Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 10s (2025)

Context-Based Emotion Recognition using the EMOTIC Dataset

Arvind Laxman Kedar[1], Prof. Dr. Monika Rokade[2], Prof. Dr. Sunil Khatal[3]

[1]Department of Computer Engineering, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune, India

arvindkedar00@gmail.com

[2]Department of Computer Engineering, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune, India

monikarokade4@gmail.com

[3]HOD, Department of Computer Engineering, Sharadchandra Pawar College of Engineering, Dumbarwadi, Otur, Pune, India

khatalsunils88gmail.com

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract:

Emotion recognition plays a vital role in human-computer interaction, enabling machines to better understand and respond to human emotions. Traditional methods primarily focus on facial expressions or speech for emotion detection. However, context plays a crucial role in accurately interpreting emotions. This paper explores a context-based approach to emotion recognition using the EMOTIC dataset, which includes images labelled with both apparent emotion and contextual information. By employing deep learning models that integrate contextual cues with traditional emotion recognition techniques, we aim to enhance the accuracy and reliability of emotion detection systems. The proposed model achieves significant improvements over baseline methods, demonstrating the importance of context in emotion recognition.

Keywords: Natural Language Processing, Customer relationship management, Machine Learning.

I. INTRODUCTION

Introduce the significance of emotion recognition in human-computer interaction, social robotics, and other fields. Discuss the limitations of traditional emotion recognition methods that rely solely on facial expressions or speech. Computerized sentiment analysis, using artificial intelligence and computer vision, has grown indispensable in recent years. Deep neural networks have come a long way, which means that this technology can now include environmental, social, and cultural factors along with facial emotions. Our objective is to develop more empathic systems for many applications, ranging from healthcare to analyzing emotional exchanges on social media. We used genuine photos from many databases, including EMOTIC, to train our models for the development of this technology. We devised two advanced algorithms using deep learning methodologies. Through the optimization of our models' hyperparameters, we examine context and body language to enhance our comprehension of human emotions in photographs. We integrate the 26 distinct emotional categories with the three continuous emotional dimensions to ascertain feelings under context. We finalize the suggested pipeline by integrating our models.

II. LITERATURE SURVEY

According to [1] method utilizes data from a camera, a wearable device, and physiological sensor signals such as photoplethysmography (PPG), electroencephalography (EEG), and galvanic skin response (GSR) to determine an individual's emotional state. This data is included as an enhancement to the music recommendation device. Sensor and facial expression data may enhance the efficacy and precision of the recommendation engine.

According to [2] a commercially available EEG Bluetooth headset that uses sensors to detect changes in brain waves (such as alpha and beta waves). A mobile device can securely transfer data using a Bluetooth transmission mechanism. The EEG signal may provide a wealth of information concerning a broad variety of cognitive disorders and illnesses. In addition to providing medical professionals with a reference for treatment, the EEG signal, categorization, and classification of depression level may be utilised as an assessment foundation for music therapy.

According to [3] developed an attention device that uses a specific electroencephalographic (EEG) method to record and assess attention based on individual music content preferences. We further filter the collected brainwave data values for inaccuracies before integrating them with an SVM classifier. The SVM classifier differentiates between two categories of brainwave data values: attention and non-attention, for computation and analysis. The construction of an attention mechanism has led to the development of a hybrid music recommendation model. It relies on an intriguing apparatus that records the user's musical preferences and electroencephalographic (EEG) data.

According to [4] The research used cerebral signals to investigate the effects of English and Urdu music on human stress levels. Twenty-seven individuals, including 14 men and 13 women, whose first language is Urdu and whose ages range from 20 to 35, participated in the research study. We record the patients' electroencephalographic (EEG) data as they listen to various music tracks utilizing a four-channel MUSE headset. The state and trait anxiety questionnaire requires participants to subjectively evaluate their stress levels. The four genres of English music that were utilised in this research are rock, metal, electronic, and rap.

According to [5] examines modified machine learning and deep learning methods, utilizing feature selection short of feature selection, to categorize six basic emotions into two groups: emotional states and non-emotional states. The emotions under consideration are anger, disgust, fear, pleasure, sorrow, and surprise. We achieve this classification by extracting physiological features and applying deep learning models to EEG data. This dataset has EEG, physiological, and video signal analyses for detecting emotions.

According to [6] ScientoPy is used by the system to do a scientometric analysis of important articles on recommendation approaches that come from scientific databases like Clarivate's Web of Science and Elsevier's Scopus. The review highlights processes including data collection, processing, and feature extraction from sensors and wearable devices to evaluate emotions. Key focus areas of the study include recommender systems, emotion detection, wearable technologies, and machine learning.

According to [7] Electroencephalography can quantify the alterations in brain activity caused by the listener's emotional reaction to music. A scoping review identified recent relevant research on the impact of music on cerebral activity and emotional conditions in digital therapeutic programs. We found 6 relevant studies from the 585 publications that met all the study's requirements.

According to [8] a brain-computer interface based on EEG that doesn't hurt the user lets them control a robotic arm and reach and grab multiple targets while getting around obstacles using hybrid control. Six individuals used a hybrid-control robotic arm device to do the online activities. The system demonstrates effectiveness by integrating MI-based EEG, computer vision, gaze tracking, & semi-autonomous navigation. This makes online work much more accurate and reduces the mental load that comes with long periods of mental work.

According to [9] after preprocessing, the preprocessed data are utilised to train the LSTM network, and then the Softmax function is used to categorise the input data into normal and seizure data. Normalization of EEG data, application of suitable filters to pick the useful sections of the data, and data management are the three phases of preprocessing. The normalisation of EEG data, the use of suitable filters to pick the important sections of the data, and data management are the three phases of preprocessing. After preprocessing, the preprocessed data are utilised to train the LSTM network, which is then used to categorise the input data into normal and seizure data using the Softmax function.

According to [10] they get the data from the user’s text on social media using an IoT. The text data will be analyzed for emotion detection. They offered two approaches for music suggestion after emotion was identified. The first strategy is an expert-based approach, in which certain experts are used to distribute music depending on emotion. The second technique is a feature-based strategy, which does not need the assistance of an expert. They employ music’s rhythm and articulation to distribute songs according to emotion. For the music suggestion, they devised a feedback mechanism. As a result, the algorithm will provide music recommendations based on user response.

III. PROPOSED SYSTEM DESIGN

In the proposed research work, evaluate the entire system using supervised learning algorithms and first collect data from the EMOTIC dataset. In order to extract various features from inputs, the CNN algorithm sets and generates the trained model appropriately. In the testing system, classify each input EMOTIC dataset with the appropriate labels and show the system's efficiency. Feature extraction aims to reduce the amount of features in a dataset by generating new features from existing features. In the development of a predictive model, feature selection entails minimizing the quantity of input variables. We can classify both structured and unstructured data, dividing a given dataset into multiple categories.

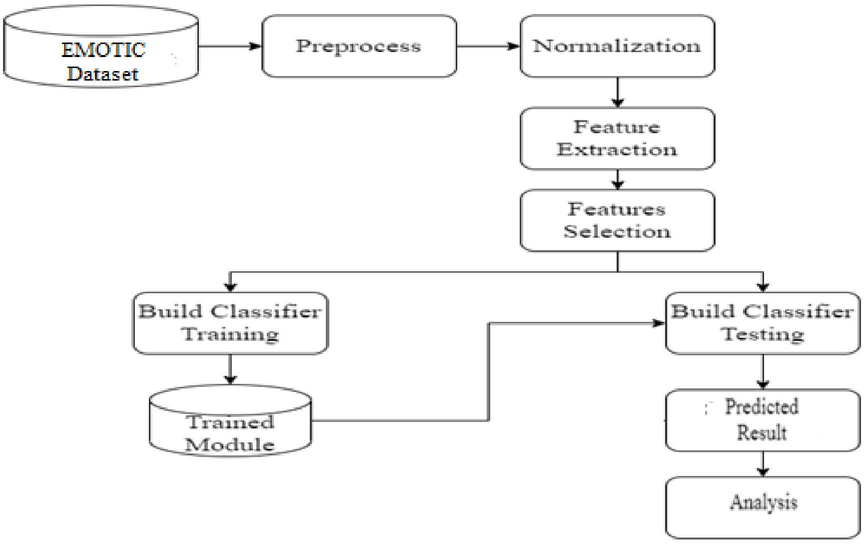

Figure 1: Proposed System Design

Implement Module

Dataset Description:

- Provide a complete description of the EMOTIC dataset, including the types of images, emotion categories, and contextual labels.
- Discuss any data preprocessing steps taken to get ready the dataset for model training.

Feature Extraction:

- Outline the features extracted from images, including both visual features (e.g., facial expressions, body language) and contextual features (e.g., background, objects).
- Various feature descriptors such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), or deep features extracted from pre-trained CNNs like VGG-Face or ResNet are used to capture the intricate details of facial expressions.

Validation & Training:

- Describe the training process, including any techniques used to optimize the model's performance.
- Discuss the validation methods and metrics used to evaluate the model.

Classification

Describe the deep learning model used for emotion recognition, such as CNNs or transformers. Explain how the model integrates both facial and contextual features to make emotion predictions. We frame emotion recognition as a multi-label classification problem. Based on both facial expressions and contextual cues, the EMOTIC dataset classifies each image into one or more of several predefined emotion categories. The EMOTIC dataset includes a variety of emotion categories, such as fear, surprise, rage, sorrow, and happiness. Each image may exhibit multiple emotions simultaneously, reflecting the complexity of human emotional states. In addition to facial expressions, the classification model considers contextual information present in the image, such as the surrounding environment, objects, and other people, which can significantly influence the perceived emotion. The proposed model employs deep learning techniques, such as CNNs or transformer-based models, to extract features from both the subject's facial expressions and the context. We then use these features to classify the image into the relevant emotion categories.

Evaluation Metrics:

We use measures such as accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC) to assess the classification model's performance. Special attention is given to how well the model handles the multi-label nature of the problem and the impact of context on classification accuracy.

RESULT AND ANALYSIS

In our proposed experimental model, we used images for detection and classification purpose. The different cross validation technique is used for module training as well as module testing. The data splitting standard approach has utilized for training and testing. The CNN is deep learning classifier has used with multiple convolutional layers. In convolutional layer we extract various features from input image while the optimization has done in pooling layer

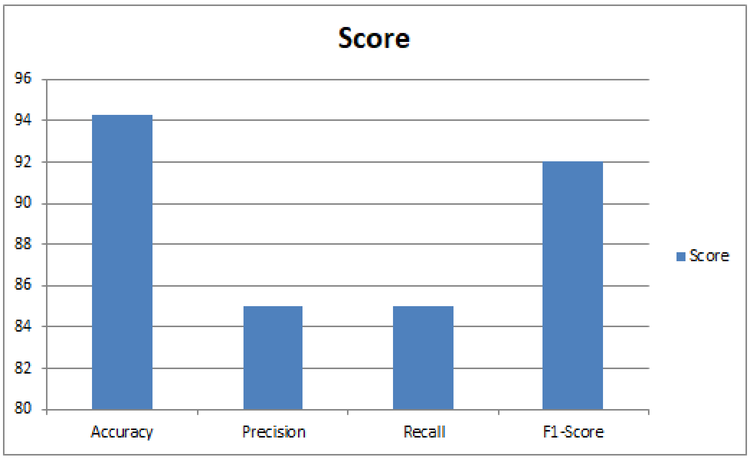

Figure 1.2: Performance analysis of proposed system

CONCLUSION

Use the proposed approach to generate substantial emotion for a diverse range of photos. The system may extract hybrid characteristics, including picture and body information, to aid in sentiment creation. The system is capable of functioning with both picture and text datasets. Detecting emotions via contextual analysis of photos presents considerable hurdles in learning methodologies. Multiple attributes, such as color, form, texture, and other encoder characteristics, might affect the creation of visual captions during the classification process. Numerous studies indicate that situational context is a significant influence in the proper interpretation of others' emotions. The absence of data has impeded a thorough study on context processing for automated emotion recognition. A potential solution is an emotion-detection system using real-time visual and contextual data. This work fundamentally used hybrid features from real-time image datasets to develop a classifier capable of identifying attributes based on those features. The hybrid deep learning classifier, using a convolutional neural network (CNN), has successfully predicted the emotion of a certain text (RNN).

REFERENCES

[1] Shreya, L., and N. Nagarathna. ”Emotion Based Music Recommendation System for Specially-Abled.” 2021 International Conference on Recent Trends on Electronics, Information, Communication Technology (RTEICT). IEEE, 2021.

[2] Lee, Shih-Hsiung, et al. ”A music recommendation system for depression therapy based on EEG.” 2020 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-Taiwan). IEEE, 2020.

[3] Zhang, Ting-Zheng, Tiffany Chang, and Ming-Hao Wu. ”A Brainwave-based Attention Diagnosis and Music Recommendation System for Reading Improvement.” 2021 IEEE International Conference on Artificial Intelligence, Robotics, and Communication (ICAIRC). IEEE, 2021.

[4] Asif, Anum, Muhammad Majid, and Syed Muhammad Anwar. ”Human stress classification using EEG signals in response to music tracks.” Computers in biology and medicine 107 (2019): 182-196.

[5] B˘alan, Oana, et al. ”Emotion classification based on biophysical signals and machine learning techniques.” Symmetry 12.1 (2019): 21.

[6] Santamaria-Granados, Luz, Juan Francisco Mendoza-Moreno, and Gustavo Ramirez-Gonzalez. ”Tourist recommender systems based on emotion recognition—a scientometric review.” Future Internet 13.1 (2020): 2.

[7] Cunha, Joana, et al. ”The Effect of Music on Brain Activity an Emotional State.” Engineering Proceedings 7.1 (2021): 19.

[8] Xu, Baoguo, et al. ”Continuous Hybrid BCI Control for Robotic Arm Using Noninvasive Electroencephalogram, Computer Vision, and Eye Tracking.” Mathematics 10.4 (2022): 618.

[9] Jaafar, Sirwan Tofiq, and Mokhtar Mohammadi. ”Epileptic seizure detection using deep learning approach.” UHD Journal of Science and Technology 3.2 (2019): 41-50.

[10] Kang, Dongwann, and Sanghyun Seo. ”Personalized smart home audio system with automatic music selection based on emotion.” Multimedia Tools and Applications 78.3 (2019): 3267-3276.