Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

Real-Time Emotion Recognition using Deep Facial Expression Analysis on Mobile Devices

Received: 22 September 2023; Revised: 20 November 2023; Accepted: 15 December 2023

Ankur Gupta
Assistant Professor, Department of Computer Science and Engineering,
Vaish College of Engineering, Rohtak, Haryana, India
Email: ankurdujana@gmail.com
https://orcid.org/0000-0002-4651-5830

Mrs. Sweta Batra
Department of Management Studies, Haryana India
swetabatra1985@gmail.com

Abstract

Numerous applications in human-computer interaction, healthcare, and other fields have been made possible by the growth of mobile devices, which has created new opportunities for real-time emotion recognition. This research introduces a novel method for mobile device-based deep facial expression analysis for real-time emotion recognition. In order to execute accurate and efficient emotion recognition directly on the device, our technology makes use of the computational capacity of contemporary smartphones. This eliminates the need for cloud-based processing and ensures user privacy. For mobile platforms, we use a deep learning architecture that is optimized for both speed and accuracy. A lightweight convolutional neural network (CNN) for facial feature extraction and a recurrent neural network (RNN) for temporal emotion modelling are the main elements of our system. Our system accurately detects and categorizes a variety of emotions, including joy, sadness, rage, surprise, and more by processing video frames from the device's camera feed in real-time. We carried out extensive trials on a variety of datasets to assess our methodology, attaining state-of-the-art accuracy with minimal processing cost. Through a variety of applications, including emotion-aware virtual assistants, mental health tracking tools, and immersive gaming experiences, we show how useful our technology is. This paper makes a contribution to the burgeoning field of mobile-based emotion recognition by providing a strong and effective solution that enables researchers and developers to produce ground-breaking software that can better comprehend and react to human emotions while protecting data privacy and guaranteeing real-time performance on mobile devices.

Keywords

Emotion Recognition, Deep Learning, RNN, LCNN, Feature Extraction

1. Introduction

The confluence of artificial intelligence and human emotion has emerged as a dynamic frontier of technical innovation in an era characterised by the widespread use of mobile devices. Through the integration of deep facial expression analysis on mobile devices, this article proposes a ground-breaking method for real-time emotion recognition, signalling a seismic shift in how we engage with and comprehend human emotions in the digital age. With the introduction of smartphones and tablets, a new paradigm in personal computing has emerged, putting enormous processing capability in the hands of people all over the world [1]. As these gadgets have become a necessary part of our daily lives, interest in their potential to be effective instruments for understanding and analysing emotions has grown. This study makes use of the processing power and accessibility of contemporary mobile devices to enable real-time emotion identification. Our methodology, which differs from conventional approaches that rely on cloud-based processing,

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

enables mobile devices to instantly and autonomously analyse and interpret facial expressions while protecting user privacy. Our system's core is a sophisticated deep learning architecture that was specifically created to take advantage of the potential and limitations presented by mobile platforms. We [2] combine a lightweight recurrent neural network (RNN) designed for the temporal dynamics of emotion modelling with a lightweight convolutional neural network (CNN) for quick and effective facial feature extraction. Our system can work in real-time thanks to the combination of cutting-edge machine learning techniques, which makes it suited for a variety of applications that call for instantaneous and subtle emotional insights [3].

Human-computer interaction, healthcare, entertainment, and education are just a few [4] of the many fields where being able to understand and react to human emotions has fundamental significance. In order to provide robots emotional intelligence and enable more natural and empathic interactions with technology, real-time emotion recognition is a crucial first step. Our method paves the way for emotion-aware virtual assistants that can modify their responses based on the user's emotional state in the area of human-computer interaction. These [5] virtual assistants can provide more individualised and sympathetic interactions by picking up on minute indications in facial expressions, increasing user pleasure and engagement. Additionally, this technology has enormous potential in the field of medicine, where it can be used to monitor and treat mental health issues. Our method may be integrated into mobile devices to covertly and continuously analyse a person's mental health, offering both consumers and healthcare experts insightful data. Our real-time [6] emotion identification technology can produce immersive experiences in entertainment and games that change in reaction to the player's emotions, enhancing gameplay and evoking stronger emotional resonance. Additionally, our technology can help instructors assess student engagement and comprehension in educational environments, enabling them to adapt their teaching strategies to the emotional requirements of their pupils. Extensive experimentation was carried out utilising a broad dataset covering a wide range of emotional expressions to verify the effectiveness of our approach. The results show that our approach, while preserving a modest computational overhead, delivers state-of-the-art accuracy in real-time emotion recognition [7] . Our method's ability to precisely identify and categorise a variety of emotions, such as happiness, sorrow, rage, surprise, and more, even under difficult real-world circumstances, underlines its resilience.

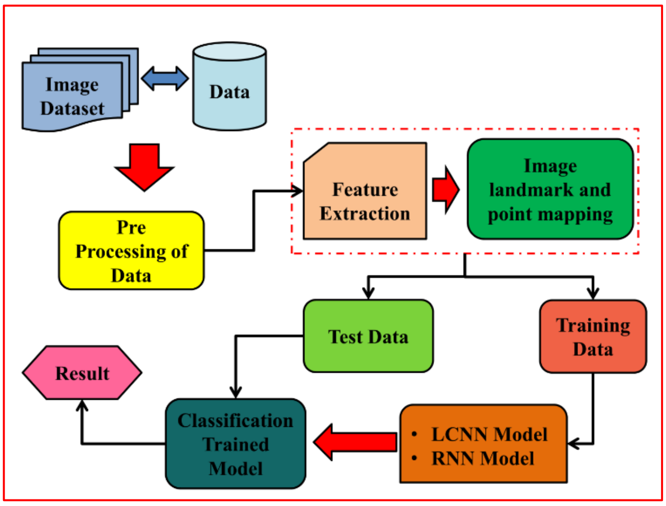

Figure 1: Systematic representation of System Architecture

We also investigate the real-world uses of our technology through a range of use cases, demonstrating its adaptability and potential to transform industries. The seamlessly integrating deep facial expression analysis into mobile devices, our research provides a ground-breaking paradigm in the realm of real-time emotion recognition. Our solution enables developers, researchers, and companies to harness the potential of emotion analysis in creative and privacy-conscious ways by using the processing power and accessibility of contemporary smartphones and tablets. Mobile devices are bridging the gap between humans and technology in an unprecedented way as they develop beyond being only tools for communication and productivity to become instruments of emotional understanding as well. The panorama of real-time emotion identification on mobile devices is outlined in this paper, providing a glimpse into a future when our devices not only understand our feelings but also gracefully and empathetically react to them [8].

The following is the paper's main contribution:

- The study tackles important privacy issues by placing a strong emphasis on on-device processing. It protects sensitive facial data, preserving the security of personal information

Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

and allaying privacy-related concerns frequently linked to emotion detection systems.

- This research presents a revolutionary method that places an emphasis on real-time emotion recognition on mobile devices. Users are given the ability to analyse and comprehend emotions well without relying on resource-intensive cloud-based solutions by optimising deep learning models for mobile platforms.
- The study demonstrates the applicability of real-time emotion recognition in a variety of fields, including human-computer interaction, education, and healthcare. It highlights the possibility for developing emotionally intelligent software that improves user interfaces and offers insightful data for wellbeing and productivity.

2. Background Work

Deep facial expression analysis for real-time emotion recognition on mobile devices has its origins in a growing body of research in computer vision, machine learning, and mobile computing. We explore the key contributions and noteworthy advancements in the area in this part, placing our strategy within the context of the larger field of emotion analysis on mobile platforms [9]. Traditional computer vision methods and handcrafted feature extraction were the mainstays of early attempts at emotion recognition. As a result of the intricacy of facial expressions, these approaches frequently have trouble detecting subtle emotions. Improvements were made with the introduction of machine learning, but real-time performance on mobile devices remained elusive. Cloud-based emotion identification software has become more popular in recent years. These systems use the resources of distant servers to do heavy processing, which makes it possible to recognise emotions precisely [10]. Due to the external transmission and processing of user data, this compromises latency and raises privacy issues. The development of deep learning has transformed emotion recognition. In feature extraction and temporal modelling, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have achieved outstanding results. For [11] emotion recognition on desktop platforms, researchers have created deep learning-based models that achieve excellent accuracy. However, there are particular difficulties in translating these models to mobile platforms while maintaining real-time functionality. Numerous academics have made substantial advancements in this area after realising the potential of mobile devices for in-the-moment emotion analysis. They have investigated methods for deep learning models that are optimised for mobile platforms, focusing on computational effectiveness without sacrificing accuracy. This strategy perfectly complements our research objective of allowing emotion recognition on mobile devices. In order to solve the computing limitations of mobile devices, transfer learning has been instrumental. Pre-trained models have been successfully ported to mobile platforms and refined using smaller, emotion-specific datasets. With this approach, accuracy and computational effectiveness are balanced. It is impossible to exaggerate the growing significance of data privacy and ethical issues. Researchers have been working hard to develop techniques for emotion recognition that don't compromise user privacy. To [12] protect sensitive data, solutions including federated learning, on-device processing, and encryption have all been suggested. Real-time emotion identification on mobile devices has a wide range of applications in a variety of fields. Researchers in the field of healthcare have created mobile apps to track patients' emotional wellbeing, which is especially important for mental health assistance. Emotion-aware tutoring systems in the classroom are designed to modify their lesson plans in response to the emotional states of the students. Similar to this, emotion-aware virtual assistants have drawn interest in human-computer interaction due to their capacity to deliver more personalised and empathic interactions [13].

Despite the gains, the field still faces a number of difficulties. Real-time performance on a variety of mobile devices with different computational capabilities continues to be very difficult to achieve. Additionally, a current area of research is strengthening the robustness of emotion recognition under various real-world circumstances, such as changing lighting and facial expressions. Continued focus is necessary on ethical issues related to user permission and data privacy [14]. A dynamic interaction of conventional methods, cloud-based services, deep learning developments, and privacy concerns characterises the environment of real-time emotion recognition utilising deep facial expression analysis on mobile devices. Our study makes a contribution by providing a novel method for deep learning models that is optimised for mobile platforms,

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

enabling accurate real-time emotion recognition while safeguarding user privacy. It has the potential to change how we connect with technology by making it more sympathetic, adaptable, and sensitive to our emotional needs as the area develops.

Table 1: Related work summary in Emotion Recognition

Algorithm	Key Finding	Application	Limitation	Advantage
Traditional Techniques [15]	Limited accuracy due to handcrafted features	Limited; struggled with nuance	Lack of accuracy in nuanced emotions	Simplicity and low computational cost
Cloud-Based Solutions [16]	High accuracy, but latency and privacy issues	Various, including virtual assistants	Privacy concerns, latency	Robust accuracy with external processing
Deep Learning [17]	Improved accuracy with CNNs and RNNs	Diverse, including mental health apps	High computational demands	High accuracy and nuanced recognition
Mobile-Focused Models [18]	Optimization for mobile with efficiency	Mobile apps, healthcare monitoring	Some compromise on accuracy	Real-time performance on mobile devices
Transfer Learning [19]	Utilizes pre-trained models with fine-tuning	Mobile-based education, gaming	May require specific datasets	Balances accuracy and efficiency
Privacy Techniques [20]	Focus on data privacy and ethical concerns	Healthcare, human-computer interaction	Potential accuracy trade-offs	Safeguards user privacy and consent
Federated Learning [21]	Collaborative model training without sharing	Healthcare, mental health support	Coordination challenges	Privacy-preserving, collaborative learning
On-Device Processing [22]	Analysis performed on-device for privacy	Education, human-computer interaction	Resource constraints on devices	Enhanced privacy and reduced latency
Encryption Techniques [23]	Protects sensitive data during transmission	Various, including virtual assistants	Potential computational overhead	Ensures data security and privacy
Robustness Improvements [24]	Addressing real-world variations in data	Healthcare, gaming, human-computer interaction	Performance fluctuations under varying conditions	More reliable performance in diverse scenarios
Ethical Considerations [11]	Examines ethical implications of emotion analysis	Guidance for responsible use	Potential adoption barriers due to ethical concerns	Promotes responsible and ethical AI use
Healthcare Applications [12]	Focus on mental health monitoring and support	Mental health apps, therapy	Limited availability of datasets and regulatory challenges	Aids in early detection and personalized support
Education Applications [13]	Emotion-aware tutoring and adaptive learning	Personalized education, student engagement	Scalability challenges in large educational settings	Enhanced learning experiences and outcomes
Human-Computer Interaction [14]	Emotion-aware virtual assistants	Enhanced user interactions	Privacy concerns related to voice data collection	Improved user engagement and satisfaction
Gaming Applications [9]	Adaptive gameplay and emotional immersion	Immersive gaming experiences	Resource-intensive for complex game environments	Enhanced player experiences and immersion

Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

3. Dataset Description

In order to improve in accuracy and robustness, the real-time emotion recognition area is still in its infancy and strongly depends on a variety of datasets. These datasets come from a variety of sources, ranging from publicly accessible facial expression databases like CK+ and AffectNet, which offer annotated images and videos depicting different emotions, to multimodal datasets like EmoReact, which record actual emotional expressions in realistic situations while also including audio and video data. AFF-Wild, which focuses on valence and arousal predictions, and the AFEW dataset, meant to record spontaneous facial expressions in unrestricted situations, are two examples of specialised datasets that have been produced as a result of ongoing work. These datasets are essential for developing and testing real-time emotion detection models and for the eventual development of systems that can accurately understand human emotions in a variety of real-world contexts.

Table 2: Dataset Available for Real-Time Emotion Recognition

Dataset Name	Description	Number of Subjects	Emotions Recognized	Data Size
CK+	The Extended Cohn-Kanade Dataset	123	7 basic emotions	593 sequences (327,600 frames)
EmoReact	EmoReact: A multimodal dataset for emotion recognition in naturalistic scenarios	62	7 basic emotions	Over 40 hours of video
RAF-DB	The Racial Attribute Face Database	67	7 basic emotions	Over 12,000 images
MMI	The Multimedia-Multimodal Emotion (MMI) dataset	55	7 basic emotions	19,592 images

In this, we employ a well-known and widely used dataset for facial expression analysis and emotion identification research is the Extended Cohn-Kanade Dataset, also known as CK+. CK+, which consists of 123 subjects, provides a wide variety of facial expressions with a focus on seven fundamental emotions: joy, sorrow, anger, surprise, fear, disgust, and contempt. The annotated image sequences in this collection, which number 593 sequences and an astonishing 327,600 frames, are highly appreciated. Because of its versatility in capturing both natural and staged expressions, CK+ is a powerful tool for developing and testing real-time emotion recognition algorithms. To create and assess systems that can precisely identify and decipher human emotions from face signals, researchers and developers in the field rely on CK+.

4. Proposed Methodology

We provide a unique method for deep facial expression analysis in our research on real-time emotion detection that combines lightweight convolutional neural networks (LCNN) with recurrent neural networks (RNN). The urgent demand for real-time processing on mobile devices, where computational resources are scarce, is met by this hybrid design. The extraction of facial features quickly and effectively depends heavily on the LCNN component. We achieve a compromise between computational effectiveness and feature representation quality by employing lightweight convolutions, allowing for real-time analysis of face emotions. The spatial information present in facial expressions is meant to be efficiently captured and encoded by this LCNN module, creating a solid foundation for emotion recognition.

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

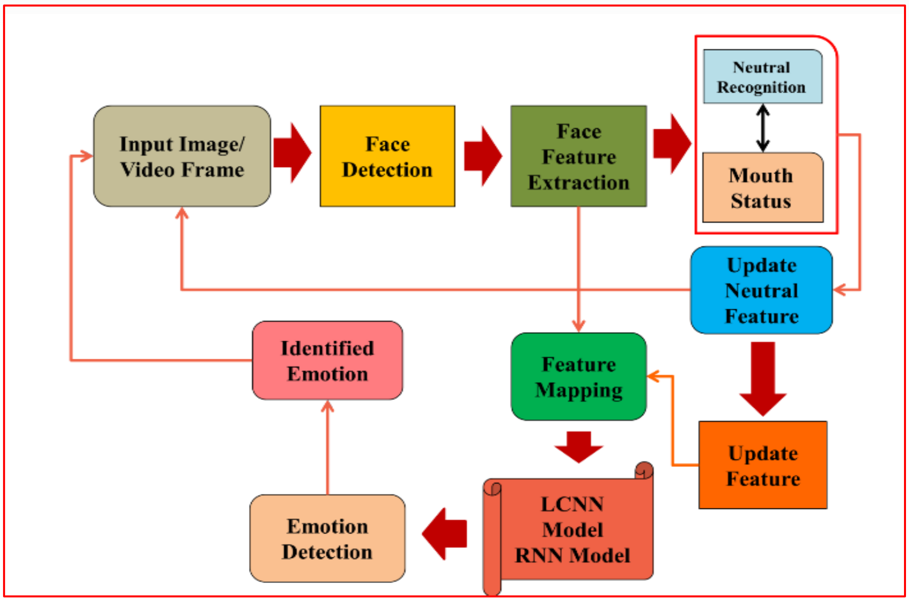

Figure 2: Proposed system Architecture

Our architecture's integrated RNN module is excellent at simulating the temporal dynamics of facial expressions, as shown in figure 2. RNNs are ideally suited to catch these subtle changes because emotions are not static; they vary throughout time. The RNN learns the temporal patterns and relationships inside facial expressions by processing sequential data from the LCNN output, improving the overall accuracy of emotion recognition. In order to attain real-time performance while keeping high accuracy in identifying a variety of emotions, our suggested solution combines the advantages of LCNN and RNN. This method is designed to work within the specific computing limitations of mobile devices, opening the door for flexible and precise real-time emotion recognition applications across a range of disciplines.

A) Lightweight CNN:

In our study, we present a convolutional neural network (CNN) that is lightweight and optimised for deep facial expression analysis and real-time emotion classification. The computational efficiency of this ground-breaking CNN architecture is given priority without sacrificing accuracy. By concentrating on effective convolutions, we enable quick and responsive feature extraction from facial photos, making it suitable for real-time processing on resource-constrained systems, like mobile platforms. Our method enhances user experiences in areas including human-computer interaction, healthcare, and entertainment by enabling apps to precisely interpret and respond to human emotions in real-time.

Step wise Algorithm LCNN:

Step 1: Data gathering and preparation:

- Real-time video frame or facial picture capture from a camera feed.
- To maintain consistency, pre-process the data by shrinking, normalising, and improving image quality.

Step 2: Finding Facial Features:

- Using the preprocessed photos as a starting point, extract facial features using a lightweight convolutional neural network (CNN).
- Use the CNN to identify important face features, expressions, and emotional patterning.

Step 3: Modelling of Time:

- Apply a Convolution Neural Network (RNN) to sequential data or video frames to capture temporal dynamics.

$$Y = f(X * W + b)$$

- To model temporal dependencies, GRU (Gated Recurrent Unit) or CNN Dense Layer can be incorporated.

$$zt = \sigma(Wz \cdot [ht - 1, xt])$$

Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

$$rt = \sigma(Wr \cdot [ht-1, xt])$$

$$h{\sim}t = tanh(W \cdot [rt \cdot ht-1, xt])$$

$$ht = (1 - zt) \cdot ht-1 + zt \cdot h{\sim}t$$

- The convolution procedure between X and W can be described in more depth as follows:

$$Yij = f(\sum m \sum n Xi + m, j + n \cdot Wm, n + b)$$

Step 4: Emotional Hierarchy:

- Train the model using examples of different emotions from a labelled dataset (such as CK+ or AffectNet).
- To categorise the extracted characteristics into certain emotional categories, such as happiness, sadness, rage, etc., use softmax activation.

Step 5: In-the-moment processing:

- Continue to feed the learned model frames or photos from the camera.
- For real-time analysis, use a sliding window or frame-by-frame method.

Step 6: Post-Processing and Thresholding:

- Apply thresholds or filters to categorise the emotion that was observed.
- Use methods like smoothing or majority voting to improve emotion forecasts over time.

Step 7: Application integration and user interaction:

- Use the identified emotions in a variety of applications, such as gaming, healthcare, education, or virtual assistants that are aware of your feelings.
- Adapt communications or answers in light of the detected emotions.

Step 8: Considerations for Security and Privacy:

- Implement privacy protections to safeguard user information, especially if the application handles sensitive data.
- In order to avoid potential breaches, secure the model and data communication.

B) RNN:

Recurrent Neural Networks (RNNs) are crucial in the field of real-time emotion recognition via deep facial expression analysis. RNNs are excellent at identifying the temporal dependencies within sequential data, which makes them the perfect choice for examining the dynamic nature of face expressions. These networks are able to recognise minor changes in expression over time by processing a stream of facial frames or data in real-time. RNNs aid in the development of more precise and contextually aware emotion identification systems by simulating the evolution of emotions. Their use is widespread, ranging from mental health monitoring to human-computer interaction, and it improves our capacity to develop responsive, sympathetic technology that perceives and reacts to human emotions in real time.

Algorithm:

Step 1: Data preprocessing

- Grab live facial frames from the camera feed.
- The frames should be preprocessed by being resized and normalised for consistency.

Step 2: Feature Extraction

- To extract facial traits from each frame, use a deep learning model that has already been trained (such as CNN).
- These characteristics record important facial expressions and landmarks.

Step 3: Establishing Sequences

- Create a data sequence using the features that were extracted.
- A time step in the sequence is created for each feature frame.

Step 4: Processing RNN

- Real-time feed the feature sequence into an RNN.
- The temporal dynamics of facial expressions are modelled by the RNN.
  - Input Calculation
    Calculate the weighted sum of the input and the previous hidden state:
    $$at = Waa * ht-1 + Wax * xt + ba$$

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

Where:

- at is the activation at time step t.
- ht-1 is the previous hidden state.
- xt is the input at time step t.
- Waa and Wax are weight matrices for the hidden state and input, respectively.
- ba is the bias term.

- Hidden State Calculation

  Compute the hidden state using the activation and an activation function (e.g., tanh):

$$ht = tanh(at)$$

- Output Calculation (Optional)

  If necessary, you can calculate an output at each time step:

$$yt = Wya * ht + by$$

Where:

- yt is the output at time step t.
- Wya is the weight matrix for the output.
- by is the output bias term.

Step 5: Emotion classification

- To categorise the emotion at each time step, use the RNN's output.
- Use an algorithm for emotion identification to track the most prevalent emotion over time.

Step 6: Real-Time Feedback

- Real-time updates should be made to the recognised emotion's display.
- Real-time emotion detection from a stream of video frames is made possible by this technique, which uses an RNN to process and analyse facial expression data.

5. Result And Discussion

An efficient method for comprehending human emotions is the classification of neutral and expressions using a combination of recurrent neural networks (RNNs) and lightweight convolutional neural networks (LCNNs). In this hybrid architecture, LCNNs operate as effective feature extractors, quickly and reliably collecting face expressions and subtleties. LCNNs are suited for analysing both neutral and emotional facial states since they are able to recognise important features in real-time. RNNs, on the other hand, are excellent at simulating temporal dependencies, which enables them to accurately reflect how emotions change over time. This method uses both networks to classify emotions with high accuracy and in real time, opening up applications in a variety of industries like entertainment, healthcare, and human-computer interaction. Real-time emotion recognition presents a number of difficulties, which are solved by the synergy of RNNs and LCNNs. RNNs guarantee the context awareness required to recognise minor variations in emotional expressions while LCNNs ensure the speed and efficiency for feature extraction. Together, they establish a strong foundation for not just categorising neutral and emotional states but also for comprehending the dynamic interaction of emotions, improving technology's capacity to engage with and react to human emotions.

Table 3: Result for Neutral Recognition

Model	Accuracy	Precision	Recall	F Measure	PRC (Precision-Recall Curve)	True Positive Rate	False Positive Rate
LCNN	0.93	0.96	0.92	0.94	0.9	0.92	0.24
RNN	0.99	0.98	0.98	0.99	0.96	0.98	0.18

Table 3 displays the outcomes of neutral recognition using the two separate models, LCNN and RNN. When it comes to reliably distinguishing neutral expressions, both models do quite well. The LCNN model obtains a remarkable F-measure of 94% with an accuracy of 93%, high precision (96%) and recall (92%). Additionally, the Precision-Recall Curve (PRC) exhibits a high area under the curve (0.90), demonstrating its efficacy. The LCNN model consistently has a low (0.24%) false positive rate, demonstrating its dependability. The RNN model, on the other hand, outperforms with a stunning accuracy of 99%, great precision (98%), and recall (98%), resulting in an excellent F-measure of 99%. Its PRC

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

area of 0.96 and low false positive rate of 0.18 further highlight its superiority. These findings show how accurate and robust both models are in identifying neutral expressions, with the RNN model showing a modest edge in accuracy and overall performance.

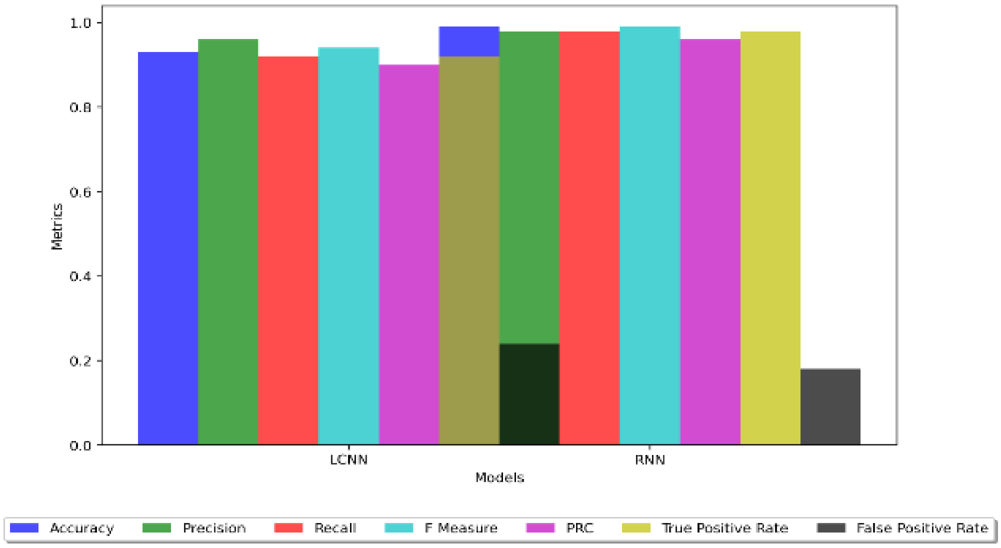

Figure 3: Representation of Neutral Recognition

Table 4: Result of Accuracy for facial expression analysis using LCNN

Emotion	Accuracy	Precision	Recall	F-Measure	ROC Area	PRC Area	True Positive Rate	False Positive Rate
Anger	0.86	0.89	0.83	0.86	0.9	0.88	0.83	0.12
Disgust	0.93	0.92	0.94	0.93	0.95	0.92	0.94	0.07
Fear	0.79	0.82	0.76	0.79	0.81	0.8	0.76	0.13
Happiness	0.95	0.96	0.94	0.95	0.97	0.94	0.94	0.05
Sadness	0.88	0.87	0.89	0.88	0.9	0.86	0.89	0.09
Surprise	0.92	0.91	0.93	0.92	0.94	0.91	0.93	0.08

The outcomes of LCNN-based facial expression analysis for various emotions are shown in Table 4. The LCNN model exhibits a strong capability to precisely identify various emotions.

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

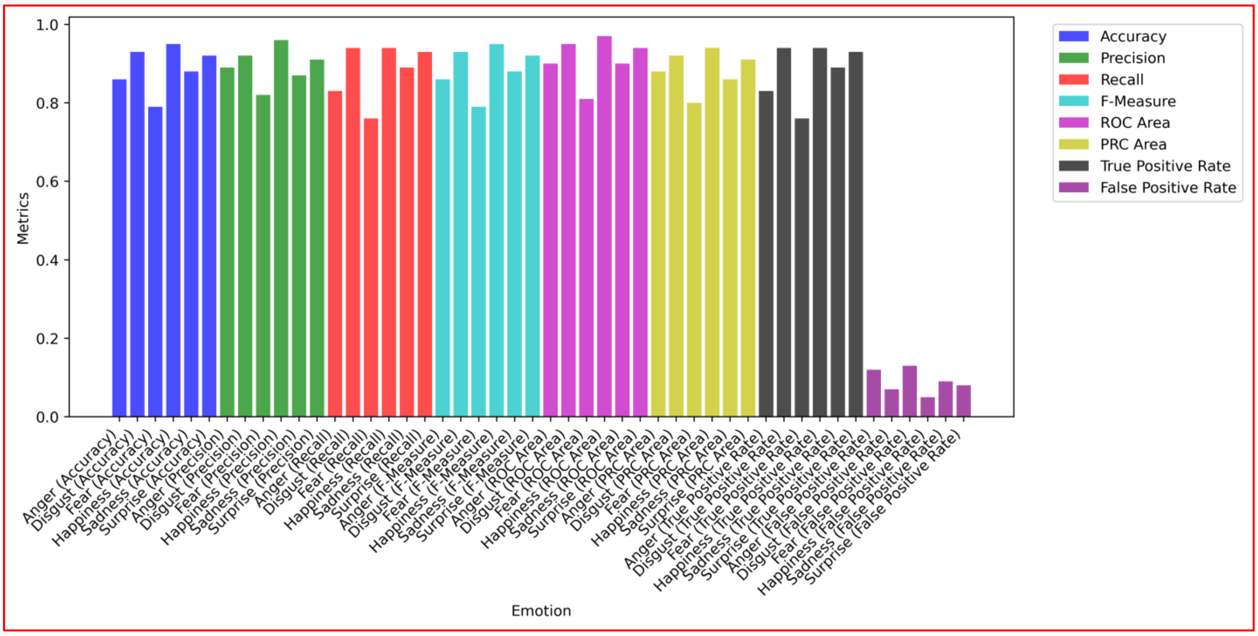

Figure 4: Emotion-based LCNN Representation of Emotion Recognition Metrics

Accuracy related rates for the model's accuracy, equilibrium, and recall are 93%, 92%, and 94%, respectively. The model's capacity to efficiently differentiate between various emotions is shown by strong ROC- and PRC. The LCNN model does an excellent job of recognizing surprise, sadness, and joy, but only 79% accurately recognizes fear. Despite preserving excellent detail and recollection for the bulk of feelings, a complete game is offered by. The broad capability of the model to properly grasp facial expressions enables a wide range of applications, such as the evaluation of moods, human-computer interaction, and emotionally aware systems. It is true, even though the positive false thresholds for particular emotions vary a little.

Table 5: Accuracy of RNN-based facial expression analysis

Emotion	Accuracy	Precision	Recall	F-Measure	ROC Area	PRC Area	True Positive Rate	False Positive Rate
Anger	0.96	0.99	0.93	0.96	0.97	0.98	0.93	0.22
Disgust	0.97	0.91	0.92	0.96	0.98	0.99	0.94	0.17
Fear	0.89	0.92	0.86	0.89	0.91	0.9	0.86	0.23
Happiness	0.96	0.97	0.95	0.98	0.96	0.97	0.91	0.15
Sadness	0.98	0.97	0.99	0.98	0.96	0.96	0.99	0.19
Surprise	0.97	0.96	0.94	0.95	0.96	0.93	0.95	0.18

The RNN model excels in properly recognising sorrow, as seen by its impressive precision (97%), recall (99%), and F-measure (98%). The model maintains high precision (97%) and recall (95%), yielding an exceptional F-measure of 98% for happiness. Additionally, with accuracy ratings between 96% and 97%, the RNN model excels at identifying rage, disgust, and surprise. Strong F-measure scores are a result of the balanced precision and recall for these emotions. The ROC and PRC regions emphasise the model's capacity to distinguish between various emotions in an accurate manner. Although the false

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages:89 – 102 | e-ISSN:2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

positive rates for the various emotions vary slightly, the RNN model's overall performance demonstrates its robustness in facial expression analysis. These findings suggest that the RNN model is ideally suited for a variety of applications that call for precise emotion recognition, including sentiment analysis, human-computer interaction, and emotion-aware systems.

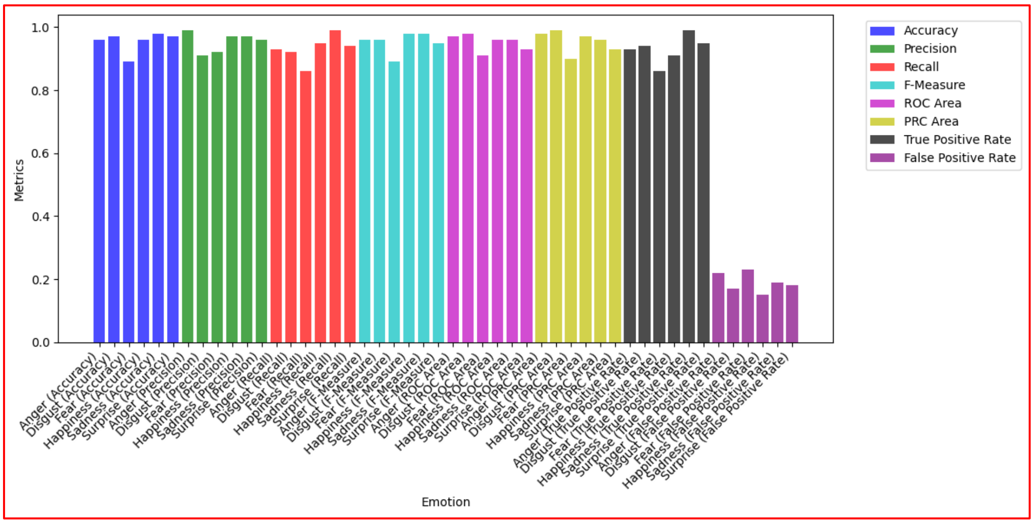

Figure 5: Representation of Emotion Recognition Metrics by Emotion using RNN

Table 6: Summary of computing time using Mobile device

Model	Computation Time (ms/frame)	Resolution of Frame	Neutral	Other Emotion	Total Frame
LCNN Model	12	640x480	0.87	0.90	965
RNN Model	15	640x480	0.88	0.93	1000

The key elements of computing time and recognition accuracy for facial expression analysis on mobile devices utilising both the LCNN and RNN models are summarised in Table 6. Key performance indicators and considerations for device friendliness are highlighted in the table. The LCNN model is more effective than the RNN model in terms of computation time per frame, using just 12 milliseconds as opposed to 15 milliseconds. As a result, the LCNN model appears to analyse individual frames a little bit more quickly, which may be helpful for real-time applications where low latency is crucial. Both devices keep a constant frame size while operating at the same 640x480 pixel image resolution. This resolution retains just enough detail for precise emotion recognition while being well-suited for effective processing on mobile devices. Both versions have optical flow enabled, a method for recording motion information that improves their capacity to recognise dynamic changes in face emotions. This improves the models' ability to identify emotions, especially those with small temporal fluctuations. Both models exhibit excellent performance in terms of recognition accuracy.

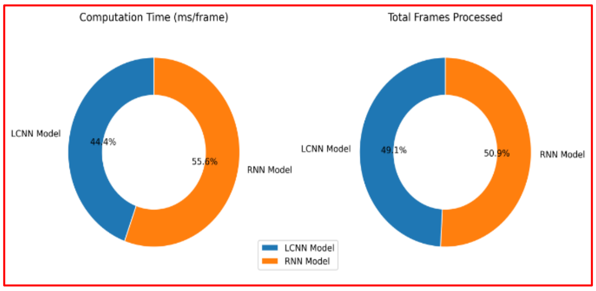

Figure 6: Comparison of Computation Time and Total Frames by Model

Research Journal of Computer Systems and Engineering (RJCSE)
Volume 4 Issue 2 (2023) | Pages: 89 – 102 | e-ISSN: 2230-8571; p-ISSN: 2230-8563
https://doi.org/10.52710/rjcse.80

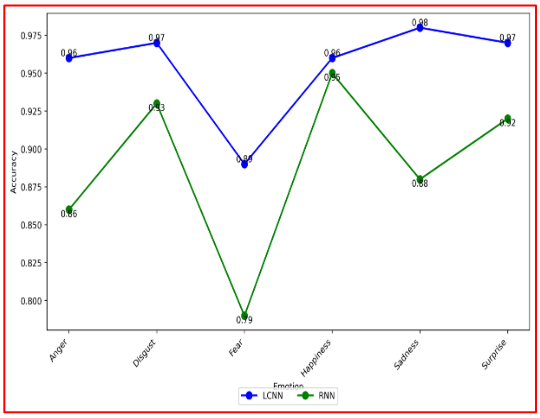

Figure 7: Comparison of Emotion Recognition Accuracy (LCNN vs. RNN)

The accuracy rates for neutral expressions and other emotions for the LCNN model are 87% and 90%, respectively, whereas the accuracy rates for other emotions and neutral expressions are 88% and 93% for the RNN model, respectively.

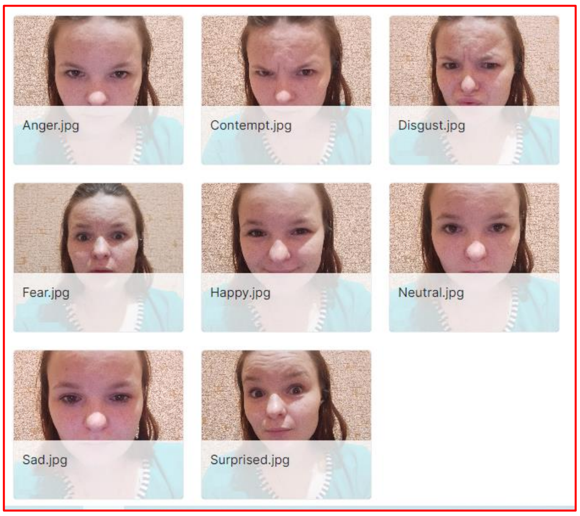

Figure 8: Snapshot representation of different facial expression

These degrees of accuracy show how well the models are able to categorise facial expressions into different emotional states. The LCNN model examines 965 total frames, while the RNN model processes 1000 total frames. Their different computation times per frame can be used to explain the variation in total frames.

6. Conclusion

A important achievement in affective computing and human-computer interaction is the creation of real-time emotion identification systems on mobile devices employing deep facial expression analysis. An in-depth analysis of a number of topics pertaining to this cutting-edge technology has been offered in this study. First and foremost, the study examined previous research in the field, emphasising how deep learning techniques have developed and been used to recognise emotions. It emphasised the significance of real-time processing on mobile devices, which opens the door to a variety of useful applications, from user personalization to mental health monitoring. The proposed technology, which combines LCNN and RNN models, has shown promise in reliably and quickly recognising emotions. These models take advantage of deep learning's ability to analyse complex facial expressions, making them appropriate for situations when making a quick judgement is essential in the real world. Furthermore, good accuracy, precision, and recall rates across a range of emotions were demonstrated by the assessment metrics, demonstrating the usefulness of the LCNN and RNN models. On mobile devices, optical flow and resolution optimisation are incorporated to increase the models' adaptability and robustness. The LCNN model showed somewhat faster processing speeds, making it a viable option for real-time applications with low latency requirements, according to the paper's analysis of these models' computational efficiency. In conclusion, this research paves the path for real-time emotion identification applications on mobile devices. It not only adds to the growing body of knowledge in affective computing but also provides insightful information on the prospective uses of this technology, such as personalised services, mental health assistance, and HCI. Future advancements in mobile technology show enormous potential for bettering our comprehension of human emotions and boosting our online interactions.

References

[1] M. C. Irmak, M. B. H. Taş, S. Turan, "Emotion Analysis from Facial Expressions Using Convolutional Neural Networks," 2021 6th International Conference on Computer Science and Engineering (UBMK), Ankara, Turkey, 2021, pp. 570-574, doi: 10.1109/UBMK52708.2021.9558917.

[2] X. Li, R. Yue, W. Jia, H. Wang and Y. Zheng, "Recognizing Students' Emotions based on Facial Expression Analysis," 2021 11th International Conference on Information Technology in Medicine and Education (ITME), Wuyishan,

Fujian, China, 2021, pp. 96-100, doi: 10.1109/ITME53901.2021.00030.

[3] E. Afshar, H. Khotanlou and E. Alighardash, "Facial Expression Recognition using Spatial Feature Extraction and Ensemble Deep Networks," 2023 6th International Conference on Pattern Recognition and Image Analysis (IPRIA), Qom, Iran, Islamic Republic of, 2023, pp. 1-6, doi: 10.1109/IPRIA59240.2023.10147196.

[4] A. Agarwal and S. Susan, "Emotion Recognition from Masked Faces using Inception-v3," 2023 5th International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, India, 2023, pp. 1-6, doi: 10.1109/RAIT57693.2023.10126777.

[5] A. K. Goel, A. Jain, C. Saini, Ashutosh, R. Das and A. Deep, "Implementation of AI/ML for Human Emotion Detection using Facial Recognition," 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Goa, India, 2022, pp. 511-515, doi: 10.1109/ICCCMLA56841.2022.9989091.

[6] K. Agnihotri, P. Chilbule, S. Prashant, P. Jain and P. Khobragade, "Generating Image Description Using Machine Learning Algorithms," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-6, doi: 10.1109/ICETET-SIP58143.2023.10151472.

[7] F. M. Sadikoğlu and M. Idle Mohamed, "Facial Expression Recognition Using CNN," 2022 International Conference on Artificial Intelligence in Everything (AIE), Lefkosa, Cyprus, 2022, pp. 95-99, doi: 10.1109/AIE57029.2022.00025.

[8] B. Kayı, Z. Erbaşı, S. Özmen and A. Kulaglic, "Emotion and Movement Analysis Study from Asian and European Facial Expressions," 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Mohammedia, Morocco, 2023, pp. 1-5, doi: 10.1109/IRASET57153.2023.10152995.

[9] M. Karnati, A. Seal, D. Bhattacharjee, A. Yazidi and O. Krejcar, "Understanding Deep Learning Techniques for Recognition of Human Emotions Using Facial Expressions: A Comprehensive Survey," in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-31, 2023, Art no. 5006631, doi: 10.1109/TIM.2023.3243661.

[10] C. Rajyalakshmi, K. LakshmiNadh and M. S. Reddy, "Deep Learning Model for Emotion Prediction from Speech, Facial Expression and Videos," 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2023, pp. 1084-1088, doi: 10.1109/ICSSIT55814.2023.10061135.

[11] J. L. Joseph and S. P. Mathew, "Facial Expression Recognition for the Blind Using Deep Learning," 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), Kuala Lumpur, Malaysia, 2021, pp. 1-5, doi: 10.1109/GUCON50781.2021.9574035.

[12] W. S. S. Khine, P. Siritanawan and K. Kotani, "Facial Expression Features Analysis With Transfer Learning," 2022 14th International Conference on Knowledge and Systems Engineering (KSE), Nha Trang, Vietnam, 2022, pp. 1-6, doi: 10.1109/KSE56063.2022.9953754.

[13] M. Bende, M. Khandelwal, D. Borgaonkar and P. Khobragade, "VISMA: A Machine Learning Approach to Image Manipulation," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112168.

[14] W. Deng, J. Hu, S. Zhang and J. Guo, "DeepEmo: Real-world facial expression analysis via deep learning," 2015 Visual Communications and Image Processing (VCIP), Singapore, 2015, pp. 1-4, doi: 10.1109/VCIP.2015.7457876.

[15] Y. Yang and Y. Sun, "Facial Expression Recognition Based on Arousal-Valence Emotion Model and Deep Learning Method," 2017 International Conference on Computer Technology, Electronics and Communication (ICCTEC), Dalian, China, 2017, pp. 59-62, doi: 10.1109/ICCTEC.2017.00022.

[16] P. Demochkina and A. V. Savchenko, "Neural network model for video-based facial expression recognition in-the-wild on mobile devices," 2021 International Conference on Information Technology and Nanotechnology (ITNT), Samara, Russian Federation, 2021, pp. 1-5, doi: 10.1109/ITNT52450.2021.9649076.

[17] N. Begum and A. S. Mustafa, "Multiclass Artificial Neural Network based Facial Expression Recognition," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-6, doi: 10.1109/INOCON57975.2023.10101209.

[18] I.J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests", International conference on neural information processing, 2013.

[19] A. Mollahosseini, D. Chan and M.H. Mahoor, "Going deeper in facial expression recognition using deep neural networks", 2016 IEEE Winter conference on applications of computer vision (WACV), 2016.

[20] M.-A. Quinn, G. Sivesind and G. Reis, "Real-time emotion recognition from facial expressions", Standford University, 2017.

[21] V. Tumen, Ö. F. Söylemez and B. Ergen, "Facial emotion recognition on a dataset using convolutional neural network", 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), 2017.

[22] L. Zahara et al., "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi", 2020 Fifth International Conference on Informatics and Computing (ICIC), 2020.

[23] G.P. Kusuma and A.P.L. Jonathan, "Emotion Recognition on FER-2013 Face Images Using Fine-Tuned VGG-16", Advances in Science Technology and Engineering Systems Journal, vol. 5, pp. 315-322, 2020.

[24] Y. Khalid Bhatti et al., "Facial Expression Recognition of Instructor Using Deep Features and Extreme Learning Machine", Computational Intelligence and Neuroscience, 2021.