. (2015)	Dynamic 91.6%	Trained AU detection. AU computation and evaluation.	2.5 D kinetic camera and RGB-D camera images used.	Wearable Social Assistant device. 2.5 D video chatting.
Zhang et al. (2016)	Dynamic 96.77%	Improved accuracy	Not focused on the geometric face images. Only uses the image not considering the video images.	Used in MR images, CR images, remote sensing images. To identify autistic person and AD.
Shojaeilangar et al. (2105)	Static 92.74%	Increases the accuracy or the acted and self-generated facial expression.. Head pose variation and occlusion images are used to detect emotion.	Failure in optical flow. Fail in face detection. Fail to detect. Reference point in face	Driver warning system. Real world application.
Qayyum et al. (2017)	Dynamic 98.83%	Better performance in accuracy and classification.	Use kinect camera	Kinect based application.
Oh et al. (2016)	Dynamic 90%	Use intrinsic two-dimensional (i2D)	i2D feature does not improve accuracy rate for all emotions.	To detect micro expressions. Signal processing.
Agarwal et al. (2017)	Static 94.2%	High accuracy. Time and space complexity has been reduced.	Difficult to detect anger and sadness.	Emotion recognize application.

Table 2.2: Performance comparison between various emotion recognition methods

The compilation in table 2.2 acts as a thorough archive including the results of recent research projects in the field of emotion recognition. This tabular resource, in particular, provides insights on response times, a parameter important for real-time applications where quick emotional recognition is essential. The highlighted properties give a view into the input dimensions impacting the identification process, while the accuracy metrics assess the dependability of the models. It also demonstrates the varied influence that emotion recognition may have across domains by including a wide range of applications.

In recent years, emotion detection from facial expressions using CNN has been extensively studied in the field of human-computer interaction and data analytics [29]. Paul Ekman's pioneering work on universal facial expressions laid the foundation for emotion detection using facial features [30]. Researchers have been able to automate the procedure thanks to CNNs by teaching deep learning models to distinguish patterns corresponding to distinct emotions. Various techniques including the utilization of artificial neural networks (ANNs) have been employed for facial expression identification by convoluting input images to create feature maps [31,32]. Deep learning algorithms, such as CNN, have been used for facial emotion recognition, achieving high accuracy rates on datasets like FER-2013 [33], similar to what we used in this study. Face identification, feature extraction, and expression classification are all steps in these algorithms.

2.1.2 Facial Features Used In Models

CNNs are particularly adept at image analysis due to their hierarchical structure inspired by the human visual system. Layers of convolutional filters and pooling operations enable CNNs to capture spatial hierarchies in facial images [34]. Massive datasets of annotated facial expressions are used to train the CNN models. Models learn to recognize distinguishing face characteristics, textures, and patterns that correspond to various emotions by feeding these annotated pictures into CNN structures. Furthermore, emotion detection using CNNs has been greatly benefited by transfer learning. Large picture datasets are used to train pre-trained CNN models like VGG, ResNet, and Inception, which are then tweaked for emotion identification tasks [35]. This method expedites model training and reduces the requirement for large, annotated emotion datasets.

To determine which characteristics are most efficient in identifying a specific facial expression, a variety of real-valued and binary parameters need to be retrieved and examined. The features which do not offer any worthwhile information of the facial expression portrayed in the dataset images are generally removed and are not used in the final study.

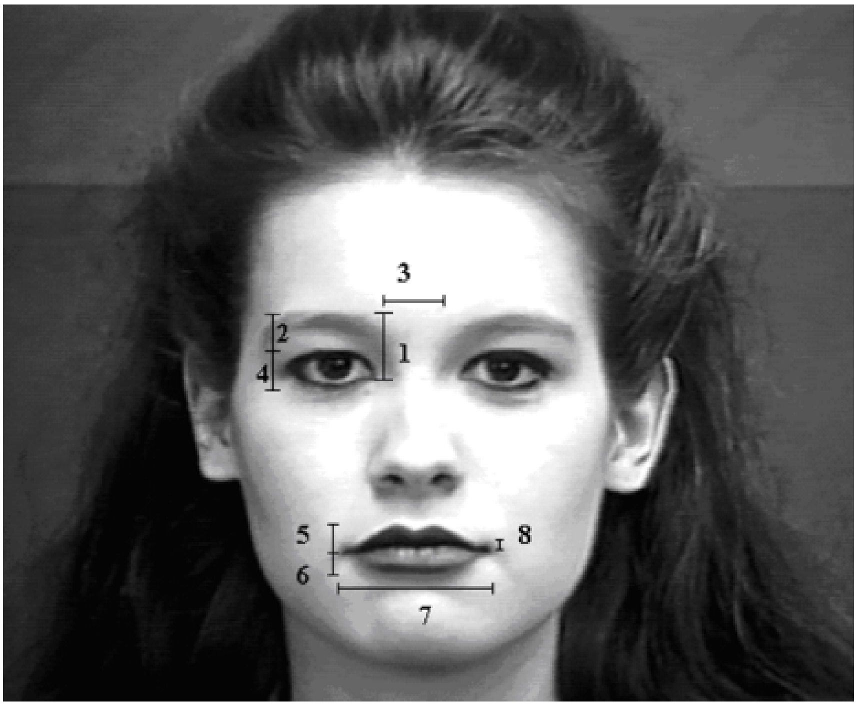

Figure 2.1: Real valued parameters in facial features [2]

The real valued parameters depicted in figure 2.1 are [2, 36]:

1. Eyebrow raise distance: The distance between the inner corner of the eyelid and the eyebrow edge on top of it.

2. Upper eyelid to eyebrow distance: The distance between the outer corner of the eyelid and the tail of the eyebrow on top of it.

3. Inter-eyebrow distance: The distance between both of the eyebrows in the middle of the forehead.

4. Upper eyelid-lower eyelid distance: The distance between the upper and lower eyelids.

5. Top lip thickness: The thickness of the top lip of the mouth.

6. Lower lip thickness: The thickness of the lower lip of the mouth.

7. Mouth width: The distance between the two corners of the mouth.

8. Mouth opening: The distance between the lower surface of top lip and upper surface of lower lip.

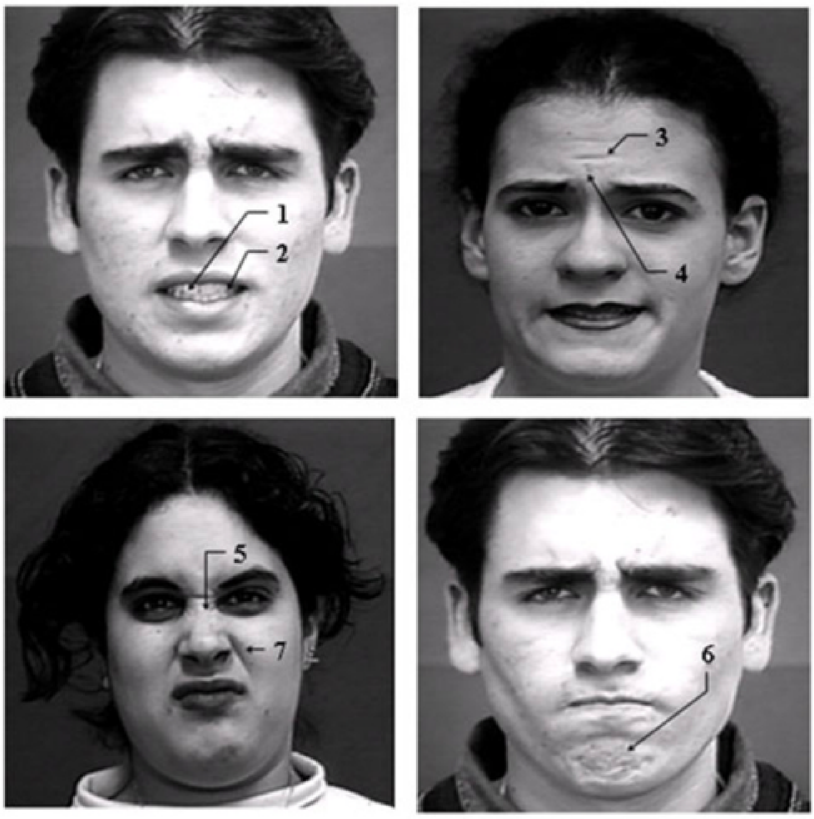

Figure 2.2: Binary parameters in facial features [2]

Also, the binary parameters depicted in figure 2.2 are [2, 36]:

1. Upper teeth visible: Whether the upper teeth are visible or not.

2. Lower teeth visible: Whether the lower teeth are visible or not.

3. Forehead lines: Presence or absence of wrinkles in the upper part of the forehead.

4. Eyebrow lines: Presence or absence of wrinkles in the space above and between the eyebrows.

5. Nose lines: Presence or absence of wrinkles in the space between the eyes extending over the nose.

6. Chin lines: Presence or absence of wrinkles or lines on the chin region just below the mouth.

7. Nasolabial lines: Presence or absence of deep lines on both sides of the nose extending down towards the upper lip.

The amalgamation of computer vision and deep learning has inaugurated a novel era of emotional intelligence. Leveraging these intricate facial expression patterns, CNNs have become formidable tools, discerning nuances previously inconspicuous.

Although CNNs have shown tremendous progress, some problems still exist. Misclassifications may result from variations in lighting, position, and facial features [37]. The identification process is further complicated by cultural variations and context-dependent expressions [38]. Researchers are continuously investigating methods to strengthen CNNs against these difficulties.

2.2 Implementation of FER

Facial Emotion Recognition (FER) is a confluence of computer vision, machine learning, and effective computing that seeks to decipher the nuanced manifestations of human emotions expressed through facial movements. Understanding emotions is a critical component of good human contact and connection, and FER systems aim to recreate this skill in artificial entities. FER algorithms are meant to detect minute signs conveying a wide range of emotional states, from pleasure and grief to amazement and repulsion, by meticulously analyzing face characteristics and their arrangements.

The advancement of FER technology has enormous promise across a wide range of fields, including human-computer interaction, healthcare, marketing, and entertainment. Deep learning approaches and advances in computer capabilities have accelerated FER in recent years, allowing for more precise, robust, and real-time emotion recognition. Despite these advances, some problems remain in FER research, including the impact of cross-cultural variations, face occlusions, and concerns about privacy and data security.

Despite these challenges, FER's attraction in improving user experiences and creating deeper relationships between humans and machines acts as a motivator for continued research and innovation in this fascinating sector. The capacity to effectively understand and respond to human emotions has important implications for a variety of applications. In human-computer interaction, FER can help create more intuitive and sympathetic interfaces, increasing user engagement and happiness. In healthcare, FER technologies have the potential to help in the early identification and monitoring of mental health issues by analysing facial expressions that indicate anguish or pain. Furthermore, in marketing and entertainment, FER allows for targeted content distribution and individualized experiences based on individuals' emotional responses.

While FER technology has great potential, its implementation is predicated on addressing the aforementioned problems. Efforts to reduce cross-cultural biases, build strong algorithms that are resistant to occlusions, and implement strict privacy measures are critical for the continuing growth and ethical deployment of FER systems. Nonetheless, FER's disruptive influence on human-machine interactions emphasizes the significance of ongoing study and innovation in this dynamic and varied subject.

Several methods have been proposed to increase the accuracy and efficiency of face expression identification. One option is to employ deep convolutional neural networks (DCNN) [39]. Another technique is to use feature points rather than action units to create a faster and more efficient recognition mechanism [40]. In addition, a strategy based on a deep sparse convolutional neural network has been presented to optimize the network weight and increase the algorithm's generalization [41]. Researchers have also utilized CNN to increase accuracy in datasets, with excellent results [42].

2.3 Challenges of various FER methods

The problem of emotion recognition in human-computer interaction is a complex one, requiring substantial amounts of emotionally labeled data for machine learning methods, which may be difficult to obtain. Model-based approaches are not available for emotion recognition, which limits the options for analysis. The paper discusses problems of data gathering, indicating potential challenges in obtaining relevant data for the study.

However, FER models trained on certain datasets may demonstrate bias towards certain demographic groups or cultural backgrounds, resulting in erroneous or unreliable predictions for persons outside the training data distribution. Keeping models fair, inclusive, and generalized to varied groups remains a problem (Buolamwini Gebru, 2018) [43]. When FER systems are implemented in real-world contexts with fluctuating lighting conditions, occlusions, and backdrops, they frequently fail (Dhall et al., 2019) [44]. These characteristics can reduce the quality of face photographs, affecting the performance and reliability of emotion identification algorithms, emphasizing the need of facial structure analysis.

Various methods have been proposed to address this issue, including the use of a Cauchy Naive Bayes classifier for recognizing emotions through facial expressions. The Cauchy assumption does not significantly improve recognition rate due to fewer outliers in the data. Morever there are subsequent difficulties in estimating the parameters of the Cauchy distribution. Recognition rates may not be sufficient for real-world use, depending on the particular application (Sebe, 2002) [45].

Ferdous Ahmed et. al [46] proposed an emotion recognition system that outperformed all of the state-of-the-art methods but it was the least explored modality. Consideration of movement bias from individual movement style limits the robustness of the system, as well as a slight decrease in performance in action-independent cases.

Many datasets used for facial expression recognition (FER) exhibit a notable deficiency in diversity across various demographic factors such as age, gender, ethnicity, and cultural backgrounds. This lack of representation can introduce bias into the datasets, potentially leading to biased models and diminished generalization capabilities. Therefore, the under-representation of diverse demographics in FER datasets poses significant challenges to the development of robust and inclusive facial expression recognition systems.

FER research necessitates collaboration across various fields, including psychology, computer vision, and machine learning (Picard et al. 2016) [47]. Bridging trans-disciplinary barriers and incorporating multiple viewpoints is vital.

Chapter 3

Methodology

3.1 Google ML Kit

The process for developing face structure and contour recognition using Google ML Kit includes the following steps:

With ML Kit's face detection API, you can detect faces in an image, identify key facial features, and get the contours of detected faces. Note that the API detects faces, it does not recognize people. [48]

Face recognition provides the information required to conduct activities such as enhancing selfies and portraits or creating avatars from a user's photo. Because ML Kit can recognize faces in real time.

Identifying and pinpointing facial characteristics includes determining the exact coordinates of the eyes, ears, cheeks, nose, and mouth of each face under observation. This approach allows for accurate mapping of facial features, such as the eyes, brows, lips, and nose, resulting in detailed outlines of identifiable faces. Furthermore, the detection of facial emotions, such as grins or closed eyes, adds deeper insight to the investigation. As faces move across successive video frames, their identities are continually monitored, preserving continuity and allowing for smooth manipulation of individual faces within the video stream. This permanent tracking technique gives a unique identity to each detected face, allowing for continuous engagement and analysis.

Moreover, the incorporation of real-time video processing capabilities, backed by effective face identification algorithms incorporated in smartphones, improves the speed and responsiveness of facial analysis jobs. This fast processing power is especially useful for applications that require instant response, such as live video editing, where prompt identification and interpretation of face characteristics is critical for providing seamless user experiences. Overall, these advances in facial analysis technology represent considerable progress in improving the accuracy, efficiency, and real-time applicability of facial recognition systems across a wide range of domains and applications.

3.2 Facial Structure Detection

Enabling face contour detection outputs a list of points for each facial feature that was recognized. These points indicate the form of the feature. The graphic below depicts how these points transfer to a face.

Figure 3.1: Face detection with ML Kit

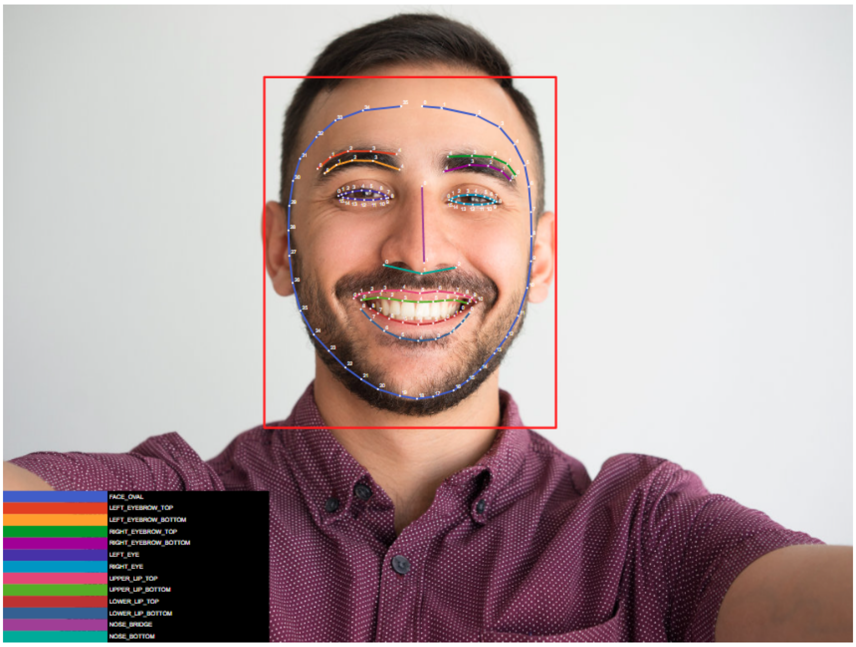

Figure 3.2: Facial structure points [3]

Here are some of the terms that we use regarding the face detection feature of ML Kit:

1. Face tracking extends face detection to video sequences. Any face that appears in a video for any length of time can be tracked from frame to frame. This means a face detected in consecutive video frames can be identified as being the same person. Note that this isn't a form of face recognition; face tracking only makes inferences based on the position and motion of the faces in a video sequence.

2. A landmark is a point of interest within a face. The left eye, right eye, and base of the nose are all examples of landmarks. ML Kit provides the ability to find landmarks on a detected face.

3. A contour is a set of points that follow the shape of a facial feature. ML Kit provides the ability to find the contours of a face.

4. Classification determines whether a certain facial characteristic is present. For example, a face can be classified by whether its eyes are open or closed, or if the face is smiling or not [49].

3.3 Tenserflow

TensorFlow and TensorFlow Lite are critical components of mobile facial emotion recognition (FER) systems. TensorFlow, a strong open-source machine learning framework, offers a solid platform for creating and training deep learning models. Using TensorFlow, a pre-trained convolutional neural network (CNN) model for FER is chosen and optimized. This model is highly accurate in recognizing emotions from facial expressions. To provide easy deployment on mobile devices, the TensorFlow model is transformed to TensorFlow Lite format. TensorFlow Lite is tailored for on-device inference, allowing the FER model to run efficiently on Android smartphones. This lightweight version of TensorFlow retains great performance while lowering memory footprint and latency, which is critical for real-time applications. They are optimized for on-device machine learning by addressing five key constraints: latency (no round-trip to a server), privacy (no personal data leaves the device), connectivity (no internet connection is required), size (reduced model and binary size), and power consumption (efficient inference and no network connections) [50].

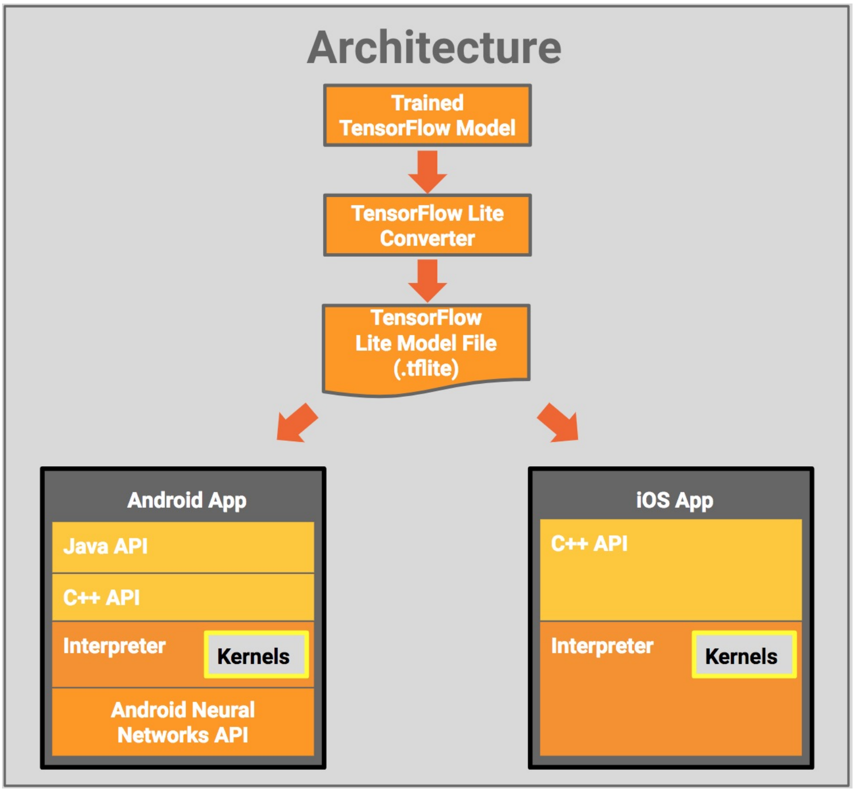

Figure 3.3: How TensorFlow models are used in mobile devices [4]

We have achieved our aim of building a scalable, efficient, and accurate FER solution appropriate for mobile device deployment by integrating TensorFlow Lite into the Android system.

Chapter 4

Implementation

4.1 Overview of the System

The Android Studio-developed technology provides a full solution for facial emotion recognition (FER) on mobile devices. The system uses machine learning and deep neural networks, as well as Google's ML Kit and TensorFlow frameworks, to discern emotions from facial expressions accurately and in real time.

At its heart, the system uses ML Kit's face detection API to recognize and locate faces in photos or live video feeds. This feature lays the groundwork for future facial structure analysis, allowing the system to detect important face landmarks and shapes required for emotion recognition.

Furthermore, TensorFlow integration improves the system's capabilities by allowing for the deployment of a pre-trained convolutional neural network (CNN) model created exclusively for FER tasks. This TensorFlow model, adapted to TensorFlow Lite for better performance on mobile devices, allows the system to accurately assess face characteristics and identify emotions.

The system's procedure starts with capturing or choosing a picture with human faces. The ML Kit face detection API recognizes and extracts facial areas from images, which are subsequently processed by the TensorFlow Lite model for emotion recognition. Real-time emotion inference is performed by efficiently executing the TensorFlow Lite model directly on the Android smartphone, resulting in reduced latency and a smooth user experience.

The system's user interface (UI) offers simple controls for taking pictures, selecting photographs from the device's collection, and viewing real-time emotion forecasts. Visual feedback, such as applying emotion labels or emoticons to recognized faces, improves user involvement and comprehension of the system's output.

Overall, the system is a complex integration of the ML Kit and TensorFlow frameworks within the Android development environment, providing a reliable and efficient solution for facial emotion recognition on mobile devices. The system's seamless mix of machine learning algorithms and mobile technologies allows users to explore the intriguing world of emotion detection straight from their smartphones or tablets.

4.2 Dataset

To train the CNN model for face emotion identification, a hybrid dataset was created by merging photos from various available datasets, such as CK+ [51], JAFFE [52], FER2013 [53], and RAF-DB [54]. The CK+ dataset was used for all photos except those tagged as disdain, and the whole JAFFE and FER2013 datasets were incorporated. The RAF-DB dataset was limited to photos from the happy class. All pictures were preprocessed to match the FER2013 dataset format, which included turning them to grey-scale and shrinking them to 48x48 pixels.

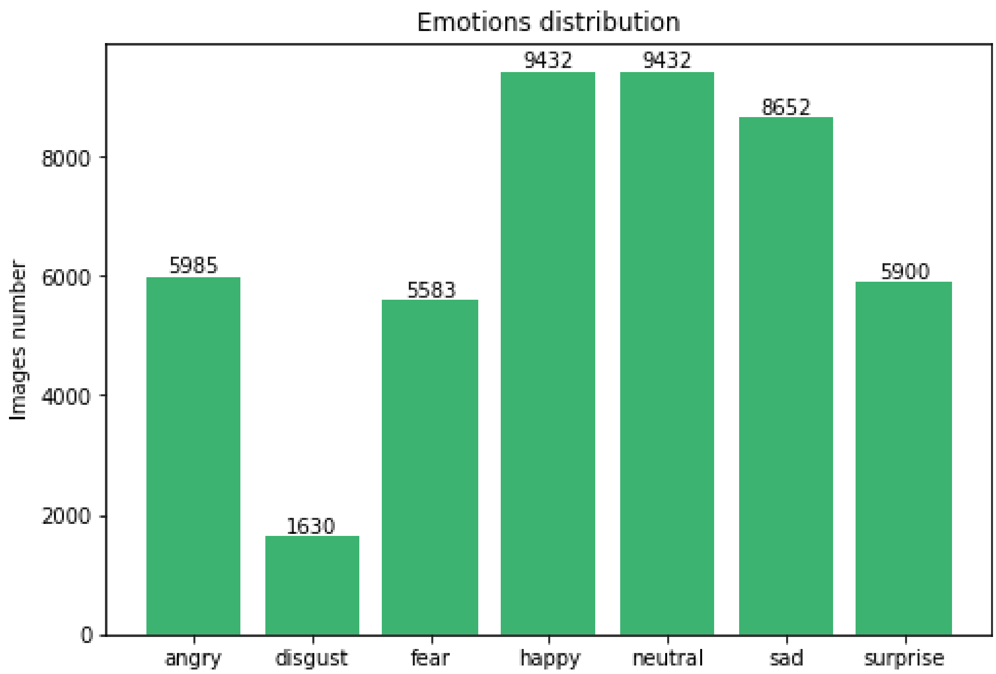

Figure 4.1: Distribution of the Dataset

The CNN model used for face emotion categorization is a deep convolutional neural network, which was trained and deployed as *simple_classifier.tflite*. This model was particularly created and refined for accurate emotion recognition tasks. As part of the preprocessing pipeline, the input pictures' pixel values were normalized from the original range of [0, 255] to [0, 1].

4.3 Training and Testing

The hybrid dataset was divided into two subsets: a training subset (with 80% of the data) and a testing subset (with the remaining 20%). This partitioning guarantees that the model gets trained on a broad variety of samples while simultaneously confirming its performance on previously unknown data during testing. By training the CNN model on a hybrid dataset drawn from several sources, the system has access to a broader range of facial expressions and emotional subtleties, improving its capacity to generalize to many real-world settings.

During the training phase, numerous parameters and settings were provided to improve the CNN model's performance and convergence. The training settings are a minimal delta of 0.0001, a patience value of 10 for early stopping, the Adam optimizer with a learning rate of 0.0001, and the categorical cross-entropy loss function. The batch size of 96 was chosen to strike a compromise between computational efficiency and model convergence.

The CNN model's training method included creating a hybrid dataset from numerous existing datasets, preparing pictures to adhere to a common format, specifying training parameters, and splitting the training and validation. This method allows the CNN model to efficiently learn and categorize facial emotions, setting the framework for accurate and robust emotion identification in the deployed system.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_178 (Conv2D)          (None, 48, 48, 32)        320
_________________________________________________________________
conv2d_179 (Conv2D)          (None, 48, 48, 32)        9248
_________________________________________________________________
conv2d_180 (Conv2D)          (None, 48, 48, 32)        9248
_________________________________________________________________
conv2d_181 (Conv2D)          (None, 48, 48, 32)        9248
_________________________________________________________________
max_pooling2d_70 (MaxPooling (None, 23, 23, 32)        0
_________________________________________________________________
dropout_90 (Dropout)         (None, 23, 23, 32)        0
_________________________________________________________________
conv2d_182 (Conv2D)          (None, 23, 23, 64)        18496
_________________________________________________________________
conv2d_183 (Conv2D)          (None, 23, 23, 64)        36928
_________________________________________________________________
conv2d_184 (Conv2D)          (None, 23, 23, 64)        36928
_________________________________________________________________
max_pooling2d_71 (MaxPooling (None, 11, 11, 64)        0
_________________________________________________________________
dropout_91 (Dropout)         (None, 11, 11, 64)        0
_________________________________________________________________
conv2d_185 (Conv2D)          (None, 11, 11, 128)       73856
_________________________________________________________________
conv2d_186 (Conv2D)          (None, 11, 11, 128)       147584
_________________________________________________________________
max_pooling2d_72 (MaxPooling (None, 5, 5, 128)         0
_________________________________________________________________
dropout_92 (Dropout)         (None, 5, 5, 128)         0
_________________________________________________________________
conv2d_187 (Conv2D)          (None, 5, 5, 256)         295168
_________________________________________________________________
max_pooling2d_73 (MaxPooling (None, 2, 2, 256)         0
_________________________________________________________________
dropout_93 (Dropout)         (None, 2, 2, 256)         0
_________________________________________________________________
flatten_18 (Flatten)         (None, 1024)              0
_________________________________________________________________
dense_36 (Dense)             (None, 1024)              1049600
_________________________________________________________________
dropout_94 (Dropout)         (None, 1024)              0
_________________________________________________________________
dense_37 (Dense)             (None, 7)                 7175
=================================================================
Total params: 1,693,799
Trainable params: 1,693,799
Non-trainable params: 0
_________________________________________________________________
```

Figure 4.2: The Neural Network Structure

Chapter 5

Results

5.1 Performance Evaluation

The performance metrics acquired from assessing the convolutional neural network (CNN) model on the specified test subset are important indicators of the model's efficacy in face emotion recognition. With an accuracy of 67.8%, the model exhibits its ability to properly categorize approximately 68% of the occurrences in the test subset, demonstrating its overall efficacy in recognizing distinct facial expressions. Furthermore, the model's accuracy score of 66.2% demonstrates its ability to properly identify real positive predictions while limiting false positives, assuring the trustworthiness of its classifications. In addition to precision, the F1 score of 64.7% gives a comprehensive assessment of the model's performance, taking into account both precision and recall. This score represents the model's ability to accurately categorize examples while also retrieving relevant instances from the dataset. When combined, these measures provide full insights into the CNN model's performance in face emotion identification tasks, leading further refining and optimization for improved accuracy and resilience in real-world deployment settings.

Figure 5.1: Confusion matrix of the tensorflow lite model

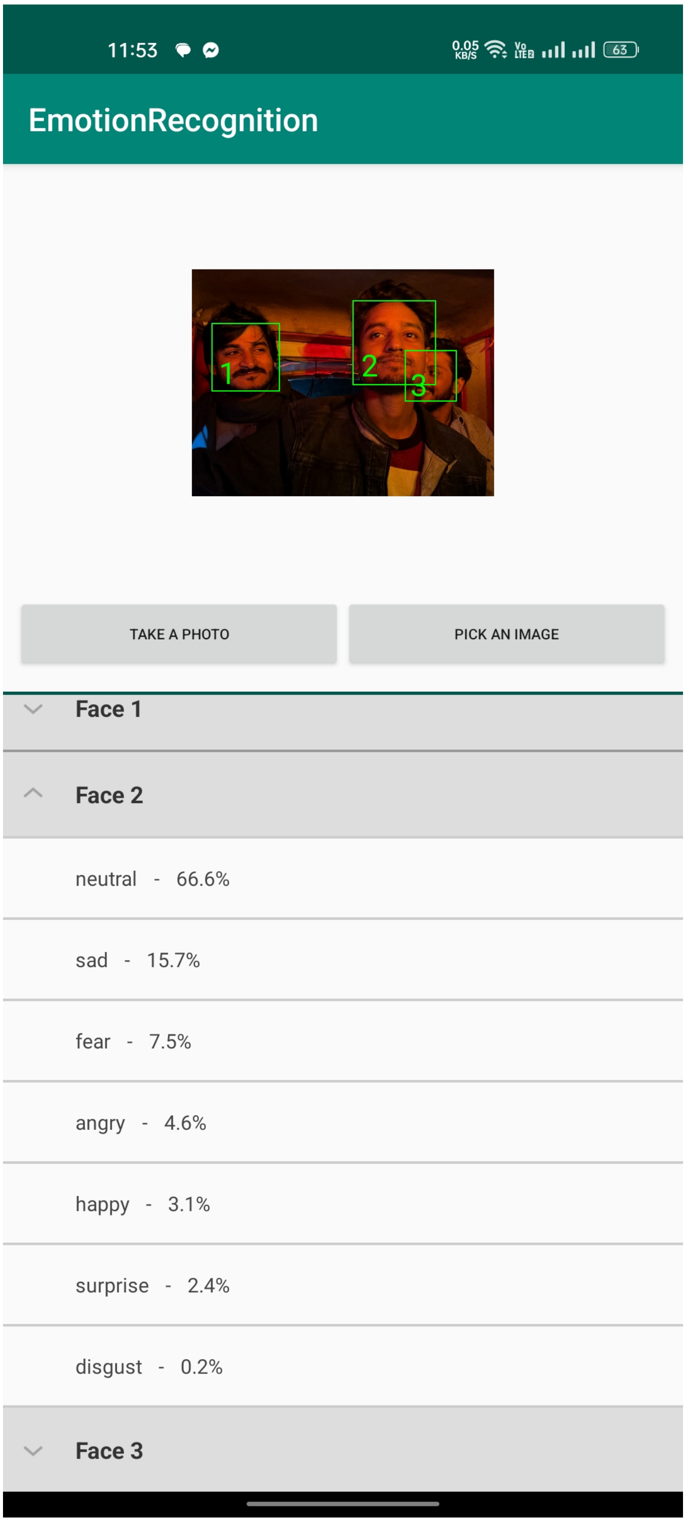

Figure 5.2: Practical emotion recognition

Chapter 6

Future Work

Continuously soliciting user feedback and conducting iterative improvements is essential for maintaining the relevance and effectiveness of the system over time. Future work could involve establishing mechanisms for collecting user feedback, such as in-app surveys, feedback forms, or user analytics. Analyzing user feedback and usage patterns can provide valuable insights into user preferences, pain points, and feature requests, which can inform future iterations and updates to the system. Additionally, adopting agile development methodologies, such as continuous integration and deployment (CI/CD), enables rapid iteration and deployment of new features and improvements based on user feedback.

To achieve this, we will continue to work on expanding the system's reach. This might include creating versions of the system for iOS, web browsers, and desktop platforms utilizing cross-platform frameworks like React Native. By assuring compatibility with a diverse variety of devices and platforms, the system may reach a bigger audience and be smoothly incorporated into a web API capable of analyzing a wide range of fresh datasets for future study.

As the subject of facial expression detection evolves, there are several opportunities for additional research and development to improve the system's capabilities and usefulness. The points below suggest potential areas for further work:

- Enhanced Emotion Recognition System: One area for future research is improving the performance and accuracy of the emotion identification system. This might include experimenting with sophisticated deep learning architectures like recurrent neural networks

(RNNs) or attention mechanisms to better capture temporal dependencies and spatial linkages in face emotions. Furthermore, combining multimodal input sources, such as audio or text data, may give additional information to improve emotion identification accuracy. Furthermore, studying approaches for dealing with unbalanced datasets and reducing bias in emotion detection algorithms might result in more robust and equitable systems.

- Improved model's generalization capability: Future research should focus on data augmentation techniques and transfer learning procedures to increase the model's generalizability across varied populations and contexts. By supplementing the training dataset with synthetic or augmented data and fine-tuning the model on domain-specific datasets, the model may better adapt to differences in facial expressions, demography, and cultural background. Furthermore, investigating approaches for domain adaptation and model calibration might assist lessen the influence of dataset biases and improve the model's performance in real-world circumstances.

- Comparative Analysis: Comparative assessments with existing face emotion detection systems and benchmarks are critical for evaluating the system's performance and identifying development opportunities. Future study might include comparing the system against cutting-edge models and datasets and assessing its performance across a variety of metrics, datasets, and situations. Comparative analysis can give useful insights into the system's strengths and limitations in comparison to current solutions, which can help guide future research and optimization.

Chapter 7

Conclusion

After investigating the potentials and limitations of Emotion Recognition on Android devices, our study has shifted to developing a system capable of leveraging Google's ML Kit for facial structure detection while smoothly incorporating TensorFlow Lite for emotion classification. The development procedure included the use of ML Kit for face structure identification and TensorFlow Lite for deep learning-based emotion categorization. A hybrid dataset containing photos from the CK+, JAFFE, FER2013, and RAF-DB datasets supplied various training examples, which improved the model's knowledge of face emotions. The convolutional neural network (CNN) model improved its ability to detect and categorize face emotions after extensive training and tuning, thus establishing a firm platform for ongoing innovation and improvement in the dynamic field of emotion analysis. Moving ahead, there are interesting avenues for further study and development. These include investigating advanced deep learning architectures, improving the model's adaptability across diverse demographics, ensuring cross-platform compatibility, collecting user feedback through another system for iterative improvements, and conducting comparative analyses with existing FER systems.

esis

References

[1] "Ml kit sdk keeps all machine learning on the device," 2020. [Online]. Available: https://medium.com/mobilepeople/ml-kit-sdk-keeps-all-machine-learning-on-the-device-4cc581696afd

[2] S. S. Kulkarni, N. P. Reddy, and S. Hariharan, "Facial expression (mood) recognition from facial images using committee neural networks," *Biomedical engineering online*, vol. 8, no. 1, pp. 1–12, 2009.

[3] "Face detection concepts," 2024. [Online]. Available: https://developers.google.com/ml-kit/vision/face-detection/face-detection-concepts

[4] "Using tensorflow lite on android," 2018. [Online]. Available: https://blog.tensorflow.org/2018/03/using-tensorflow-lite-on-android.html

[5] B. Montagne, R. P. Kessels, E. H. De Haan, and D. I. Perrett, "The emotion recognition task: A paradigm to measure the perception of facial emotional expressions at different intensities," *Perceptual and motor skills*, vol. 104, no. 2, pp. 589–598, 2007.

[6] S. N. Nagalkar and S. Sali, "Emotion recognition using facial expressions," *International Journal of Advance Research and Innovative Ideas in Education*, vol. 5, pp. 3023–3028, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:149660265

[7] C. Maaoui and A. Pruski, "Emotion recognition through physiological signals for human-machine communication," *Cutting edge robotics*, vol. 2010, no. 317-332, p. 11, 2010.

[8] P. Tarnowski, M. Kołodziej, A. Majkowski, and R. J. Rak, "Emotion recognition using facial expressions," *Procedia Computer Science*, vol. 108, pp. 1175–1184, 2017.

[9] N. Silberman, "Tf-slim: A lightweight library for defining, training and evaluating complex models in tensorflow," 2017.

[10] A. Pavithra and S. M. Prakash, "Tensorflow in deep learning," *International Journal of Innovative Research in Technology*, 2019.

[11] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "{TensorFlow}: a system for {Large-Scale} machine learning," in *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 2016, pp. 265–283.

[12] L. Sun, Y. Zhang, Z. Li, Z. Zhou, and Z. Zhou, "inml kit: empowering the prototyping of ml-enhanced products by involving designers in the ml lifecycle," *AI EDAM*, vol. 36, p. e8, 2022.

[13] S. Miret, K. L. K. Lee, C. Gonzales, M. Nassar, and M. Spellings, "The open matsci ml toolkit: A flexible framework for machine learning in materials science," *arXiv preprint arXiv:2210.17484*, 2022.

[14] N. Abdullah and M. Yaacob, "I-virtual lab kit with microcomputer-based laboratory (mbl) system for basic electronics," in *Journal of Physics: Conference Series*, vol. 1185, no. 1. IOP Publishing, 2019, p. 012037.

[15] E. Bisong and E. Bisong, "Google cloud machine learning engine (cloud mle)," *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, pp. 545–579, 2019.

[16] Q. Gu, A. Kumar, S. Bray, A. Creason, A. Khanteymoori, V. Jalili, B. Grüning, and J. Goecks, "Galaxy-ml: An accessible, reproducible, and scalable machine learning toolkit for biomedicine," *PLoS computational biology*, vol. 17, no. 6, p. e1009014, 2021.

[17] D. E. King, "Dlib-ml: A machine learning toolkit," *The Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[18] S. S. Narayanan, "Emotion recognition system," 2012, uS Patent 8,209,182. [Online]. Available: https://patents.google.com/patent/US8209182B2/en

[19] Y. N. Y. H. Yumiko Kato, Takahiro Kamai, "Emotion recognition apparatus," 2012. [Online]. Available: https://patents.google.com/patent/US8204747B2/en

[20] Y.-H. Yang, Y.-F. Su, Y.-C. Lin, and H. H. Chen, "Music emotion recognition: The role of individuality," in *Proceedings of the international workshop on Human-centered multimedia*, 2007, pp. 13–22.

[21] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," 1971.

[22] W. V. F. Paul Ekman, "Facial action coding system: a technique for the measurement of facial movement," 1978.

[23] R. M. Neal, "Pattern recognition and machine learning," 2007. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1198/tech.2007.s518

[24] T. Burr *et al.*, "Pattern recognition and machine learning. christopher m. bishop," *Journal of the American Statistical Association*, vol. 103, pp. 886–887, 2008.

[25] Y. Anzai, *Pattern Recognition Machine Learning*, 2016. [Online]. Available: https://dl.acm.org/doi/10.5555/3002487

[26] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[27] W. Wei, Q. Jia, Y. Feng, and G. Chen, "Emotion recognition based on weighted fusion strategy of multichannel physiological signals." 2018.

[28] Y. LeCun, Y. LeCun, Y. Bengio, G. E. Hinton, and G. E. Hinton, "Deep learning," 2015.

[29] F. Julin, "Vision based facial emotion detection using deep convolutional neural networks," 2019.

[30] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," 1971.

[31] E. Yagis and M. Unel, *Facial Expression Based Emotion Recognition Using Neural Networks*, 2018.

[32] P. Shetty and D. Jagli, "Human emotion recognition using machine learning," 2019.

[33] *Face Emotion Detection Using Deep Learning*, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9563343

[34] Y. LeCun, L. Bottou, L. Bottou, Y. Bengio, Y. Bengio, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," 1998.

[35] *Very Deep Convolutional Networks for Large-Scale Image Recognition*, 2015.

[36] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings fourth IEEE international conference on automatic face and gesture recognition (cat. No. PR00580)*. IEEE, 2000, pp. 46–53.

[37] R. Ranjan, V. M. Patel, and R. Chellappa, "A deep pyramid deformable part model for face detection," in *2015 IEEE 7th international conference on biometrics theory, applications and systems (BTAS)*. IEEE, 2015, pp. 1–8.

[38] R. A. Calvo and S. K. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," 2010.

[39] s. P. M. S. R. P. R. T. Ramya, Y. Anurag, "Facial emotion recognition," *YMER*, 2022.

[40] C. D. Kashyap and P. R. Vishnu, "Facial emotion recognition," *International Journal of Engineering and Future Technology*, vol. 7, no. 7, pp. 18–29, 2016.

[41] W. Su, L. Chen, M. Wu, M. Zhou, Z. Liu, and W. Cao, "Nesterov accelerated gradient descent-based convolution neural network with dropout for facial expression recognition," in *2017 11th Asian Control Conference (ASCC)*. IEEE, 2017, pp. 1063–1068.

[42] E. S. M. Tiwari and E. M. S. Alam, "Facial emotion recognition."

[43] J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in *Conference on fairness, accountability and transparency*. PMLR, 2018, pp. 77–91.

[44] A. Dhall, R. Goecke, J. Joshi, K. Sikka, and T. Gedeon, "Emotion recognition in the wild challenge 2014: Baseline, data and protocol," in *Proceedings of the 16th international conference on multimodal interaction*, 2014, pp. 461–466.

[45] N. Sebe, M. S. Lew, I. Cohen, A. Garg, and T. S. Huang, "Emotion recognition using a cauchy naive bayes classifier," in *2002 International conference on pattern recognition*, vol. 1. IEEE, 2002, pp. 17–20.

[46] F. Ahmed, A. H. Bari, and M. L. Gavrilova, "Emotion recognition from body movement," *IEEE Access*, vol. 8, pp. 11 761–11 781, 2019.

[47] R. W. Picard, "Affective computing: from laughter to ieee," *IEEE transactions on affective computing*, vol. 1, no. 1, pp. 11–17, 2010.

[48] "Face detection," 2024. [Online]. Available: https://developers.google.com/ml-kit/vision/face-detection

[49] "Face detection concepts," 2024. [Online]. Available: https://developers.google.com/ml-kit/vision/face-detection/face-detection-concepts

[50] "Tensorflow lite," 2022. [Online]. Available: https://www.tensorflow.org/lite/guide

[51] "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," 2010. [Online]. Available: https://www.researchgate.net/publication/224165246_The_Extended_Cohn-Kanade_Dataset_CK_A_complete_dataset_for_action_unit_and_emotion-specified_expression

[52] "The japanese female facial expression (jaffe) dataset," 1998. [Online]. Available: https://zenodo.org/records/3451524.XuHa20UzZPY

[53] "fer2013," 2018. [Online]. Available: https://www.kaggle.com/datasets/deadskull7/fer2013

[54] "Real-world affective faces database." [Online]. Available: http://whdeng.cn/RAF/model1.html