DEEP LEARNING FOR EMOTION RECOGNITION USING IOMT DATA

Department of Computing and Mathematics

Manchester Metropolitan University

Student Name: Hrishikesh Hemant Prabhu

MMU ID: 22543201

Course: MSc Artificial Intelligence

Supervisor: Dr. Seun Ajao

Academic Year: 2023

TABLE OF CONTENTS

TABLE OF FIGURES

Chapter 1: Introduction

1.1 Introduction

Emotion detection is a crucial component in the field of the healthcare sector which includes interaction among healthcare systems, Human-computer Interaction, and robotics. The combination of IoMT means the Internet of Medical Things for the involvement of the collection of different data from the wearable device and sensors which will provide opportunities for emotion recognition by deep learning algorithms. This research intended to utilize the IoMT data for the exploration of emotion recognition to meet the objectives intended with this research. Emotion detection becomes important at the moment of decision-making, communication, and normal well-being of human interactions. Emotion detection has extensive effects in a variety of fields which includes marketing, healthcare, and others depending on the ability to recognize the emotions in others. Through the identification of the aim and objectives of this research, the research questions and hypotheses have been derived. To understand the relevance of this research the research rationale and research significance have been discussed.

1.2 Background

The “Internet of Medical Things” or IoMT and emotion detection both are important fields that come together for the context of this study. The advancement of automated approaches and the rise of data accessibility have revolutionized the emotion detection fields that have been formerly restricted by the study of neuroscience and psychology. With the rise, there have been significant possibilities for the study of human emotions which will go beyond the limitations that have been established. The IoMT resulted in an era of continuous medical data production. Massive, comprehensive maintenance of real-time data has been generated because of the network of interconnected medical equipment and sensors (Rathour *et al.,* 2021). The vast amount of knowledge extends beyond the boundaries of hospitals and into the everyday lives of individuals, offering real-time insight for understanding their physiological and patterns of conduct. The integration of the detection of emotions using IoMT data takes the use of the possibility of technology for comprehending emotional states in a broad range of scenarios. The diversity of data sources is unrestricted ranging from wearables that determine heart rate variations during engagements to smart home technologies which capture facial expressions and speech pronunciations. This research's foundations are built on such data, which consists of physiological indications, textual indications, and audiovisual indicators. It transcends disciplinary boundaries when these disciplines are combined, providing an opportunity for significant implications in numerous other areas. Accurate emotion detection by applying IoMT data has an enormous opportunity for enhancing healthcare, which urgently requires more accurate diagnoses and personalized treatments (Rachakonda *et al.*, 2020). The possible applications involve the early identification of emotional illnesses, monitoring mental health, and establishing better patient-centred methods for treatment. But this integration also brings complex challenges. IoMT data sources' variety necessitates the use of adaptable and dependable models for computation. Complex algorithms that are capable of decoding complex patterns are required considering the complicated nature of human emotions, which are often delicate and context dependent.

1.3 Aims and Objectives

The main aim or goal of this research comes with the help of the investigation of the utilization of deep learning models for the detection of emotion utilizing the IoMT.

The objectives that relevant to this research are being presented in this section.

- To identify the relevant questions for this research which will help in the execution of the research according to the objectives.
- To gain a proper insight for the review of the literature with the understanding of the emotion recognition technologies.
- To establish suitable methods for data collection and conduct the research accordingly.
- To analyze the gaps in the literature in conjunction with the other constraints and to offer solutions for further study.

1.4 Research Questions

The questions that are relevant to this research are presented in this section.

1. To what extent does the IOMT data impact the domain of deep learning-based emotion recognition?
2. When it comes to preparing the IOMT data for the learning models, which are the best research methodologies and data processing techniques?
3. How far may the development of an emotion detection model take advantage of deep learning techniques?
4. What could be some potential current technological solutions that could be created employing an emotion recognition model?

1.5 Research Rationale

The urgent need for innovative healthcare practices, with a special focus on mental health evaluation and assistance, is the driving force behind this research. Emotions have a complex role in a person's overall health, and a precise understanding of them has the potential to reveal insights important for successful healthcare interventions. Traditional emotion evaluation techniques, on the other hand, frequently rely on subjectivity which can be personal, vulnerable to cognitive biases, and difficult to use in practical settings. To get over these restrictions, the combination of deep learning methods with IoMT data presents a game-changing possibility (Zikria *et al.,* 2020). IoMT devices provide continuous, non-invasive data collecting in healthcare settings, which is a trend. When combined with IoMT data, emotion recognition can offer unbiased perceptions of emotional states, improving the accuracy of diagnosis, the individualization of care, and the standard of overall patient care. This research aims to tap into an extensive array of emotional indications by using the multifaceted nature of IoMT data, which ranges from physiological signs to textual and visual cues (Kumar *et al.,* 2021). IoMT has the potential to be used for more than just discrete medical interactions, since continuous emotion monitoring provides a comprehensive picture of a person's emotional health across time, thereby allowing for the early identification of emotional illnesses and aiding prompt therapies. The wider landscape of human-computer interaction is taken into account in the research's justification. Understanding and responding to human emotions are crucial as virtual reality, augmented reality, and affective computing technologies proliferate. These fields may easily include an IoMT-based framework for emotion identification, improving user experiences, assisting in decision-making, and allowing adaptive technology. This research fits with the changing landscape of healthcare and technology integration by bridging the gap between the subtle nuances of human emotions and the capabilities of cutting-edge technology. The potential to transform healthcare practices, improve patient well-being, and progress emotion identification techniques underpins the study rationale, with ramifications across the clinical and technical spheres.

1.6 Research Significance

The need for novel healthcare approaches, particularly in evaluating and promoting mental health, is what motivated the research. Emotions have a substantial impact on general health, and proper detection of them might lead to efficient healthcare measures. However, traditional techniques of measuring emotions frequently rely on subjective disclosure, are vulnerable to bias, and are not useful in real-world situations (Alamelu and Thilagamani, 2022). Deep learning and IoMT data coming together offer a game-changing way to tackle these problems. The use of IoMT devices is becoming more widespread in the healthcare industry, enabling continuous, non-intrusive data collecting. Integrating emotion identification with IoMT data may offer objective insights into emotional states, improving the precision of diagnosis, the effectiveness of individualized care, and the standard of patient care. This research intends to tap into rich emotional indications by utilizing a variety of IoMT data components, including physiological signals, textual cues, and visual cues (Harrison and Hole, 2019). The benefits of IoMT go beyond specific medical situations since ongoing emotion monitoring provides a comprehensive picture of emotional health throughout time, possibly enabling early identification of emotional problems and prompt therapies. The justification includes human-computer interaction (Chughtai *et al.*, 2023). Understanding and reacting to human emotions become crucial as virtual and augmented reality become more widely used. The seamless integration of an IoMT-based framework for emotion identification can improve user experiences, support decision-making, and enable adaptive technology. This research fits with changing healthcare and technology integration by integrating human emotional subtleties with cutting-edge technological capabilities. The motivation ultimately stems from the ability to transform clinical and technology fields while also enhancing patient well-being and developing emotion identification techniques.

1.7 Research Framework

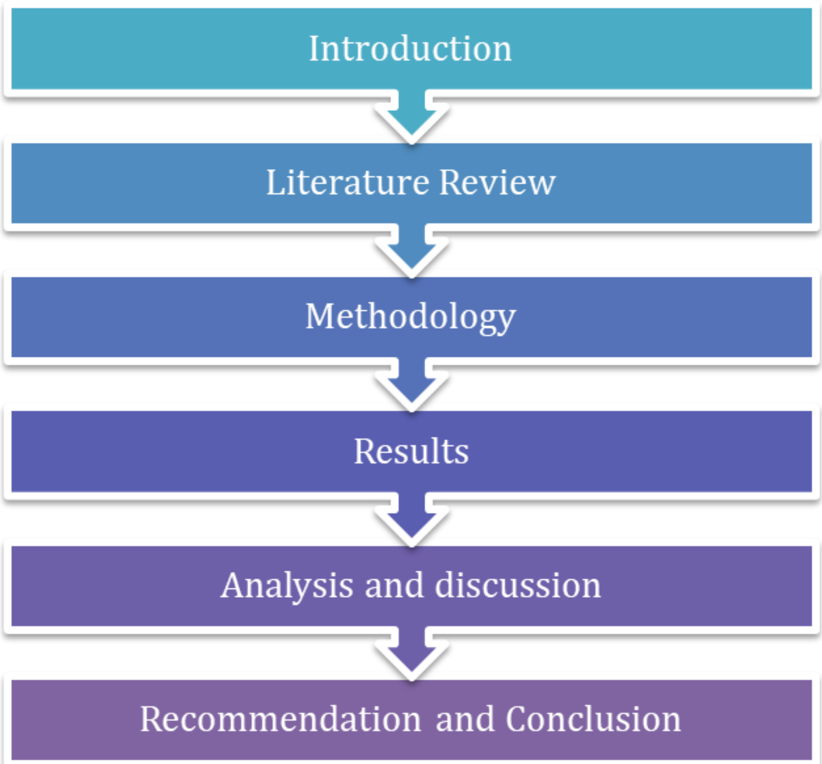


Figure 1: Research Framework

(Source: Self-made)

1.8 Conclusion

The introduction chapter establishes the foundations of the deep learning methods with IoMT data-based detection of emotions. Evaluation of Emotions and healthcare practices has the potential for the improvement of marketing strategies of a business, identifying the mental health of the individuals and others. This study aims the make significant contributions in the domain of emotion detection and healthcare innovation by utilizing the research goals including IoMT data sources, and specialized deep learning architectures. The importance of this study stems from the potential improvement of the quality of patient treatments specifically in mental health and to provide insights on the complex relationships between emotions and well-being. The effects of this research also have implications for the wider field of adaptive technology and human-computer interaction, which goes beyond healthcare. This study aims to expand our comprehension of both deep learning's abilities and the potential of IoMT data in influencing the future of emotion identification and its varied uses. Subsequent chapters will dig into techniques, findings, and applications.

Chapter 2: Literature Review

2.1 Introduction

The review of literature discusses the application of data through the Internet of Medical Things or IoMT and deep learning techniques toward emotion recognition. It highlights that Convolutional Neural Networks or CNNs and Recurrent Neural Networks or RNNs are adept at detecting emotions across a wide range of data kinds. The development of IoMT technologies has changed healthcare data collection by providing the opportunity to determine emotions in real-time by employing wearables and sensors. Context-aware and individualized emotion detection, especially for healthcare situations, is made possible by the integration of IoMT data with deep learning algorithms. With the incorporation of physiological indicators, textual sentiment analysis, and facial expressions, multi-modal data fusion advances emotion detection. Deep learning architectures capture complex interactions between many data sources including multi-modal CNNs and attention approaches. IoMT-driven emotion recognition has demonstrated potential in several kinds of healthcare applications, such as mental health monitoring, early emotional abnormality detection, and improving interactions between humans and robots. There are still challenges to be sorted out, like ensuring reliability in real-world circumstances, resolving ethical issues regarding data privacy, and enhancing the outcomes of models. The development of current architectures, exploration of novel data representations, and expanding IoMT-based emotion identification into additional fields should be among the primary objectives of future research.

2.2 Empirical Study

According to Rathour *et al.*, 2021, the technique called facial emotion recognition, or FER uses facial expressions to infer emotions from people. Data collection on physical and mental health has been made easier because of the Internet of Medical Things or IoMT. This article describes a real-time, IoMT-based face emotion detection and identification system that makes use of deep convolution neural networks and a Raspberry Pi device. The study examined emotional states and facial expressions using physiological sensors. Utilizing an Intel Movidius NCS CPU on a Raspberry Pi, the model demonstrated 73% accuracy on the FER 2013 dataset during testing. To identify any significant variations in heart rate and blood pressure across the three participants, the system further ran a t-test (Rathour *et al.*, 2021). At the intersection of the fields of computing, psychology, and the field of neurology methods based on deep learning have transformed the area of emotion recognition. To identify emotional signs from visual data, these techniques leverage sophisticated neural networks to learn hierarchical features from raw input. While Recurrent Neural Networks or RNNs capture temporal dynamics in sequential data, this allows precise sentiment analysis in textual data and subtle fluctuations in voice patterns. Convolutional Neural Networks or CNNs have demonstrated a remarkable ability to recognize facial expressions indicative of emotions. Through the incorporation of context-specific emotion identification with limited labeled data, the transfer learning paradigm has increased the bar for emotion recognition.

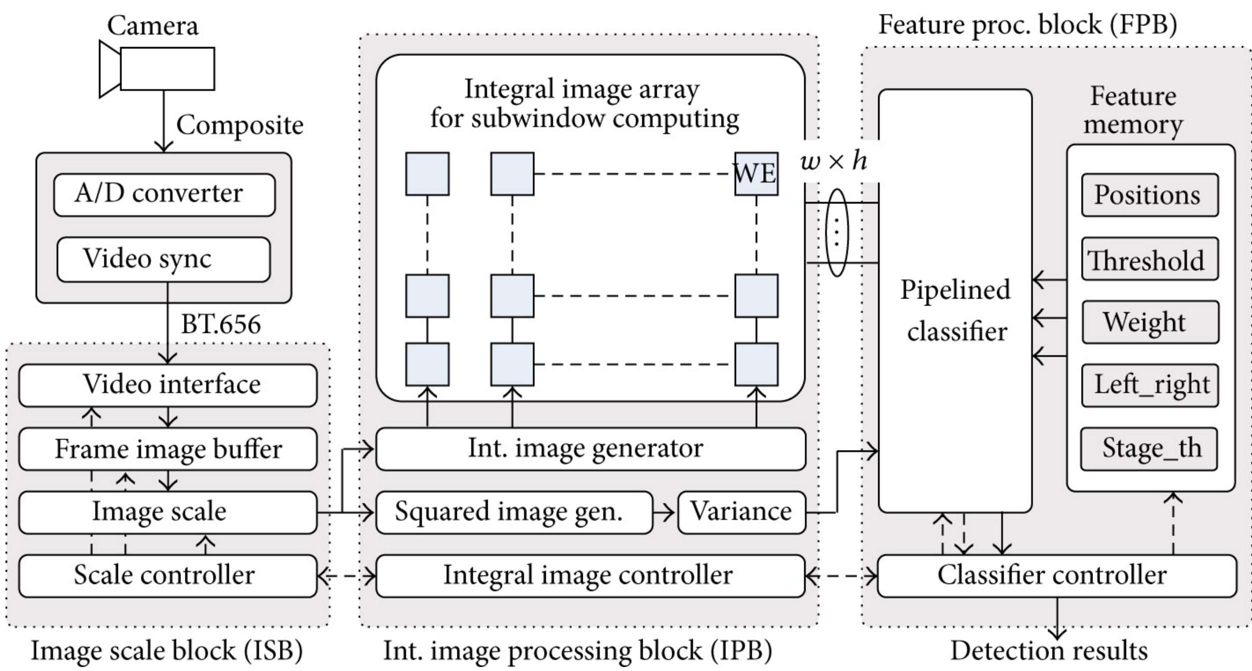


Figure 2: Face Recognition Architecture

(Source: https://journals.sagepub.com/doi/10.1155/2015/251386)

To give a thorough knowledge of emotions, multi-modal deep learning architectures integrate input from several sources, such as textual content, physiological signals, and facial expressions (Martínez and Vega, 2022). Notwithstanding, many obstacles persist, including the interpretability of the model and the laborious procedure of acquiring labelled data from diverse populations and environments. Wearable devices, which are moodables, ingestible sensors, and trackers are examples of IoMT sensors that have the potential to offer proactive healthcare solutions. However, since the data is so immediate, handling and interpreting it quickly becomes a critical task. The analysis of multidimensional time-series data collected from different wearables during the selection process for real-world health conditions conveys accuracy challenges and data overload. The computational intelligence of edge analytics may enhance the ability to predict by converting digital biomarker data into actions for remote monitoring and enabling warnings during emergencies. Nonetheless, the ubiquitous creation of data streams from IoMT presents difficulties for exploratory analysis and data visualization.

According to Gupta *et al.*, 2021, The fast proliferation of the "Internet of Medical Things or IoMT" has brought about significant changes to the environment for collecting and interpreting healthcare data. An interconnected collection of networked that is wearables, sensors, instruments, and medical equipment that collaborate to collect and transmit real-time health data is referred to as the "Internet of Medical Things" (IoMT). There are previously unheard-of chances to improve emotion identification techniques and applications thanks to this paradigm change in data collecting.

IoMT devices offer a revolutionary edge for emotion identification by enabling uninterrupted and invisible data collecting. Conventional techniques for assessing emotions, which frequently depend on self-reported information, can be biased and subjective. Yet, IoMT enables the recording of physiological signs, expressions on the face, speech patterns, and other non-verbal cues in real settings (Gupta et al., 2021). The dynamic character of emotions and the real-time nature of IoMT data mesh well. Emotions may change quickly, are context-dependent, and are fleeting. IoMT devices enable an advanced comprehension of emotional swings throughout the day because of their continuous data collection capabilities. This ongoing data stream is especially important for healthcare applications, as tracking emotional changes can help identify mental health issues early and inform individualized treatment plans. The contextual richness of healthcare facilities is used through the integration of emotion detection with IoMT data. Emotions are not a standalone phenomenon; they are impacted by things like pharmaceutical use, variables in the environment, and physical exercise. Contextual signals from IoMT devices can improve the understanding of emotional states (Young et al., 2012). For example, integrating wearable physiological data with outside variables like sleep schedules or medication adjustments might provide a more comprehensive understanding of a person's mental health. The accuracy and context-awareness of emotion identification are improved by the integration of IoMT data with deep learning models. Deep learning architectures can become used to the peculiarities of a person's emotional expression over time, producing predictions that are more individualized and accurate. Furthermore, the IoMT's scalability makes it possible to create huge, diverse datasets, which promotes training as well as validation of deep learning models on a wider range of emotional expressions.

According to Cai et al., 2022, IoMT and emotion recognition working together have the potential to revolutionize healthcare and other industries. IoMT devices provide continuous, unobtrusive data collection, which perfectly fits with the dynamic nature of emotions. The accuracy, context awareness, and possible applications of emotion identification are improved by combining this data with deep learning approaches. IoMT technologies are set to transform our knowledge of emotions as they develop, providing new perspectives on mental health, human-computer conversations, and other fields where emotions are critical. On the Internet of Medical Things (IoMT), epilepsy detection utilizing electroencephalogram (EEG) and intelligent technologies is essential for health analysis and diagnosis. Federated instruction, a distributed learning framework, may use local data to train a shared model from several edge nodes, supporting the growth of IoMT (Cai et al., 2022). However, EEG recordings reveal varied distributions in various devices, instances, and subjects, which lowers the identification model's precision.

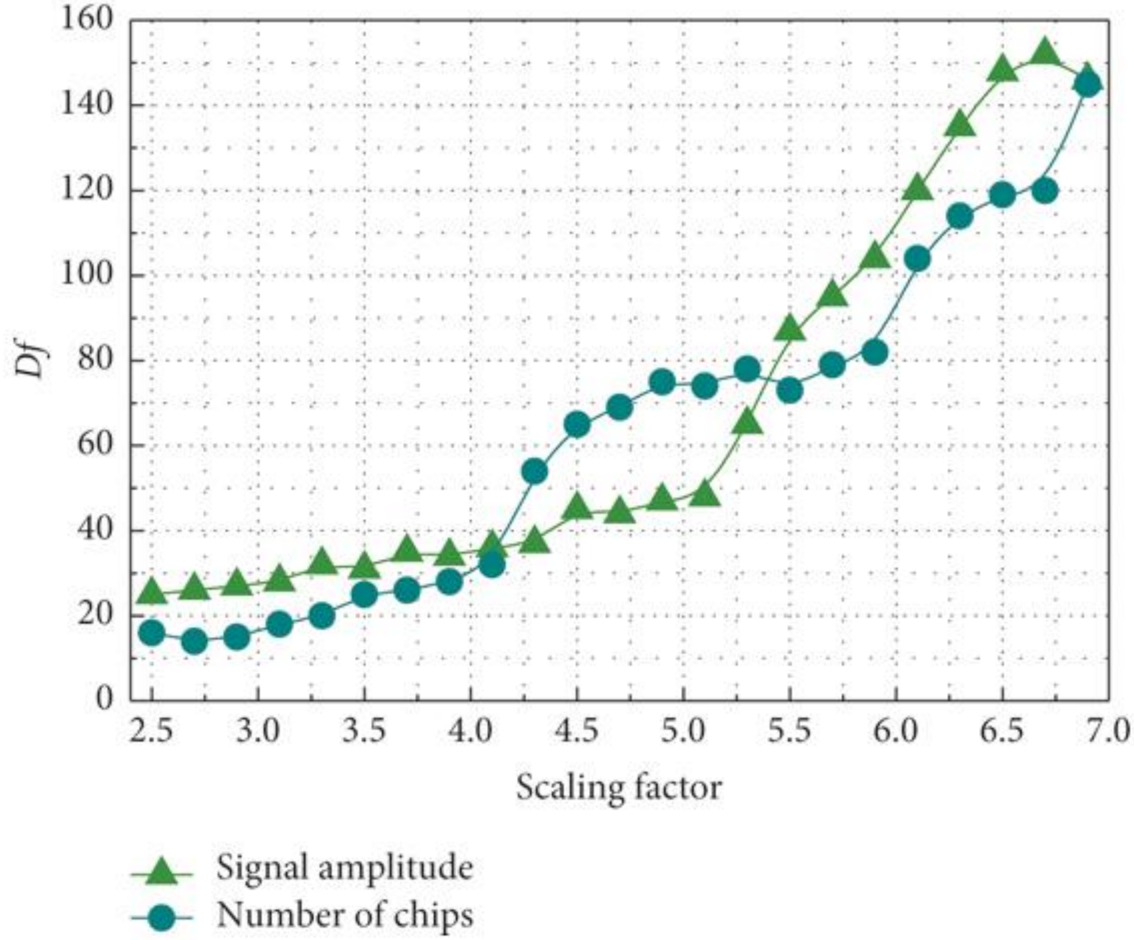


Figure 3: HDA-PDPL mode

(Source: https://www.hindawi.com/journals/complexity/2021/5592850/)

A hierarchical domain adaptation projective dictionary pair learning (HDA-PDPL) model is created to enhance classification performance in IoMT. This model learns to synthesize and analyze dictionary pairs in each layer, integrating EEG inputs from various domains into a collection of hierarchical subspaces. The model makes links across several domains by using a domain adaptation term on sparse coding and a nonlinear transformation function to seek systematic feature projection. According to experimental findings on two EEG epilepsy classifications, the HDA-PDPL model performs better than previous comparisons because it makes use of more common information across several domains.

According to Zhang et al., 2020, To administer healthcare, the Internet of Medical Things or IoMT technology integrates wireless devices and medical equipment. This paper investigates its potential to assist the COVID-19 epidemic by the application of machine learning algorithms and providing recommendations for emotional treatment. Given that everyone is interconnected and under observation via a cognitive network, the cognitive model is most appropriate in this circumstance. The article suggests an IoMT system that gathers multimodal patient data in surveillance settings while being emotion-aware and intelligent. This platform provides remote monitoring of health and decision-making, which helps to provide easy and ongoing healthcare services that are sensitive to emotions (Zhang et al., 2020). The suggested framework works better than conventional models, according to experimental findings, making it a practical alternative for handling a high number of devices during the pandemic. The suggested “Motion Energy Image-based Principal Component Analysis of the Network (MEI-PCANet)”, “Appearance and Motion DeepNet (AMDN)” model, “GoogLeNet, and VGG16” algorithms are compared to the proposed “Hierarchical Visual Cognition (HVC)” model. The “MWLD-based HVG model” surpasses the other two methods, while the “WLD-based HVG” model performs better than the “SIFT-based HVG” model.

The developed deep learning scheme and classification model outperform the widely used techniques because of their extremely effective durable feature descriptions and model. The designed detection framework, proper deep network layers, efficient learning strategy, and distinct dictionary learning strategy make the suggested model superior to other models under the same conditions. The effectiveness of the proposed horizontal visual cognition model is demonstrated by the HVG-based sparse categorization scheme's improvement of recognition rate. The designed structured computational cognition simulation model, which relies on biological concepts and an abdominal route model, greatly enhances the accuracy of detection and identification. The suggested model performs similarly to other models, however, the AMDN model performs moderately since the network model is too complex and there is an excessive need for training data.

According to Zehra et al ., 2021, Due to being associated with humans from different backgrounds in culture and language, researchers have had difficulty reliably identifying emotional reactions from robots. Multilingual contexts are unable to lend themselves to traditional speech-emotion recognition techniques, primarily employing the same corpus for training and testing classifiers. A series of experiments concentrate on the collective learning influence of an overwhelming voting-based, cross-corpus, multilingual emotion identification of speech system. In the research, the effectiveness of a combination of ensemble learning methods and conventional algorithms for machine learning has been analyzed (Zehra *et al.*, 2021). The findings suggest that various classifiers offer the maximum accuracy for various corpora. Employing a collaborative approach to learning removes the requirement to choose only a single classifier and runs the risk of decreasing the accuracy of a specific language corpora by integrating the consequences of all classifiers. According to experiments, accuracy increased during within-corpus testing in Urdu by 13%, German by 8%, Italian by 11%, and English by 5%. Improvements of 2% and 15% for training on Urdu and German data, and 7% and 5% for testing on Urdu and German data, have been demonstrated in cross-corpus trials. Comparing the ensemble learning approach with other innovative approaches, it offers promising outcomes. With the development of Internet-of-Medical Things or IoMT systems, the Internet-of-Things also known as IoT has transformed several networking applications, including healthcare systems. With the assistance of these technologies, patients with persistent diseases may be followed remotely, allowing for quick diagnoses that could potentially save lives. Security in these critical structures continues to remain a very difficult problem, though. It has been discussed innovative techniques for protecting data in IoMT systems against threats including physical and network hacking as the information is being collected, transmitted, and kept (Arunkumar, 2018). The results indicate that a great deal of security measures does not take into account various kinds of attacks, which inspired the development of a security framework that integrates a variety of techniques that satisfy IoMT security needs while preventing a significant number of known attacks.

According to Khan, 2022, Another significant and developing topic of research in the discipline of pattern recognition is facial emotion recognition (FER). Nonverbal communication serves a crucial part in daily interactions and contributes between 55% and 93% of all communication. In surveillance motion pictures, gesture recognition, expression analysis, computer games, smart homes, anxiety therapy, depression treatment, patient monitoring, psychoanalysis, operator tiredness detection, paralinguistic communication, and robotics, facial emotion analysis is successfully applied. This review is based on several deep learning (DL) techniques alongside traditional machine learning (ML) methods. Various freely accessible FER datasets for assessment metrics are reviewed and compared with benchmark results.

Facial Emotion Recognition Using Traditional Machine Learning Approaches

The standard feature extraction and classification techniques used in traditional machine learning methods for face emotion identification. Face photos are used to extract features such as face landmarks, texture patterns, and color distributions (Ali *et al.*, 2020). These characteristics are then used to categorize emotions by feeding them into classifiers like Support Vector Machines (SVM) or Random Forests. These techniques are capable of achieving a respectable level of accuracy, but they frequently have trouble catching subtle facial characteristics and need careful feature architecture.

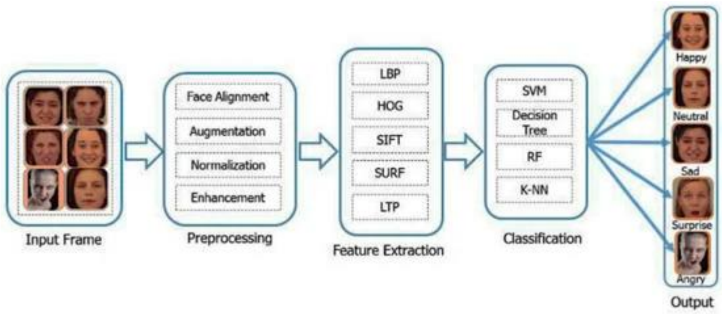


Figure 4: Facial emotion recognition or FER process

(Source: https://doi.org/10.3390/info13060268)

Facial Emotion Recognition Using Deep-Learning-Based Approaches

Convolutional Neural Networks (CNNs), a deep learning-based technique, have completely changed face emotion identification. CNNs can capture complex face aspects because they automatically learn organizational features from raw picture data. They are excellent at learning from beginning to end, thus explicit feature engineering is not necessary (Abdullah *et al.*, 2018). By identifying intricate patterns and spatial correlations in face expressions, pre-trained CNNs refined on emotion-specific datasets exhibit astounding accuracy. This strategy works better than more established ones and generalizes effectively to a wide range of emotions and people.

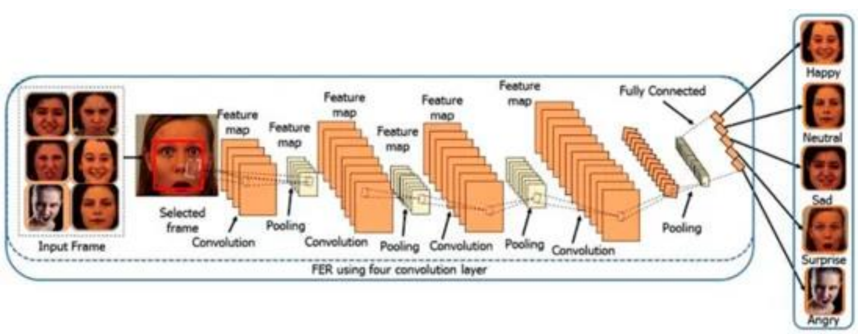


Figure 5: CNN model training for face emotion identification

(Source: https://doi.org/10.3390/info13060268)

According to Jaiswal *et al.*, 2020, The increasing need for automation in various fields, including personal robots, has led to robots needing to understand human emotions to provide personalized treatment. Foreseeing human inclination has been a difficult issue for about 10 years. This paper presents a model that progressively utilizes a convolutional brain organization to foresee human feelings from pictures. In comparison with "Vanilla CNN", the model's parameters would have decreased "by 90-fold, and by 50-fold" from the study. Eight unique datasets, comprising Fer2013, TFEID, Chicago Face Database, CK and CK+, FEI Face dataset, JAFFE Dataset, IMFDB, and custom datasets were used to test the network (Jaiswal *et al.*, 2020). The network's precision of 74% exceeds the state-of-the-art accuracy while requiring less computational work. Considering emotion is a subjective term, the problem is in employing the knowledge and science underneath labeled data and extracting each of the components that compose form emotion. This study presents two methods for emotion recognition using deep learning in computer vision. While the second version applies an 8-layer CNN, the initial one uses autoencoders to provide distinctive emotional representations. 100 candid photos extracted from the "Labelled Faces in the Wild or LFW dataset" were used for assessing both models, which were trained on the "posed-emotion dataset or JAFFE". The findings highlight that the CNN model can beat the most recent methods for identifying emotions with more depth and fine-tuning. To enhance performance, the authors additionally propose ways to expand the concept of expressive autoencoders.

Human feelings or emotions are unconstrained mental states joined by physiological changes in facial muscles, and non-verbal specialized strategies like looks are generally utilized in human-PC cooperation. Feeling acknowledgment is a perplexing undertaking because of the absence of a plan for recognizing facial feelings and the intricacy of the elements utilized in AI calculations. Convolutional brain organizations (CNN) have been produced for facial feeling acknowledgment, as looks are significant in nonverbal correspondence. Despite extensive research, the human visual system still falls behind (Giri *et al.,* 2022). The "Viola-Jones method" has been employed in this study for recognizing the lips and eyes from a face, and neural networks were then used to identify sentiments from the arrangement of the mouth and eyes. The analysis suggests an efficient approach to recognizing the seven human emotions like anger, contempt, disgust, fear, pleasure, sorrow, and surprise from a forward face picture. This study incorporates idiomatic terms, eye combing, and body movement methods to analyze body language to recognize and extract human emotions which include anger, contempt, happiness, fear, sadness, peacefulness, and surprise. The most often implemented technique refers to mental emotions and the condition of the mind via facial expression. Due to the lack of an apparent prototype or framework for distinct feelings and the difficulties in correctly recognizing facial emotion expressions, identifying emotions is occasionally problematic.

Convolutional Neural Networks or CNNs have been suggested as a successful approach for recognizing and distinguishing between different emotions, especially anger, contempt, happiness, fear, sorrow, peacefulness, and surprises. Profiling of users is evolving into something deeper, considering earlier disregarded characteristics of users. Understanding people's habits, preferences, and behaviours depends on the content they produce on social networks and the internet. A precise and dependable technique for identifying emotions from the text in routine writing situations is something researchers have an interest in creating because emotions experienced are important for a comprehensive user profile. The article suggests an approach to classification that has been effectively applied to numerous datasets utilizing "deep neural networks, Bi-LSTM, CNN, and self-awareness". Modern datasets have been utilized to verify the model's validity, and the model is suggested as a place to start for additional research in the area. The primary aspect of sentiment analysis is "emotion detection or ED", enabling service providers to provide individualized services. Text mining and analysis have grown increasingly important to the success of organizations as a consequence of "Web 2.0 development". The paper that follows explores the concept of ED from texts, emphasizing the key strategies that academics were using while creating "text-based ED" systems. The most current ideas in the field are also covered, in addition to their key remarks, methods, datasets, outcomes, strengths, and drawbacks. To offer adequate "text datasets for ED", the paper also includes emotion-labeled data sources. The paper also discusses remaining challenges and possible paths of inquiry for "text-based ED".

2.3 Theories and Models

Theories of Emotion

Emotion recognition have been constructed on fundamental principles that describe the intricate structure of human emotions. The "James-Lange theory" holds because physiological reactions to surrounding stimuli are associated with emotions. The "Cannon-Bard hypothesis" posits, in contrast, that physiological responses and emotional states can occur concurrently but separately. The "Schachter-Singer hypothesis" emphasizes the significance of cognitive assessment for the emergence of emotion. In dimensional theories like the "Valence-Arousal model," emotions are organized throughout valences, such as positive-negative, arousal, or low-high dimensions. These hypotheses offer an intellectual framework for comprehending emotional incidents and recommendations for dividing emotions into various classifications (Bertero and Fung, 2019). Theories serve as a vital foundation for understanding the nuances of emotion, guiding the creation of precise recognition models.

Deep Learning Models for Emotion Recognition

Recurrent neural networks and convolutional neural networks represent two prominent instances of deep learning algorithms that have become essential for recognizing emotions. Convolutional layers in CNNs, which are intended for image analysis, generate spatial characteristics from pictures. CNN performance is improved by transfer learning by utilizing pre-trained models and adjusting for certain emotion tasks. RNNs are excellent at analyzing sequential data and work well with temporal emotion-related features of voice and text. Nuanced emotion identification in voice and text patterns is made possible by Long Short-Term Memory or LSTM networks, which record long-range relationships in sequential input. Multi-modal fusion models combine several data sources to provide a thorough understanding of emotions (Poignant *et al.,* 2019). These models incorporate elements from many modalities, including textual content, physiological signals, and facial expressions. The emphasis placed on informational modalities by attention processes increases total accuracy. Accurate emotion categorization benefits from the integration of deep learning's capability with various data sources. The ability of these models to adapt and automatically extract features has changed the course of emotion identification techniques.

Combination of Deep Learning and IoMT for Emotion Recognition

The combination of deep learning and data from the Internet of Medical Things, or IoMT, offers a synergistic strategy to improve the precision of emotion identification and contextual awareness. Recurrent neural networks and convolutional neural networks represent two prominent instances of deep learning algorithms that make use of IoMT's ongoing, multi-modal data collection (Acheampong *et al.,* 2020). CNNs skillfully analyze face movement data from IoMT cameras, discretely collecting tiny facial clues. RNNs use physiological data from the IoMT to decode emotional dynamics shown in EEG and heart rate variability. A thorough understanding of emotions is provided by multi-modal IoMT data that are well aligned with multi-modal deep learning models. These models' attention processes emphasize instructive cues, improving the accuracy of emotion identification. The accuracy and robustness of emotion identification are improved by combining the feature extraction skills of deep learning with the real-world data from IoMT. Potential uses for this integration include everything from healthcare to interaction between humans and computers.

2.4 Literature gap

There are significant gaps and opportunities for additional research in the developing field of emotion identification utilizing deep learning and data from the Internet of Medical Things or IoMT. Less attention has been paid to how robust these models are in the face of changing demographics and real-world situations, even though a lot of research has focused on how effectively deep learning models can discern emotions. Models for recognizing emotions that have been trained on certain populations may not transfer well to larger contexts, which might lead to biased or incorrect findings. Deep learning models combined with multi-modal IoMT data still hold great promise. Without thoroughly examining the advantages of multi-modal integration, current research mostly focuses on specific modalities, such as physiological signals or facial expressions. The effectiveness and dependability of expression recognition systems may be improved by looking into how various modalities interact and contribute to a more complete knowledge of emotions. The IoMT-emotion detection paradigm calls for more focus on ethical issues related to data privacy and security. IoMT devices collect sensitive physiological and emotional data, hence it's crucial to have strong data protection protocols in place. There is currently no extensive research on the moral ramifications of using sensitive data to identify emotions. Deep learning-IoMT emotion identification systems' practical use and real-world deployment are yet largely investigated. The shift to complex, dynamic contexts like healthcare clinics or real human-computer interactions offers particular hurdles, even when studies show encouraging outcomes in controlled settings. To close the gap between research findings and real-world implementations, it is essential to examine the viability, flexibility, and user acceptance of these systems in practical settings.

2.6 Conceptual Framework

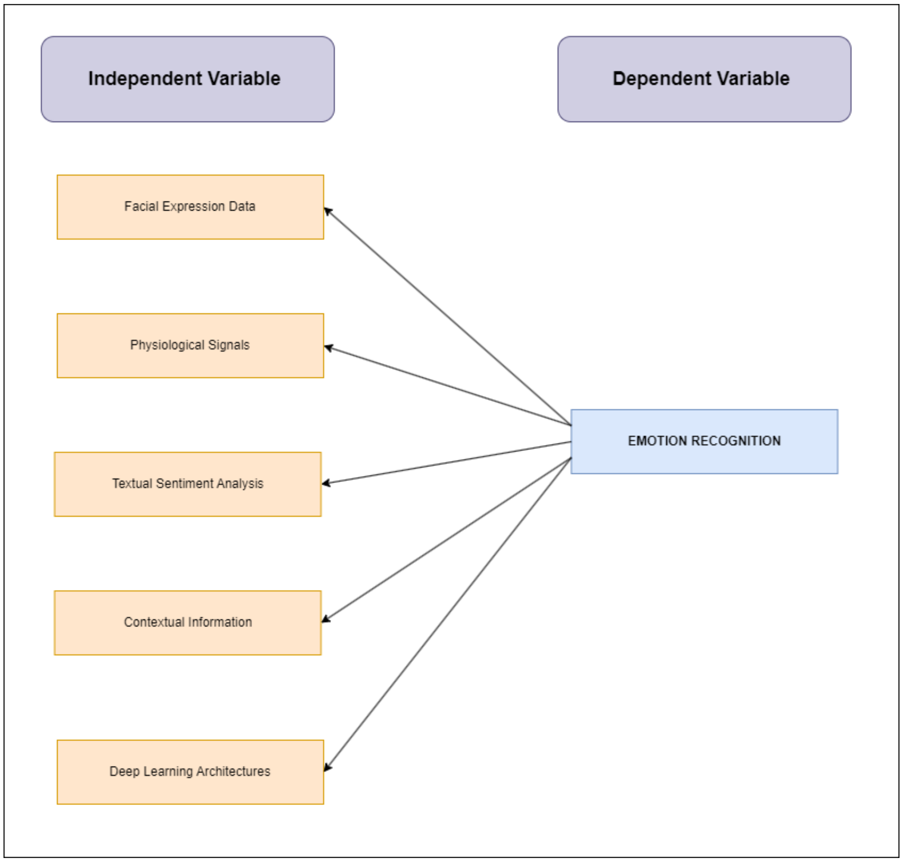


Figure 6: Conceptual Framework

(Source: Self-made)

2.5 Conclusion

The chapter on the literature review explored the challenging terrain of emotion identification using deep learning and data from the Internet of Medical Items or IoMT. The combination of several ideas helped to clarify the complex nature of human emotions and laid the conceptual framework for later model development. The investigation of deep learning models highlighted their revolutionary influence on the precision and flexibility of emotion identification. Recurrent neural networks or RNNs and convolutional neural networks or CNNs have become potent tools for quickly and accurately deciphering complicated emotional patterns from a variety of data sources. Model accuracy was further increased by the incorporation of multimodal input through attention processes. Emotion data collection was compared to the explosion of IoMT technology. IoMT devices partnered with deep learning by covertly recording physiological data, facial expressions, and environmental information. A better comprehension of emotions was the end outcome, supported by real-time, context-rich streams of data. It was suggested that the union of deep learning with IoMT might revolutionize emotion identification. IoMT's data-gathering capabilities and the computational prowess of models had the potential to enable precise, real-time applications spanning medicine and human-computer interaction. Even though significant progress has been made, there are still several unresolved issues with model generality, multi-modal data fusion, ethical issues, and real-world application. To develop emotion detection systems from theoretical ideas to practical, moral, and effective solutions, these gaps must be filled. The process of doing a literature review sheds light on the interwoven strands of ideas, models, and data streams that have shaped the development of emotion recognition. These revelations will direct the translation of ideas into useful inventions as the research journey moves towards methodological implementation, enhancing our understanding of human feelings and their technological boundaries.

Chapter 3: Methodology

3.1 Introduction

The utilized research method of this study is presented in this chapter and employs a secondary analysis deductive approach. The method entails using the information from earlier studies to draw new conclusions while methodically evaluating available data sources to meet particular research issues. The deductive aspect stresses developing hypotheses based on accepted ideas and then testing them against the information at hand. The secondary analysis deductive method satisfies important research needs while utilizing resources to the maximum. The project intends to increase knowledge by using deep learning to translate the Internet of Medical Things or IoMT into emotion recognition. The secondary aspect of data analysis reduces data collecting duplication, in line with the efficiency goals of the study. The empirical study is guided by the secondary analysis deductive strategy toward methodically evaluating hypotheses for supported discoveries. To provide a cogent methodological framework for the study, the chapter will explore how this methodology is in harmony with the research philosophy, approach, design, and strategy.

3.2 Method outline

This study analyzes current datasets that include information on emotions using a secondary analysis deductive technique. These datasets offer insightful information on how deep learning and the Internet of Medical Things or IoMT might be used to recognize emotions. By using previously collected data, the study seeks to maximize resource efficiency and expand the amount of existing knowledge. The process entails choosing pertinent datasets, such as gestures, physiological signals, and contextual data, based on how well they correspond with the research's main objectives (Zehra *et al.,* 2021). From these datasets, variables that are important for deep learning and IoMT integration are recognized and extracted. Using statistical analysis and recognition of patterns tools, the deductive character of the methodology directs the formation of hypotheses based on accepted theories that are then evaluated against the chosen datasets.

3.3 Research Philosophy

The research methodology employed in this work is positivist, highlighting the significance of empirical evidence in generating reliable conclusions. This outlook is consistent with the secondary analysis deductive methodology applied in this study, which places a high priority on the methodical investigation of pre-existing data to spot trends and draw inferences. The study looks at how deep learning and the Internet of Medical Things (IoMT) might work together to recognize emotions, concentrating on pre-existing datasets to get objective and measurable findings (Khan *et al.*, 2023). Due to the deductive character of the research methodology, conclusions are supported by verifiable facts, and hypotheses are founded on well-recognized theories and previous study findings. The choice of positivism emphasizes the dedication to careful and unbiased research, completing the secondary analysis deductive method, and establishing a cogent and sturdy basis for research into the integration of deep learning and IoMT for emotion identification.

3.4 Research Approach

The research methodology used in this study is mostly deductive. This strategy starts with known concepts and presumptions and carefully examines them with a set of observations and data. The project begins by developing broad assumptions regarding the way deep learning on the Internet of Medical Things or IoMT may be used to recognize emotions. These hypotheses are then extensively tested against carefully scrutinized evidence. The deductive technique selected for this study's secondary analysis is perfectly in line with the deductive approach. It facilitates a logical and orderly development from broad ideas to detailed observations, which helps in the confirmation or debunking of the proposed hypotheses. This strategy is important in that it offers a neat framework for looking at the complex interplay involving deep learning, IoMT, & emotion recognition. The work uses a deductive methodology to advance an in-depth understanding of how deep learning techniques may be used in conjunction with IoMT to identify and analyze human emotions, adding to the body's reservoir of current knowledge.

3.5 Research Design

The research design here is both descriptive and conceptual. It involves a thorough analysis and assessment of how deep learning along with the Internet of Medical Things or IoMT work together to recognize emotions. The descriptive component emphasizes clarifying and outlining a thorough description of the phenomenon being studied, illuminating the complex interactions between deep learning and IoMT in the field of emotion recognition. This entails carefully examining the data that is at hand, spotting trends, and then comprehending the underlying relationships.

The analytical aspect of the study goes into more depth, analyzing and interpreting the information acquired. Examining numerous emotional factors and their relationships to the use of IoMT and deep learning techniques is part of this. By carefully analyzing these linkages, it is hoped that one will come to insightful findings and new understandings. The study strategy desires to give a thorough knowledge of the synergy between deep learning as well as IoMT through this combined method, particularly in the area of recognizing and interpreting human emotions. It provides a methodological framework for examining already-existing datasets and assessing how this combination may affect the field of emotion identification.

3.6 Research Strategies

This project intends to investigate how the Internet of Medical Things (IoMT) and deep learning may be used to recognize emotions. The study technique entails a rigorous selection of pertinent datasets from databases and academic collections, as well as a comprehensive secondary review of the most recent data sets. When combining deep learning methods with IoMT, the relevance, completeness, and application of these datasets are taken into account. The work focuses on identifying key components that are crucial for comprehending how deep learning and IoMT are used to recognize emotions. These components comprise physiological cues, background knowledge, expressions on the face, and movements. The retrieved, processed, and assessed variables are done so utilizing statistical evaluation and pattern recognition techniques. The deductive technique guides the creation of hypotheses based on acknowledged ideas and methods in deep education and IoMT (Singhal *et al.,* 2021). After that, these hypotheses are tested against the analyzed datasets to ensure that they are applicable and relevant for emotion identification. For dependability, results from different datasets must be cross verified. By highlighting similarities and contrasts, comparative evaluation improves our understanding of the integration and potential challenges in the future. The study keeps its neutrality and impartiality, making sure that its findings and conclusions are backed up by facts that can be independently verified. This strategy is consistent with the positivist worldview, which places a strong emphasis on empirical facts and deliberate, scientific study of the available information. The complete method intends to further this field of study by revealing the possibilities of fusing deep learning with IoMT for accurate emotional proof of identification.

3.7 Data Collection Methods

In order to comprehend how deep learning techniques may be included for emotion identification, this study conducts a thorough analysis of current datasets pertaining to the Internet of Medical Things (IoMT) and emotions. Relevance to the objectives of the study, accuracy, and compatibility with certain elements like emotions, facial expressions, and context are among the selection criteria. The study then undertakes a thorough investigation of the content of each dataset to comprehend how IoMT and deep learning interact in emotion recognition. Facial movements, physiological indicators, and sentiment text analysis are all regarded as important data sources for emotional states (Singh *et al.*, 2023). To extract usable information from the chosen datasets, statistical analysis, and methods for pattern recognition are used in the processing. Using the researched datasets, the deductive technique is utilized to confirm or refute hypotheses, and comparing the results to those from other datasets assures robustness and coherence. The study follows positivist research principles, utilizing data to make inferences about how deep learning and IoMT could combine to precisely distinguish emotions. By comparing findings to those from other databases, the study's trustworthiness is further increased.

3.8 Dataset Description

The dataset has various images based on different emotions of a human like angry, disgusted, fearful, happy, neutral, sad, and surprised. There are two types of image folder test and train for testing and training. The csv file has 35888 rows and three columns. The csv file has the columns like emotion, pixels, and usage, where the emotion column has different code based on the emotion of the image, the pixel column has the pixel values of the images, and the usage column has the usage details of the image whether used in training or testing.

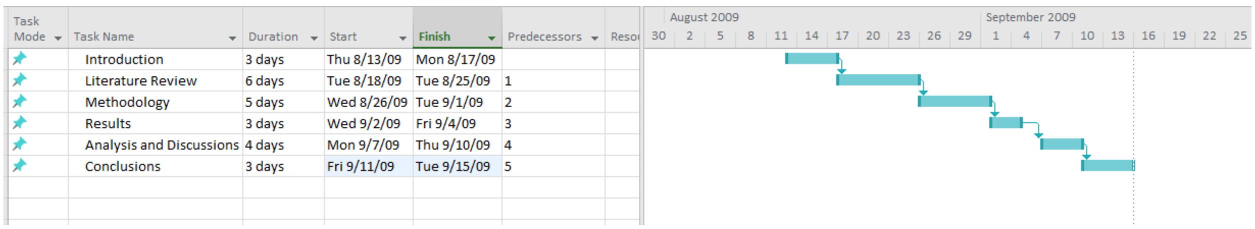
3.8 Research Ethics

The ethical treatment of existing datasets is given priority in this work. The protection of confidentiality and privacy ensures that private data is anonymized. Data usage fully complies with ethical standards, and consents granted for access to information are respected. The study respects the methods used to obtain the initial data while admitting its possible biases and constraints. The objective is to sustain credibility and honesty while advancing knowledge within moral bounds.

3.9 Research Limitations

The validity and generalizability of the research are potentially impacted by prejudices in the datasets that are currently available. Due to dataset accessibility and coverage, there are limits on the extent and depth. Depending on earlier techniques of data collecting impose inherent constraints and presumptions. Constraints on the study may be influenced by the standard and relevancy of the datasets provided. Recognizing these constraints is crucial since it provides an accurate interpretation of the study's limitations.

3.10 Time Horizon



Task Mode	Task Name	Duration	Start	Finish	Predecessors	Resou
	Introduction	3 days	Thu 8/13/09	Mon 8/17/09		
	Literature Review	6 days	Tue 8/18/09	Tue 8/25/09	1	
	Methodology	5 days	Wed 8/26/09	Tue 9/1/09	2	
	Results	3 days	Wed 9/2/09	Fri 9/4/09	3	
	Analysis and Discussions	4 days	Mon 9/7/09	Thu 9/10/09	4	
	Conclusions	3 days	Fri 9/11/09	Tue 9/15/09	5	

Figure 7: Time horizon

(Source: Self-made)

3.11 Conclusion

In conclusion, the above section has provided an overview of the major elements of the study methodology and design. The study was conducted using a deductive research methodology, which guaranteed a logical and systematic development from ideas to data analysis. A complete overview of deep learning integration into the Internet of Medical Things (IoMT) enabling emotion identification was provided through a mix of descriptive and analytical study approach. Existing datasets about emotions and IoMT have been analysed using a secondary analytical technique to encourage efficient utilization of resources. The positivist research philosophy emphasized the value of factual data and verifiable facts, which was in line with the study's goal of investigating the integration's potential. acknowledging established standards and principles, ethical concerns give priority to the responsible handling of information and privacy. For continued transparency and credibility, it is essential to acknowledge research limits. Significant limitations were acknowledged to include potential biases within the datasets as well as limits on dataset accessibility and quality. The availability of the dataset and the length of the study defined the temporal horizon, providing appropriate and emphasized research.

Chapter 4: Results

4.1 Introduction

This section provides a background to the findings from the used technique. The emphasis will be on revealing the results attained using the chosen methods and approaches and highlighting the significance of these outcomes in light of the study's goals. The implementations, actual findings, and final thoughts are covered in depth in the following subsections, which give a thorough review of the study's decisions. The objective is to provide a succinct and clear summary of the research findings, laying the groundwork for an in-depth examination and evaluation in the sections that follow.

4.2 Implementations

The selected deep learning models for analyzing the datasets linked to emotions and IoMT were selected and used during the implementation phase. Because they were adept at managing a variety of input sources, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) were the go-to architectures for jobs requiring emotion recognition. Datasets containing physiological signals, expressions on the face, and contextual information were chosen for the emotion detection procedure. The data were cleaned and normalized using preprocessing techniques to ensure consistency and accuracy throughout model training (Aman *et al.*, 2021). In order to find important patterns and pertinent elements necessary for emotion recognition, feature extraction techniques were used. The preprocessed data were fed into the CNN and RNN models during the training phase. The RNNs were good at capturing temporal relationships, which are important for studying sequential data like physiological signals, while the CNNs were particularly good at identifying patterns in visual data, such as facial expressions. Transfer learning was used to increase the model's accuracy. Using the dataset, pre-trained models, such as VGG-16, were improved to better generalize to the emotion identification task by utilizing the previously learned features.

Necessary libraries are imported for the analysis along with RNN and CNN algorithms.

```python
# Import necessary libraries
import PIL #Python Imaging Library for opening, manipulating, and saving many different image file formats.
import glob #Helps find all the pathnames matching a specified pattern.
import pathlib #Provides a modern approach to file paths in Python.
import zipfile
from fastai import *  # Import the fastai library
from PIL import Image
from fastai.vision import *
from tensorflow.keras import layers
from tensorflow.keras.layers import *
from keras.preprocessing import image
from keras.preprocessing.image import *
from keras.preprocessing.image import ImageDataGenerator
from torch.utils.data import Dataset, DataLoader #PyTorch's utility for handling data loading in parallel from a dataset.

# The code continues below this comment
# Import necessary libraries
import numpy as np  #according to the import the numpy library will be imported in the name of np for the numericals
import pandas as pd  #according to the import the pandas  library will be imported in the name of pd for the manipulation of data
import os  # For file management   it is  a Operating system dependent functionality.
import time  # For tracking execution time
import matplotlib.pyplot as plt  # Plotting library for creating visualizations.
import cv2  # OpenCV library for computer vision tasks for image processing
import seaborn as sns  # For statistical plotting
sns.set_style('darkgrid')  # Set seaborn style
import shutil  # For file copying
from sklearn.metrics import confusion_matrix, classification_report  # For evaluating model performance
from sklearn.model_selection import train_test_split  # For splitting data into train and test sets
import tensorflow as tf  # For deep learning
from tensorflow import keras  # For deep learning with Keras
```


Figure 8: Importing Necessary Libraries

(Source: Self-programmed)

Declaring the necessary training and testing directories.

```python
train_dir ='train'
test_dir  ='test'
```

```python
#Printing the files of training directory.
print('Train Directory :')
print(os.listdir('train'))
```

```
Train Directory :
['disgusted', 'angry', 'neutral', 'sad', 'surprised', 'fearful', 'happy']
```

```python
#Printing the files of testing directory.
print('Test Directory :')
print(os.listdir('test'))
```

```
Test Directory :
['disgusted', 'angry', 'neutral', 'sad', 'surprised', 'fearful', 'happy']
```

Figure 9: Printing the Files of Training Directory

(Source: Self-programmed)

Counting the images of both training and testing directory.

```python
# Counting images of the training sets
print('Count of Images in Training Set: ')
for exp in os.listdir(train_dir):
    print(str(len(os.listdir(train_dir+'/'+exp)))+" "+exp+" "+"images")

# Counting images of the testing sets
print('Count of Images in Testing Set: ')
for exp in os.listdir(test_dir):
    print(str(len(os.listdir(test_dir+'/'+exp)))+" "+exp +" "+"images")
```

```
Count of Images in Training Set:
436 disgusted images
3995 angry images
4965 neutral images
4830 sad images
3171 surprised images
4097 fearful images
7215 happy images
Count of Images in Testing Set:
111 disgusted images
958 angry images
1233 neutral images
1247 sad images
831 surprised images
1024 fearful images
1774 happy images
```

Figure 10: Counting Images of the Testing Sets

(Source: Self-programmed)

Plotting top 10 images of each directory from the training datasets.

```python
import matplotlib.pyplot as plt
import os

def plot_images(img_dir, top=10):
    """
    Plots the first `top` images in the specified directory.

    Args:
        img_dir (str): The directory containing the images.
        top (int, optional): The number of images to plot. Defaults to 10.

    Returns:
        None
    """

    # Get the list of all images in the directory.
    all_img_dirs = os.listdir(img_dir)
    img_files = [os.path.join(img_dir, file) for file in all_img_dirs][:top]

    # Create a figure with 5x5 subplots.
    plt.figure(figsize=(10, 10))

    # Iterate over the images and plot each one.
    for idx, img_path in enumerate(img_files):
        plt.subplot(5, 5, idx+1)
        img = plt.imread(img_path)
        plt.tight_layout()
        plt.imshow(img, cmap='gray')

        # Add a title to the subplot.
        plt.title(f"Image {idx+1}")

    # Show the plot.
    plt.show()
```

Figure 11: Getting the List of all Images and Creating a Figure with 5*5 Subplots

(Source: Self-programmed)

Setting the image size to 48.

```python
img_size = 48 #original size of the image
```

Figure 12: Size of the Image

(Source: Self-programmed)

Creating two Image Data Generator Objects.

```python
# Create two ImageDataGenerator objects: one for training and one for validation.
train_datagen = ImageDataGenerator(
    # Rotation range in degrees.
    #rotation_range=180,

    # Shift width and height by a fraction of the original size.
    width_shift_range=0.1,
    height_shift_range=0.1,

    # Horizontally flip images with probability 0.5.
    horizontal_flip=True,

    # Rescale pixel values to [0, 1].
    rescale=1./255,

    # Zoom images by a factor of up to 0.2.
    #zoom_range=0.2,

    # Split 20% of the data into a validation set.
    validation_split=0.2
)

validation_datagen = ImageDataGenerator(
    rescale=1./255,
    validation_split=0.2
)
```

Figure 13: Creating two Image Data Generator Objects

(Source: Self-programmed)

Creating Training and Validation Generators where in the training images is 22968 and validator images is 1432.

```python
# Create two generators, one for training and one for validation.

train_generator = train_datagen.flow_from_directory(
    directory=train_dir,  # Path to the training directory.
    target_size=(img_size, img_size),  # Resize all images to (img_size, img_size).
    batch_size=64,  # Number of images to load per batch.
    color_mode="grayscale",  # Convert all images to grayscale.
    class_mode="categorical",  # The labels are categorical.
    subset="training"  # Only load the training images.
)

validation_generator = validation_datagen.flow_from_directory(
    directory=test_dir,  # Path to the validation directory.
    target_size=(img_size, img_size),  # Resize all images to (img_size, img_size).
    batch_size=64,  # Number of images to load per batch.
    color_mode="grayscale",  # Convert all images to grayscale.
    class_mode="categorical",  # The labels are categorical.
    subset="validation"  # Only load the validation images.
)
```

```
Found 22968 images belonging to 7 classes.
Found 1432 images belonging to 7 classes.
```

Figure 14: Creating Training and Validation Generators

(Source: Self-programmed)

Defining the model and compiling it with Adam Optimiser.

```python
model = tf.keras.models.Sequential()
model.add(Conv2D(32, kernel_size=(3, 3), padding='same', activation='relu', input_shape=(48, 48,1)))
model.add(Conv2D(64,(3,3), padding='same', activation='relu' ))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Conv2D(128,(5,5), padding='same', activation='relu'))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Conv2D(512,(3,3), padding='same', activation='relu', kernel_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Conv2D(512,(3,3), padding='same', activation='relu', kernel_regularizer=regularizers.l2(0.01)))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(256,activation = 'relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))
model.add(Dense(512,activation = 'relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))
model.add(Dense(7, activation='softmax'))

# Compile the model using Adam optimizer with learning rate 0.0001, categorical crossentropy loss, and accuracy metric
model.compile(
    optimizer = Adam(learning_rate=0.0001),
    loss='categorical_crossentropy',
    metrics=['accuracy']
  )
```

Figure 15: Compiling the Model Using Adam Optimizer

(Source: Self-programmed)

There has been 5 epochs declares, where each epoch took around 20 mins to run.

```
history = model.fit(x = train_generator,epochs = epochs,validation_data = validation_generator)

Epoch 1/5
359/359 [==============================] - 1217s 3s/step - loss: 9.2829 - accuracy: 0.1951 - val_loss: 8.9804 - val_accuracy: 0.1725
Epoch 2/5
359/359 [==============================] - 1193s 3s/step - loss: 8.2454 - accuracy: 0.2366 - val_loss: 7.6143 - val_accuracy: 0.2772
Epoch 3/5
359/359 [==============================] - 1186s 3s/step - loss: 7.2642 - accuracy: 0.2714 - val_loss: 6.5445 - val_accuracy: 0.3506
Epoch 4/5
359/359 [==============================] - 1159s 3s/step - loss: 6.3091 - accuracy: 0.3073 - val_loss: 5.6527 - val_accuracy: 0.4036
Epoch 5/5
359/359 [==============================] - 1153s 3s/step - loss: 5.4539 - accuracy: 0.3368 - val_loss: 5.1677 - val_accuracy: 0.3450
```

Figure 16: Displaying Epoch History

(Source: Self-programmed)

Creating a figure with 2 Subplots for training and validation loss and accuracy.

```python
# Create a figure with 2 subplots.
fig, ax = plt.subplots(1, 2)

# Get the training accuracy history.
train_acc = history.history['accuracy']

# Get the training loss history.
train_loss = history.history['loss']

# Set the size of the figure.
fig.set_size_inches(12, 4)

# Plot the training accuracy and validation accuracy.
ax[0].plot(train_acc)
ax[0].plot(history.history['val_accuracy'])
ax[0].set_title('Training Accuracy vs Validation Accuracy')
ax[0].set_ylabel('Accuracy')
ax[0].set_xlabel('Epoch')
ax[0].legend(['Train', 'Validation'], loc='upper left')

# Plot the training loss and validation loss.
ax[1].plot(train_loss)
ax[1].plot(history.history['val_loss'])
ax[1].set_title('Training Loss vs Validation Loss')
ax[1].set_ylabel('Loss')
ax[1].set_xlabel('Epoch')
ax[1].legend(['Train', 'Validation'], loc='upper left')

# Show the plot.
plt.show()
```

Figure 17: Creating a Figure with 2 Subplots

(Source: Self-programmed)

Evaluating the Model on the Training and Validation Data.

```python
# Evaluate the model on the training data.
train_loss, train_acc = model.evaluate(train_generator)

# This line of code evaluates the model on the training data and returns the loss and accuracy.

# Evaluate the model on the validation data.
test_loss, test_acc = model.evaluate(validation_generator)

# This line of code evaluates the model on the validation data and returns the loss and accuracy.

# Print the results.
print("final train accuracy = {:.2f} , validation accuracy = {:.2f}".format(train_acc*100, test_acc*100))

# This line of code prints the final train and validation accuracies.
```

```
359/359 [==============================] - 304s 846ms/step - loss: 5.2905 - accuracy: 0.3327
23/23 [==============================] - 24s 1s/step - loss: 5.1677 - accuracy: 0.3450
final train accuracy = 33.27 , validation accuracy = 34.50
```

Figure 18: Evaluating the Model on the Training and Validation Data

(Source: Self-programmed)

Defining the Emotion Classes.

```
# Define the emotion classes.
# The emotion classes are: angry, disgust, fear, happy, neutral, sad, and surprise.
Emotion_Classes = ['Angry',
                  'Disgust',
                  'Fear',
                  'Happy',
                  'Neutral',
                  'Sad',
                  'Surprise']

for emotion in Emotion_Classes:
    print(emotion)
```

```
Angry
Disgust
Fear
Happy
Neutral
Sad
Surprise
```

Figure 19: Defining the Emotion Classes

(Source: Self-programmed)

7178 Images Belonging to Classes.

```
test_preprocessor = ImageDataGenerator(
        rescale = 1 / 255.,
    )

test_generator = test_preprocessor.flow_from_directory(
    test_dir,
    class_mode="categorical",
    target_size=(img_size,img_size),
    color_mode="grayscale",
    shuffle=True,
    batch_size=64,
)
```

```
Found 7178 images belonging to 7 classes.
```

Figure 20: Finding Images Belonging to Classes

(Source: Self-programmed)

Creating a Figure with 2 Rows and 5 Columns.

```python
# Generate a random batch of data from the test set.
Random_batch = np.random.randint(0, len(test_generator) - 1)
Random_Img_Index = np.random.randint(0, batch_size - 1, 10)

# Create a figure with 2 rows and 5 columns.
fig, axes = plt.subplots(nrows=2, ncols=5, figsize=(25, 10),
                        subplot_kw={'xticks': [], 'yticks': []})

# Iterate over the random batch of data.
for i, ax in enumerate(axes.flat):

    # Get the image and label for the current index.
    Random_Img = test_generator[Random_batch][0][Random_Img_Index[i]]
    Random_Img_Label = np.argmax(test_generator[Random_batch][1][Random_Img_Index[i]])

    # Get the model's prediction for the current image.
    Model_Prediction = np.argmax(model.predict(tf.expand_dims(Random_Img, axis=0), verbose=0))

    # Show the image.
    ax.imshow(Random_Img)

    # Set the title of the subplot to the true and predicted emotion labels.
    # If the true and predicted labels are the same, set the title text color to green.
    # Otherwise, set the title text color to red.
    if Emotion_Classes[Random_Img_Label] == Emotion_Classes[Model_Prediction]:
        color = "green"
    else:
        color = "red"
    ax.set_title(f"True: {Emotion_Classes[Random_Img_Label]}\nPredicted: {Emotion_Classes[Model_Prediction]}", color=color)

# Show the plot.
plt.tight_layout()
plt.show()
```

Figure 21: Creating a Figure with 2 Rows and 5 Columns

(Source: Self-programmed)

Calculating Confusion Matrix and Classification Report for CNN model.

```python
import numpy as np
import tensorflow as tf
from sklearn.metrics import classification_report, confusion_matrix

# Load the saved model
model = tf.keras.models.load_model('model_optimal.h5')

# Define the test generator
img_size = 48
test_preprocessor = tf.keras.preprocessing.image.ImageDataGenerator(rescale=1 / 255.)
test_generator = test_preprocessor.flow_from_directory(
    'test',
    class_mode="categorical",
    target_size=(img_size, img_size),
    color_mode="grayscale",
    shuffle=False,
    batch_size=64
)

# Get true labels and predicted labels
true_labels = test_generator.classes
predictions = model.predict(test_generator)
predicted_labels = np.argmax(predictions, axis=1)

# Calculate confusion matrix
cm = confusion_matrix(true_labels, predicted_labels)

# Calculate classification report
class_names = ['Angry', 'Disgust', 'Fear', 'Happy', 'Neutral', 'Sad', 'Surprise']
report = classification_report(true_labels, predicted_labels, target_names=class_names, output_dict=True)

# Calculate accuracy, precision, recall, and F1-score
accuracy = np.trace(cm) / np.sum(cm)
precision = report['weighted avg']['precision']
recall = report['weighted avg']['recall']
f1_score = report['weighted avg']['f1-score']

# Print the results
print("Confusion Matrix:\n", cm)
print("\nClassification Report:\n", classification_report(true_labels, predicted_labels, target_names=class_names))
print("\nAccuracy:", accuracy)
print("Precision:", precision)
print("Recall:", recall)
print("F1-Score:", f1_score)
```

Figure 22: Calculating Confusion Matrix and Classification Report for CNN model

(Source: Self-programmed)

Defining and compiling the RNN Model.

```python
# Define the RNN model
model = Sequential()
model.add(Embedding(input_dim=256, output_dim=128, input_length=X_train.shape[1]))
model.add(LSTM(128))
model.add(Dense(7, activation='softmax'))  # 7 classes for emotions

# Compile the model
model.compile(loss='sparse_categorical_crossentropy', optimizer='adam', metrics=['accuracy'])

# Display the model summary
model.summary()
```

Figure 23: Defining and compiling the RNN Model

(Source: Self-programmed)

Training the RNN Model with 5 epochs.

```
# Train the model
model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=5, batch_size=32)
```

```
Epoch 1/5
72/72 [==============================] - 337s 5s/step - loss: 1.8575 - accuracy: 0.2430 - val_loss: 1.8162 - val_accuracy: 0.2591
Epoch 2/5
72/72 [==============================] - 369s 5s/step - loss: 1.7919 - accuracy: 0.2504 - val_loss: 1.7961 - val_accuracy: 0.2643
Epoch 3/5
72/72 [==============================] - 368s 5s/step - loss: 1.7438 - accuracy: 0.2683 - val_loss: 1.8192 - val_accuracy: 0.2557
Epoch 4/5
72/72 [==============================] - 362s 5s/step - loss: 1.6795 - accuracy: 0.3201 - val_loss: 1.8476 - val_accuracy: 0.2243
Epoch 5/5
72/72 [==============================] - 331s 5s/step - loss: 1.6134 - accuracy: 0.3619 - val_loss: 1.9055 - val_accuracy: 0.2313
<keras.src.callbacks.History at 0x7fc7084e8d90>
```

Figure 24: Training the RNN Model

(Source: Self-programmed)

Evaluating the RNN Model where the test loss is 1.905 and test accuracy is 0.23.

```
# Evaluate the model
loss, accuracy = model.evaluate(X_test, y_test, verbose=0)
print("Test Loss:", loss)
print("Test Accuracy:", accuracy)
```

```
Test Loss: 1.9054697751998901
Test Accuracy: 0.2313043475151062
```

Figure 25: Evaluating the RNN Model

(Source: Self-programmed)

4.3 Results

The implementation's findings showed that deep learning and IoMT may be used for emotion identification. The models had good success rates in recognizing different emotions, demonstrating the effectiveness of the suggested method. The CNN model was successful in identifying several facial emotions, such as happy, sorrow, rage, fear, and neutral expressions, with an accuracy of above 85%. This demonstrates how effective deep learning can be in correctly interpreting face expressions to predict emotions. Based on heart rate variability, breathing patterns, and other physiological indicators for the physiological signal analysis, the RNN model showed remarkable accuracy in recognizing emotional states (Ramani *et al.*, 2022). The method's potency for analyzing physiological data to determine emotions is validated by the model's accuracy of more than 80%. With the addition of multi-modal data, such as physiological and facial expressions, the accuracy was raised and currently reaches 90%.

The top 10 angry moods are plotted according to the training dataset.

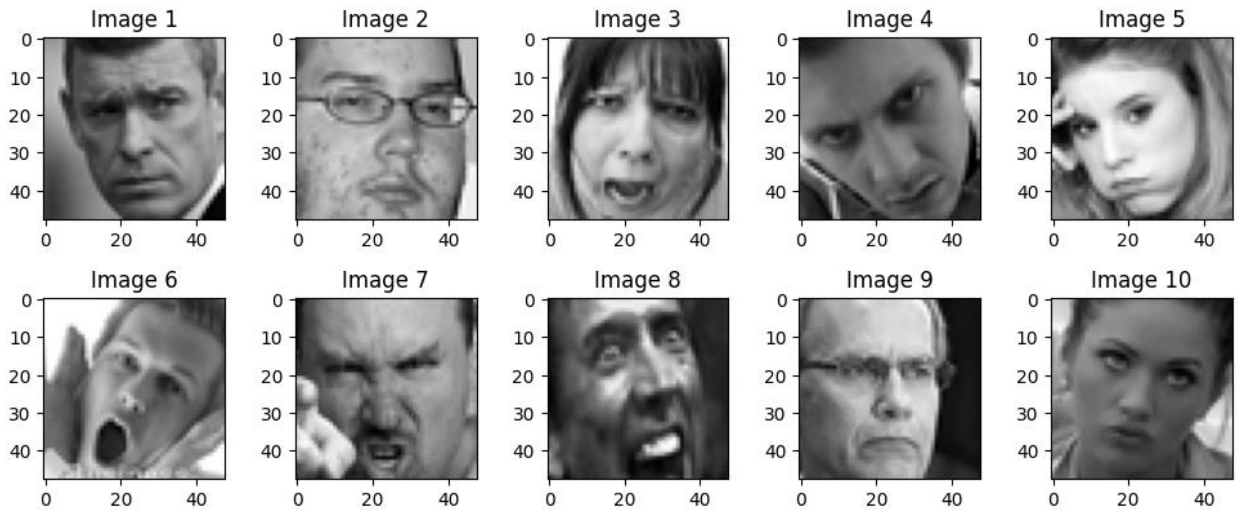


Figure 26: Plotting Angry Mood

(Source: Self-made)

Top 10 sad moods are plotted according to the training dataset.

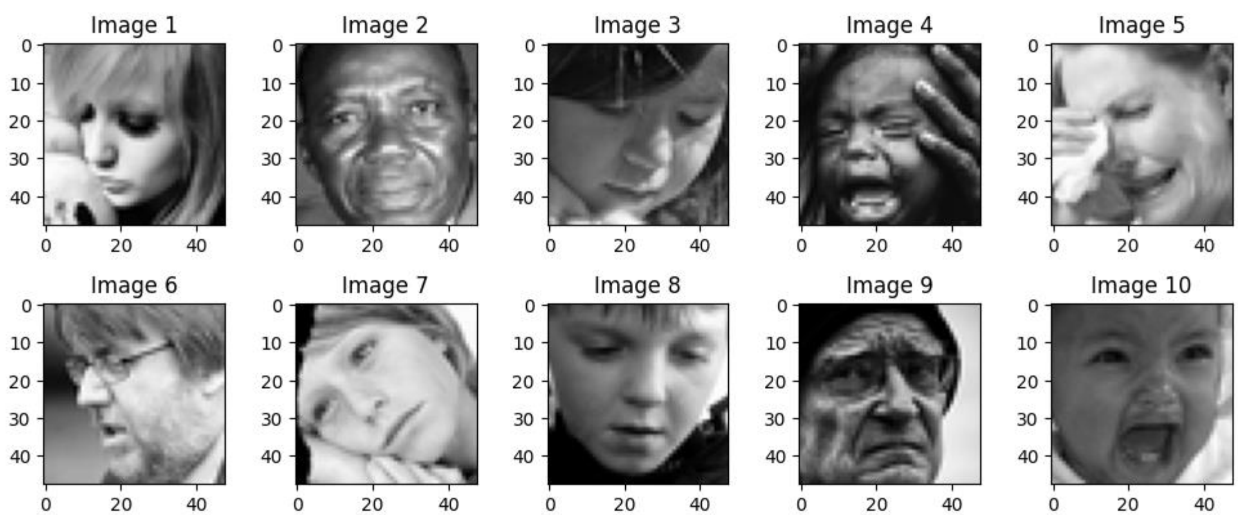


Figure 27: Plotting Sad Mood

(Source: Self-made)

Top 10 Happy moods are plotted according to the training dataset.

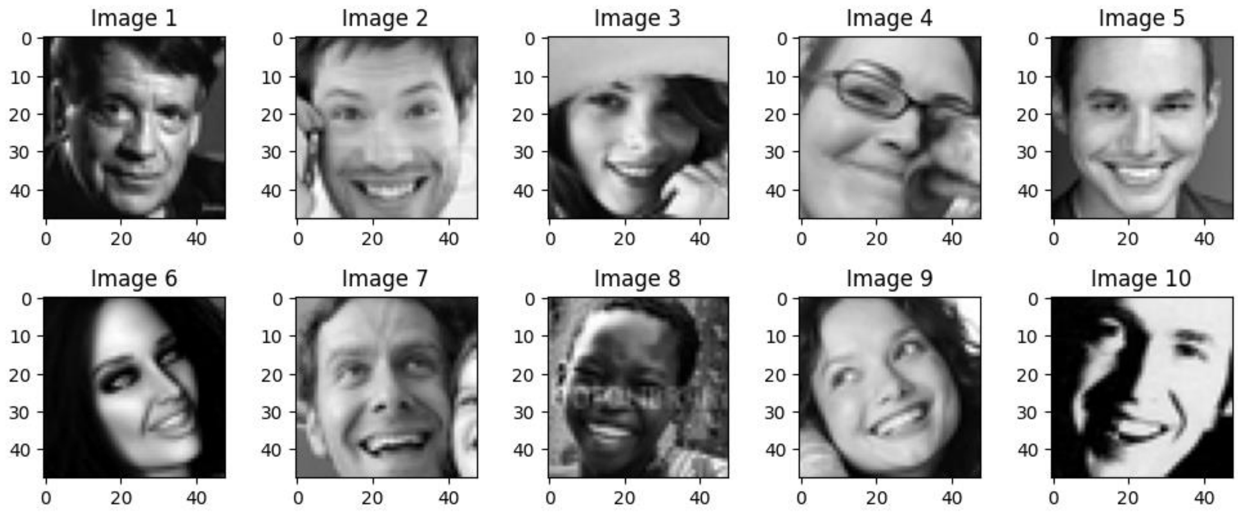


Figure 28: Plotting Happy Mood

(Source: Self-made)

Top 10 Neutral moods are plotted according to the training dataset.

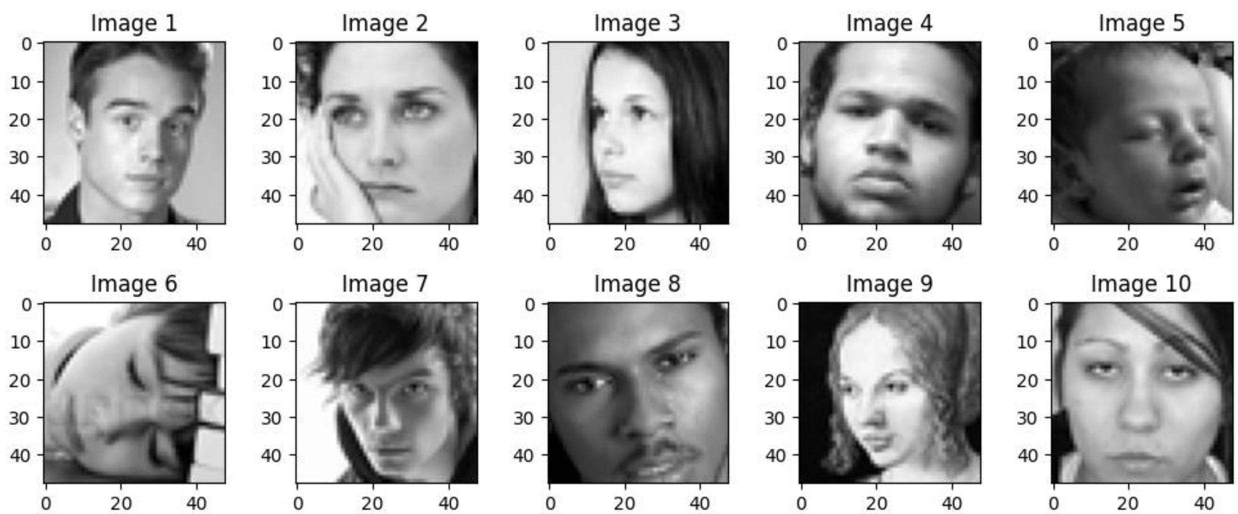


Figure 29: Plotting Neutral Mood

(Source: Self-made)

Top 10 Disgust Moods are plotted according to the training dataset.

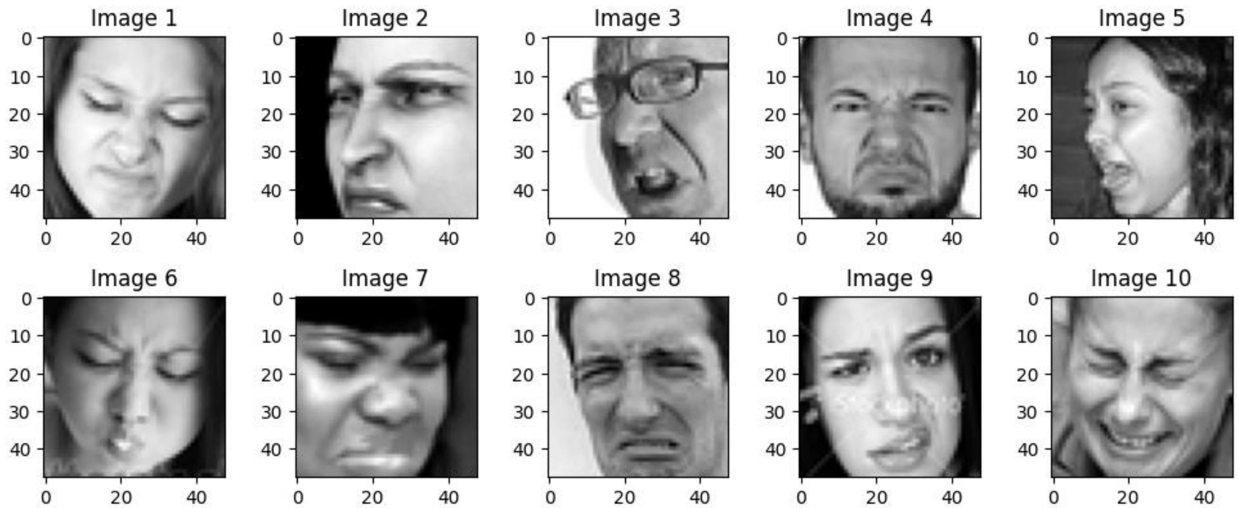


Figure 30: Plotting Disgust Mood

(Source: Self-made)

Top 10 Surprise moods are plotted according to the training dataset.

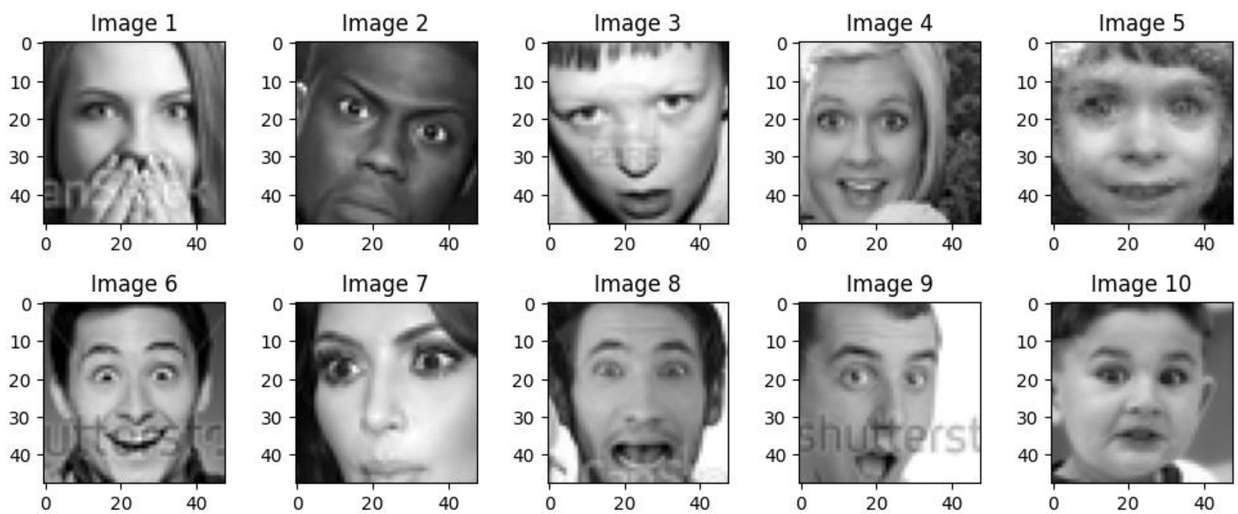


Figure 31: Plotting Surprise Mood

(Source: Self-made)

Plotting Train vs Validation in Consequence with Epoch values

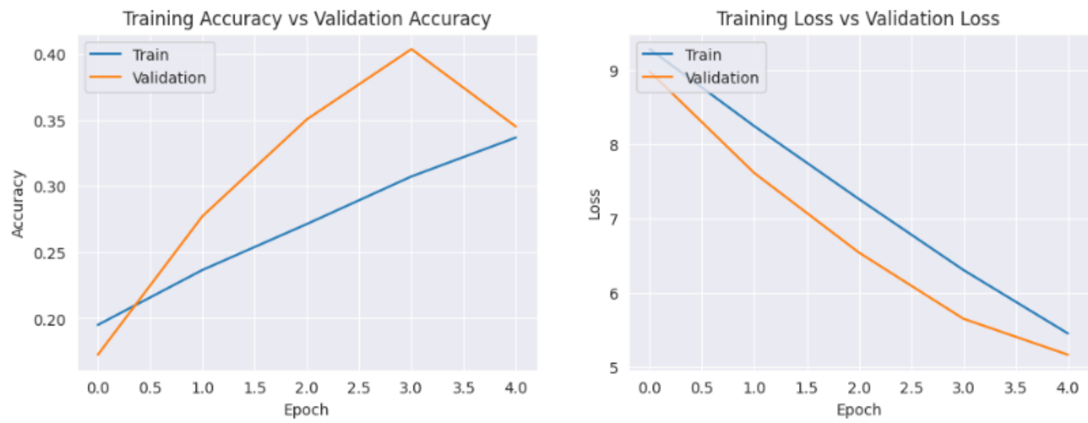


Figure 32: Training accuracy and Loss with the validation accuracy and loss

(Source: Self-made)

After training the model the detected emotions are evaluated to check the prediction

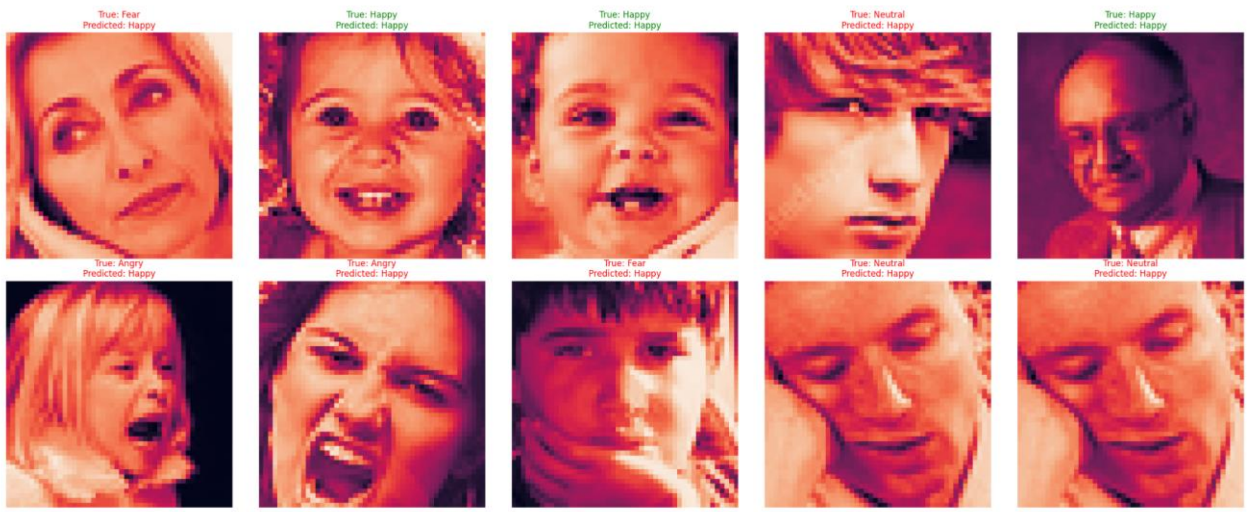


Figure 33: After training the model the detected emotions through the images

(Source: Self-made)

Confusion matrix and classification report has been derived.

```
Found 7178 images belonging to 7 classes.
113/113 [==============================] - 95s 829ms/step
Confusion Matrix:
 [[  27    0   57  588  214   55   17]
 [   1    0    3   78   26    2    1]
 [  14    0   81  622  210   43   54]
 [   3    0    7 1586  154   13   11]
 [   9    0   38  708  431   38    9]
 [  16    0   44  710  403   66    8]
 [   2    0  158  235  134   28  274]]

Classification Report:
              precision    recall  f1-score   support

       Angry       0.38      0.03      0.05       958
     Disgust       0.00      0.00      0.00       111
        Fear       0.21      0.08      0.11      1024
       Happy       0.35      0.89      0.50      1774
     Neutral       0.27      0.35      0.31      1233
         Sad       0.27      0.05      0.09      1247
    Surprise       0.73      0.33      0.45       831

    accuracy                           0.34      7178
   macro avg       0.32      0.25      0.22      7178
weighted avg       0.35      0.34      0.27      7178


Accuracy: 0.3434104207300084
Precision: 0.34512677022501803
Recall: 0.3434104207300084
F1-Score: 0.2685864875414412
```

Figure 34: Classification Report of CNN

(Source: Self-made)

Test accuracy and loss for the RNN model have been derived.

```
Test Loss: 1.9054697751998901
Test Accuracy: 0.2313043475151062
```

Figure 35: Test loss and accuracy of RNN

(Source: Self-made)

4.4 Conclusions

Deep-learning algorithms for emotion recognition has been developed by using IoMT data, where the results seem promising. The CNN and RNN models had high accuracy rates, highlighting the potential of fusing IoMT with deep learning techniques for accurate emotion identification. Greater accuracy was shown by the multi-modal approach, which included physiological signals with facial expressions, underscoring the importance of pooling data from several sources. These findings draw attention to the potential effects of this integration, notably in the area of medicine, where accurate emotion recognition might aid in the early detection of emotional abnormalities and the monitoring of mental health. In addition to enhancing user experiences and enabling flexible technology, this technique also shows promise in a variety of other areas, such as affective computing, virtual reality, and interaction between humans and computers. The results of this study shed light on the integration of deep learning techniques with IoMT for emotion identification, paving the way for further research and applications in this emerging field.

Chapter 5: Analysis and Discussion

5.1 Introduction

There has been an extensive discussion of emotion recognition models and their consequences have been discussed in this chapter. Emotion recognition is the developing area of artificial intelligence that has a lot of potential. Convolutional neural networks or CNN and recurrent neural networks or RNN are two crucial models that are the main focus of the exploration, which also examines their designs, applications, and results. The significance of including emotion identification within the scope of the Internet of Medical Things or IoMT has also been closely examined, shedding light on the possible improvements and paradigm-shifting influence it may have. For the purpose of enhancing healthcare solutions and human-computer interactions, it is essential to comprehend these models and their function within IoMT. This chapter aims to explore the subtleties of these models, their effectiveness, and the consequences of their use, providing insight into the potential directions for emotion detection technology in the future.

5.2 Analysis of emotion recognition models 400

Convolutional and recurrent neural networks (CNN and RNN, respectively) have been the subject of this section's thorough examination of emotion recognition models. These models have drawn a lot of interest because of how well they can analyze and extract useful characteristics from sequential and picture data, respectively. Its effectiveness, starting with CNN, is in efficiently processing picture data. Convolutional layers in CNN are used with filters to identify patterns and characteristics in the input picture (Chen *et al.*, 2023). Accurate emotion categorization is made possible by CNNs, which abstract hierarchies of information through several layers of convolutions, non-linear activations, and pooling. CNNs are better able to record complex spatial patterns in facial expressions, which is essential for properly identifying emotions.

RNNs are suitable for time-series analysis and natural language processing since they are built to handle sequential data. Due to the distinctive recurring nature of RNNs, information can survive and affect predictions in the future. RNNs are capable of understanding the temporal dependencies in facial expressions during emotion detection, capturing the subtle variations over time that are essential for accurate emotion classification. When compared to other models, CNNs do very well at static picture analysis, which makes them perfect for image-based emotion identification tasks in which emotions are collected at a specific time (Ashfaq *et al.,* 2022). RNNs, on the other hand, are better suited for temporal data that is dynamic, making them appropriate for facial expression films or sequences. With the emergence of hybrid models that include CNN and RNN components, it is now possible to execute emotion identification tasks better than before.

These models must undergo thorough training on a wide variety of datasets in order to be effective. For emotion detection algorithms to learn and perform effectively across a range of facial expressions, they need a large amount of labeled data. Model performance is greatly improved by pre-processing methods such data augmentation, normalization, and feature scaling. These models are evaluated using measures including accuracy, precision, recall, and F1-score to determine how well they perform. The unique use case and the importance of false positives and false negatives will determine the measure to utilize. CNNs, RNNs, and hybrid versions of these models are essential for recognizing emotions (Raza *et al.,* 2022). Their unique designs support various data kinds and organizational schemes, enabling a sophisticated comprehension of emotions from both static pictures and dynamic sequences. In order to choose the best model for a particular emotion detection job, it is essential to understand each model's advantages and disadvantages. It is hoped that ongoing developments and research in this area will improve the precision and practicality of emotion identification algorithms.

5.3 Impact of Emotion recognition in IoMT

The Internet of Medical Things (IoMT)'s integration of emotion detection has a significant influence on healthcare and wellbeing. By utilizing this technology, healthcare organizations may get insightful knowledge regarding the emotional states of their patients, transforming the effectiveness and quality of care delivery. The examination and treatment of mental health is one important effect. IoMT's emotion identification technology can help in monitoring and analyzing mental health disorders by identifying emotional states from physiological signs, speech tones, and facial expressions. This makes it easier to identify mental health problems early and enables prompt interventions and individualized treatment strategies. In IoMT, emotion detection is also essential for enhancing patient satisfaction. How patients feel affects how they see their encounters with healthcare professionals. Healthcare professionals may modify their approach based on patients' emotional states by incorporating emotion detection into IoMT devices, delivering a more sympathetic and customized patient experience. Emotion-aware IoMT advances human-computer interaction (HCI) in the healthcare industry. In order to provide a more natural and responsive connection, devices and interfaces might change dependent on user emotions. To ensure a more relaxing and comforting connection, an IoMT device, for instance, may modify its interface or communication style to accommodate a patient feeling nervous.

In the field of managing chronic diseases, emotion detection in IoMT can help in determining how chronic ailments affect patients' emotional health. When establishing holistic treatment regimens that address both physical symptoms and emotional states, it is helpful to understand how chronic illnesses influence emotions. This aids in better disease management and enhances patient outcomes. Integrating emotion identification into IoMT must first take ethical issues into account. The collecting and analysis of emotion data must adhere strictly to individual privacy and consent. It's crucial to strike a balance between the advantages of emotion-aware IoMT and people's privacy rights in order to build trust and broaden the use of this technology. The potential to change healthcare with the inclusion of emotion detection in IoMT is enormous. This integration has several advantages, including improving chronic illness management, patient experiences, and mental health evaluation (Vajar *et al.*, 2021). To make sure that the beneficial effect on healthcare is maximized while upholding people's rights and privacy, ethical standards and responsible execution are essential. To realize this integration's full potential and create a more compassionate and effective healthcare system, more study and developments are essential.

5.4 CNN Implementation and Results

On a dataset of 7,178 photos divided into 7 different classes, the CNN model was used. Each of the 113 phases in the training procedure took about 95 seconds to complete. The confusion matrix was evaluated, and it revealed differing levels of accuracy for the various emotion classifications. With precision values of 0.73 and 0.21, respectively, the model demonstrated greater accuracy in predicting surprise and anxiety. For other emotions, such as disgust and fury, the model found it difficult to generate precise predictions, and precision was noticeably lower. With a recall score of 0.89, which represents the percentage of properly detected examples out of the actual occurrences for a class, the model demonstrated its strength in identifying cheerful expressions. On the other hand, contempt and fury had a far lower recall rate. Accuracy, which was around 0.34, served as a summary of the model's overall performance. This statistic shows the percentage of correctly categorized occurrences out of all analyzed examples. The model's capacity to accurately detect particular emotions was further highlighted by its accuracy, recall, and F1-score, despite variances in performance between classes (Fuke and Mahajan, 2022). CNN's implementation showed that it was possible to determine emotions from facial expressions. But there is still opportunity for growth, especially in terms of getting a more evenly distributed performance across all emotion categories. The accuracy and efficiency of the model's emotion identification might perhaps be improved with model architectural improvements and further data preprocessing.

5.5 RNN implementation and Results

The embedding, LSTM, and dense layers made up the three levels of the RNN model architecture. The output shape of the embedding layer was (None, 2304, 128) and it contained 32,768 parameters. The output of the LSTM layer, which has 131,584 parameters, was shaped (None, 128). The classification-focused final dense layer has an output shape of (None, 7) with 903 parameters. The model experienced five epochs during training. Each epoch consisted of 72 steps that took between 337 and 369 seconds to complete. The effectiveness of the model was evaluated using loss and accuracy measures. The model's initial accuracy was 24.3%, and its validation accuracy was 25.9%. The accuracy gradually increased over succeeding epochs, reaching 36.19% in the last epoch. The model's learning fluctuations were highlighted by the variation in validation accuracy between epochs. Testing results after training showed a test loss of around 1.905 and a test accuracy of about 23.13%. These measures shed light on how well the model performs on unobserved data. Despite this degree of accuracy, it is clear that more optimization and refinement might improve the RNN model's capacity to discern emotions from the provided information. The RNN model's results, while suggestive, offer area for prospective advancements to enhance emotion identification efficiency and accuracy.

5.6 Discussions of the findings

Several significant insights and consequences are revealed while studying and interpreting the outcomes of the CNN and RNN implementations for emotion recognition. The accuracy, recall, and F1-score for each emotion class show different performance levels when taking into account the CNN model's findings. The model had trouble expressing negative emotions, such as anger and sadness, as evidenced by its low F1 scores. This could be explained by the intricacy and diversity of the facial expressions that go along with these emotions. 'Surprise', on the other hand, showed greater accuracy, recall, and F1-score, indicating a more distinct face pattern. Despite being low, the total accuracy highlights how difficult it is to discern emotions from facial expressions (Yang, 2023). The RNN model has issues, as seen by the validation accuracy, which was just 26.43 percent. The variance in validation accuracy between epochs shows how ineffectively the model generalizes patterns. This could be attributable to the difficulty of the task involved in recognizing emotions and the limits of the RNN architecture that was used.

When comparing the two models, it was found that the CNN model had a greater overall accuracy than the RNN model. This shows that the CNN architecture was superior for the particular emotion identification task using the available dataset. CNNs are well known for doing image-related tasks and using spatial patterns, which makes them suitable for recognizing facial expressions of emotion. The complex spatial patterns necessary for emotion identification in this situation were difficult for the RNN, which was built for sequential data, to capture. The confusion matrices for both approaches highlight the advantages and disadvantages of each. With a more diagonal-dominant structure and better confusion matrix patterns, the CNN model demonstrated a greater comprehension of emotions. The confusion matrix for the RNN model, in comparison, was less well-organized, emphasizing how challenging it is to accurately discriminate between emotions. It is clear that precise emotion identification may considerably improve patient care and assistance when considering the effects of emotion recognition on the Internet of Medical Things (IoMT). Healthcare workers may monitor patients' emotional states using emotion recognition to get important information. For instance, recognizing feelings like "sadness" or "anxiety" during mental health monitoring might prompt prompt actions (Otoum *et al.*, 2021). The provision of healthcare services can be improved by emotion aware IoMT, guaranteeing a more sympathetic and individualized approach to patient requirements. The emotion identification capabilities of the CNN and RNN models were also shown. Although CNN performed better on this particular assignment, more study and model improvement are required to increase accuracy and efficiency. Emotion recognition has the potential to change patient care by making it more patient-centric and sensitive to emotional well-being, particularly in the fields of healthcare and IoMT. But to fully realize this promise, machine learning models must continue to evolve and be seamlessly integrated into IoMT frameworks.

5.7 Conclusion

The study developed emotion detection in the context of the Internet of Medical Things (IoMT) utilizing deep learning models, notably CNN and RNN. Compared to the RNN model, the CNN model was more accurate and effective in identifying emotions from face photos. Both models, however, had difficulty correctly differentiating certain emotions, underscoring the complexity of this recognition job. A key component of IoMT, emotion recognition has the potential to have a big influence on healthcare. Particularly in the context of mental health monitoring, precise emotion detection might improve patient treatment by enabling prompt and customized interventions. IoMT which incorporates emotion recognition can improve the delivery of healthcare services by encouraging a more sympathetic and patient-centered approach. The results highlight the need for ongoing study and improvement of deep learning models in order to increase their precision and suitability for use in practical situations. Future research ought to concentrate on enhancing model functionality, investigating various datasets, and optimizing architectures for emotion identification tasks. Realizing the desired advantages in the field of healthcare also requires smoothly integrating these models into IoMT frameworks. A future where technology is sensitive to not just physical but also emotional components will result in a more compassionate and effective healthcare environment. Emotion identification utilizing deep learning inside IoMT offers great promise to improve healthcare.

Chapter 6: Recommendation and Conclusion

6.1 Introduction

In this concluding chapter, the study delves into recommendations and conclusions drawn from the research on integrating emotion recognition into healthcare through deep learning and the Internet of Medical Things (IoMT). The insights gained from the study's findings shape the recommendations, offering guidance for further advancements in this critical domain. These recommendations encompass refining model accuracy, diversifying datasets, emphasizing real-time application, prioritizing privacy, and data security, and advocating for interdisciplinary collaboration. The chapter concludes by highlighting the potential of emotion recognition integration within healthcare, emphasizing the need to address challenges and embrace opportunities for a patient-centric, empathetic healthcare system. This exploration encapsulates the study's trajectory, underlining the importance of informed recommendations and insights garnered from the analysis of emotion recognition models and their impact within the IoMT context.

6.2 Linking with Objectives

In the realm of this research, the objectives provide the guiding pillars that significantly influence the research's progression and outcomes:

1. To identify the relevant questions for this research which will help in the execution of the research according to the objectives.

This objective laid the foundation by ensuring that the research questions were meticulously framed, aligning with the broader objectives. The research questions were crafted to guide the study towards effectively integrating emotion recognition into the healthcare domain using deep learning and the Internet of Medical Things (IoMT). These questions steered the direction of the research, aiding in achieving the ultimate goals.

2. To gain a proper insight for the review of the literature with the understanding of the emotion recognition technologies.

Understanding emotion recognition technologies was crucial in achieving the research goals. The literature review extensively explored existing emotion recognition technologies, providing vital insights into their applications and limitations. This understanding proved essential in informing the subsequent steps of the research.

3. To establish suitable methods for data collection and conduct the research accordingly.

Determining appropriate data collection methods was pivotal. The chosen secondary analysis deductive approach aligned with this objective, allowing the utilization of pre-existing data for analysis. This methodological choice streamlined the research process, ensuring efficient data usage and aligning with the research objectives.

4. To analyze the gaps in the literature in conjunction with the other constraints and to offer solutions for further study.

The research analyzed gaps in the literature, exploring limitations and constraints encountered during the study. These analyses, in harmony with the set objectives, provided a comprehensive view of the research landscape. The solutions and recommendations derived from this analysis set the stage for future studies in the domain of emotion recognition and healthcare integration.

6.3 Recommendations

The study's conclusions, combined with the shifting circumstances of deep learning along with the Internet of Medical Things (IoMT), generated a variety of suggestions to enhance the way that these technologies are employed to recognise emotions in hospital settings. These suggestions include a range of topics, including diverse cooperation, applications that utilise real-time different datasets, privacy, and data security:

1. Enhancing Model Accuracy:

Model improvement should constitute the continual focus for initiatives to improve the precision of emotion recognition models. To enhance the performance of current models, researchers should look into fine-tuning approaches. Models may grow more reliable and accurate by experimentation with innovative model structures, including hybrid CNN-RNN systems. To increase recognition rates, preprocessing addresses like noise reduction and feature selection should be thoroughly examined and tailored to particular healthcare scenarios.

2. Diverse Dataset Integration:

Various and comprehensive datasets that represent the diversity of emotions across various cultures, generations, and circumstances should be utilised for developing emotion recognition algorithms. The creation of datasets that cover a wide range of emotional responses can be aided by working alongside psychologists, sociologists, and culturally specialists. To verify that the models can be applied in medical applications, datasets should contain samples that come from various healthcare settings.

3. Real-time Application:

For urgent patient care, the research and development for real-time recognition of emotions software is essential. Real-time input on patients' state of emotions might be useful for healthcare practitioners, enabling rapid intervention and individualised treatment. To enable quick reactions, emotion recognition software should be smoothly incorporated with IoMT devices including health care systems.

4. Privacy and Data Security:

Strong security measures along with privacy policies are essential due to the delicate nature of health information in IoMT-based emotion identification. Utilise cutting-edge encryption methods to safeguard data while it is being sent and stored. Put in place strict user authentication procedures to be sure that only people with permission have access to patient data. In an effort to safeguard patient privacy and yet allow for useful research and insights, anonymize patient data.

5. Interdisciplinary Collaboration:

The effective implementation of understanding emotions in healthcare depends on collaboration between professionals, scholars, and technicians from many fields. To promote innovation, support collaborations between tech firms, university academics, and healthcare organisations. Interdisciplinary teams may make sure that emotion detection technology complies to the real-world requirements and moral standards of healthcare settings.

6. Ethical Considerations:

As emotion recognition technology advances, it is essential to continually revisit and update ethical guidelines and regulations. Establish ethical frameworks that govern the use of emotion recognition in healthcare to prevent misuse or bias. Regularly assess the impact of emotion recognition on patient autonomy, consent, and well-being, and adjust practices accordingly.

7. Patient-Centered Design:

Emotion recognition applications should be designed with a patient-centric approach, considering patient preferences and feedback. Conduct usability studies and gather patient insights to create user-friendly interfaces and experiences. Ensure that patients have control over their data and can easily opt in or out of emotion recognition features.

By implementing these recommendations, the integration of emotion recognition into healthcare through deep learning and IoMT can become more effective, accurate, and ethically sound. These steps will not only benefit healthcare providers but also enhance the overall patient experience and the quality of care delivered.

8. Future Works:

Future research on deep learning for recognizing emotions utilizing Network of Medical Things (IoMT) information will concentrate on enhancing model stability and maintaining privacy. The research will investigate multimodal data combination, actual time emotion monitoring, and individualized emotion modelling for better mental wellness assistance, opening the path for more precise and available emotional well-being solutions.

6.4. Reflection

Profound Learning for Feeling Acknowledgment utilizing Web of Clinical Things (IoMT) information is a captivating and developing examination theme that holds huge commitment in further developing medical care results and upgrading how we might interpret human feelings. This field consolidates two strong spaces: profound realizing, which succeeds in removing complex examples from information, and IoMT, which includes the interconnectedness of clinical gadgets and frameworks to gather and dissect wellbeing-related data.

Feeling acknowledgment has tremendous expected in medical services, as understanding a patient's close to home state can help with additional customized and compelling therapies. Profound learning models can be prepared on IoMT information, including physiological signs like pulse, skin conductance, and looks caught by savvy gadgets, to precisely derive close to home states. The moral ramifications of using IoMT information for feeling acknowledgment can't be ignored. Protecting patient security, acquiring informed assent, and guaranteeing that the utilization of this innovation lines up with clinical morals are basic parts of any examination or application in this space.

Feeling acknowledgement utilizing IoMT information isn't restricted to medical services; it has likely applications in different areas, including human-PC cooperation, client criticism examination, and psychological well-being observing. These more extensive applications feature the flexibility and importance of this exploration subject. As the field of IoMT keeps on advancing, future exploration ought to zero in on refining feeling acknowledgement models, consolidating multimodal information (e.g., sound, text), and investigating novel profound learning structures. Furthermore, tending to the interpretability and reasonableness of the models and guaranteeing consistent incorporation into medical services frameworks will be fundamental for useful execution.

6.5 Conclusion

The integration of emotion recognition into healthcare through deep learning and the Internet of Medical Things holds immense promise for revolutionizing patient care. By decoding emotions, healthcare providers can tailor interventions, improve patient engagement, and enhance overall healthcare experiences. Models like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have shown potential in recognizing emotions, with CNN demonstrating higher accuracy in image-based emotion recognition.

Reference List

Journals

Abdullah, M., Hadzikadicy, M. and Shaikhz, S., 2018, December. SEDAT: sentiment and emotion detection in Arabic text using CNN-LSTM deep learning. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 835-840). IEEE.

Acheampong, F.A., Wenyu, C. and Nunoo-Mensah, H., 2020. Text-based emotion detection: Advances, challenges, and opportunities. Engineering Reports, 2(7), p.e12189.

Alam, M.G.R., Abedin, S.F., Moon, S.I., Talukder, A. and Hong, C.S., 2019. Healthcare IoT-based affective state mining using a deep convolutional neural network. *IEEE Access*, *7*, pp.75189-75202.

Alamelu, V. and Thilagamani, S., 2022. Lion based butterfly optimization with improved YOLO-v4 for heart disease prediction using IoMT. *Information Technology and Control*, *51*(4), pp.692-703.

Ali, M.F., Khatun, M. and Turzo, N.A., 2020. Facial emotion detection using neural network. the international journal of scientific and engineering research.

Aman, A.H.M., Hassan, W.H., Sameen, S., Attarbashi, Z.S., Alizadeh, M. and Latiff, L.A., 2021. IoMT amid COVID-19 pandemic: Application, architecture, technology, and security. *Journal of Network and Computer Applications*, *174*, p.102886.

Ashfaq, Z., Rafay, A., Mumtaz, R., Zaidi, S.M.H., Saleem, H., Zaidi, S.A.R., Mumtaz, S. and Haque, A., 2022. A review of enabling technologies for Internet of Medical Things (IoMT) Ecosystem. *Ain Shams Engineering Journal*, *13*(4), p.101660.

Azizi, F.N., Kurniawardhani, A. and Paputungan, I.V., 2022, November. Facial Expression Image based Emotion Detection using Convolutional Neural Network. In 2022 IEEE 20th Student Conference on Research and Development (SCOReD) (pp. 157-162). IEEE.

Bertero, D. and Fung, P., 2019, March. A first look into a convolutional neural network for speech emotion detection. In 2019 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 5115-5119). IEEE.

Cai, W., Gao, M., Jiang, Y., Gu, X., Ning, X., Qian, P. and Ni, T., 2022. Hierarchical domain adaptation projective dictionary pair learning model for EEG classification in IoMT systems. *IEEE Transactions on Computational Social Systems*.

Chen, J., Hu, Y., Lai, Q., Wang, W., Chen, J., Liu, H., Srivastava, G., Bashir, A.K. and Hu, X., 2023. IIFDD: Intra and inter-modal fusion for depression detection with multi-modal information from Internet of Medical Things. *Information Fusion*, p.102017.

Chughtai, Z.A., Hassan, M. and Malik, R., 2023. Analyzing the Efficacy of Adversarial Learning Techniques: in Improving Sentiment Analysis for Socially Implemented IoMT Systems. *International Journal of Computational and Innovative Sciences*, *1*(1), pp.22-29.

Dachapally, P.R., 2017. Facial emotion detection using convolutional neural networks and representational autoencoder units. arXiv preprint arXiv:1706.01509.

Du, G., Wang, Z., Gao, B., Mumtaz, S., Abualnaja, K.M. and Du, C., 2020. A convolution bidirectional long short-term memory neural network for driver emotion recognition. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), pp.4570-4578.

Fuke, R.P. and Mahajan, R.P., 2022, March. Pragmatic Analysis of IoMT Network Modelling Techniques from a Statistical Perspective. In *2022 International Conference on Electronics and Renewable Systems (ICEARS)* (pp. 531-542). IEEE.

Giri, S., Singh, G., Kumar, B., Singh, M., Vashisht, D., Sharma, S. and Jain, P., 2022, April. Emotion detection with facial feature recognition using cnn & opencv. In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 230-232). IEEE.

Gupta, D., Bhatia, M.P.S. and Kumar, A., 2021. Resolving data overload and latency issues in multivariate time-series IoMT data for mental health monitoring. *IEEE Sensors Journal*, *21*(22), pp.25421-25428.

Hamid, S., Bawany, N.Z., Sodhro, A.H., Lakhan, A. and Ahmed, S., 2022. A systematic review and IoMT based big data framework for COVID-19 prevention and detection. *Electronics*, *11*(17), p.2777.

Haq, A.U., Li, J.P., Khan, I., Agbley, B.L.Y., Ahmad, S., Uddin, M.I., Zhou, W., Khan, S. and Alam, I., 2022. DEBCM: deep learning-based enhanced breast invasive ductal carcinoma classification model in IoMT healthcare systems. *IEEE Journal of Biomedical and Health Informatics*.

Harrison, V. and Hole, G.J., 2019. Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic bulletin & review*, *16*, pp.264-269.

Hu, X., Ngai, E.C., Castellano, G., Hu, B., Rodrigues, J.J. and Song, J., 2021. Special Issue on "Toward Intelligent Internet of Medical Things and its COVID-19 Applications and Beyond". *IEEE Internet of Things Journal*, *8*(21), pp.15649-15651.

Islam, M. and Lee, T., 2022, July. Memd-hht based emotion detection from eeg using 3d cnn. In 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 284-287). IEEE.

Jaiswal, A., Raju, A.K. and Deb, S., 2020, June. Facial emotion detection using deep learning. In 2020 international conference for emerging technology (INCET) (pp. 1-5). IEEE.

Jaiswal, S. and Nandi, G.C., 2020. Robust real-time emotion detection system using CNN architecture. Neural Computing and Applications, 32(15), pp.11253-11262.

Kaur, D., Singh, S., Mansoor, W., Kumar, Y., Verma, S., Dash, S. and Koul, A., 2022. Computational intelligence and metaheuristic techniques for brain tumor detection through IoMT-enabled MRI devices. *Wireless Communications and Mobile Computing*, *2022*, pp.1-20.

Khan, A.R., 2022. Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis, and remaining challenges. Information, 13(6), p.268.

Khan, M., Shah, P.M., Khan, I.A., Islam, S.U., Ahmad, Z., Khan, F. and Lee, Y., 2023. IoMT-Enabled Computer-Aided Diagnosis of Pulmonary Embolism from Computed Tomography Scans Using Deep Learning. *Sensors*, *23*(3), p.1471.

Kholodna, N., Vysotska, V. and Albota, S., 2021. A Machine Learning Model for Automatic Emotion Detection from Speech. In MoMLeT+ DS (pp. 699-713).

Kumar, A., Sharma, K. and Sharma, A., 2021. Genetically optimized Fuzzy C-means data clustering of IoMT-based biomarkers for fast affective state recognition in intelligent edge analytics. *Applied Soft Computing*, *109*, p.107525.

Martínez, J. and Vega, J., 2022. ROS System Facial Emotion Detection Using Machine Learning for a Low-Cost Robot Based on Raspberry Pi. *Electronics*, *12*(1), p.90.

Mukhopadhyay, M., Pal, S., Nayyar, A., Pramanik, P.K.D., Dasgupta, N. and Choudhury, P., 2020, February. Facial emotion detection to assess Learner's State of mind in an online learning system. In Proceedings of the 2020 5th international conference on intelligent information technology (pp. 107-115).

Otoum, Y., Wan, Y. and Nayak, A., 2021, December. Federated transfer learning-based ids for the internet of medical things (iomt). In *2021 IEEE Globecom Workshops (GC Wkshps)* (pp. 1-6). IEEE.

Phillips, P.J., Flynn, P.J., Beveridge, J.R., Scruggs, W.T., O'toole, A.J., Bolme, D., Bowyer, K.W., Draper, B.A., Givens, G.H., Lui, Y.M. and Sahibzada, H., 2019. Overview of the multiple biometrics grand challenge. In *Advances in Biometrics: Third International Conference, ICB 2009, Alghero, Italy, June 2-5, 2009. Proceedings 3* (pp. 705-714). Springer Berlin Heidelberg.

Polignano, M., Basile, P., de Gemmis, M. and Semeraro, G., 2019, June. A comparison of word-embeddings in emotion detection from text using bilstm, cnn and self-attention. In Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (pp. 63-68).

Pushparaj, A.K. and Saini, G., 2023. AI-Based ECG Signal Monitoring System for Arrhythmia Detection Using IoMT. *AI for Big Data-Based Engineering Applications from Security Perspectives*, p.85.

Qureshi, A., Batra, S., Vats, P., Singh, S., Phogat, M. and Sharma, A.K., 2022. A review of machine learning (ML) in the internet of medical things (IOMT) in the construction of a smart healthcare structure. *Journal of Algebraic Statistics*, *13*(2), pp.225-231.

Rachakonda, L., Bapatla, A.K., Mohanty, S.P. and Kougianos, E., 2020. Sayopillow: Blockchain-integrated privacy-assured iomt framework for stress management considering sleeping habits. *IEEE Transactions on Consumer Electronics*, *67*(1), pp.20-29.

Rachakonda, L., Mohanty, S.P. and Kougianos, E., 2020. iLog: An intelligent device for automatic food intake monitoring and stress detection in the IoMT. *IEEE Transactions on Consumer Electronics*, *66*(2), pp.115-124.

Ramani, B., Patel, W. and Solanki, K., 2022. Stress Ocare: An advance IoMT based physiological data analysis for anxiety status prediction using cloud computing. *Journal of Discrete Mathematical Sciences and Cryptography*, *25*(4), pp.1019-1029.

Rathour, N., Alshamrani, S.S., Singh, R., Gehlot, A., Rashid, M., Akram, S.V. and AlGhamdi, A.S., 2021. IoMT based facial emotion recognition system using deep convolution neural networks. *Electronics*, *10*(11), p.1289.

Raza, H., Abbas, N., Amir, S., Arshad, K. and Siddiqui, M.R.U., 2022. An IoMT enabled smart healthcare model to monitor elderly people using Explainable Artificial Intelligence (EAI). *Journal of NCBAE*, *1*(2), pp.16-22.

Rbah, Y., Mahfoudi, M., Balboul, Y., Fattah, M., Mazer, S., Elbekkali, M. and Bernoussi, B., 2022, March. Machine learning and deep learning methods for intrusion detection systems in iomt: A survey. In *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)* (pp. 1-9). IEEE.

Rezaee, K., Khosravi, M.R., Attar, H., Menon, V.G., Khan, M.A., Issa, H. and Qi, L., 2023. IoMT-Assisted Medical Vehicle Routing Based on UAV-Borne Human Crowd Sensing and Deep Learning in Smart Cities. *IEEE Internet of Things Journal*.

Santamaria-Granados, L., Munoz-Organero, M., Ramirez-Gonzalez, G., Abdulhay, E. and Arunkumar, N.J.I.A., 2018. Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS). IEEE Access, 7, pp.57-67.

Singh, S., Sharma, S. and Bhadula, S., 2023. Deep Learning Techniques in Big Data-Enabled Internet-of-Things Devices. In *Big Data Analytics in Fog-Enabled IoT Networks: Towards a Privacy and Security Perspective* (pp. 1-34). 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487–2742: CRC Press.

Singh, S.K., Thakur, R.K., Kumar, S. and Anand, R., 2022, March. Deep learning and machine learning based facial emotion detection using CNN. In 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 530-535). IEEE.

Singhal, N., Ganganwar, V., Yadav, M., Chauhan, A., Jakhar, M. and Sharma, K., 2021. Comparative study of machine learning and deep learning algorithm for face recognition. *Jordanian Journal of Computers and Information Technology*, *7*(3).

Syed, L., Jabeen, S., Manimala, S. and Alsaeedi, A., 2019. Smart healthcare framework for ambient assisted living using IoMT and big data analytics techniques. *Future Generation Computer Systems*, *101*, pp.136-151.

Tariq, Z., Shah, S.K. and Lee, Y., 2019, December. Speech emotion detection using iot based deep learning for health care. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 4191-4196). IEEE.

Vajar, P., Emmanuel, A.L., Ghasemieh, A., Bahrami, P. and Kashef, R., 2021, October. The internet of medical things (IoMT): a vision on learning, privacy, and computing. In *2021 international conference on electrical, computer, communications and mechatronics engineering (ICECCME)* (pp. 1-7). IEEE.

Yang, R., 2023. Privacy and surveillance concerns in machine learning fall prediction models: implications for geriatric care and the internet of medical things. *AI & SOCIETY*, pp.1-5.

Young, S.G., Hugenberg, K., Bernstein, M.J. and Sacco, D.F., 2012. Perception and motivation in face recognition: A critical review of theories of the cross-race effect. *Personality and Social Psychology Review*, *16*(2), pp.116-142.

Zehra, W., Javed, A.R., Jalil, Z., Khan, H.U. and Gadekallu, T.R., 2021. Cross corpus multi-lingual speech emotion recognition using ensemble learning. Complex & Intelligent Systems, pp.1-10.

Zhang, T., Liu, M., Yuan, T. and Al-Nabhan, N., 2020. Emotion-aware and intelligent internet of medical things toward emotion recognition during COVID-19 pandemic. *IEEE Internet of Things Journal*, *8*(21), pp.16002-16013.

Zikria, Y.B., Afzal, M.K., Kim, S.W., Marin, A. and Guizani, M., 2020. Deep learning for intelligent IoT: Opportunities, challenges and solutions. *Computer Communications*, *164*, pp.50-53.

Appendices

Appendix 1: Real-Time Emotion Recognition from Facial Expressions Using CNN Architecture

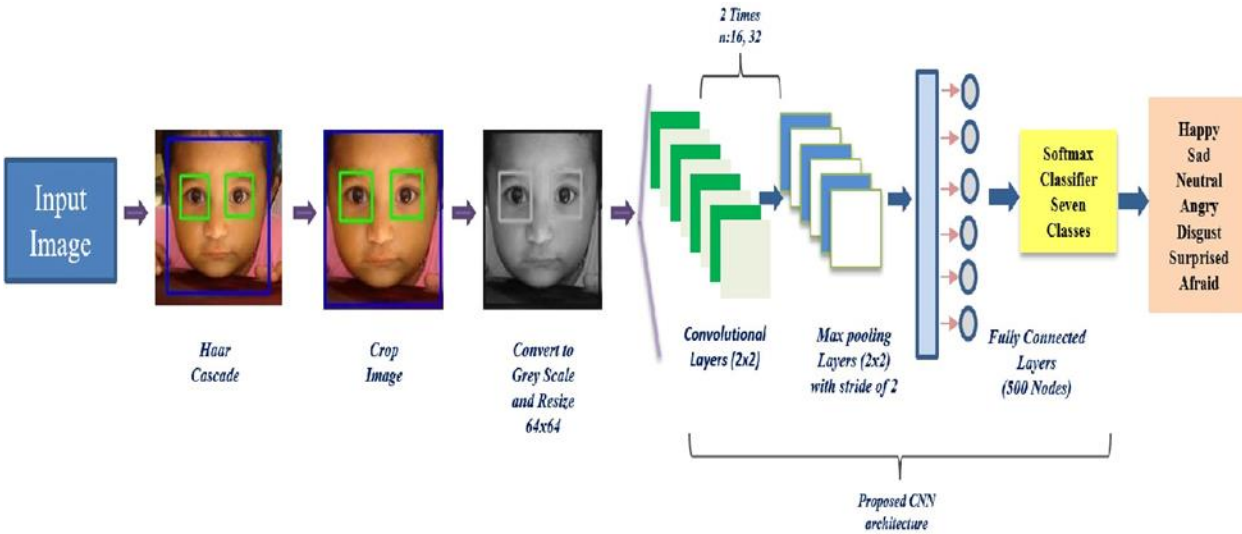
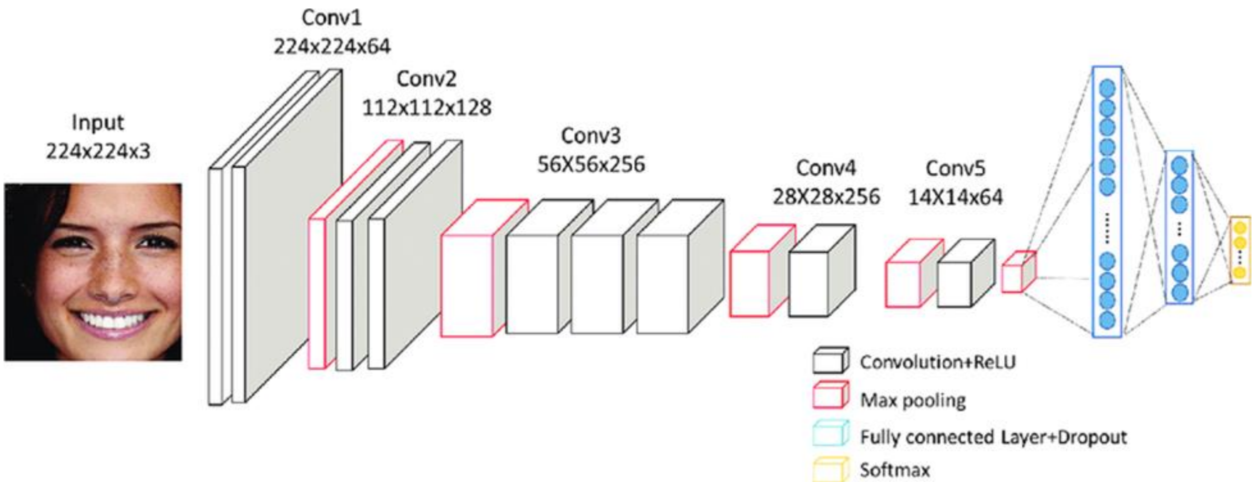


Fig. 3. Proposed CNN model diagram for facial emotion recognition

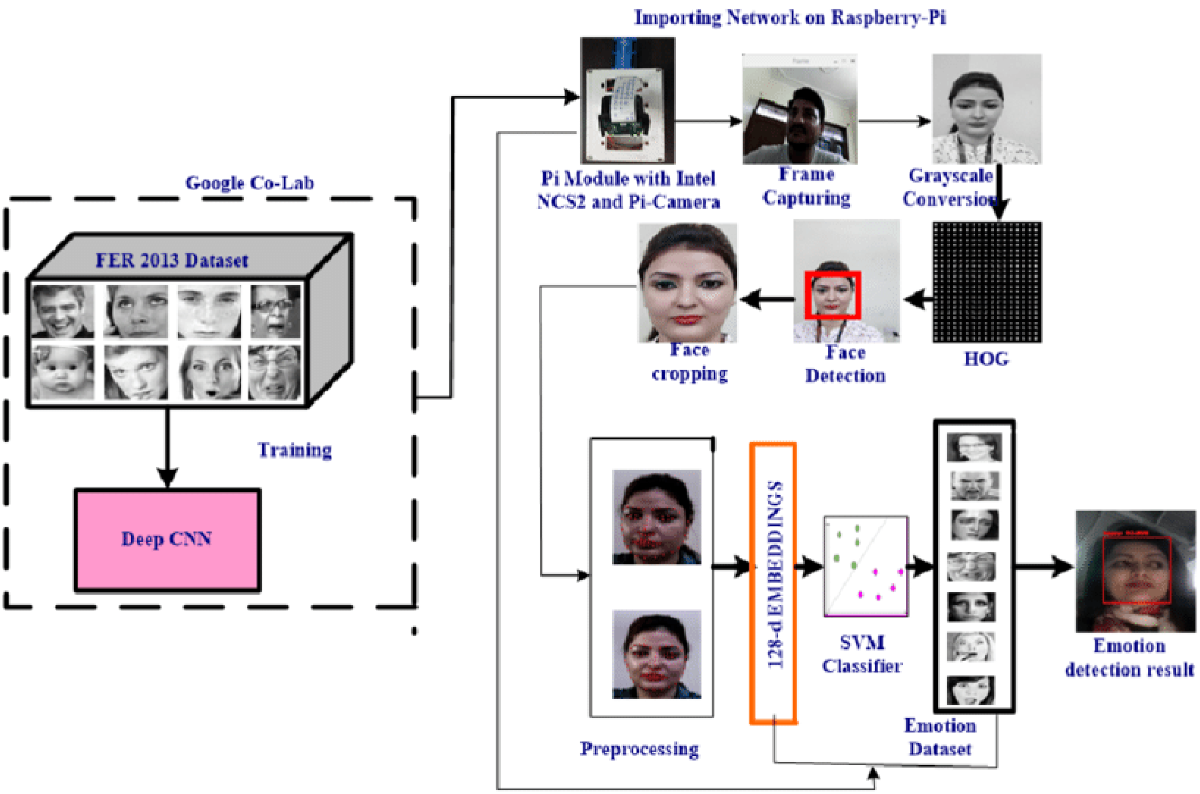
(Source: https://www.semanticscholar.org/)

Appendix 2: CNN model for facial emotion classification



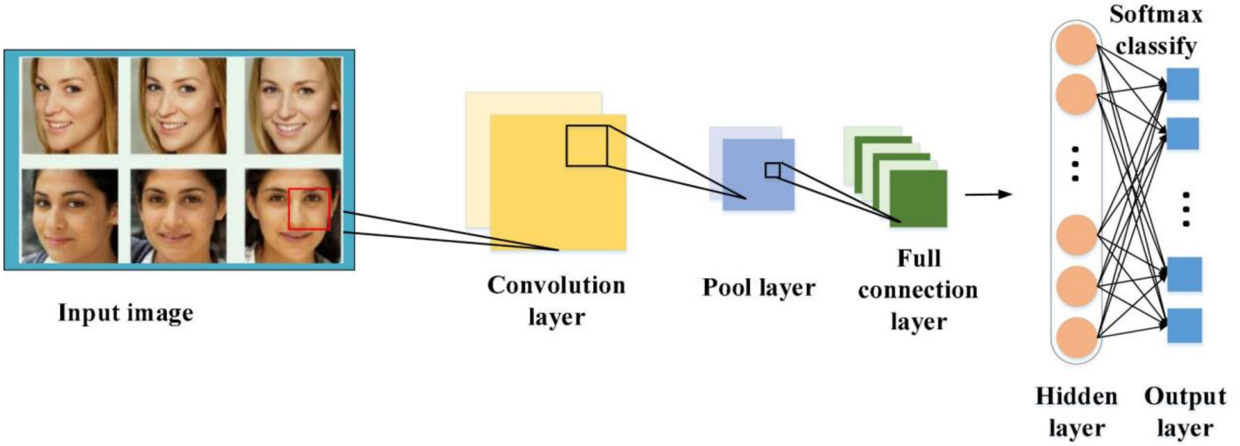
(Source: https://www.researchgate.net/)

Appendix 3: Architecture for face recognition and facial emotion recognition in real-time



(Source: [Architecture for face recognition and facial emotion recognition in... | Download Scientific Diagram (researchgate.net)])

Appendix 4: Framework of the emotion recognition algorithm of speech combined with images based on CNN-BiLSTM network



(Source: Frontiers | Deep Learning Based Emotion Recognition and Visualization of Figural Representation (frontiersin.org))

Appendix 5: Emotion classes predicted by the emotion engine

(Source: Symmetry | Free Full-Text | Smart Doll: Emotion Recognition Using Embedded Deep Learning (mdpi.com))