ISSN: 2633-4828 Vol. 5 No.2, June, 2023

International Journal of Applied Engineering & Technology

A MARKET ANALYSIS OF ARTIFICIAL INTELLIGENCE APPLICATIONS OF EMOTIONAL RECOGNITION

Serdar Engin Koç [1]

[1] Faculty of Education, Department of Computer Education and Instructional Technologies,
Başkent University, Ankara, Türkiye

[1] ekothekoch@gmail.com

Abstract

The recognition of emotions employs many artificial intelligence methods with different data that come from facial expressions, electroencephalogram, voice analysis, music emotion, speech emotion, micro expressions, heart rate, heart rate variability and other formats. Especially neural networks have shown great potential in the last few years owing to their efficient analysis and classification capabilities. This paper's scope is to present historical and conceptual perspectives in the artificial intelligence techniques and methods that proved significant results in emotion recognition. This paper highlights the key elements and performance of each architecture/model, and the advantages and limitations of the proposed models in the analyzed papers. Additionally, this paper presents the available datasets that are currently used for emotion recognition from different data. Another highlight is at which phases of emotion recognition artificial intelligence techniques are used. Open issues and future possible developments in emotion recognition are identified. It is expected that this research can put clear criteria to separate techniques that are considered as artificial intelligence or not.

***Keywords**: Artificial intelligence, electroencephalogram, emotion recognition, facial emotion, micro-expression, voice analysis.*

INTRODUCTION

The training of artificial neural networks with data from different sources such as skeletal movements, eye gazing, audio files, textual segments to get wider recognition and combining different results with fusion techniques to get better recognition of emotions in subject-independent way, all point to moving towards rising trend of multimodal approaches in emotion recognition.

Facial expression is the fastest reaction to events and as data to understand emotions of a human. In recent years real-time emotion recognition has seen some work that would be beneficial for education, market analysis, health issues and it remains to be a trend in artificial intelligence domain. That is why more sensitive equipment and methods are being used to fuse different kinds of data to get the most accurate emotion recognition. Neural networks analyze facial structures through facial analysis which can vary from full-face processing to facial landmarks [1] where different images are compared to the subjects face or geometric relationships are used for analyzation.

While some of the data can be faked, generally physiological signals such as electroencephalography represent real changes hence although being weak signals, are more reliable for emotion recognition. [2] used Convolutional Neural Network to extract EEG information and used Deep Neural Network to classify emotions. EEG is also used to get information about attention levels of learners [3, 4], while EEG is mixed with peripheral physiological signals, and AI algorithms to classify emotions [5].

While speech/voice analysis has limited database for training of artificial neural networks, there are promising accuracy results with emotion recognition from voice (81%,[6]: 85%, [7]) with machine learning algorithms. Technology today is also able to extract useful emotional information from textual data via natural language processing with sentiment analysis and deep learning. Textual data remain as supplementary when it comes to recognizing emotions throughout most of the related research, however skill-based assessment such as writing, creativity, etc. use textual data as their main source of information. [8] used BERT-CNN to analyze 8867 learners' discussion to find the effect on emotional and cognitive engagement on learning achivement.[9] used Emotnizer application to classify text rewriting behaviors as signs emotion.

Galvanic skin response (GSR) is also a useful indicator of physiological or psychological arousal as such when combined with AI and other forms of emotional data can create a large area of research in a lot of fields as education, user experience, games, and human computer interaction. [10] used GSR with facial expressions to classify emotions in the valence arousal dimension.

3D Skeletal movement data is gaining its importance in the field of AI as being important signals of non-verbal communication due to lack of databases covering skeletal movement data. Like text, skeletal data is also considered a supplementary base for emotion recognition, and it is meaningful in research to fuse with audio and textual data to get better understanding of emotions. Most of the research reviewed sees fusion strategies as a solution to increase the emotion recognition efficiency with a few exceptions.

EEG signals have some disadvantages while being easy to use in a non- invasive way, making real-time emotion recognition difficult and interpretation depends on concentration of subjects and body movements. So, some researchers [11] concentrated on photoplethysmography (PPG) and electromyography (EMG) to perform better results. [12]. Wrote a literature review about cardiac signals used in emotion recognition including Respiration Rate (RR) and Skin Temperature (SKT).

There are also. self-report measures of emotion recognition that are verbal (PAD dimensions), visual (SAM) and moment to moment reported in [13]. These will be skipped here as the primary objective of this paper is automatic recognition of emotions based on artificial intelligence but used together with automatic methods, self-report measures will be mentioned.

So how does emotion recognition happen and what are the stages of automatic recognition?

Phases of automatic emotion recognition generally start with the acquisition of data. Researchers generally use the publicly available datasets for facial expressions, speech analysis, etc... however, some researchers use their own data to create these databases or picture sets to train their network system. On this data, normalization methods are applied where the data is adjusted to a standard scale. After that, background noise in audio or visual form is removed then data is segmented into smaller pieces for easy management. The next stage is called as feature extraction and in this stage, relevant properties such as geometric facial features, pitch and energy of sound, peaks and lows of brain wave patterns and word embeddings, contextual embeddings of textual data are extracted. Based on these properties, emotions are classified using machine learning, deep learning and ensemble methods. In post-processing, smoothing and contextual analysis methods are applied to refine predictions and reduce fluctuations depending on other data or supplementary cues. Fig. 1 shows the explanation of these processes for EEG signal.

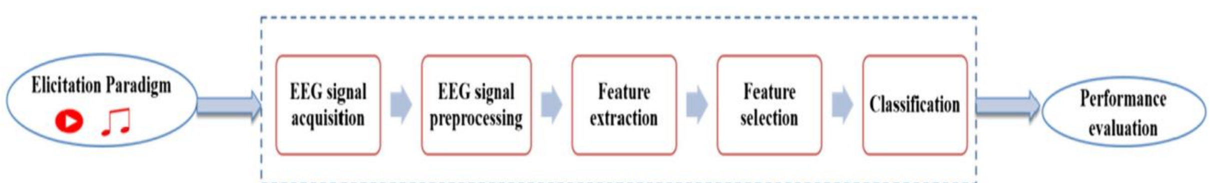

Fig 1. Phases of emotion recognition. Taken from [14], p. 8.

This paper is a comprehensive survey of artificial intelligence solutions and techniques used in emotion recognition. In this context, it aims to provide recent research done in emotion recognition using artificial intelligence methods. That is why most of the methods that are not considered as artificial intelligence will not be listed but if they are consistent with the AI systems then they will be listed with the methods. This study is a first to focus on artificial intelligence methods in the field of emotion recognition as out of all the articles reviewed they either focused on the type of data or the proposed framework/method they used in their research. Also, it is expected that this research will draw clear lines whether the methods or techniques that belong to the artificial intelligence family or not and try to redefine the phases of emotion recognition and how important it is to state what method is used in each phase.

The rest of the article is organized as follows. Section 2 represents the methodology used and the criteria of the articles that are surveyed. Section 3 shows the list of the artificial intelligence techniques with some methods that are used with AI systems, but they are not AI themselves. Section 4 shows the architecture, feature extraction, classification, feature selection methods used in surveyed articles followed by databases used in emotion recognition with their accuracy values and emotion types with dimensions in Section 5. Preprocessing was not involved as most of the time it is done by conventional means. Moreover, some challenges, opportunities and limitations of emotion recognition systems are discussed. The last section presents conclusions.

SECTION 2.

This review focuses on the artificial intelligence techniques used in emotion recognition systems. In this sense Web of Science database is used to identify relevant papers and the results were reported in the form of Preferred Reporting Items for Systematic Reviews and Meta-Analysis [46].

There were different searches made in WOS database, however, a search in the titles did not give good results as those articles that contained artificial intelligence did not have any techniques belonging to that genre so only documents that had "artificial intelligence" and "emotion recognition" in their abstract were searched. The resulting documents were filtered by the years 2020-2024. As none of the artificial intelligence techniques were preferred on other, the search was not filtered to another parameter, but the chosen documents were checked whether the methods they used were artificial intelligence or not. Fig. 2 and 3 show the status of the articles viewed and selected.

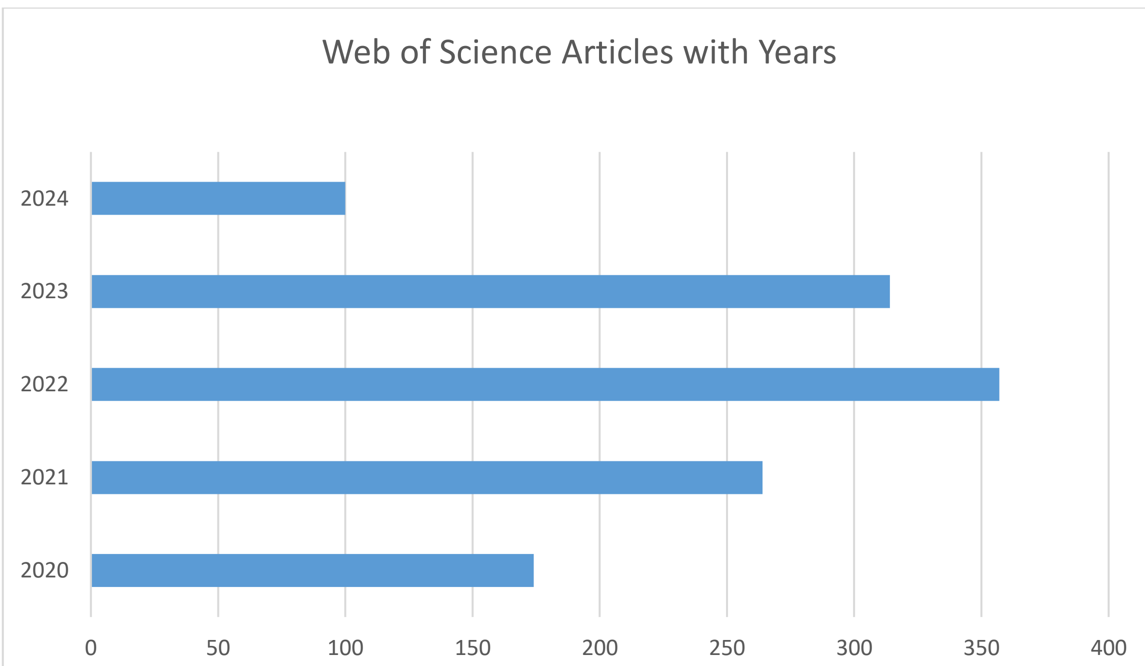

Fig. 2. Web of Science Articles

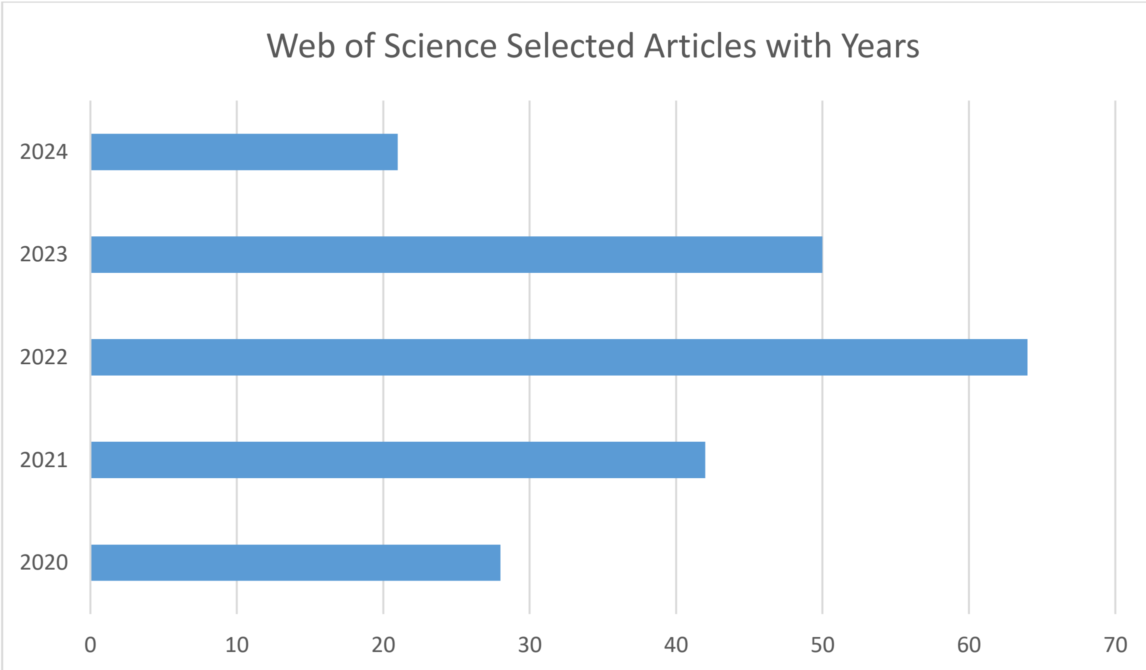

Fig. 3. Web of Science Reviewed Articles

After an initial set of 1209 articles, 1105 articles were removed before screening. 204 records were screened for abstract, and title and 112 articles were excluded. 14 articles were unreachable, 40 records were too technical, 54 articles were unrelated within the scope and 4 articles were retracted. 92 articles were selected for eligibility, and these were assessed by reading full text out of which 10 were found to be unrelated, 7 were too technical and 8 documents were reviews, so they were saved for meta-analysis. The eligibility criteria also contained that chosen article should use techniques/calculations/models of parts of artificial intelligence. Fig. 4 shows the Prisma flow diagram of the search. Fig. 5 shows the implementation stages of the Prisma Flow Diagram.

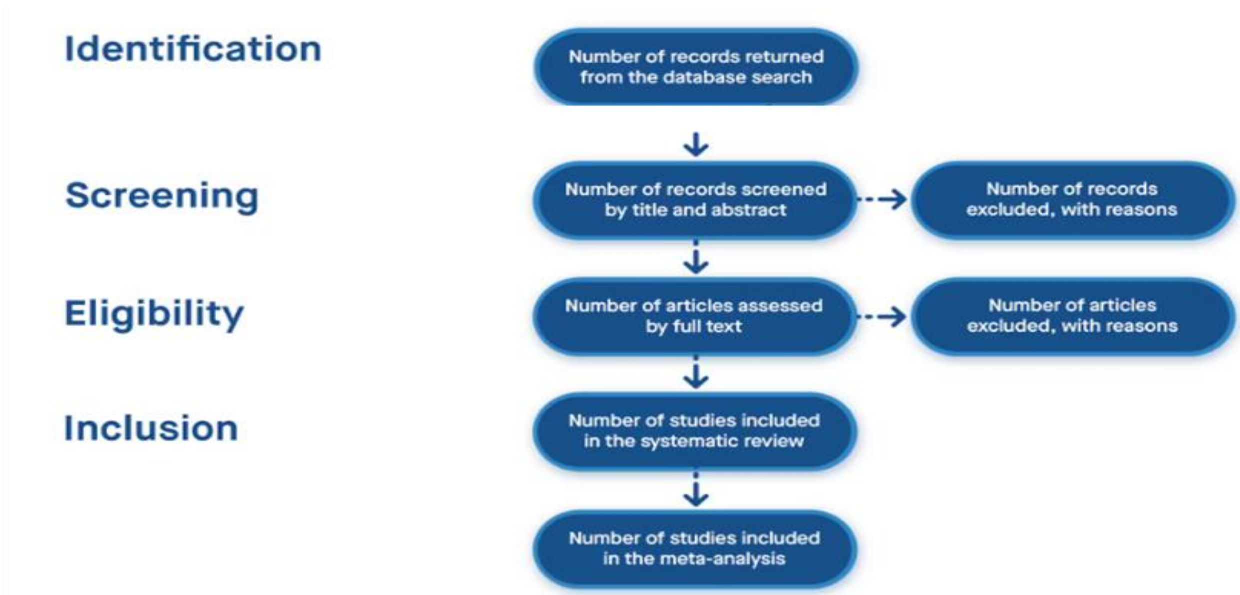

Fig. 4. Prisma Flow Diagram.

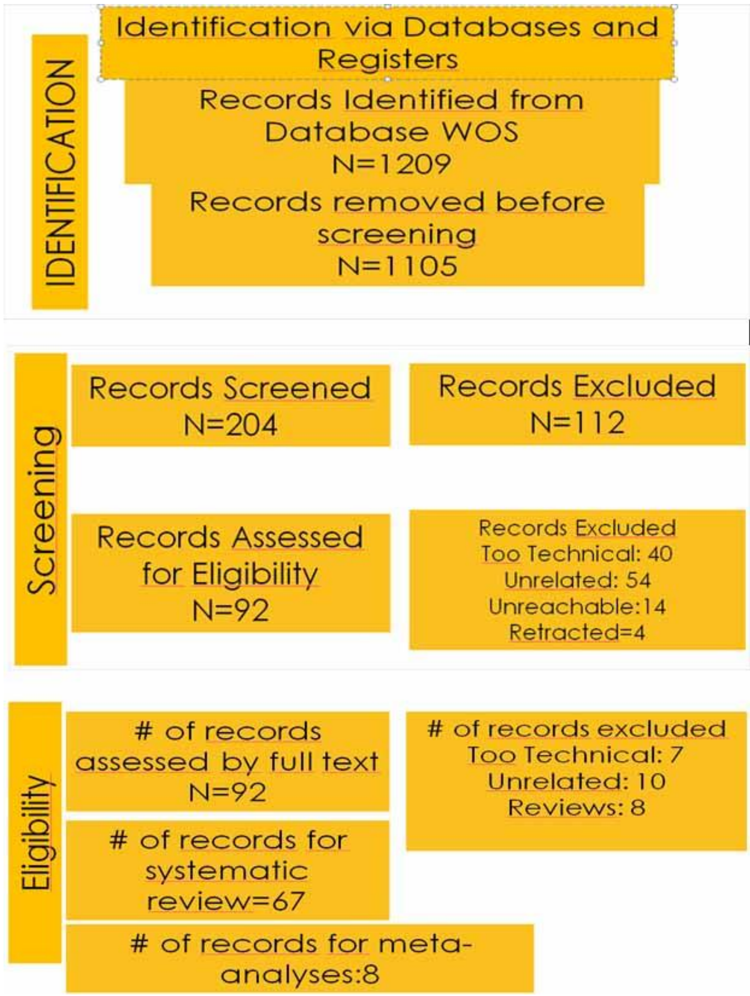

Fig. 5. Implementation stages of Prisma Flow Diagram.

SECTION 3. DEFINITION OF ARTIFICIAL INTELLIGENCE TECHNIQUES AND METHODS

These methods are all AI algorithms but in research many other non-AI methods are also used for various purposes. For the sake of the research only AI methods are included.

Deep Neural Networks (DNNs) are multi-layered artificial neural networks that excel at capturing intricate data patterns, making them well-suited for tasks such as emotion recognition. They have input layer to receive raw data like images, audio or text. They have hidden layers and recurrent layers for hierarchical representations and for temporal dependencies respectively. The outer layer is to produce classification or regression outputs in the form of labels such as emotion labels. Deep Neural Networks (DNNs) are powerful tools for emotion recognition, leveraging their ability to learn complex patterns from various data types. Different types of DNNs, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), and Transformers, excel at different tasks. CNNs are great for recognizing spatial hierarchies in images and videos, RNNs are effective for sequential data like speech and text, LSTMs handle long-term dependencies better than standard RNNs, and Transformers efficiently manage long-range dependencies and parallelize training, making them suitable for text and sequential data.

Ensemble methods combine multiple models to improve classification performance by aggregating predictions through voting and using techniques like boosting and bagging. They are effective for multimodal emotion recognition, integrating facial, vocal, and textual data. By combining these methods, models can accurately detect and classify emotions, enabling a wide range of applications from human-computer interaction to mental health monitoring.

Support Vector Machines (SVMs) are generally used for for binary and multi-class classification, finding optimal class boundaries for emotions from physiological signals whereas *K-Nearest Neighbors (KNN)* are used with small datasets and non-linear classification, using nearest neighbors for emotion recognition from facial landmarks or text features. *Decision Trees and Random Forests* are used to classify structured data where tree structures and ensemble learning methods are added to increase accuracy emotion classification. *Naive Bayes* is a method used for emotion recognition from textual data and uses probability theorems with sentiment analysis

AI employs a variety of classification methods for emotion recognition, each suited to different types of data and applications. These methods leverage the power of deep learning, statistical learning, and ensemble techniques to accurately identify and classify human emotions, enabling advancements in human-computer interaction, mental health monitoring, and more.

Optimization methods in artificial intelligence (AI) are crucial for enhancing the performance of emotion recognition models by finding optimal parameters or configurations. Table I shows the optimization methods mostly used in research and their working mechanisms shortly.

Table I. AI Optimization Methods

Method	Mechanism/advantage
Gradient Descent	İteratively updates parameters, used in training networks
Stochastic Gradient Descent	Efficient with large datasets, computes gradients using subsets of training data
Adam Optimizer	Combines methods that automatically adjusts the learning rate and helps to reduce training time
Generic algorithms	Provides solutions that improve over timewith feature selection and parameter tuning
Bayesian Optimization	Uses a probabilistic model and acquisition function for hyperparameter tuning.

These methods enhance the accuracy and efficiency of emotion recognition systems, with the choice of method depending on task requirements and data nature.

SECTION 4.

Architectures:

The research in emotion recognition on various platforms with different kinds of data requires the use of different frameworks, models and architectures. In some cases, it is more important to understand the depth to capture complex features as in [15], whereas in some experiments accuracy or real-time recognition takes precedence in [16], [17] and [18]. That's why researchers use and evaluate best methods for their own needs with their tasks. For the sake of this review, it is thought that the architectures are important parts of these recognition systems and the whole process of this task, so they will be reported with their general advantages, and this will reveal the reasons of developing different architectures for different purposes. Some of the selected articles do not show or give examples of their architecture so they are not included, hence only if the architecture is shown in detail was included.

Table II. Architectures used and advantages.

Article #	AI Architectures	Advantage
[19]	Spatial Temporal Graph Convolutional Network	By leveraging both spatial and temporal information, capture complex patterns in videos
[20]	Music emotion classification fuzzy model	Handles the inherent ambiguity and subjectivity in music emotion perception.
[21]	Deep Neural Network	Ability to learn hierarchical representations of data
[4]	Temporal Convolutions Residual Neural Network (TC-Resnet)	Effective for tasks that require understanding of both spatial and temporal information
[22]	6 Self-Organizing Neural Network with Ensemble Layer	faster convergence and better generalization performance.
[6]	Multiple Layer Perceptron	Flexible, scalable, usable, interpretable, universal approximation
[23]	Xception	For image classification tasks, effective in terms of both accuracy and efficiency
[24]	YOLO5	Balance and accuracy over real time detection.
[25]	Multifeature fusion architecture	Improve the robustness and accuracy of machine learning models by combining accurate predictions.
[7]	Spatial multi-head attention-based convolutional network	By incorporating both convolutional and attention mechanisms leads to increased performance in understanding spatial relationships
[14]	EEG based BCI system for emotion recognition	non-invasiveness, high temporal resolution, and relatively low cost.
[26]	Resnet50	Transfer learning, balanced complexity, state of the art performance, depth and performance.
[10]	IMOTIONS Recognition System	Integrate and synchronize multimodal biometric research.
[27]	Resnet (Residual Network)	Ease of training, improved performance, scalability, feature reuse, interpretability
[28]	Temporal Difference Feature Network	improve emotion recognition by integrating information and capturing temporal dependencies
[11]	Bimodal Stacked Sparse Auto-Encoder	Feature extraction, noise ignoring, flexible

Article #	AI Architectures	Advantage
[29]	Own architecture using CNN	Good at capturing spatial hierarchies and recognizing patterns, reduces number of parameters and computation,
[30]	3D CNN	Enhanced feature representation, accuracy, depth information
[31]	VGG13	Simple and uniform, speed and accuracy
[15]	CapsNet	Reduced information loss, object recognition

Training and Testing Sets

To validate the emotion recognition model, researchers split the dataset for training, validation and testing purposes followed by model comparison, where model is compared to other models in terms of accuracy, possible subjective evaluation with human evaluators and real-world testing. For the sake of this research, complex statistical models will not be reported but the ratios of training/validation/testing dataset of the model and validation methods will be covered shortly as that will give more information of the techniques used in this area of research. In addition to this, there are three strategies named as Leave-One-Subject-Out (LOSO), Leave-One-Person-Out (LOPO), Leave-One-Person-One-Session-Out(LOPOS) and Random Assignment (RA) splits data into training and testing sets. These strategies help to make data less frequent-dependent and generalize well across different conditions. When we look at Table III it is apparent that most of the researchers have a training/testing ratio of 80-90/20-10. Ratio of the data for validation is not given most of the time except with [29] where we can see testing ratio and validation ratio is close to each other. It would appear that researchers used LOSO in [32] when they have different models, LOPO in [33] when they don't have class or data balancing technique and RA methods [28] when they have data imbalance problems caused by dataset and LOSO in [3] as a continuation of previous research. These methods are not considered old when compared to AI methods and they are more useful when data comes from different individuals. Further checking reveals that [31] has different individuals for a set of negative, positive and neutral images as dataset but on [33], 8 emotions set of data come from the same individuals while in [32] dataset is not mentioned. It is worth mentioning that there are a lot of variables that could effect accuracy with training/validation/testing ratio. For instance, while [21] uses 80/20 ratio with training/testing, the accuracy was 50-60 with speech recognition, while [29] used a ratio of 1:0.5:0.41 for training/testing and validation and had very high accuracy with facial emotion recognition.

Table III. Training, Testing and Validation.

Article	Training %/Ratio/#	Testing %/Ratio/#	Validation %/Ratio/#
[21]	80%	20%	-
[34]	9	1	-
[4]	8	1	1
[6]	75%	25%	-
[36]	85%	15%	-
[24]	5701	-	-
[35]	80%	20%	-
[7]	8	1.5	-

Article	Training %/Ratio/#	Testing %/Ratio/#	Validation %/Ratio/#
[20]	LOSO		
[9]	9	1	-
[11]	80%	20%	-
[29]	34K	17K	14K
[30]	80%	20%	-
[15]	70%	30%	-
[28]	LOSO, LOPO, RA		
[33]	LOSO, LOPO		
[31]	LOSO		

Feature Extraction

Feature extraction is one of the most important stages covered in emotion recognition. The objective is to capture the most relevant information that distinguishes different emotions. Feature extraction involves extracting relevant information from input data, such as facial expressions, voice recordings, or text, to represent the emotional content effectively. The general process for feature extraction with data from different sources are as follows:

Facial Expression Recognition: Key facial landmarks, such as eyes, nose, and mouth, are detected to understand facial expressions. Features like the distances between facial landmarks, angles of facial components, or intensity patterns are extracted to represent different emotions. Speech Emotion Recognition: Features such as Mel-frequency cepstral coefficients (MFCCs), pitch, energy, and formants are extracted from the speech signal. Text-Based Emotion Recognition: Features like word frequency, n-grams (sequences of adjacent words), and sentiment scores are extracted from the text.

Table IV Feature Extraction Methods.

Article	Feature Extraction/AI OR NOT
[34]	Energy threshold-based multicommon spatial pattern
[4]	TCResnet/AI
[6]	Mel-Frequency Cepstral Coefficient, Short-time Fourier Transform
[37]	Adam Optimizer, DenseNet121 /AI
[36]	Fast Fourier Transform, Mel-Frequency Cepstral Coefficient,
[35]	CNN-Bi-LSTM-Attention/AI
[23]	Xception based CNN/AI
[25]	Viola Jones
[7]	(Speech) OpenSmile Toolkit, (text) GloVe Vector
[26]	Conv2D, Residual Blocks
[5]	Power Spectral Density

Article	Feature Extraction/AI OR NOT
[38]	Res2net50/AI
[27]	Multitask Cascade CNN/AI
[33]	Residual Attention Module
[39]	LSTM + CNN/AI
[28]	TDFNet, deep scale transformer
[11]	Stacked Encoder
[30]	3D CNN/AI
[15]	CapsNet/AI
[31]	Landmark Feature Map/ Compact LFM
[24]	YOLO5/AI
[20]55	Welch

When we look at the feature extraction methods, most methods rely on the framework/model used for emotion recognition. Also, the methods used for feature extraction only 9 out of 22 articles used artificial intelligence methods for this job. Other articles do not mention how they performed feature extraction. We see that 4 articles that used AI methods for feature extraction looked for EEG data and 6 articles that used AI methods for feature extraction looked for facial emotion data. If we compare 50 articles about FER with 8 research about EEG, we can see that AI methods for feature extraction are not popular or they are not efficient because they are not used for this purpose much.

Classification

Classification in emotion recognition involves using machine learning models to categorize input data (such as facial expressions, voice recordings, or text) into different emotion classes. First a model is selected than this model is trained on splitted data using extracted features. This model is used to predict emotion classes and final results are passed through filtering or smoothing operations to be improved. This model is compared to others in terms of accuracy and evaluation matrix. Iterative refinement of feature extraction and model selection are required mostly to get better results with accuracy. For this review, classification phase seems to be the category where AI methods are mostly used. Over 34 articles, 29 used AI methods for classification which is significant when compared to other phases of emotion recognition. Also 12 articles used AI methods for optimization of classification. There is still the problem of reporting studies that do not mention their optimization or classification methods directly but that is believed to be due to the focus of the research and also optimization is not a must-phase of emotion recognition. It can easily be seen that 8 articles employed multiple data for classification but it cannot be said that they all have greater accuracy than single source of data for classification.[4] achieves greater accuracy by using EEG and, functional near-infrared spectroscopy data but there are a lot of factors that could give this result. Fig. 5 and 6 show results better in terms of data vs. accuracy and also dataset vs. accuracy. In Fig. 6 we can clearly say that FERPLUS has enhanced quality and granularity of the labels in FER2013 that is why it may give better accuracy. The databases were not intended to be compared to each other as they all have different variables and different number of emotions they are intended for training. The comparison of databases or emotions requires much more data and maybe include all studies that give sufficient detail. However, by looking at multiple usage of the same database, we can make a few conclusions. SEED database was used three times and gave 99.5%, 82% and 97% accuracy in [35], [34] and [30] studies respectively. [35] used Adam optimizer and non-AI techniques, [34] used AI techniques without an optimizer and [30] used AI techniques without an optimizer. DEAP dataset was used

4 times and gave 73%, 96%, 94% and 72% accuracy in [5], [30], [40] and [34] respectively with AI classification without an optimizer in all of them. RAVDESS dataset was used 3 times with 50%, 81% and 85% accuracy in [21], [6] and [36] respectively. All used AI for classification and only [36] did not use an optimizer. IEMOCAP database was used 2 times giving 76 % and 78% accuracy in [28] and [7] respectively. [7] used AI classification and included multiple data with Adam optimizer whereas [28] used non-AI classification without an optimizer. Looking at all datasets used it is not possible to explain the accuracy of emotion recognition as it would need more data to be significant and with these numbers there doesn't seem to be a pattern over using AI classification or optimizer. At this point, it would require new research for databases used to be able to do that distinction and come up with factors responsible for the accuracy. In this regard, this research is seen as a first step to analyze databases for this purpose.

Articles [11], [10], [5],[26], [7], [20], and [4] used multifeature data as input. Except article [4] we cannot say that they achieved greater accuracies in their research, however throughout most of the research screened, multidata usage usually ended in greater accuracy of models. The reasons can be tied to introducing a new method of emotion recognition, using a new or own database, the number of emotions included. Real-time recognition of emotions is also a challenging task for this field of research. Those articles reviewed that achieved good results with real-time emotion recognition used facial emotions ([38], [27], [17]), Photoplestimogtraphy with heart rate variability ([41]), movement with facial emotions ([42]), behavior ([24]), content creation ([43]), and EEG ([4], [34]) for real time recognition of emotions. However, [41], [42], [24] and [43] were not pure emotion recognition studies. [42] included a study of a robot and it had an emotion recognition model that was not mentioned in terms of details of working. Also [41], [24] and [43] were directed at discovering other outputs. Also, in [41] the model was trained using learning activities in LMS course, [24] used actions and attention in addition to emotions and [43] intended to create a practical application for "…future content creators to better control an intended emotional response delivered and received by the audience." ([43], p.1).

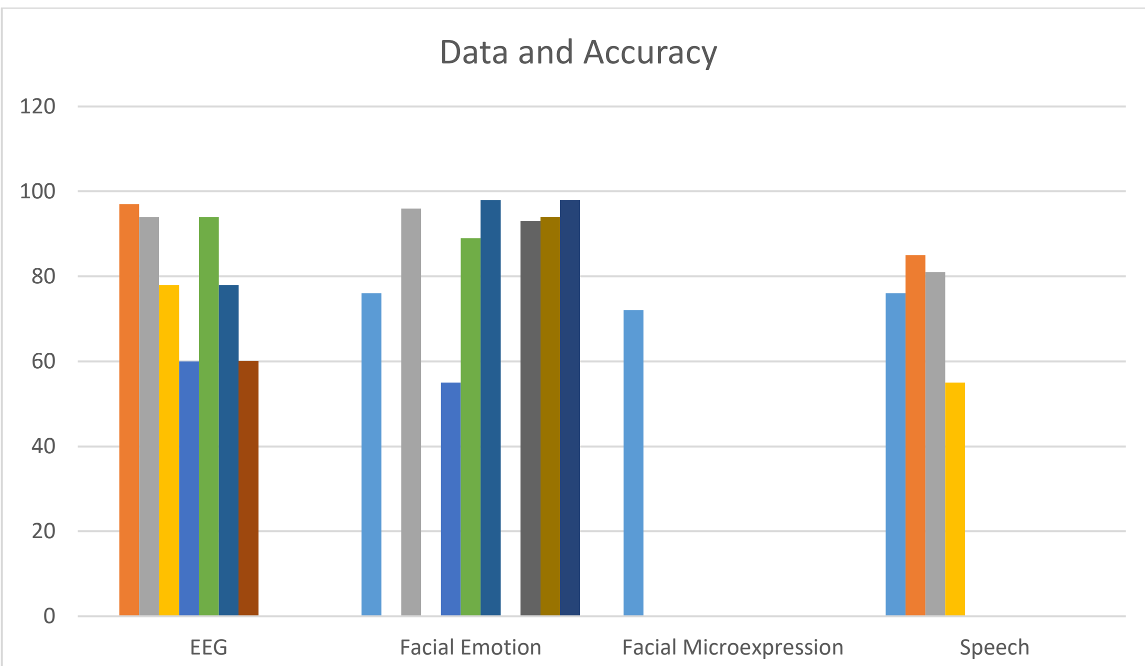

Fig. 5. Data Type and Accuracy of the model

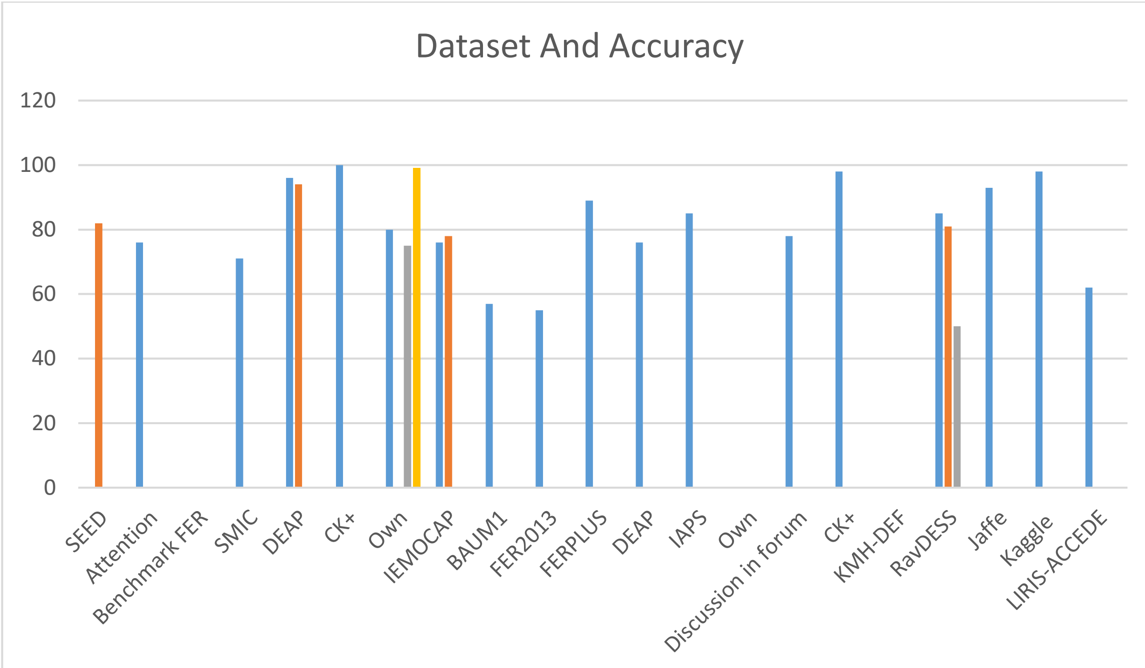

Table V Classification Methods.

Article	Classification/AI OR NOT	Accuracy %	Data	Dataset	Optimizer
[35]	Softmax	99.55	EEG	SEED	Adam /AI
[24]	YOLO5 /AI	76	Facial Emotion	Attention-NoAttention, Affect Dex	
[15]	BİLSTM /AI	99.5	Facial Emotion	Benchmark FER	Pelican Optimizer/AI
[31]	CNN-LSTM/AI	71-74	Facial Microexpression	SMIC, SAMM CASME II	
[44]	AI Emotion Software /AI				
[30]	3DCNN, Mask RCNN, Support Vector Machine /AI	96-97	EEG	DEAP, SEED, and MAHNOBHCI	
[29]	Improved NAS/AI	100,99,95	Facial Emotion	CK+, JAFFE, FERG	
[11]	Bimodal Stacked Sparse Auto-Encoder/AI	80, 75	PPG, EMG	Own dataset	Adam/AI

Article	Classification/AI OR NOT	Accuracy %	Data	Dataset	Optimizer
[28]	Multimodal Fusion Classification in TDFNET	76	Speech	IEMOCAP	
[39]	LSLTM, CNN/AI	Classification time less (not declared)	EEG	-	Adam/AI
[33]	Multiple Layer Perceptron/AI	57, 70	Facial Emotion	Baum-1, RML	
[27]	CNN/AI	55	Facial Emotion	FER2013	
[38]	NAgNet/AI	89	Facial Emotion	FERPLUS	
[5]	Naïve Bayes, Linear Regression, Support Vector Machine/AI	76-71-73	EEG+Peripheral physiological signals	DEAP	
[10]	AFFDEX Module/AI	85	GSR+ Facial Emotion	IAPS	
[26]	SoftMax	-	facial expression and speech sentiment	AffectNet	
[28]	Decision Tree, SAM/AI	77.4, 84	Keyboard and mouse movements	Own dataset	
[7]	Multi-Head Attention based CNN/AI	78	Skeletal movements, audio, text	IMEOCAP	Adam/AI
[45]	Bert-CNN/AI	78	Text	Discussion in MOOC forum	backpropagation algorithm
[25]	SoftMax	98	Facial Emotion	CK+, Yale Face, Karolinska	Adam/AI
[23]	LTSM /AI	99.7	Facial Emotion	KMH-FED	jellyfish search
[36]	Gaussian Mixture Model/AI	85	Speech	RAVDESS	
[20]	C3DCNN, 2DCNN, SVM/AI	-	Eye gaze, body pose, voice	MONGODB	

Article	Classification/AI OR NOT	Accuracy %	Data	Dataset	Optimizer
[37]	Weighted Kernel Extreme Learning Machine/AI	93	Facial Emotion	JAFFE, FER2013	Adam/AI
[40]	Dual Attention 3D Convolutional Neural Network/AI	94	EEG	DEAP	
[6]	Multiple Layer Perceptron/AI	81	Voice	RAVDESS	Adam/AI
[19]	Spatial-Temporal Graph Convolutional Network/AI	75	Skeletal movement data	Own dataset of 360 videos	
[4]	TCResnet/AI	99	EEG, functional near-infrared spectroscopy	Own dataset	
[17]	Reinforcement learning classifier/AI	94	Facial Emotion	-	
[20]	Pleasure Arousal Music Fuzzy Model/AI	-	Music Emotion	Own dataset 18 music clips	
[21]	Convolutional Deep Neural Networks/AI	50 , 60	Speech	RAVDESS, EMODB	Bayesian/AI
[34]	LTSM/AI	82, 72,81	EEG	SEED, DEAP, DREAMER	Adam/AI
[22]	Self-Normalizing Neural Network/AI	98	Facial Emotion	KAGGLE	Adam/AI
[43]	EEG-based recognition system	62,57	EEG	LIRIS-ACCEDE	Adam/AI

CONCLUSION

This research carried out a comprehensive report on findings of usage of artificial intelligence methods in classification, feature extraction and optimization of emotion recognition processes. It is also viewed as a step-stone in showing the power of artificial intelligence technology over other methods and the advantages and disadvantages of these methods with specific databases used in emotion recognition. It is expected that research build on top of this review will include more detailed information about phases of emotion recognition and AI technologies used for different phases so that there might be a link between accuracy of models, technology and database used. It is believed that this could change the direction of emotion recognition research if this detailed information is involved in all research carried out. The researcher is also determined to find related research that uses the same database to clarify the reasons of accuracy and complex relations between used methods for all phases of recognition. There is also another question that awaits to be revealed concerning the superiority of multiple data used in emotion recognition. Just as the models need more databases to be trained on, researchers need more data to be able to create reasons and develop better models for better accuracy and benefit from automatic emotion recognition.

REFERENCES

[1] Cao J, Zhao Y, Lai X, Ong MEH, Yin C, Koh ZX, Liu N (2015) Landmark recognition with sparse representation classification and extreme learning machine. J Franklin Inst 352(10):4528–4545

[2] Liu, J.; Wu, G.; Luo, Y.; Qiu, S.; Yang, S.; Li, W.; Bi, Y. EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder. Front Syst. Neurosci. 2020, 14, 43. [CrossRef]

[3] Liu, N.H.; Chiang, C.Y.; Chu, H.C. Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors. Sensors 2013, 13, 10273–10286. [CrossRef]

[4] Chen, C.M.; Wang, J.Y.; Yu, C.M. Assessing the Attention Levels of Students by Using a Novel Attention Aware System Based on Brainwave Signals. Br. J. Educ. Technol. 2017, 48, 348–369. [CrossRef]

[5] Zhang Q, Zhang H, Zhou K, Zhang L. Developing a physiological signal-based, mean threshold and decision-level fusion algorithm (PMD) for emotion recognition. Tsinghua Science and Technology. 2023 Jan 6;28(4):673-85.

[6] Alnuaim AA, Zakariah M, Shukla PK, Alhadlaq A, Hatamleh WA, Tarazi H, Sureshbabu R, Ratna R. Human-computer interaction for recognizing speech emotions using multilayer perceptron classifier. Journal of Healthcare Engineering. 2022 Mar 28;2022.

[7] Shi J, Liu C, Ishi CT, Ishiguro H. 3D skeletal movement-enhanced emotion recognition networks. APSIPA Transactions on Signal and Information Processing. 2021 Jan;10: e12.

[8] Liu S, Liu S, Liu Z, Peng X, Yang Z. Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. Computers & Education. 2022 May 1; 181:104461.

[9] Magdin M, Kapusta J, Koprda Š. Investigation of behavioral characteristic parameters intended for emotional state classification. IEEE Access. 2022 Dec 8; 10:129029-41.

[10] Flynn M, Effraimidis D, Angelopoulou A, Kapetanios E, Williams D, Hemanth J, Towell T. Assessing the effectiveness of automated emotion recognition in adults and children for clinical investigation. Frontiers in human neuroscience. 2020 Apr 7;14:70.

[11] Lee YK, Pae DS, Hong DK, Lim MT, Kang TK. Emotion recognition with short-period physiological signals using bimodal sparse autoencoders. Intelligent Automation & Soft Computing. 2022 May 1;32(2).

[12] Dzedzickis A, Kaklauskas A, Bucinskas V. Human emotion recognition: Review of sensors and methods. Sensors. 2020 Jan 21;20(3):592.

[13] Poels K, Dewitte S. How to capture the heart? Reviewing 20 years of emotion measurement in advertising. Journal of advertising research. 2006 Mar 1;46(1):18-37.

[14] Houssein EH, Hammad A, Ali AA. Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review. Neural Computing and Applications. 2022 Aug;34(15):12527-57.

[15] Alonazi M, Alshahrani HJ, Alotaibi FA, Maray M, Alghamdi M, Sayed A. Automated Facial Emotion Recognition Using the Pelican Optimization Algorithm with a Deep Convolutional Neural Network. Electronics. 2023 Nov 11;12(22):4608.

[16] Pinto LV, Alves AV, Medeiros AM, da Silva Costa SW, Pires YP, Costa FA, da Rocha Seruffo MC. A Systematic Review of Facial Expression Detection Methods. IEEE Access. 2023 Jun 16.

[17] Chen Y. Design and simulation of AI remote terminal user identity recognition system based on reinforcement learning. International Journal of Modeling, Simulation, and Scientific Computing. 2023 Feb 31;14(01):2341005.

[18] Chen J, Yu K, Wang F, Zhou Z, Bi Y, Zhuang S, Zhang D. Temporal convolutional network-enhanced real-time implicit emotion recognition with an innovative wearable fNIRS-EEG dual-modal system. Electronics. 2024 Mar 31;13(7):1310.

[19] Tsai MF, Chen CH. Spatial temporal variation graph convolutional networks (STV-GCN) for skeleton-based emotional action recognition. IEEE Access. 2021 Jan 18;9:13870-7.

[20] He JX, Zhou L, Liu ZT, Hu XY. Digital empirical research of influencing factors of musical emotion classification based on pleasure-arousal musical emotion fuzzy model. Journal of Advanced Computational Intelligence and Intelligent Informatics. 2020 Dec 20;24(7):872-81.

[21] Diemerling H, Stresemann L, Braun T, Von Oertzen T. Implementing machine learning techniques for continuous emotion prediction from uniformly segmented voice recordings. Frontiers in Psychology. 2024 Mar 20;15:1300996.

[22] Bibbo' L, Cotroneo F, Vellasco M. Emotional Health Detection in HAR: New Approach Using Ensemble SNN. Applied Sciences. 2023 Mar 3;13(5):3259.

[23] PrabhakaraRao T, Patnala S, Raghavendran CV, Lydia EL, Lee Y, Acharya S, Hwang JY. Oppositional Brain Storm Optimization with Deep Learning based Facial Emotion Recognition for Autonomous Intelligent Systems. IEEE Access. 2024 Mar 7.

[24] Trabelsi Z, Alnajjar F, Parambil MM, Gochoo M, Ali L. Real-time attention monitoring system for classroom: A deep learning approach for student's behavior recognition. Big Data and Cognitive Computing. 2023 Mar 9;7(1):48.

[25] Ahadit AB, Jatoth RK. A novel multi-feature fusion deep neural network using HOG and VGG-Face for facial expression classification. Machine Vision and Applications. 2022 Jul;33(4):55.

[26] Abdollahi H, Mahoor MH, Zandie R, Siewierski J, Qualls SH. Artificial emotional intelligence in socially assistive robots for older adults: a pilot study. IEEE Transactions on Affective Computing. 2022 Jan 18;14(3):2020-32.

[27] Ballesteros JA, Ramírez V GM, Moreira F, Solano A, Pelaez CA. Facial emotion recognition through artificial intelligence. Frontiers in Computer Science. 2024 Jan 31;6:1359471.

[28] Zhao Z, Wang Y, Xu Y, Zhang J. TDFNet: Transformer-based Deep-scale Fusion Network for Multimodal Emotion Recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2023 Sep 18.

[29] Deng S, Sun Y, Galvan E. Neural architecture search using genetic algorithm for facial expression recognition. InProceedings of the Genetic and Evolutionary Computation Conference Companion 2022 Jul 9 (pp. 423-426).

[30] Salama ES, El-Khoribi RA, Shoman ME, Shalaby MA. A 3D-convolutional neural network framework with ensemble learning techniques for multi-modal emotion recognition. Egyptian Informatics Journal. 2021 Jul 1;22(2):167-76.

[31] Choi DY, Song BC. Facial micro-expression recognition using two-dimensional landmark feature maps. IEEE Access. 2020 Jul 3;8:121549-63.

[32] Standen PJ, Brown DJ, Taheri M, Galvez Trigo MJ, Boulton H, Burton A, Hallewell MJ, Lathe JG, Shopland N, Blanco Gonzalez MA, Kwiatkowska GM. An evaluation of an adaptive learning system based on multimodal affect recognition for learners with intellectual disabilities. British Journal of Educational Technology. 2020 Sep;51(5):1748-65.

[33] Zhao X, Chen G, Chuang Y, Tao X, Zhang S. Learning expression features via deep residual attention networks for facial expression recognition from video sequences. IETE Technical Review. 2021 Nov 2;38(6):602-10.

[34] Kim SH, Nguyen NA, Yang HJ, Lee SW. eRAD-Fe: Emotion recognition-assisted deep learning framework. IEEE Transactions on Instrumentation and Measurement. 2021 Sep 23;70:1-2.

[35] Huang Z, Ma Y, Wang R, Li W, Dai Y. A model for EEG-based emotion recognition: CNN-BI-LSTM with attention mechanism. Electronics. 2023 Jul 22;12(14):3188.

[36] Alemu Y, Chen H, Duan C, Caulley D, Arriaga RI, Sezgin E. Detecting clinically relevant emotional distress and functional impairment in children and adolescents: protocol for an automated speech analysis algorithm development study. JMIR Research Protocols. 2023 Jun 23;12(1):e46970.

[37] Albraikan AA, Alzahrani JS, Alshahrani R, Yafoz A, Alsini R, Hilal AM, Alkhayyat A, Gupta D. Intelligent facial expression recognition and classification using optimal deep transfer learning model. Image and Vision Computing. 2022 Dec 1;128:104583.

[38] Zhu H, Hu P, Tang X, Xia D, Huang H. NAGNet: A novel framework for real-time students' sentiment analysis in the wisdom classroom. Concurrency and Computation: Practice and Experience. 2023 Oct 10;35(22):e7727.

[39] Jiang P, He H. EEG Signal Analysis based on machine learning in psychological balance detection of athletes. Evolutionary Intelligence. 2023 Oct;16(5):1533-8.

[40] Chang W, Xu L, Yang Q, Ma Y. Eeg signal-driven human–computer interaction emotion recognition model using an attentional neural network algorithm. Journal of Mechanics in Medicine and Biology. 2023 Oct 31;23(08):2340080.

[41] Ciolacu MI, Svasta P. Education 4.0: AI empowers smart blended learning process with biofeedback. In2021 IEEE Global Engineering Education Conference (EDUCON) 2021 Apr 21 (pp. 1443-1448). IEEE.

[42] Kim JW, Choi YL, Jeong SH, Han J. A care robot with ethical sensing system for older adults at home. Sensors. 2022 Oct 3;22(19):7515.

[43] Rico Garcia OD, Fernandez Fernandez J, Becerra Saldana RA, Witkowski O. Emotion-Driven Interactive Storytelling: Let Me Tell You How to Feel. InInternational Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar) 2022 Apr 15 (pp. 259-274). Cham: Springer International Publishing.

[44] Hsieh YZ, Lin SS, Luo YC, Jeng YL, Tan SW, Chen CR, Chiang PY. ARCS-assisted teaching robots based on anticipatory computing and emotional big data for improving sustainable learning efficiency and motivation. Sustainability. 2020 Jul 12;12(14):5605.

[45] Liu S, Liu S, Liu Z, Peng X, Yang Z. Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement. Computers & Education. 2022 May 1;181.

[46] Jafari M, Shoeibi A, Khodatars M, Bagherzadeh S, Shalbaf A, García DL, Gorriz JM, Acharya UR. Emotion recognition in EEG signals using deep learning methods: A review. Computers in Biology and Medicine. 2023 Sep 9:107450.