

Detection of range and degrees of emotion through facial Image analysis using Machine Learning for understanding students' feelings

**Shatadru Sengupta¹, Chayan Paul², Apratim Mitra³,
Somsubhra Gupta⁴**

¹Swami Vivekananda University
shatadru.sengupta@gmail.com

²Swami Vivekananda University
chayanp@svu.ac.in

³Swami Vivekananda University
apratim777@yahoo.com

⁴Swami Vivekananda University
gsomsubhra@gmail.com

Abstract

This work concentrates on two issues. Firstly, to develop an emotion detection system based on Machine Learning algorithms applied to facial and allied analyses, in particular, detecting the dilation features of the pupil of the eye, so that various degrees of emotions and attentiveness can be detected. This would answer questions like, in a classroom teaching environment, how attentive is a student – that is to what degree or percentage – in a class, how happy, sad or scared – once again to what degree or percentage – a person is in response to an event or in an environment, and questions like that. Secondly, the system should be able to detect these emotions as displayed by a group of people, and not just an individual. As it is seen in existing literature, facial analysis research has concentrated on many aspects like social norms, refinement of analysis, reduction of noise etc. However, the problem that this research is addressing has dealt with in very few instances of research. It may thus be concluded that in the issue of mass scale emotion detection, a lot of research is

yet to come. It is hoped that the findings of this research will pave way to solving problems like understanding student behaviour in a teaching environment, understanding crowd behaviour in an open environment, understanding audience emotions in response to a piece of art, and so on.

Keywords : *Emotion detection, facial analysis, pupil dilation, pupil dilation analysis, Machine Learning, Neural Networks.*

ICTIEE Track: Technology Enhanced Learning

ICTIEE Sub-Track: Transforming Education Through Technology: Best Practices and Case Studies

Introduction

Emotion detection is both quite instinctive and rather inherent to human beings when they communicate with one another. In fact, effective communication between two or more human beings is very much founded in understanding and reacting to the reactions of one another, and it is seen that those who can understand others' emotions can easily steer a conversation in their willed way.

However intelligent and instinctive human beings may be, they more often than not fall short of skills when it comes to understanding the masses. If a classical example of a teacher teaching a number of students in a class is considered, it is impossible for the teacher to catch the reactions of each and every student because it is awfully hard, even impossible, to keep track of how each and every student is emoting. It is to be noted here, though, that students will react through emoting to the content of the lecture on the basis of how they feel about the content, the lecture delivery, the time of lecture, etc, which are all indeed important for the teacher to keep track of if s/he wishes to connect and reach out to the students in order to truly help them learn. Yet, in large samples of students, it is found to be difficult to keep track of their reactions.

This is where the use of Machine Learning (ML) in emotion detection will come into play, as has been indicated in the following section, Background, Literature Review and Analysis.

Background Study, Literature Review And Analysis

Background

Historically, The study of facial expressions and their connection to emotions has roots in the work of Charles Darwin and Paul Ekman, who identified universal emotions and their corresponding facial expressions. Darwin's seminal work, "The Expression of the Emotions in Man and Animals" (1872), laid the groundwork for understanding facial expressions as a means of non-verbal communication. Ekman's identification of six basic emotions (happiness, sadness, fear, disgust, anger, and surprise) further advanced this field by providing a framework for categorizing facial expressions.

Literature Review

(Saxena, Khanna and Gupta, 2020) have accounted for the performance of various computing techniques based on whether emotion detection was done on the basis of, apart from text records, facial changes, physiological properties, as well as speech. Apart from this several other pieces of research exist that talk about emotion detection in general. (Alslaity and Orgi, 2022) have also dealt with such issues in general.

The results also showed that a significant amount of work exists in the field of emotion detection from the point of view of analysis of dialogues represented textually.

One of the most general treatises on emotion detection has been presented (Khan et al., 2020) by Saxena et al. writing about the multifarious ways and modes in which researchers have tried to capture and analyse human emotion.

(Zad et al, 2021) have stated that using certain psychological models, automatic Machine Learning (ML) based detection of hidden emotions in dialogues, in the form of texts, is an important trend in Natural Language Processing (NLP). (Nandwani et al. 2021) say that, for various purposes, sentiment analysis of people - regarding their happiness, frustration, sorrow etc towards any topic of discussion must be captured, and where sentiment analysis falls

short, emotion detection can be done to predict with some authority what peoples' moods are, And these analyses can be done through analysing textual information, as found in written dialogues over the Internet, for example. Citing the need for understanding and predicting customer moods, (Acheampong et al. 2021) recommend text-based detection of customer emotions from feedbacks etc.

(Guo, 2022) demonstrated a high degree of accuracy in predicting emotion embedded in texts using a combination of deep learning and natural language processing. Yet again, Acheampong et al. venture to list, describe and evaluate critically several BERT-based approaches to text-based emotion detection (Acheampong et al., 2021). Sifting through more than 800 scopus and more than 80 Web of Science papers, Kusal et al. have described various AI driven approaches to emotion detection in big data especially from social media textual conversations (Kusal et al, 2021). (Zhu and his associates 2021) also talk about transformer-driven approaches to emotion detection from dialogues.

(Mao and his companions 2022) have critically studied Pre-Trained Language Models (PLMs) in detection of emotion from prompts. Many other papers including (Ali et al, 2022; Deng and Rewn, 2023; Imran et al, 2020; Kabir and Madria, 2021; Heredia et al, 2022; Li et al, 2020; Krommyda et el, 2021; Zhang et al, 2020; Bashir et al, 2023; Aslam et al, 2022), to name just a few, readily speak in the lines described just - dealing with textual or conversational data. (Tsfagergish et al. 2022) have done the same thing, using a two pronged approach.

Some of the research, however, dipped hands in providing a multimodal view – in describing how many different ways or sources can be used in the analysis of emotions. For example,], (Zhang, Pin and their colleagues2022) have dealt with multimodal techniques in order to do emotion detection, and so have (Shenoy et al. 2020) who have advocated that in order to detect the emotion of persons, a multimodal approach is best chosen, and they have devised a solution involving end to end RNNs to that end, which, so they claim, outdid existing multimodal methods at the time of the paper by addressing existing drawbacks. (Abdullah et al. 2021) have

also dealt with the topic of multimodal emotion detection. (Huddar et al., 2020) advocate an RNN based approach to deal with interlocutor and contextual states in conversations; their work, again, focus on the multimodal approach. This also extends to (Tu et al., 2024).

A few other searches worth mentioning are those that involved EEG (Electro Encephalogram) recordings. For example, (Ding et al. 2020) had come up with “TCeption” – their own deep learning framework for analysing emotions from subjects’ EEG data, and it was claimed in this paper that TCeption was performing better than similar approaches. (Liu et al., 2024) have worked on multimodal as well as EEG approaches to emotion detection.

Since the main focus of our research was face-based emotion detection, the next string that was used was “emotion detection on a mass scale through facial analysis”. Many research paper exist for facial detection, of which a very few are listed below as examples.

(Shahzad et al. 2023) go on to prove that while doing emotion detection through Deep Convolutional Neural Networks, if the classifiers of each layer is substituted by an SVM or ensemble classifier, the results improve by 7-9% per layer.

(Chavali and others 2023) propose a system of telling a person’s gender and age through facial analysis and emotion recognition.

(Kalyta et al. 2023) have devised a model to predict the sudden emotional changes in a subject under study.

The mass capture of emotions seems to be a not-so-treaded area. However, a part of the papers that seem to be very relevant for our research are as follows.

(Mukhopadhyay et al., 2020) have worked on continuous image frames from online classes to detect what emotion and what state of mind is exhibited by the students, and have respectively achieved 62% and 65% accuracy.

(Gupta, Kumar and Tekchandani 2023) performed real-time facial analysis on students of online classes to see whether

or not they were engaging in the learning activities. Using various deep learning models on a number of datasets in a system of their own devise, the authors have reported significant gains in accuracy.

(Savchenko, et al.2023) addressed the difficulty of emotion detection in videos for want of proper and clean-enough labels in databases for this purpose, and has proposed a combination of multi-layered perceptron and LightAutoML classifier to tackle the situation.

(Savchenko,2022) looks into the issue of students' emotion detection in online teaching-learning environments. The authors refer to their work on recognising student emotions through a single neural network, after a certain level of processing. In this paper, it is claimed that the system is able to detect (a) the degree of involvement in the lectures, (b) the type of emotion and (c) nature of effect to the study group.

Although (Zou et al.2023) worked on large-scale videos in order to establish context-to-facial expression correlation in order to tackle confusions in emotion detection, this work does not fit the bill when it comes to detection of emotion of the masses. However, their ideas can be borrowed and perhaps can be scaled up in establishing emotions in the reactions of the masses.

(Castellano et al., 2023), has prepared an FER based system for emotion detection on masked faces. This study perhaps can be used in our research as masses will include masked faces.

(Alisawi et al., 2023) provide a valuable insight into the research on real-time emotion recognition done in the last few years.

(Bisogni et al.,2023) present a comparative study and conducts experiments that veer towards recognition of emotions from a distance.

(Atymtayeva et al.,2023) describe methods for rapid detection of emotions from candidates' facial contours.

Analysis

It is seen in the Literature Review section that, a few ML techniques have been used for detecting a range of emotions in general, analysing facial dynamics of the human subjects, electro encephalographs, textual information as evident from conversations and the like. There is a good scope of improving on the facial analyses aspects to increase the range and the degree of emotions that can be detected for individual subjects. It can then be explored if this same technology can be extended to include groups of subjects too.

On the other hand, if it is enquired: how reliable is facial expression on its own, in telling emotions correctly, then it is found to be somewhat lacking in the sense, as per common knowledge goes, people sometimes make all sorts of expressions that do not really match with the situation at hand. For example, the loss of one's home cricket team ought to make people make a sad face, but many a times people laugh or smile sarcastically and joke about it to lighten the grief or even to emphasise the ignominy in the loss. It would then be impossible for a machine learning algorithm to decide if the people were not laughing out of joy or happiness or merriment. (Huddar et al. 2021) presented a multimodal model by referring to interlocutor and context base emotion detection, albeit in textual data. It would then seem that the context of the home team loosing being put into view, the laughter would be deemed to have happened out of sarcasm and not out of joy.

However, in a classroom environment, when the task ITSELF is to determine the responsiveness and attention students are paying to a particular topic, or teaching style and mode and environment, or time of the day, etc, such contextual considerations are irrelevant.

In this paper we look at pupil dilation as an allied means to supplement facial analysis finding. To this effect, one or two noteworthy pieces of work may be cited.

It is noted by (Hershman, 2023) that pupil dilation is a direct function of cognitive engagement, a finding that is very relevant to the research that is being undertaken.

(Baltacti et al., 2016), resorting to an amalgamation of decision tree and Adaboost Random Forest (ABRF), have worked on stress detection in computer users, where they have used pupil dilation as one of the methods.

(Khan and others, 2020) report using pretrained models and multiple strata of convolution to achieve pupil detection in video pictures, and claim to have gained more significant results than other efforts of that time.

Thus, these three pieces of work gives enough incentives to pursue pupil dilation as a supplementary way to be double sure that our emotion detection, with respect to a student reacting to the prevalent teaching environment, is as far accurate as can be.

The accuracy shown and reported by the most relevant pieces of work for the purposes of this paper is given below:

Author and research referred to	What to measure	Technology	Accuracy Measure and/or performance comments
Shahzad et al. 2023	Emotion detection	SVM or Ensemble classifier	7 – 9% more than relevant technology of the day
Chavali et al 2023	Emotion detection	SVM	>36%
		KNN	>36%, <35%
		CNN	>84%
		VGG-16	>97%
Kalyta et. al 2023	Sudden emotional changes	Their own classifier	> 82%, on an average 0.9 to 2% on common indicators than state of the art models
Mukhopadhyay et al., 2020	Emotion detection	CNN	62%
	State of mind	CNN	65%

Gupta, Kumar and Tekchandani 2023	Student states online: engaged or disengaged	Various classifiers: InceptionV3, VGG19, ResNet-50	>89%, >90%, >92% respectively
Castellano et al., 2023	Emotion detection from masked faces	FER Based search, deep learning techniques	14% greater than if the same task was done by humans
Bisogni et al., 2023	Recognising emotions from far away	Manually aided ML model, DL model	Accuracy of manually aided ML model is 0.15 higher than DL model
Hershman et al., 2023	Pupillometry	Temporal methods over traditional single value methods	Improvement seen in temporal methods is greater
Baltci et al. 2016	Computer related stress using pupillometry and rise in temperature	Decision Tree, Adaboost with Random Forest	Stress detection accuracy – 65 – 83.8%, sensitivity 83.9% and specificity 83.8%
Khan and others, 2020	Pupil localisation and eye centre location	DL based model to extract eye frame; Image processing based iris and pupil detection	Their proposed model has worked better than other models

We can see from Table 2.1 that emotion detection can be largely thought of as best done through Convolutional Neural Networks. Although a fair amount of image processing is needed for the detection of pupils and the degree of dilation of the pupils, there will be need to classify the various degrees of dilation that can be detected, in order to pair them with facial emotion detected, to predict the actual feeling of the student.

Proposed Model

It is clear that this study aims to give accurate feedback regarding classes to the concerned teachers and mentors. Before a proposed model is designed, the following assumptions are to be made:

(i) *The videos acquired during classroom teaching NEED NOT be analysed in real-time.* The feedback of the class is to be taken after the class is over, and instead of feedback forms, the class video is to be perused.

(ii) *Moving pictures need not be analysed for feedback, only frame by frame analyses are sufficient.* The data for emotion detection and analysis can be understood by slicing the video feed in frames and then analysing them.

(iii) *The classes into which the final results can be assigned can be made such that they are mutually exclusive and exhaustive.* This is done so that there is no confusion in the classification. There may be several overlapping features in each of the results. However, for avoiding clashes and mixes, each of the predictions will be deemed separate and not overlapping. For example, if A, B, C, D are features, and we consider two predictions: P1(A-20%, B-40%, C-40%, D-20%) and P2(A-30%, B-30%, C-30%, D-10%), then they will be considered as separate groups. This makes sense: although every frame of emotion that a human being exhibits contains bits and pieces of other frames possible, each frame refers to (or results in) a particular state of mind, and hence may easily be counted as separate from any other frame of emotion. In fact, leading on, each of the system results will form an Entropy of Information = 0 (pure) by design itself.

As such, the system is proposed as follows:

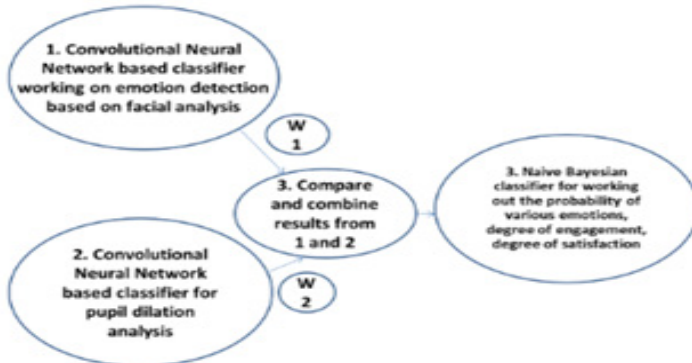


Fig. 3.1: Network diagram of the proposed emotion detection system

Explanation of (Figure 3.1):

Before explaining 3.1, an important concept regarding this work has to be explained. That concept is *CLASSES OF FEELINGS* (CF): These refer to the conclusion about students' feelings in a particular frame. These will have to be manually constructed by analysing the stress and emotion levels of the students and predicting the possible state the students must be in.

Subsystem 1:

- (i) Convolutional Neural Network based classifier
- (ii) Trains on available data sets as well as datasets that may be prepared from video frames of classroom footages
- (iii) Detects various degrees of various facial expressions, as possible indicators of emotions.
- (iv) Its output is combined with a weight W_1 indicative of how important its output is in the scheme of things.
- (v) Algorithm: The abstract idea or algorithm of how the Subsystem 1 should work is given in the following figure:

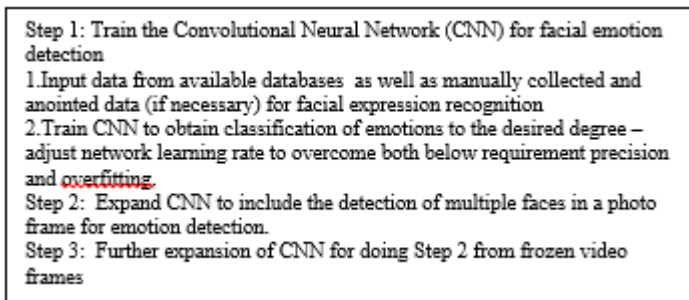


Figure 3.2: Abstracted algorithm for Subsystem 1

Subsystem 2:

- (i) Convolutional Neural Network based model to detect and analyse various degrees of pupil dilation.
- (ii) Its output is combined with a weight W_2 to determine its importance in the total calculation.
- (iii) Algorithm

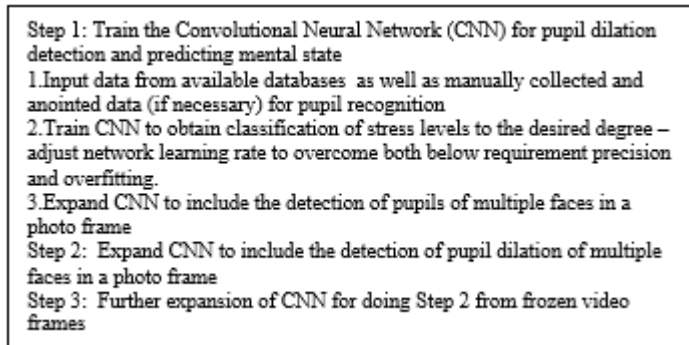


Figure 3.3: Abstracted algorithm for Subsystem 2

Subsystem 3:

- (i) Compares and combines (1, W_1) and (2, W_2) to form dataset D of tuples 1 to N . Let each tuple be designated as $T_1, T_2, \dots, T_n$ such that: Each T_i consists of values to attributes $A_1, \dots, A_m$ of the dataset. There are separate classes of results as per assumption (iii). For all $i \leq k$, C_i is mutually exclusive

and exhaustive to C_{i+1} . K = number of classes.

(ii) Algorithm:

Step 1: Take a result from Subsystem 1 ($ss1$) and pair it with the weight W_1 ($ss1, W_1$) signifying how important it is in the final result.
Step 2: Take a result from Subsystem 2 ($ss2$) and pair it with the weight W_2 ($ss2, W_2$) signifying how important it is in the final result.
Step 3: Form a Conditions Table (CT)
Step 4: For each unique combination of $((ss1, W_1), (ss2, W_2))$ form a table row entry in the CT.

Figure 3.4: Abstracted algorithm for Subsystem 3

Subsystem 4:

(i) A Naïve Bayesian Classifier that will map on the observed attributes to a particular state of mind in the class, with regards to attentiveness, liveliness, engaged, somewhat engaged, disengaged, etc

(ii) Algorithm

Step 1: From the Conditions Table (CT) (refer to Figure 3.4) pick each row, use a Naïve Bayesian Classifier and classify each of the $((ss1, W_1), (ss2, W_2))$ to a class in the CF set.
Step 2: Count the number of instances of each class of feelings
Step 3: Calculating the average of the types of feelings, judge whether the frame presents a generally happy picture, or a generally sad picture, or whether the picture represents indifference, or so on.

Figure 3.5: Abstracted algorithm for Subsystem 4

System Testing

The System is to be tested in the following way with the help of confusion matrices:

(i) The primary checking is to be done using the Classes of Feelings (CF) that is to be made using human analysis. If the predictions are correct according to the CF it can be considered as a relatively well-working system that conforms to the human concepts of the feelings.

(ii) There also needs to be some checking done on data manually acted out by students who would be deployed to participate in the experiments. The students would be asked

to opine on the particular classes, their feelings analysed from the set of emotions laid down in questionnaires. These feelings will be the same as those that will be used to train the CNNs and the Naïve Bayesian Classifier.

(iii) Training of the subsystems 1 and 2 will be done by usual methods, forming confusion matrices against the training data for the concerned CNNs.

Design Analysis

The following points are to be observed:

(i) Use of convolutional neural networks renders the classifications very accurate, leaving room for fine tuning of the networks as per performance,

(ii) Weighted outputs and weight-based combination of 1 and 2 further allows “the mixing of the right ingredients”. The weights W_1 and W_2 maybe started off as 0.5 and 0.5 respectively, to show equal weightage to both facial analysis and pupil dilation. However, as the experiments keep running, it will become more and more decidable as to which subsystem is poised to have a better say in the predictions, and here, the weights will play a crucial role.

(iii) The nature of forming the final data set before final classification at Subsystem 4 ensures least amount of entropy at the dataset before classification.

(iv) Use of a naïve Bayesian classifier ensures that all the probabilities have been taken into consideration before a particular emotion and state of attentiveness is declared.

(v) Once the two subsystems 1 & 2 have been trained on frozen frames from the classroom footage, they may further be developed to train on real-time data.

A valid question also is on the quality or resolution of the CCTV cameras that will enable the capture of the pupil as well as the facial emotions. It must be noted that in order to develop, train and refine the technology preliminarily, it is not required to use extremely high end frames. Development can occur in experimental settings, with the last stages of development done with data frames from live feeds.

Conclusion

In this paper, a preliminary design of an emotion detecting and classifying unit is provided that is aimed at aiding in understanding students' emotions and attentiveness in classes better by analysing facial expressions and pupil dilation of the eye. The approach uses Convolutional Neural Networks as primary classifiers individually for facial expressions and pupil dilation, combines the results into low entropy datasets, and finally uses a Naïve Bayesian Classifier to formulate the results. Although without actually concretising the set up and running the experiments, only guesses maybe made on the technology, the design appears to be theoretically sound and the problem it targets is definitely worth pursuing.

References:

S. Zad, M. Heidari, J. H. J. Jones and O. Uzuner, "Emotion Detection of Textual Data: An Interdisciplinary Survey," 2021 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2021, pp. 0255-0261, doi: 10.1109/AIIoT52608.2021.9454192.

Nandwani, P., Verma, R. "A review on sentiment analysis and emotion detection from text" Soc. Netw. Anal. Min. 11, 81 (2021). <https://doi.org/10.1007/s13278-021-00776-6>

Francisca Adoma Acheampong, Chen Wenyu, Henry Nunoo-Mensah. "Text-based emotion detection: Advances, challenges, and opportunities". Wiley Online Library, Engineering Reports, 28 May 2020 <https://doi.org/10.1002/eng2.12189>

Guo, Jia. "Deep learning approach to text analysis for human emotion detection from big data" Journal of Intelligent Systems, vol. 31, no. 1, 2022, pp. 113-126. <https://doi.org/10.1515/jisys-2022-0001>

Acheampong, F.A., Nunoo-Mensah, H. & Chen, W. Transformer models for text-based emotion detection: a review of BERT-based approaches. Artif Intell Rev 54, 5789-5829 (2021). <https://doi.org/10.1007/s10462-021-09958-2>

Kusal, S.; Patil, S.; Kotecha, K.; Aluvalu, R.; Varadarajan, V. "AI Based Emotion Detection for Textual Big Data: Techniques and Contribution". Big Data Cogn. Comput. 2021, 5, 43. <https://doi.org/10.3390/bdcc5030043>

Lixing Zhu, Gabriele Pergola, Lin Gui, Deyu Zhou, Yulan He. "Topic-Driven and Knowledge-Aware Transformer for Dialogue Emotion Detection". arXiv:2106.01071, 2 Jun 2021. <https://doi.org/10.48550/arXiv.2106.01071>

R. Mao, Q. Liu, K. He, W. Li and E. Cambria, "The Biases of Pre-Trained Language Models: An Empirical Study on Prompt-Based Sentiment Analysis and Emotion Detection," in *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 1743-1753, 1 July-Sept. 2023, doi: 10.1109/TAFFC.2022.3204972.

Anvita Saxena, Ashish Khanna, Deepak Gupta (2020). "Emotion Recognition and Detection Methods: A Comprehensive Survey". *Journal of Artificial Intelligence and Systems*, 2, 53–79. <https://doi.org/10.33969/AIS.2020.21005>

Ali I. Siam, Naglaa F. Soliman, Abeer D. Algarni, Fathi E. Abd El-Samie, Ahmed Sedik. "Deploying Machine Learning Techniques for Human Emotion Detection". *Hindawi, Computational Intelligence and Neuroscience*, Volume 2022, Article ID 8032673, 16 pages, <https://doi.org/10.1155/2022/8032673>

J. Deng and F. Ren, "Multi-Label Emotion Detection via Emotion-Specified Feature Extraction and Emotion Correlation Learning," in *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 475-486, 1 Jan.-March 2023, doi: 10.1109/TAFFC.2020.3034215.

W. Nie, Y. Bao, Y. Zhao and A. Liu, "Long Dialogue Emotion Detection Based on Commonsense Knowledge Graph Guidance," in *IEEE Transactions on Multimedia*, vol. 26, pp. 514-528, 2024, doi: 10.1109/TMM.2023.3267295.

A. S. Imran, S. M. Daudpota, Z. Kastrati and R. Batra, "Cross-Cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets," in *IEEE Access*, vol. 8, pp. 181074-181090, 2020, doi: 10.1109/ACCESS.2020.3027350.

Md. Yasin Kabir, Sanjay Madria. "EMOCOV: Machine learning for emotion detection, analysis and visualization using COVID-19 tweets". *Online Social Networks and Media*, Volume 23, 2021, 100135, ISSN 2468-6964, <https://doi.org/10.1016/j.osnem.2021.100135>. (<https://www.sciencedirect.com/science/article/pii/S2468696421000197>)

J. Heredia et al., "Adaptive Multimodal Emotion Detection Architecture for Social Robots," in *IEEE Access*, vol. 10, pp. 20727-20744, 2022, doi: 10.1109/ACCESS.2022.3149214.

Jingye Li, Donghong Ji, Fei Li, Meishan Zhang, and Yijiang Liu. 2020. HiTrans: A Transformer-Based Context- and Speaker-Sensitive Model for Emotion Detection in Conversations. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4190-4200, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Krommyda, M.; Rigos, A.; Bouklas, K.; Amditis, A. An Experimental Analysis of Data Annotation Methodologies for Emotion Detection in Short Text Posted on Social Media. *Informatics* 2021, 8, 19. <https://doi.org/10.3390/informatics8010019>

X. Zhang, W. Li, H. Ying, F. Li, S. Tang and S. Lu, "Emotion Detection in Online Social Networks: A Multilabel Learning Approach," in *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8133-8143, Sept. 2020, doi: 10.1109/JIOT.2020.3004376.

Muhammad Farrukh Bashir, Abdul Rehman Javed, Muhammad Umair Arshad, Thippa Reddy Gadekallu, Waseem Shahzad, and Mirza Omer Beg. 2023. Context-aware Emotion Detection from Low-resource Urdu Language Using Deep Neural Network. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* 22, 5, Article 131 (May 2023), 30 pages. <https://doi.org/10.1145/3528576>

N. Aslam, F. Rustam, E. Lee, P. B. Washington and I. Ashraf, "Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model," in *IEEE Access*, vol. 10, pp. 39313-39324, 2022, doi: 10.1109/ACCESS.2022.3165621.

Aman Shenoy, Ashish Sardana. "Multilogue-Net: A Context Aware RNN for Multi-modal Emotion Detection and Sentiment Analysis in Conversation". *Challenge-HML, ACL* 2020, 19-28. <https://arxiv.org/abs/2002.08267v3> <https://doi.org/10.18653/v1/2020.challengehml-1.3>

Jianhua Zhang, Zhong Yin, Peng Chen, Stefano Nichele. "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review". *Information Fusion*, Volume 59, 2020, Pages 103-126, ISSN 1566-2535, <https://doi.org/10.1016/j.inffus.2020.01.011>. (<https://www.sciencedirect.com/science/article/pii/S1566253519302532>)

Abdullah, Sharmeen M. Saleem Abdullah, et al. "Multimodal emotion recognition using deep learning." *Journal of Applied Science and Technology Trends* 2.01 (2021): 73-79.

Huddar, Mahesh G.; Sannakki, Sanjeev S.; Rajpurohit, Vijay S. "Attention-based Multi-modal Sentiment Analysis and Emotion Detection in Conversation using RNN". 06/2021. <https://doi.org/10.9781/ijimai.2020.07.004>

Gannouni, S., Aledaily, A., Belwafi, K. et al. "Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification". *Sci Rep* 11, 7071 (2021). <https://doi.org/10.1038/s41598-021-86345-5>

Tesfagergish, S.G.; Kapočiūtė-Dzikiene, J.; Damaševičius, R. Zero-Shot Emotion Detection for Semi-Supervised Sentiment Analysis Using Sentence Transformers and Ensemble Learning. *Appl. Sci.* 2022, 12, 8662. <https://doi.org/10.3390/app12178662>

Y. Ding et al., "TSception: A Deep Learning Framework for Emotion Detection Using EEG," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, pp. 1-7, doi: 10.1109/IJCNN48605.2020.9206750.

Alsiaity, A., & Orji, R. (2022). Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions. *Behaviour & Information Technology*, 43(1), 139-164. <https://doi.org/10.1080/0144929X.2022.2156387>

Tu, G., Xie, T., Liang, B., Wang, H., & Xu, R. (2024). Adaptive Graph Learning for Multimodal Conversational Emotion Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(17), 19089-19097. <https://doi.org/10.1609/aaai.v38i17.29876>

Zhiwei Liu, Tianlin Zhang, Kailai Yang, Paul Thompson, Zeping Yu, Sophia Ananiadou. "Emotion detection for misinformation: A review". *Information Fusion*, Volume 107, 2024, 102300, ISSN 1566-2535, <https://doi.org/10.1016/j.inffus.2024.102300>. (<https://www.sciencedirect.com/science/article/pii/S1566253524000782>)

Shahzad, H.M.; Bhatti, S.M.; Jaffar, A.; Akram, S.; Alhajlah, M.; Mahmood, A. Hybrid Facial Emotion Recognition Using CNN-Based Features. *Appl. Sci.* 2023, 13, 5572. <https://doi.org/10.3390/app13095572>

Surya Teja Chavali, Charan Tej Kandavalli, T M Sugash, R Subramani. "Smart Facial Emotion Recognition With Gender and Age Factor Estimation". *Procedia Computer Science*, Volume 218, 2023, Pages 113-123, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2022.12.407>. (<https://www.sciencedirect.com/science/article/pii/S1877050922025017>)

Kalyta, O.; Barmak, O.; Radiuk, P.; Krak, I. Facial Emotion Recognition for Photo and Video Surveillance Based on Machine Learning and Visual Analytics. *Appl. Sci.* 2023, 13, 9890. <https://doi.org/10.3390/app13179890>

Moutan Mukhopadhyay, Saurabh Pal, Anand Nayyar, Pijush Kanti Dutta Pramanik, Niloy Dasgupta, and Prasenjit Choudhury. 2020. "Facial Emotion Detection to Assess Learner's State of Mind in an Online Learning System." In *Proceedings of the 2020 5th International Conference on Intelligent Information Technology (ICIIT '20)*. Association for Computing Machinery, New York, NY, USA, 107-115. <https://doi.org/10.1145/3385209.3385231>

Gupta, S., Kumar, P. & Tekchandani, R.K. "Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models". *Multimed Tools Appl* 82, 11365-11394 (2023). <https://doi.org/10.1007/s11042-022-13558-9>

Andrey V. Savchenko. "EmotiEffNet Facial Features in Uni-task Emotion Recognition in Video at ABAW-5 competition". *arXiv:2303.09162*, Mar 2023. doi: <https://doi.org/10.48550/arXiv.2303.09162>

A. V. Savchenko, L. V. Savchenko and I. Makarov, "Classifying Emotions and Engagement in Online Learning Based on a Single Facial Expression Recognition Neural Network," in *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 2132-2143, 1 Oct.-Dec. 2022, doi: 10.1109/TAFFC.2022.3188390.

Zhou, S.; Wu, X.; Jiang, F.; Huang, Q.; Huang, C. Emotion Recognition from Large-Scale Video Clips with Cross-Attention and Hybrid Feature Weighting Neural Networks. *Int. J. Environ. Res. Public Health* 2023, 20, 1400. <https://doi.org/10.3390/ijerph200000021400>

Castellano, G., De Carolis, B. & Macchiarulo, N. Automatic facial emotion recognition at the COVID-19 pandemic time. *Multimed Tools Appl* 82, 12751-12769 (2023). <https://doi.org/10.1007/s11042-022-14050-0>

M. ALISAWI and N. YALÇIN, "Real-Time Emotion Recognition Using Deep Learning Methods: Systematic Review", *Intell Methods Eng Sci*, vol. 2, no. 1, pp. 5–21, Mar. 2023. DOI: <https://doi.org/10.58190/imiens.2023.7>

Carmen Bisogni, Lucia Cimmino, Maria De Marsico, Fei Hao, Fabio Narducci. "Emotion recognition at a distance: The robustness of machine learning based on hand-crafted facial features vs deep learning models". *Image and Vision Computing*, Volume 136, 2023, 104724, ISSN 0262-8856, <https://doi.org/10.1016/j.imavis.2023.104724>. (<https://www.sciencedirect.com/science/article/pii/S0262885623000987>)

Atymtayeva, L., M. Kanatov, A. M. A. Musleh, and G. Tulemissova. "Fast facial expression recognition system: selection of models." *Appl. Math* 17, no. 2 (2023): 375-383.

Hershman, R., Milshtein, D. & Henik, A. The contribution of temporal analysis of pupillometry measurements to cognitive research. *Psychological Research* 87, 28–42 (2023). <https://doi.org/10.1007/s00426-022-01656-0>

Baltaci, S., & Gokcay, D. (2016). Stress Detection in Human-Computer Interaction: Fusion of Pupil Dilation and Facial Temperature Features. *International Journal of Human-Computer Interaction*, 32(12), 956–966. <https://doi.org/10.1080/10447318.2016.1220069>

Khan W, Hussain A, Kuru K, Al-askar H. Pupil Localisation and Eye Centre Estimation Using Machine Learning and Computer Vision. *Sensors*. 2020; 20(13):3785. <https://doi.org/10.3390/s20133785>