Journal of Theoretical and Applied Information Technology

30th November 2023. Vol.101. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A DETAILED REVIEW ON MULTI-MODALITY BASED EMOTION DETECTION WITH PHYSIOLOGICAL SIGNALS USING DEEP LEARNING

[1]T L DEEPIKA ROY, [2]D NAGA MALLESWARI

[1]Assistant Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Green Fileds, Vaddeswaram, A.P. – 522302, India.

[2]Associate Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Green Fileds, Vaddeswaram, A.P. – 522302, India.

E-mail: [1]thotadeepika001@gmail.com, [2] nagamalleswary@kluniversity.in

ABSTRACT

In several fields, such as healthcare, human-computer interaction, and affective computing, emotion recognition has attracted a lot of attention. This review paper investigates the new area of multi-modality emotion detection, which uses deep learning methods and physiological signals to improve the precision and resilience of emotion identification. We give a summary of the basic ideas, approaches, difficulties, and possible uses in this multidisciplinary field. Additionally, we examine the difficulties and constraints encountered in multi-modal emotion detection, encompassing the gathering of data, the extraction of features, and the interpretability of deep learning models. We highlight trends and new directions in the field in this study, including real-time applications, cross-modal emotion recognition, and transfer learning. A summary of the state of the art is provided in the paper's conclusion, highlighting the potential significance and real-world applications of multi-modality-based emotion detection with physiological signals through deep learning. Future directions for research in this exciting and rapidly developing field are also outlined.

Keywords: *EEG, ECG, FACS, HRV, AU…*

1. INTRODUCTION

The core of the human experience is emotion, which weaves a complicated web of beliefs, deeds, and interpersonal relationships. With the increasing integration of technology in our daily lives, comprehending and precisely identifying these emotions has become a top priority for scientific and technical progress. The ability to identify and understand emotions is a transformational and transcendent objective that has applications in healthcare, human-computer connection, and entertainment. Combining multi-modal data—which focuses on physiological signals—with deep learning's computational capabilities is a promising new direction in this direction.

With so many uses, emotion identification has evolved from a straightforward facial expression analysis task to a multi-dimensional, multi-modal one. The idea that facial expressions are the only way to convey emotions has been replaced by a more nuanced one. It recognizes that emotions are complex, dynamic states impacted by a wide range of behavioral and physiological circumstances rather than distinct, static entities. These variables go beyond outward manifestations and include the body's subtle reactions, both internal and external.

This paper delves into the newly-emerging field of multi-modal emotion detection, which uses deep learning techniques and a variety of data sources to deliver an enhanced, context-aware knowledge of human emotions. An integrated framework for achieving higher

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

precision and adaptability in emotion recognition is created by integrating physiological information with deep learning models' computational sophistication.

Setting the scene, the first chapter describes the extent and importance of multi-modal emotion detection in modern culture. It sheds light on the always changing landscape of how emotions affect our ability to make decisions, engage with technology, and maintain our wellbeing. The introduction establishes the profound implications and transformative potential of multi-modal emotion detection by examining the critical domains where emotion recognition holds paramount importance, from human-computer interaction and immersive gaming experiences to mental health assessment and therapeutic interventions.

At this nexus of technology, biology, and psychology, the combination of deep learning techniques and physiological signals is nothing short of revolutionary. Electroencephalography (EEG) offers a wealth of information about cognitive and emotional states by giving a window into the electric dance of brain activity. The rhythm and variability of the heartbeat are monitored by electrocardiography (ECG), which provides information on stress levels and emotional reactions. In the meanwhile, facial expressions continue to be an effective means of identifying a range of emotions, which are frequently subtly and expressively conveyed in the human face.

The pursuit of multi-modal emotion detection raises a number of complexities and difficulties. It takes accuracy and collaboration to gather data from these various sources, each with distinct temporal features and processing needs. However, these data modalities go beyond individual emotions and physiological reactions to provide a more comprehensive picture of the emotional landscape once they are synchronized and coordinated.

This review paper's introduction emphasizes the critical role that data preparation plays in enabling a coherent, trustworthy analysis. It does this by highlighting the necessity of noise removal, data standardization, and the management of various physiological signals. In order to transform raw physiological data into a format that deep learning models can understand, several preparatory procedures are crucial. In this transformation, the key to bridging the gap between data capture and machine learning's predictive power is the selection and extraction of relevant characteristics.

For emotion detection to undergo a paradigm change, deep learning models are essential. We explore the architectural nuances of recurrent neural networks (RNNs), multi-input neural networks, and the use of pre-trained models such as ResNet and VGG in this paper. These neural architectures offer the framework needed to combine data from several sources into a coherent representation that goes beyond the bounds of separate modalities.

The cornerstone of this multi-modal architecture is fusion techniques, which intertwine the lines of feature extraction and prediction into a seamless emotional understanding fabric. In order to provide a unique method of combining data from several modalities, this research examines the ideas of early fusion, late fusion, and mid-level fusion. The foundation for a thorough understanding of emotions is laid by the fusion techniques, which capture both the physiological reactions and the outward manifestations of emotion.

The introduction ends with a broad overview of the many uses of multi-modal emotion detection, ranging from improving gaming and human-computer interactions to transforming healthcare diagnosis and therapies. It acknowledges the ethical issues that are still raised by data collection and use, as well as the exciting potential to increase the variety of physiological signals that are used to identify emotions. With its roots in physiological signals and its incorporation of deep learning techniques, multi-modal emotion detection is a dynamic, multidisciplinary field that has the potential to completely transform our knowledge of human emotions. It offers a persistent invitation to go

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

off on this life-changing adventure, driven by curiosity, morality, and the desire to discover the deep complexity of human emotions.

There are various sorts of emotion models, including evaluation, dimensional, category, and action tendency models. Whereas dimensional models contend that emotions are systematically related to one another, categorical models contend that there are a few fundamental emotions that are unaffected by race. Action tendency models accentuate the behavioral reactions linked to particular emotions, whereas appraisal models concentrate on the assessment of events and their significance to a person's objectives. Emotional states have been studied and analyzed using these models in affective computing research.

Known by many as the "basic emotion theory," the categorical theory of emotions is a psychological and philosophical viewpoint that asserts that individuals experience a limited number of unique, innate, and universal emotions that are shared by people of all cultural backgrounds. The vast array of emotional experiences that people go through is believed to be based on these fundamental emotions, which are believed to be biologically defined. According to the notion, these fundamental emotions are distinguished by particular and distinct psychological and physiological patterns. A psychologist named Robert Plutchik created a complete theory of emotions, which is commonly depicted as a wheel called "Plutchik's Wheel of Emotions." As seen in Fig. 1, Plutchik's model offers a more sophisticated understanding of emotions than the fundamental ones suggested by some other theories.

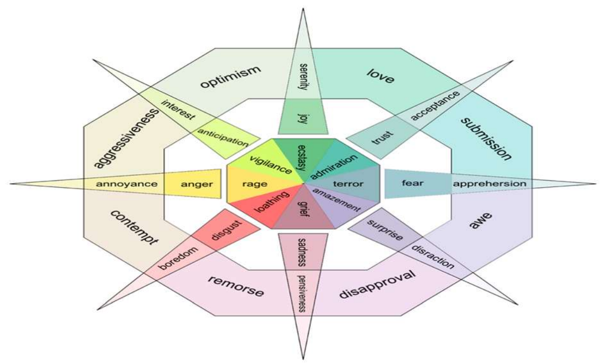


Figure 1: Plutchik's Wheel Of Emotions

Different kinds of emotions generated in Human Body: Human body generates different signals from different parts based on the input information those details are marked in the below diagram.

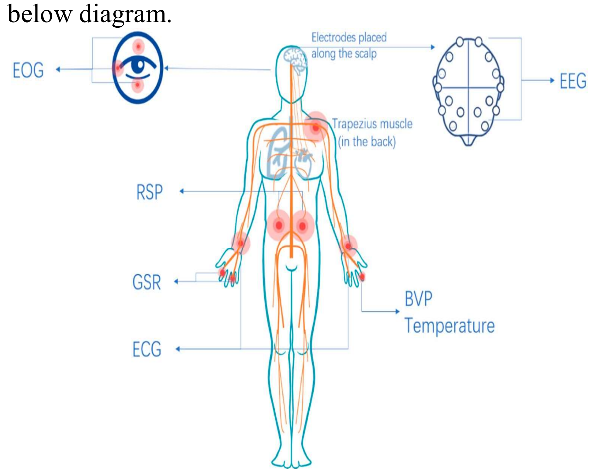
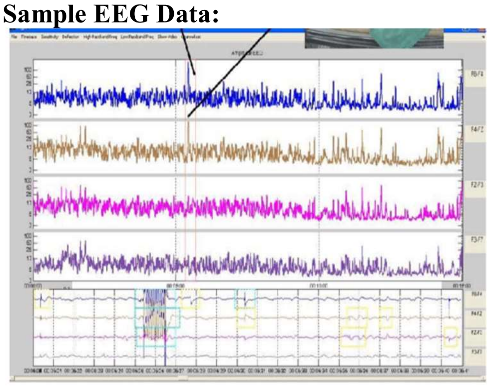


Figure 2: Human Body With Different Signal Sources

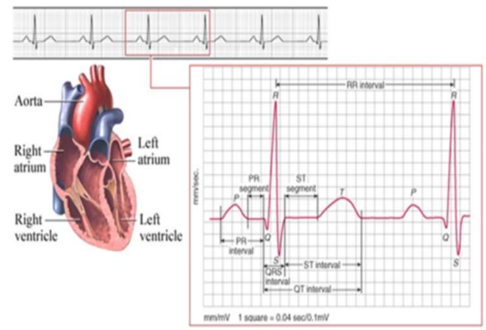
Journal of Theoretical and Applied Information Technology
30th November 2023. Vol.101. No 22

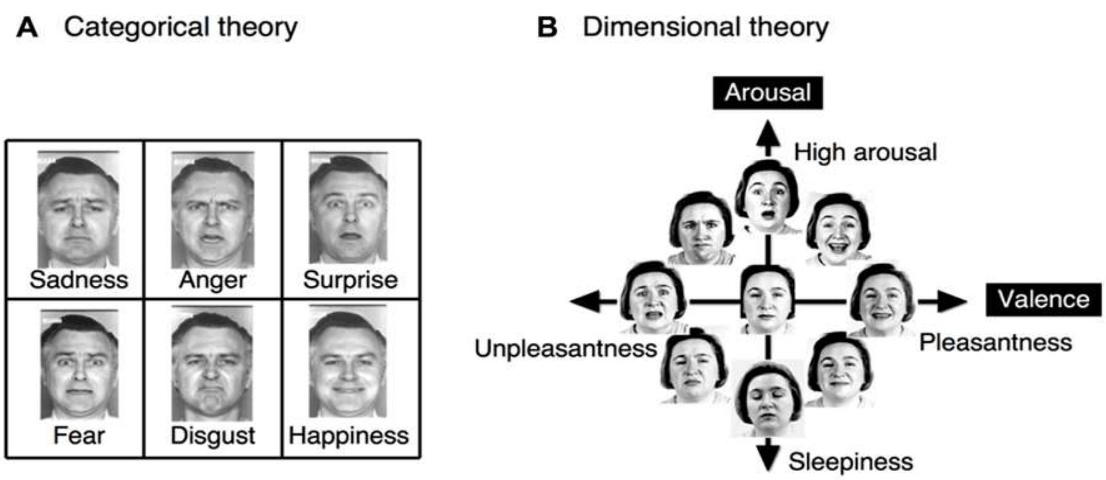
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Sample EEG Data:



Sample ECG Data:





2. DATA ACQUISITION AND PREPROCESSING

The first step in emotion detection, especially in multi-modal settings, is gathering a variety of data sources. These sources, such as bodily cues and facial expressions, offer a lot of data to reveal people's complex emotional states.

2.1 Physiological Data Acquisition

Physiological data collection frequently uses non-invasive sensors and equipment that can track a person's physiological reactions. Through electrodes affixed to the scalp, electroencephalography (EEG) offers a view into neural activity. It provides information on emotional and cognitive states by measuring the electrical potentials that the brain produces. In contrast, electrocardiography (ECG) measures the heart's electrical activity and records heart rate and heart rate variability (HRV), both of which are markers of emotional states. These sources of physiological data offer substantial, current information.

2.2 Facial Expression Data Acquisition

Expressions on the face are the most basic means of communicating emotions. High-resolution cameras or depth sensors are frequently used to record facial movements and expressions for acquiring facial expression data. Based on the Facial Action Coding System (FACS), which measures different facial muscle movements connected to particular emotions, this data can be divided into action units (AUs). Even though they are non-verbal, these expressions frequently include a richness of emotional information.

2.3 Data Synchronization and Alignment

Precision synchronization and alignment are essential for the integration of data from various sources, which is the foundation of multi-modal emotion identification. Make sure that all EEG and ECG recordings have the same timestamp when dealing with physiological data. Establishing temporal coherence for cross-modal analysis requires this alignment. Furthermore, for a thorough comprehension of the emotional context, matching the physiological data with the accompanying facial expressions becomes essential.

2.4 Noise Removal and Data Normalization

Meaningful emotional signals might be obscured by noise and artifacts present in raw physiological data. One essential element in ensuring data quality is effective preprocessing. Filtering and artifact rejection are two methods used to remove noise from pure physiological data. Baseline correction and amplitude scaling are two steps in the data normalization process that guarantee data consistency and inter-individual comparability.

2.5 Challenges in Data Acquisition and Preprocessing

Preprocessing and data gathering are not without difficulties. Despite the advancements in portability and user-friendliness of physiological sensors, there are still obstacles in guaranteeing subjects' comfort and cooperation during data collection. Careful data cleaning is required to remove motion artifacts, electrode impedance, and other noise sources. Furthermore, temporal complexities are introduced when aligning data from many modalities, necessitating the use of advanced algorithms for exact synchronization. In order to preserve data accuracy and reliability, data preparation is an essential procedure that requires rigorous methodologies and specialist subject expertise.

To put it simply, preprocessing and data collection set the stage for reliable multi-modal emotion detection. A window into the intricate interaction of emotions is provided by the integration of physiological data and facial expressions. Following data synchronization, cleaning, and normalization, feature extraction and deep learning. The remainder of the article is structured as follows: In section 3 we discuss Literature survey. Next, we describe research gaps in section 4. Finally, concluding remarks and future works are discussed in section 5.

3. LITERATURE SURVEY

[1] A hybrid CNN-LSTM model for multi-class emotion identification from EEG signals is proposed in this research. The DEAP dataset, a well-known benchmark dataset for emotion recognition, is used in the experiment. On the valence and arousal dimensions, the suggested model obtains high test accuracies of 96.87% and 97.31%, respectively. Additionally, the model reaches cutting-edge accuracy in the valence and arousal domains.

Three modalities are used in this research to identify emotions: facial photos, auditory indications, and text identification using Twitter tweets, RAVDESS, FER and CK+ databases. A CNN model that achieves 66.67% accuracy on the FER-2013 dataset and 98.4% accuracy on the CK+ dataset is used in the suggested method. [2] The research investigates several fusion strategies that produce varied outcomes when combining the modalities. By bridging the emotional gap between technology and humans, the research hopes to advance interactive systems development.

This research presents a hybridized model that combines visual and auditory signals to recognize speech emotions using a multimodal database. The model uses a deep neural network (NN) classifier that has been optimized to identify emotions from multimodal input data. [3] For audio-video signals, a hybrid texture feature descriptor is suggested in order to enhance classification results. The deep NN classifier's hyperparameters are adjusted using the learner memorizing optimization strategy, which improves accuracy. With the Enterface'05 database, the suggested method's performance is assessed, and high values for accuracy, sensitivity, and specificity are obtained.

[4] The study presents a deep neural network architecture for EEG emotion recognition that integrates graph signal processing methods for spatial domain smoothing with spatio-temporal encoding and recurrent attention network blocks. On the DEAP dataset, the suggested architecture produces emotion classification results that are at the cutting edge, and on the DREAMER and EEWD datasets, it shows promise in transfer learning for emotion classification tasks. By combining graph signal processing for spatial smoothing,

investigating transfer learning, investigating spatial information encoding in the brain, carrying out a more thorough comparison with cutting-edge techniques, and broadening the experimental analysis to three datasets, the paper builds on the preliminary work that was published in an earlier publication.

[5]In order to more accurately represent human emotional states, physiological signals are employed. In an effort to overcome the shortcomings of single-modal emotion recognition, multimodal physiological signal emotion recognition has drawn more attention. The suggested classification model takes into account the various distributions of numerous signals and efficiently captures complementary features, thereby mitigating the shortcomings of the original fusion models. From EEG signals and peripheral physiological signals (PPS) including ECG, EMG, and other physiological signals, the model extracts differential entropy properties. In order to extract frequency spatial dimension information, the features of the EEG signal are converted into a three-dimensional feature map and fed into a neural network. A long and short-term memory network is used to extract temporal information from peripheral physiological inputs. Features from the peripheral physiological signal and the EEG are combined and fed into a

[6] Emotion identification based on Electroencephalogram(EEG) data has becoming popular with the majority of research being undertaken in laboratory settings using medical grade equipment. The (GA) algorithms able to find out subset of EEG signals for emotion categorization is proposed in this study. In this work GA algorithms are used to analyse publicly available EEG data with 2548 features reduced to 49 features Emotion identification based on EEG data using KNN,RF,ANN for effective classification[7].

[8] Using LSTM-based models, the research provides an integrated IoT framework for wireless exchange of physiological inputs for emotion recognition.During pandemics, the suggested architecture allows for real-time communication and emotion identification, as well as health monitoring and distant learning support.The suggested IoT protocols achieve ultralow latency of 1 ms and better dependability when compared to the state of the art.In emotion recognition, the suggested deep learning strategy obtains a high performance (f-score) of 95%.The study emphasizes the relevance of employing physiological signals for emotion recognition because they are more dependable and wearable than EEG signals.In the IoT framework, the proposed cloud server-based data processing system allows for the gathering and analysis of physiological signals for emotion recognition.

[9] One potential disadvantage of the publication "Feature extraction based on microstate sequences for EEG- emotion recognition" is the dependency on the DEAP dataset for evaluating the proposed features. The results may be limited in their applicability to other datasets or real-world circumstances." adaptation in EEG emotion recognition. However, reliance on certain benchmark databases (SEED, SEED-IV, and DEAP) for evaluating performance could be a possible problem. The approach's generalizability to other datasets or real-world circumstances may require more examination. "

[10] The study assessed the classification accuracy of multiple EEG feature sets for identifying emotional states in a comprehensive analysis of feature extraction methods for Emotion Recognition from Multichannel EEG Recordings. The study looked at five different feature sets: According to the findings, FD characteristics obtained from EEG data are reliable for emotion recognition and might potentially be used in the creation of a real-time online EEG-based emotion recognition system. This study compared the performance of several feature sets and algorithms.

[11] The research offers a multidimensional feature extraction method for emotion identification from EEG signals based on MSTBN and EEMD-WPT.[12] The research provides four feature extraction strategies for emotion detection on EEG data (STFT, Wavelet Entropy, Hjorth, and Statistical features). [13] In this study, many approaches for effectively extracting characteristics from EEG data are proposed. The filtered data assisted in achieving high accuracy while retaining the substance of the data.

[14] This research offers an IoT and cloud-based healthcare system for extracting

ECG signal parameters and doing machine learning-based analysis to assess the risk of cardiac disease. [15] The suggested EEG emotion recognition model categorizes emotions as negative, neutral, positive. It cleans the EEG data and extracts DE and PSD features as well as frequency information from each feature. Based on position of the 62 electrodes.

[16] The suggested EEG emotion recognition model categorizes emotions as negative, neutral, or positive. Based on the position of the 62 electrodes, the information from each 1S is assembled into a brain map of size 10 9 9.it ensures that model learn all the features and produce accurate results using ResNet.

[17] EMERSK is a broad and modular framework for recognizing and explaining human emotions using visual input. It is adaptable and modular enough to handle many modalities, such as facial expressions, posture, and gait. The network is made up of many modules that can be added or withdrawn based on the available data. To extract deep features from face photos, it employs a two-stream network architecture with CNNs and encoder-decoder attention methods. To extract characteristics from posture and gait data, CNNs and RNNs with LSTM are used.

[18] Deep learning algorithms, such as the fusion approach that combines VGG 16 and ResNet 50, have been utilized for visual emotion recognition, most notably in the healthcare business for patient emotion recognition. Deep residuals are used in the proposed model.

[19] The study investigates model-level fusion in order to obtain the best multimodal model for emotion recognition utilizing audio and video modalities. Separate innovative feature extractor networks are proposed for audio and video data, and an optimal multimodal emotion detection model is developed at the model level by merging audio and video features.

[20] The suggested work employs the HEPLM technique for emotion classification based on multimodal physiological inputs. To minimize dimensionality to enhance the results. HEPLM combines (POA) and Hierarchical Extreme learning machine.

[21] This research compares deep learning algorithms for emotion logical signals for robot based intervention studies. The research combines CNN, LSTM for classifying the input signals from various sources.

[22] Deep learning approaches, notably CNN and LSTM networks, have been frequently used in robot-based intervention research for emotion identification utilizing physiological inputs. which combines a hybrid network with an SVM classifier, performs best in the pleasant-unpleasant (PU) classification, whereas the hybrid model Dec-CLS, which uses several convolutional layers, performs best in the neutral-unpleasant (NU) classification.

[23] Emotion charting with multimodal physiological signals like ECG, EEG, GSR has gained popularity in a variety of fields, including stroke patients, psychiatrists examining patients, and neuromarketing applications. Physiological signals are unbiased and generated autonomously by the human central nervous system, making them suited for emotion detection. Because of their nonlinear character and the inclusion of noise while recording, classifying these signals for emotion detection is difficult. This study describes a unique deep learning-based ensemble learning strategy for classification.

[24] Emotion recognition using physiological signals is becoming increasingly popular in a variety of disciplines, including stroke patients, psychiatrists examining patients, and neuromarketing applications. Deep learning approaches, notably CNN and LSTM networks, have been frequently used in robot-based intervention research for emotion identification utilizing physiological inputs. The research is centred on evaluating and categorizing physiological data.

[25] Emotion recognition using physiological signals is becoming increasingly popular in a variety of disciplines, including stroke patients, psychiatrists examining patients, and neuromarketing applications. Deep learning approaches, notably CNN and LSTM networks, have been frequently used in robot-based intervention research for emotion identification utilizing physiological inputs. conductance (SC). The effects of hyperparameters such as filter size, number of filters, and dropout on emotion recognition model classification performance have been investigated. Based on the detailed

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

literature survey we are presenting the best performing architectures in Table.1

4. RESEARCH GAPS

Deep learning research in emotion identification with physiological cues has made great progress in recent years, however there are still some research gaps that need to be solved. Some of these gaps are as follows:

Variability between cultures and languages: Emotions and their physiological expressions can differ among cultures and languages. To produce more globally applicable models, research should take into account these cross-cultural and cross-linguistic variations.

Robustness and Generalization: Deep learning models for emotion detection frequently confront robustness and generalization issues. Models that have been trained on a single group or demography may not perform well on another. Techniques for making these models more robust and generalizable across varied populations should be researched.

Inter-subject Variability: People's physiological reactions to emotions can differ greatly. The difficulty of inter-subject variability should be addressed in research, as should methods for personalizing emotion detection algorithms for individuals. Transfer learning, subject-specific models, or adaptable deep learning approaches may be used.**5. CONCLUSION** The use of deep learning in multimodal emotion identification via physiological markers is a promising growing subject. This method provides a more complete comprehension of human emotions than facial expressions and voice signals. Deep learning models excel at recognizing and classifying emotions by integrating data from sources such as heart rate, skin conductance, and EEG. The practical uses range from monitoring mental health to improving human-computer interactions. Nonetheless, issues like data diversity, privacy, and real-world deployment must be addressed. Ongoing research should aim to improve models and overcome these difficulties. We should expect further advances in multimodal emotion recognition in the near future can able to equip with cultural based sensitivity where emotions of the people belongs to different cultures are different in future personalized emotion detection systems very much required apart from this context aware emotion detection systems also very much required where contextual information is required to estimate the emotions of the individuals which may revolutionize different areas of our lives and interactions with technology as deep learning techniques improve.

Table 1: Best Performing Architectures Based On Literature Survey

Sno	Models used	Data set	Accuracy
1	CNN-LSTM	DEAP	97.31
2	CNN	CK+	98.21
3	LSTM	EEG	95

REFERENCES:

[1] Pawan, Kumar, Singh. "Emotion Recognition from EEG Data Using Hybrid Deep Learning Approach." undefined (2023). doi: 10.1007/978-981-19-5191-6_15

[2] Lija, Jacob. "Multimodal Emotion Recognition Using Deep Learning Techniques." undefined (2022). doi: 10.1109/ICAC3N56670.2022.10074512

[3] Sweta, Padman., Dhiraj, Magare. "Multi-modal speech emotion detection using optimised deep neural network classifier." Computer methods in biomechanics and biomedical engineering. Imaging & visualization, undefined (2023). doi: 0.1080/21681163.2023.2212082

[4] "A Hybrid End-to-End Spatio-Temporal Attention Neural Network with Graph-Smooth Signals for EEG Emotion Recognition." undefined (2023). doi: 10.48550/arxiv.2307.03068

[5] Christian, Marx. "Emotion recognition based on multiple physiological signals." Biomedical Signal Processing and Control, undefined (2023). doi: 10.1016/j.bspc.2023.104989

[6] José, M., Celaya-Padilla., Huizilopoztli, Luna-García., Alejandra, García-Hernández., Carlos, E., Galván-Tejada., Jorge, I., Galván-Tejada., Hamurabi, Gamboa-Rosales., David, Rondon., Klinge,

Orlando, Villalba-Condori. (2023). Emotional State Detection Using Electroencephalogram Signals: A Genetic Algorithm Approach. Applied Sciences, doi: 10.3390/app13116394

[7] José, M., Celaya-Padilla., Huizilopoztli, Luna-García., Alejandra, García-Hernández., Carlos, E., Galván-Tejada., Jorge, I., Galván-Tejada., Hamurabi, Gamboa-Rosales., David, Rondon., Klinge, Orlando, Villalba-Condori. "Emotional State Detection Using Electroencephalogram Signals: A Genetic Algorithm Approach." Applied Sciences, undefined (2023). doi: 10.3390/app13116394

[8] Muhammad, Awais., Mohsin, Raza., Nishant, Singh., Kiran, Bashir., Umar, Manzoor., Saiful, Islam., Joel, J., P., C., Rodrigues. "LSTM-Based Emotion Detection Using Physiological Signals: IoT Framework for Healthcare and Distance Learning in COVID-19." IEEE Internet of Things Journal, undefined (2021). doi: 10.1109/JIOT.2020.3044031Emotions have a strong connection with the physical and mental health of individuals, and physiological signals can be used as an indirect measure of emotions .

[9] Shailaja, Kotte. "EEG Signal in Emotion Detection Feature Extraction and Classification using Fuzzy Based Feature Search Algorithm and Deep Q Neural Network in Deep Learning Architectures." SSRG international journal of electronics and communication engineering, undefined (2023). doi: 10.14445/23488549/ijece-v10i5p108.

[10] Rajamanickam, Yuvaraj., Prasanth, Thagavel., John, Thomas., Jack, S., Fogarty., Farhan, Ali. "Comprehensive Analysis of Feature Extraction Methods for Emotion Recognition from Multichannel EEG Recordings." Sensors, undefined (2023). doi: 10.3390/s23020915.

[11] Shilin, Zhang., Qingcheng, Zhang. "A Multidimensional Feature Extraction Method Based on MSTBN and EEMD-WPT for Emotion Recognition from EEG Signals." undefined (2022). doi: 10.1109/BIBM55620.2022.9995251

[12] Magdalena, Marsovszky. "Implementation of an Automatic EEG Feature Extraction with Gated Recurrent Neural Network for Emotion Recognition." undefined (2023). doi: 10.1007/978-981-19-7169-3_13

[13] "Analysis of Feature Extraction Models for Emotion Recognition using EEG Signals." undefined (2022). doi: 10.1109/gcat55367.2022.9972159

[14] S., Satheeskumaran., K., Sasikala., Kumar, Neeraj., A., V., Senthilkumar., N., Sharath, Babu. "IoT based ECG Signal Feature Extraction and Analysis for Heart Disease Risk Assessment." undefined (2023). doi: 10.1109/ICSCSS57650.2023.10169414.

[15] Sumit, Jain. "Emotion Recognition from Multi-channel EEG via an Attention-Based CNN Model." undefined (2023). doi: 10.1007/978-3-031-20738-9_33

[16] KM, Shahin, Bano., Prachet, Bhuyan., Abhishek, Ray. "EEG-Based Brain Computer Interface for Emotion Recognition." undefined (2022). doi: 10.1109/CINE56307.2022.10037255

[17] EMERSK -- Explainable Multimodal Emotion Recognition with Situational Knowledge." undefined (2023). doi: 10.48550/arxiv.2306.08657

[18] Zohauddin, Ahmad. "Emotion Detection Using Deep Fusion Model." undefined (2023). doi: 10.1007/978-3-031-31164-2_40

[19] Asif, Iqbal, Middya., Baibhav, Nag., Sarbani, Roy. "Deep learning based multimodal emotion recognition using model-level fusion of audio-visual modalities." Knowledge Based Systems, undefined (2022). doi: 10.1016/j.knosys.2022.108580

[20] Anushka, Pradhan., Subodh, Srivastava. "Hierarchical extreme puzzle learning machine-based emotion recognition using multimodal physiological signals." Biomedical Signal Processing and Control, undefined (2023). doi: 10.1016/j.bspc.2023.104624

[21] Yusuf, Can, Semerci., Gökhan, Akgün., Elif, Toprak., Duygun, Erol, Barkana. "A Comparative Analysis of Deep Learning Methods for Emotion Recognition using Physiological Signals for Robot-Based

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195
Intervention Studies." undefined (2022). doi: 10.1109/TIPTEKNO56568.2022.9960200

[22] Xuan-Nam, Cao., Ming, Sun. "An Emotion Recognition Method Based On Feature Fusion and Self-Supervised Learning." undefined (2023). doi: 10.1145/3590003.3590041

[23] Amna, Waheed, Awan., Syed, Muhammad, Usman., Shehzad, Khalid., Aamir, Anwar., Roobaea, Alroobaea., Saddam, Hussain., Jasem, Almotiri., Syed, Sajid, Ullah., Muhammad, Akram. "An Ensemble Learning Method for Emotion Charting Using Multimodal Physiological Signals." Sensors, undefined (2022). doi: 10.3390/s22239480

[24] Anushka, Pradhan., Subodh, Srivastava. "Hierarchical extreme puzzle learning machine-based emotion recognition using multimodal physiological signals." Biomedical Signal Processing and Control, undefined (2023). doi: 10.1016/j.bspc.2023.104624

[25] Sharmeen, M.Saleem, Abdullah, Abdullah., Siddeeq, Y., Ameen., Mohammed, A., M., Sadeeq., Subhi, R., M., Zeebaree. "Multimodal Emotion Recognition using Deep Learning." undefined (2021). doi: 10.38094/JASTT20291