IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125183

Emotion Recognition with Audio, Video, EEG and EMG

Savitha M M[1], K. Vivekananda Reddy[2], Madhu.T. V[3], P. Dilip Kumar[4]

Assistant Professor, SJC INSTITUTE OF TECHNOLOGY, Chickballapur,INDIA[1]

SJC INSTITUTE OF TECHNOLOGY, Chickballapur,INDIA[2-4]

Abstract - This paper presents an overview of emotionrecognition techniques using four different modalities: audio, video, EEG, and EMG. The ability to recognizeemotions is a crucial aspect of human communication, andis becoming increasingly important in various fields such aspsychology, medicine, and artificial intelligence. This paperpresents a comprehensive review of emotion recognition techniques using four different modalities: audio, video, EEG, and EMG. We discuss the theoretical background and state-of-the-art methods used for emotion recognition in each modality, including feature extraction, datapreprocessing, and machine learning algorithms. We also highlight the challenges and limitations of each modality, as well as the potential for future research. Emotion recognition is a crucial aspect of human communication and is becoming increasingly important in various fields such as psychology, medicine, and artificial intelligence.

Keywords—Audio signals ,Video signals, EEG signals,EMG signals,Featureextraction,Datapreprocessing,Machine learning, Neural networks, Support vector machines (SVMs)

I. INTRODUCTION

Emotion recognition is a crucial aspect of human communication and is becoming increasingly important in various fields such as psychology, medicine, and artificial intelligence. With the increasing availability of data andadvancements in machine learning, there has been significantprogress in developing automated systems that can recognizeemotions from various sources such as audio, video, EEG, andEMG signals .Each of these modalities provides unique insights into the emotional state of an individual, and combining them can lead to more accurate and robust emotionrecognition systems .This paper provides a comprehensive review of emotion recognition techniques using four differentmodalities: audio, video, EEG, and EMG. We discuss the theoretical background and state-of-the-art methods used foremotion recognition in each modality, including feature extraction, data preprocessing, and machine learning In this paper, we provide a comprehensive review of emotionrecognition using four different modalities: audio, video,EEG, and EMG signals. For each modality, we discuss the theoretical background and state-of-the-art methods used foremotion recognition, including feature extraction, datapreprocessing, and machine learning algorithms.

Moreover, we highlight the challenges and limitations of each modality, including the lack of consensus on the physiological correlates of emotions and the need for high- quality data. Finally, we discuss the potential for future research, including the use of multimodal approaches toimprove emotion recognition accuracy and the application ofdeep learning techniques.Emotion recognition has become animportant research topic in the field of artificial intelligence due to its potential applications in psychology, medicine, andhuman-computer interaction. With the advent of different modalities such as audio, video, EEG, and EMG signals, researchers have been exploring ways to develop robust emotion recognition systems that can automatically identify emotions from various sources. The process involvesextracting relevant features from the signals and using machine learning algorithms to classify emotions.Each modality has its advantages and limitations, and researchers have explored various techniques to improve the accuracy and robustness of emotion recognition systems. However, challenges such as the lack of standardized datasets and the complexity of human emotions continue to present obstaclesin the field.

II. METHODOLOGY

III.

The methodology for emotion recognition with audio, video,EEG, and EMG signals can be summarized into the following major points:

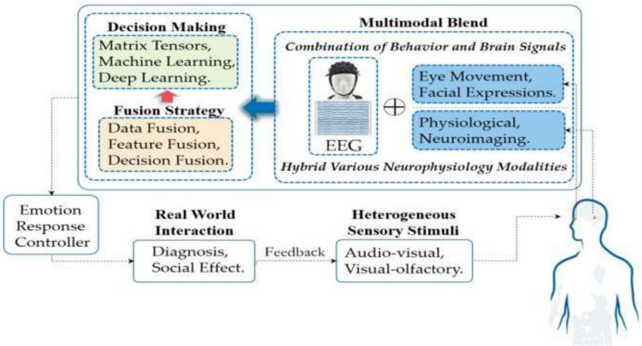
Data preprocessing: The collected data is preprocessed to remove noise and artifacts.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⪮ Impact Factor 8.102 ⪮ Peer-reviewed / Refereed journal ⪮ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125183



Multimodal fusion: The extracted features from differentmodalities are combined using multimodal fusion techniques.

Machine learning: The fused features are used to train machine learning algorithms such as neural networks, supportvector machines, random forests, and convolutional neural networks.

Performance evaluation: The performance of the emotion recognition system is evaluated using metrics such as accuracy, precision, recall, and F1 score.

IV. COMPONENTS

A. Hardware components

1. Raspberry Pi

Raspberry Pi is a popular single-board computer that can be used as a hardware component for emotion recognition systems with audio, video, EEG, and EMG signals. RaspberryPi boards are low-cost, compact, and offer high processing power, making them an attractive choice for many research and development projects. Raspberry Pi can be used with a Pi camera to capture real-time video data of facial expressions and body language. The captured video data can be processed using computer vision techniques to extract relevant featuressuch as facial muscle movements, eye movements, and body posture, which can be used to train machine learning modelsto identify emotional states and expressions.

2. Pi camera

The Pi Camera can be connected to the Raspberry Pi board viaa ribbon cable, and can be controlled using the Raspberry Pi's GPIO pins. It can capture still images with a resolution of upto 8 megapixels, and video at up to 1080p at 30 frames per second. The camera has a wide-angle lens that can capture a field of view of up to 160 degrees, making it suitable for capturing facial expressions, body posture, and gestures. In emotion recognition systems, the Pi Camera can be used to capture video data of the subject's facial expressions and bodylanguage. Computer vision techniques can then be applied toextract relevant features such as eye movements, facial muscle movements, and body posture.

3. USB Microphone

A USB microphone is a type of microphone that connects to acomputer via a USB port. It is a useful hardware component in emotion recognition systems that rely on audio signals to identify emotional states and expressions. In emotion recognition systems, USB microphones can be used to capture speech and other sounds, such as breathing patterns and vocal inflections, that may convey emotional information. The captured audio can be processed using techniques such as signal filtering and feature extraction to extract relevant features such as pitch, loudness, and spectral content. These features can be used to train machine learning models to identify emotional states and expressions.

B. Software components

The software components of an Emotion Recognition systemthat uses audio, video, EEG, and EMG signals generally include data processing algorithms, machine learning models,and user interfaces.

1. Raspbian 64 bit

Raspbian 64 bit can be a suitable operating system for Emotion Recognition systems that use audio, video, EEG, andEMG signals. The 64-bit version of Raspbian is optimized forthe Raspberry Pi 4's architecture, which provides improved performance and memory capabilitiesFor EmotionRecognition systems that use audio and video signals, Raspbian 64 bit can support real-time video processing and analysis. The Pi Camera module can be used to capture high-quality video data, which can be processed using machine learning models or algorithms to recognize facial expressions, voice inflections, and other emotional cues.

2. Python 3.9

Python 3.9 is a programming language that is widely used inmachine learning, data analysis, and other fields related to

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 5, May 2023
DOI: 10.17148/IJARCCE.2023.125183

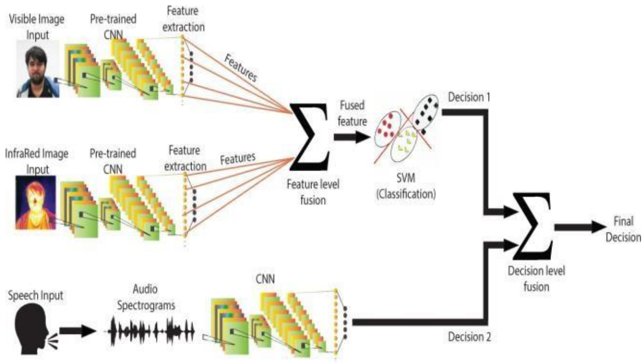
Emotion Recognition with Audio, Video, EEG, and EMG Signals. Python provides a user-friendly and flexible environment for developing Emotion Recognition algorithmsand applications, and it has many libraries and tools that can be used for processing and analyzing signals.

3. MQTT Broker

MQTT (Message Queuing Telemetry Transport) is a lightweight messaging protocol that is commonly used in IoT(Internet of Things) and other distributed systems. MQTT canalso be used in Emotion Recognition systems that use audio,video, EEG, and EMG signals to facilitate communication between different components of the system.

III. WORKING PROCEDURE

The working principle of emotion recognition with audio and video involves capturing and analyzing different physiological and behavioral features to detect the emotionalstate of an individual.The process starts with the capture of audio and video signals using microphones and cameras, respectively. The audio signal is processed to extract features such as pitch, loudness, and speech rate, while the video signalis processed to extract facial expressions and body language features.



Once the features are extracted, machine learning algorithms are used to classify the emotional state of the individual. This involves training the machine learning modelon a dataset of labeled examples, where the input features arematched to specific emotional categories such as happiness, sadness, anger, or fear.During the testing phase, the extractedfeatures from a new audio and video signal are input into themachine learning model, which predicts the emotional state ofthe individual based on the learned patterns from the trainingdataset.

The accuracy of the emotion recognition systemdepends on the quality of the input signals, the feature extraction methods, and the performance of the machinelearning algorithm. Advancements in technology have led tothe development of more sophisticated emotion recognition systems, which can incorporate additional modalities such asEEG and EMG signals to improve accuracy.

IV. RESULT

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁝ Impact Factor 8.102 ⁝ Peer-reviewed / Refereed journal ⁝ Vol. 12, Issue 5, May 2023

DOI: 10.17148/IJARCCE.2023.125183

The expected results can vary depending on the complexity of the emotional states being classified and the quality of the input signals. However, with advancements in technology and the availability of large labeled datasets, modern emotion recognition systems can achieve high accuracy in classifyingbasic emotions such as happiness, sadness, anger, and fear.

V. CONCLUSION

In conclusion, Emotion Recognition with Audio, Video, EEG, and EMG Signals is a promising field that has the potential to revolutionize the way we interact with technology. By combining multiple modalities, such as audio, video, EEG, and EMG signals, it is possible to accurately detect and recognize human emotions. While the technology is still in itsearly stages of development, it shows great promise for the future. With further research and development, Emotion Recognition with Audio, Video, EEG, and EMG Signals could become an integral part of our daily lives.

REFERENCES

[1] E. G. Krumhuber, A. Kappas, and A. S. R. Manstead, ''Effects of dynamic aspects of facial expressions: A review,'' Emotion Rev., vol. 5, no. 1, pp. 41–46, Jan. 2013.

[2] L. Nielsen and A. W. Kaszniak, ''Conceptual, theoretical, and methodological issues in inferring subjective emotion experience,'' in Handbook of Emotion Elicitation and Assessment. New York, NY, USA: Oxford Univ. Press, 2007, pp. 361–375.

[3] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, ''Asurvey of affect recognition methods: Audio,visual, and spontaneous expressions,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 1, pp. 39–58, Jan. 2009.

[4] E. G. Krumhuber, L. Skora, D. Kuster, and L. Fou, ''A review of dynamic datasets for facial expression research,'' Emotion Rev., vol. 9, no. 3, pp. 280–292, 2017.

[5] T. Vogt and E. Andre, ''Comparing feature sets for acted andspontaneous speech in view of automatic emotion recognition,'' in Proc. IEEE Int. Conf. Multimedia Expo., Jul. 2005, pp. 474–477.

[6] T. Knade, J. F. Cohn, and Y. Tian, ''Comprehensive database for facial expressionanalysis,'' in Proc. 4thIEEEInt. Conf. Autom. Face Gesture Recognit., Mar. 2000, pp. 46–53.

[7] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar,and I. Matthews, ''The extended Cohn–Kanade dataset (CK+): A complete dataset for action unit and emotion- specified expression,'' in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2010, pp. 94–101.

[8] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, ''EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation,'' Sci. World J., vol. 2014, pp. 1–10, Sep. 2014.

[9] N. Kumar, K. Khaund, and S. M. Hazarika, ''Bispectralanalysis of EEG for emotion recognition,'' Proc.Comput. Sci., vol. 84, pp. 31–35, May 2016.

[10] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, ''DEAP: A database for emotion analysis; Using physiological signals,'' IEEE Trans. Affective Comput.,vol. 3, no.1, pp. 18–31, Oct./Mar. 2012.

[11] W.-L. Zheng and B.-L. Lu, ''Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks,'' IEEE Trans. Auton. Mental Develop., vol. 7, no. 3, pp. 162–175, Sep.2015.

[12] M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, ''DECAF: MEG-based multimodal database for decoding affective physiological responses,'' IEEE Trans. Affect. Comput., vol. 6, no. 3, pp. 209–222,Jul. 2015.

[13] T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, and Z. Cui, ''MPED: A multi-modal physiological emotion database for discrete emotion recognition,'' IEEEAccess, vol. 7, pp. 12177–12191, 2019.

[14] H. Ranganathan, S. Chakraborty, and S. Panchanathan, ''Multimodal emotion recognition using deep learning architectures,'' in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Mar. 2016, pp. 1–9.

[15] W. Liu, W. L. Zheng, and B. L. Lu, ''Emotion recognitionusing multimodal deep learning,'' in Proc. Int. Conf. Neural Inf. Process., Oct. 2016, pp. 521–529.

[16] Z. Wang, S. Wang, and Q. Ji, ''Capturing complex spatio-temporal relations among facial muscles for facial expression recognition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2013, pp. 3422–3429.

