International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 07 | July 2024 www.irjet.net p-ISSN: 2395-0072

Study of Futuristic Trends for Real-Time Emotion Recognition

Roshani Atar[1],, Pratiksha Gharge[2]

[1] *Assistant Professor, Dept. of Computer Science, PVG's college of Science and Commerce, Pune, Maharashtra, India*
[2] *Assistant Professor, Dept. of Computer Science, Yashavantrao Chavan Institute of Science, Satara (Autonomous), Maharashtra, India*

Abstract - *Feelings are very powerful term in human life. Some people are unable to manage their reactions to their emotions. Also, people act in ways that seem completely out of their control. Both intrapersonal and interpersonal values of emotion which refers to the role those emotions play in maintaining social order within a society. Different models having emotion analysis is a technology that includes modalities like audio and visual data. Different fusion techniques and the classification algorithms applied are influenced by the type of audio, visual and textual features used in the recognition. Emotions are expressed through various modalities such as text, speech, facial expressions, and physiological signals. Combining multiple emotional recognition modalities in real time is an exciting and challenging research problem.*

Key Words: Emotions, fusion techniques, classification algorithms, emotional recognition, modalities

1. INTRODUCTION

As a baby, emotions are how you learn to communicate, even before you can speak. Although intense emotions can seem devastating, without them, life can seem cloudy, cloudy and empty. We cannot imagine life without our feelings. Positive as well as negative feelings are important, such as sadness when a loved one dies, anger when frustration, fear that overcomes us in a scary or unfamiliar situation, or guilt or shame toward others. Emotions color life experiences and give these experiences meaning and flavor. In fact, people who know how to perceive what they are feeling and who can calm down or adjust their behavior are more likely to succeed in life, have healthy relationships, and handle difficulties and setbacks. Once you calmed down, you realized that your emotions were getting the better of you and you wish you could handle them differently. We have all behaved this way, but as offspring grow, it's important that they learn to control these emotions in order to better handle situations. We need to understand the function of emotions, and this module does that below by dividing the discussion into three sections. The first concerns the intrapersonal functions of emotion, which refers to the role that emotions play in each of us individually. The second concerns the interpersonal functions of emotion, which refers to the role that emotions play between individuals within a group. The role of emotions plays in maintaining social order within humanity with other way of concerns the social and cultural functions of emotion. Overall, we will see that emotions inform us about who we are, what our relationships with others are like, and how to behave in social interactions. Emotions give meaning to events, without emotions, these events would be mere facts. Emotions help coordinate interpersonal relationships and emotions play an important role in the cultural functioning of maintaining the cohesion of human societies.

1.1 Purposes of emotions

Emotions are vital in our lives because they fulfill important purpose. The section on intrapersonal functions of emotions explains how emotions affect us individually. The section on the interpersonal functions discusses how emotions impact our relationships with others. The section on the social and cultural functions of emotion describes how emotions help, maintain and support our societies and cultures.

A. Intrapersonal purposes of emotion: Emotions help us act quickly with minimal awareness, emotions prepare the body for immediate action, emotions influence thoughts and emotions inspire future performance.

B. Interpersonal purposes of emotion: Emotions are expressed verbally through words and nonverbally through facial expressions, voices, gestures, body postures and movements, emotional expressions facilitate specific behaviors in perceivers, emotional terminologies signal the nature of interpersonal relationships, emotional expressions prompt desired social performance.

1.2 Applications

Multimodal sentiment analysis can be applied in the development of different forms of recommender systems, such as in the analysis of user-generated videos of movie reviews and reviews product generalities, to predict customer sentiments and create product or service recommendation. Multimodal sentiment analysis plays an important role in the advancement of virtual assistants through the application of natural language processing (NLP) and machine learning techniques. In healthcare, multimodal sentiment analysis can be used to detect certain medical conditions such as nervousness, pressure or unhappiness. Multimodal sentiment analysis can also be applied to understand the sentiments contained in video news programs, which is considered a complex and difficult

area because the sentiments expressed by journalists tend to be less obvious or neutral.

2. RELATED WORK

Understanding and reacting to how others feels is crucial for everyday interactions and relationships because social communication plays a key role in our daily lives. Grasping the subject's emotions through their eyes is particularly useful with portrait photography. A great way to ensure you get the results you want when you can severely change the entrance of your image.

In [1] proposed an improved hybrid approach for real time emotion recognition. Their method combines various classification techniques to enhance the performance of emotion recognition systems. They used dataset consisting of images and data that captured different facial expression and emotions.

Opinion mining of artificial intelligence is a natural language processing (NLP) technique that identifies important information patterns and characteristics from a broad quantity of text. It examines the thought, attitude, views, opinions, beliefs, comments, requests, questions and preferences expressed by an author on the basis of emotion rather than reason in the form of text towards entities such as services, issues, individuals, products, events, topics, organizations and their attributes. It finds the author's overall emotion for a text where the text can be blogging posts, product reviews, online forums, speeches, database sources, social media data and documents.

Author[2] proposed an algorithm for facial expression detection along with FER 2013 dataset which indicated that Haar Cascades and convolutional neural network (CNN) algorithm useful in real world application where understood seven types of human emotions are crucial. The researchers [3] were developed a system using CNN to accurately classified the emotions from facial expressions in real time. The system showed high accuracy rates and could be integrated into human computer interaction, security and entertainment systems.

In [4] highlighted that using transformers with multimodal electronic health record (EHR) data can effectively predict depression. Their method shows better identifying and managing depression in the patients through advanced machine learning techniques.

The employed variety of dataset such as Sentiment140 (dataset with tweets annotated for positive, negative and neutral), IMDB reviews (movie related sentiment), SST (Stanford Sentiment Treebank), Emotion datasets like the ISEAR (International Survey on Emotion Antecedents and Reactions) to analyze sentiment and detect the emotion from text. They were concluded that machine learning and deep learning approaches have made advances in sentiment analysis and emotion recognition[5].

Author[6] studied on real time facial emotion recognition demonstrates a machine learning system that accurately distinguishes facial emotions and demonstrated its potential in human computer interaction and behavioral analysis. They were used FER 2013 and CK+ (Cohn-Kanade) database in research for emotion detection.

Author[7] provided comprehensive overview of sentiment analysis which covered various methods, application, challenges and trends. Multimodal sentiment analysis is a traditional text-based sentiment analysis technology that includes modalities such as audio and visual data. It can be bimodal with the large amount of social media data available online in different forms such as videos and images.

In [8] developed an image and video based application using CNN algorithm, which was tested through image and video streaming with face inputs. Predictable text-based sentiment analysis has evolved into more complex multimodal sentiment analysis models which can be applied to the development of virtual assistants, analysis of social media reviews, analysis of news videos, and emotion detection such as unhappiness monitoring among others.

The gathered data from variety of sources, including public archives and research programs to provide varied range of emotional expression and circumstances for thorough analysis[9].

Author[10] utilized a dataset of facial photos annotated with various emotions for real time emotion recognition training and testing. They were developed system using image processing and machine learning to accurately recognize and classify the emotions from facial expression in real time.

The difficulty of analyzing textual, audio, and visual features to perform such a task requires the application of different fusion techniques, such as feature-level, decision-level, and hybrid fusion. The performance of these fusion techniques and the classification algorithms applied are influenced by the type of textual, audio and visual features used in the analysis.

3. ARCHITECTURE OF REAL TIME EMOTION RECOGNITION

General architecture of real time emotion recognition involves multistep process. The below workflow of real time emotion detection usually consists of several key components. First, data is collected by employing cameras to record face photos or video frames. To maintain consistency, the collected photos are preprocessed with stages such as face detection, alignment and normalization. Following that feature extraction techniques are used to create meaningful representation of face emotions from photos. These

characteristics are loaded into a pre-trained machine learning model, which often employs deep learning architecture such as CNN, to categorize the emotion presented in face photos. Finally, the identified emotions are displayed for additional processing in real time applications, ensuring that the entire process runs smoothly and efficiently.

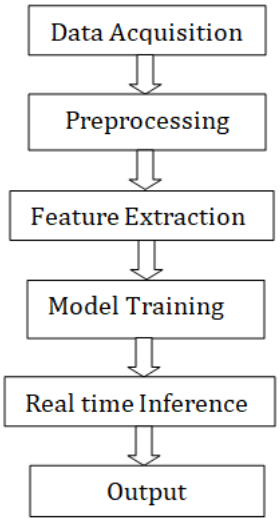

Fig -1: General Architecture of Real Time Emotion Recognition

4. TECHNIQUES FOR EMOTION CLASSIFICATION

Sentiment Analysis, which includes extracting emotions or sentiments from text, has been extensively researched in the field of natural language processing. Though, most existing methods mainly focus on examining textual data. Many researchers were explored it using machine learning techniques to combine data from text, audio, images, and physiological signals to create a robust multimodal sentiment analysis system. The system is expected to be able to accurately recognize and understand expressed emotions in real time, enabling applications in the areas of human-computer interaction, affective computing, mental health assessment, etc.

There are many methodologies to detect real time emotion features. Emotion detection also known as face expression identification, is crucial field in artificial intelligence and computer vision that involves recognizing and understanding human emotions through facial expressions. CNN, like VGG, ResNet and inception, learn hierarchical features from structured data, aiding in computer vision applications like image recognition and object detection, addressing specific challenges. Combining various models, such as RNNs and CNNs or employing both standard and deep learning-based techniques, can typically results in superior performance in real-time sentiment analysis. Transformer based models, such as BERT, GPT and their derivatives, have demonstrated cutting edge performance in variety of natural language applications, including sentiment analysis[11]. Transformer variations might be more suited for real-time applications. Fuzzy classifications like FLEC classifier and oscillation algorithm useful for real time sentiment detection.

Here are some techniques are as follows-

A. Structures**:** Feature engineering involves the selection of features that are fed into machine learning algorithms plays a key role in sentiment classification performance. In traditional sentiment analysis, one of the most fundamental tasks of multimodal sentiment analysis is sentiment sorting, which categorizes different sentiments such as positive, negative and neutral.

B. Pictorial features: Examining text-only videos is the presence of rich sentiment indications in the visual data. Visual features include facial expressions, which are of utmost importance in capturing feelings and emotions, the main channel for shaping a person's current state of mind. Smiling is considered one of the most predictive visual cues in multimodal sentiment analysis. OpenFace is an open-source facial analysis toolkit available to extract and understand these visual features.

C. Fusion procedures: Unlike traditional textual sentiment analysis, multimodal sentiment analysis undergoes a fusion process in which data from different modalities (textual, audio, or visual) are merged and analyzed together. Sentiment classification performance depends on the type of technique used.

- Early fusion: Feature-level fusion folds all the features from each modality like text, audio, or visual and combines them into a single feature vector, which is finally fed into a classification algorithm. One of the difficulties in implementing this technique is the integration of mixed characteristics.

- Late fusion: Decision-level fusion independently feeds data from each modality (text, audio, or visual) into its own classification algorithm and obtains the final sentiment classification results by merging each result into a single sentiment. Fusion technique is that it eliminates the need to merge heterogeneous data and each modality can use its most appropriate classification algorithm.

- Hybrid Fusion: Hybrid fusion is a combination of early fusion and late fusion techniques, which exploits complementary information from both methods during the classification process.

5. FUTURE SCOPE

Real time emotion recognition has broad future scope. Here are some spanning several domains and applications with possible areas of study and application being investigated.

1. Advances in natural language processing and computer vision.
2. Multimodal emotion analysis for real-time mental health assessment.
3. Criminal linguistic deception detection model Offensive language detection.
4. A human robot sensitive to emotions.

6. CONCLUSION

We have emotions which are a crucially important aspect of our psychological makeup, having meaning and function for each of us individually, for our relationships with others in groups and for our societies as a whole. From this paper we understand the importance of emotion in human life. Just as your own emotions provide valuable information to others, the emotional expressions of those around you also provide. This paper provides valuable insights into human emotions and interactions with technology. In today's era, various models combining multimodal information data for sentiment analysis can effectively improve the accuracy of sentiment analysis. In future, we can build a multimodal emotion recognition model that can complement hidden emotions well in real time.

REFERENCES

[1] C. Loconsole, D. Chiaradia, V. Bevilacqua, and A. Frisoli, "Real-time emotion recognition: An improved hybrid approach for classification performance," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8588 LNCS, pp. 320–331, 2014, doi: 10.1007/978-3-319-09333-8_35.

[2] M. O. Kolhe Durgesh, M. S. Metkar Sameer, and A. Amarja, "Emotion detection," vol. 1, no. 1, pp. 1–8, 2017, doi: 10.1145/3123818.3123852.

[3] N. Jaymon, S. Nagdeote, A. Yadav, and R. Rodrigues, "Real time emotion detection using deep learning," *Proc. 2021 1st Int. Conf. Adv. Electr. Comput. Commun. Sustain. Technol. ICAECT 2021*, 2021, doi: 10.1109/ICAECT49130.2021.9392584.

[4] Y. Meng, W. Speier, M. K. Ong, and C. W. Arnold, "Bidirectional Representation Learning from Transformers Using Multimodal Electronic Health Record Data to Predict Depression," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 8, pp. 3121–3129, 2021, doi: 10.1109/JBHI.2021.3063721.

[5] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Soc. Netw. Anal. Min.*, vol. 11, no. 1, pp. 1–19, 2021, doi: 10.1007/s13278-021-00776-6.

[6] D. Shah, K. Chavan, S. Shah, and P. Kanani, "Real-Time Facial Emotion Recognition," *2021 2nd Glob. Conf. Adv. Technol. GCAT 2021*, no. May 2022, pp. 1–4, 2021, doi: 10.1109/GCAT52182.2021.9587707.

[7] M. Wankhade, A. C. S. Rao, and C. Kulkarni, *A survey on sentiment analysis methods, applications, and challenges*, vol. 55, no. 7. Springer Netherlands, 2022. doi: 10.1007/s10462-022-10144-1.

[8] L. SIVAPRASAD, G. E. GIRIBABU, V. HARSHITHA, C. P. SAI, and P. T. KUMAR, "REAL-TIME FACIAL EXPRESSION RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK," *Int. J. Creat. Res. Thoughts(IJCRT)*, vol. 11, no. 4, pp. 550–564, 2023.

[9] A. Singh, M. Shoaib, S. Singh, and S. Sharma, "Analysis of Real-Time Emotion Detection Techniques," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4623338.

[10] Prof. Bina Rewatkar *et al.*, "Research On Real Time Emotion Recognition Using Digital Image Processing Using ML," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 376–382, 2024, doi: 10.48175/ijarsct-17052.

[11] D. Liu and T. Miller, "Federated pretraining and fine tuning of BERT using clinical notes from multiple silos," 2020, [Online]. Available: http://arxiv.org/abs/2002.08562