Anusandhanvallari
Vol 2024, No.1
December 2024
ISSN 2229-3388

Real-time Facial Emotion Recognition using CNN

[1]Ms. S. Bhuvaneshwari, [2]Dr. Pravin Ramesh Gundalwar

[1]MCA Student, Dept of IT, Anurag University,

[2]Asso. Professor, Dept of IT, Anurag University,

Abstract: Visual sentiment analysis explores how people emotionally respond to visual content like images and videos. It aims to decipher the underlying emotional context by examining visual data. The success of this field largely stems from advances in computer vision technology, leading to robust algorithms that extract meaningful information from visuals. Many models in this field focus on whole-image features or develop intricate architectures to analyze visual sentiment. However, these approaches often overlook the significance of local areas within images, which can be crucial in determining emotional responses. This research paper addresses this gap by applying Convolutional Neural Networks (CNNs) to detect human faces in images and analyze their emotions. When an image is provided, the CNN model identifies facial regions, outlines them with rectangles, and assesses the corresponding emotions. Based on this analysis, a suitable emoji is displayed to represent the detected emotion. This approach brings a more nuanced understanding of visual sentiment by concentrating on human expressions, providing a more intuitive depiction of the emotional content in images. With this work, visual sentiment enhancement analysis can be by emphasized the role of facial emotions, leading to more accurate interpretations and potential applications in various domains.

Keywords: Visual Sentiment Analysis; Facial Emotion Recognition; Convolutional Neural Networks; Face Detection; Facial Expressions; Emotion Classification.

Introduction

Visual sentiment analysis explores the emotional responses people have to visual stimuli like images and videos. This field diverges from traditional textual sentiment analysis, which gauges emotional reactions in written content. The rise of deep learning in computer vision has propelled visual sentiment analysis forward, making it possible to understand high-level visual semantics, including image aesthetic analysis and sentiment detection. Historically, visual sentiment analysis has focused on developing robust features, transitioning from pixel-level data to mid-level attributes, and eventually to deeper visual features through neural networks. These improvements have enhanced the performance of sentiment analysis systems. Yet, most approaches concentrate on whole-image analysis, ignoring the influence of specific local regions that might contribute significantly to overall sentiment. In our work, we address this gap by utilizing attention models to pinpoint local image regions that carry emotional cues, allowing us to build a sentiment classifier that incorporates these localized visual features. Beyond general sentiment analysis, facial expression recognition has garnered interest for its applications in various fields, including clinical practice, sociable robotics, and education. Understanding facial expressions provides real-time feedback on human emotions, which can be valuable in educational settings. Traditional feedback methods like exams and observations are often time-consuming and inefficient. In contrast, analyzing facial expressions can offer a dynamic approach to gauge student engagement and adapt teaching strategies. Our proposed system employs a Convolutional Neural Network (CNN) to recognize facial expressions. It detects faces in images, crops and normalizes them, and then feeds them into a CNN for emotion classification. This approach to sentiment analysis, with its focus on both global and local visual features, represents a significant step toward more accurate and nuanced interpretations of visual data.

Scope

The scope of this project encompasses the development of a facial emotion recognition system for educational

settings, with a focus on improving student engagement and learning outcomes. Using Convolutional Neural Networks (CNN), the system detects and analyzes students' facial expressions to classify emotions. This framework allows educators to gain real-time insights into students' emotional states during classroom instruction or e-learning. The project aims to enhance the learning environment by enabling adaptive teaching methods based on emotional feedback.

Literature Survey

A system is created to analyze students' facial expressions to assess classroom teaching effectiveness. The system consists of five phases: data acquisition, face detection, face recognition, facial expression recognition, and post-processing. It uses the K-nearest neighbor (KNN) algorithm for classification and Uniform Local Gabor Binary Pattern Histogram Sequence (ULGBPHS) for pattern analysis.[1],[2]

A proposed a web application that analyzes the emotions of students participating in face-to-face classroom instruction. Webcams installed in classrooms capture live recordings, which are then analyzed using machine learning algorithms. [3],[4]

An approach is developed to recognize student engagement based on their facial expressions. The approach employs Gabor features and a Support Vector Machine (SVM) algorithm to identify engagement as students interact with cognitive skills training software. Labels were derived from videos annotated by human judges. [5]

Another approach [6] employed computer vision and machine learning techniques to identify the emotions of students in a school computer laboratory. This analysis occurred while students engaged with an educational game designed to explain classical mechanics.

A different system was designed to identify and monitor students' emotions in real-time to improve the e-learning environment. This system uses eye and head movement patterns to infer students' moods during online learning. [7],[9]

A Facial Emotion Recognition System (FERS) is developed to detect the emotional states and motivations of students in videoconferencing-based e-learning. Their system uses four machine learning algorithms—SVM, KNN, Random Forest, and Classification & Regression Trees—with the aim of achieving high accuracy. [8]

These studies collectively showcase a range of methodologies and technologies used to leverage FER in educational contexts, highlighting the potential of such systems to improve teaching and learning outcomes.

Overview of System

- **Existing System**

The existing system for visual sentiment analysis and facial expression recognition encompasses several stages of development. Initially, approaches to visual sentiment analysis relied on pixel-level features, later advancing to mid-level attributes and more recently to deep visual features and unsupervised frameworks. However, these methods primarily focused on either textual analysis or image analysis, neglecting the significance of local visual regions induced by sentiment-related visual attributes. Concurrently, facial expression recognition has garnered attention in various fields such as clinical practice, sociable robotics, and education, where emotion plays a pivotal role. Current methods for assessing students' emotions often involve traditional means like exams, questionnaires, and observations, which can be inefficient. Recognizing the potential of facial expression analysis to provide real-time feedback to educators for adjusting teaching strategies and instructional materials, there is a growing interest in leveraging facial expression recognition technologies in educational contexts.

- **Disadvantages of Existing System**

The potential disadvantages of the proposed system:

- **Limited Focus on Local Visual Regions**

While the proposed system aims to consider local visual regions induced by sentiment-related visual attributes, it may still not fully exploit the potential of such regions. The emphasis on whole-image or whole-face analysis in existing methods might overlook nuanced emotional cues present in specific areas of an image or face.

- **Dependency on Facial Detection Accuracy**

The accuracy of facial detection and recognition algorithms, particularly in real-world scenarios with varying lighting conditions, facial orientations, and occlusions, could significantly impact the performance of the proposed system. False positives or failures in detecting facial expressions accurately could lead to erroneous feedback in educational settings.

- **Emotion Classification Accuracy**

Despite utilizing Convolutional Neural Network (CNN) architecture, the proposed system's accuracy in classifying emotions from facial expressions might still face challenges. Emotions are complex and subtle, and misclassifications could occur, especially for emotions that share similar facial expressions or for individuals with diverse cultural expressions.

- **Potential Bias in Educational Feedback**

Relying solely on facial expressions for feedback in educational settings might introduce biases or inaccuracies. Emotions are influenced by various factors beyond academic understanding, such as personal experiences, cultural backgrounds, and individual differences, which the proposed system may not fully account for.

- **Ethical Considerations**

The deployment of facial expression recognition technologies in educational settings raises ethical concerns regarding privacy, consent, and data security. Students' facial data collection and analysis must adhere to strict privacy regulations and ensure transparent consent procedures to mitigate potential risks of misuse or infringement of students' rights.

- **Resource Intensiveness**

Implementing and maintaining the proposed system, especially in real-time applications, may require significant computational resources and infrastructure, potentially posing challenges in resource-constrained environments such as classrooms with limited access to high-performance computing resources.

These disadvantages highlight potential challenges and limitations that need to be addressed in the design and implementation of the proposed system for analyzing students' facial expressions in educational contexts.

- **Proposed System**

The proposed system aims to address the limitations of existing methods in visual sentiment analysis and facial expression recognition by introducing several key innovations. Firstly, it shifts the focus from whole-image or whole-face analysis to considering local visual regions induced by sentiment-related visual attributes, thereby capturing more nuanced emotional cues. This approach involves leveraging recent advancements in attention models, which enable the system to learn contextual semantics and extract relevant features from specific regions of visual data. Additionally, the proposed system utilizes a Convolutional Neural Network (CNN) architecture for analyzing students' facial expressions in educational settings. Upon detecting faces in input images, the system crops and normalizes them to a standardized size before feeding them into the CNN. The CNN then processes the facial images to recognize various emotions such as anger, happiness, sadness, disgust, surprise, or neutrality. By combining attention-based feature extraction with CNN-based emotion recognition, the proposed system aims to

provide more accurate and contextually relevant analyses of students' emotional responses, thereby enabling educators to adapt their teaching strategies and instructional materials effectively.

- **Advantages of Proposed System**

Advantages of Proposed System are:

- **Enhanced Emotional Understanding**

By focusing on local visual regions induced by sentiment-related visual attributes, the proposed system can capture nuanced emotional cues that might be overlooked by traditional whole-image or whole-face analysis methods. This approach could lead to a more comprehensive understanding of human emotional responses to visual stimuli.

- **Contextual Semantics Learning**

Leveraging attention models allows the system to learn contextual semantics, enabling it to extract relevant features from specific regions of visual data. This contextual understanding could improve the accuracy and relevance of the system's analyses of emotional responses.

- **Improved Facial Expression Recognition**

Utilizing CNN architecture for facial expression recognition offers several advantages, including robustness to variations in facial orientation, lighting conditions, and occlusions. This could result in more accurate and reliable recognition of emotions such as anger, happiness, sadness, disgust, surprise, or neutrality.

- **Real-time Feedback for Educators**

In educational settings, the proposed system could provide educators with real-time feedback on students' emotional responses. By analyzing facial expressions, educators can adapt their teaching strategies and instructional materials to better meet the emotional needs of students, potentially enhancing the learning experience and academic outcomes.

- **Potential for Personalized Education**

By understanding students' emotional responses, educators can tailor their instructional approach to individual students, offering personalized support and guidance. This could foster a more inclusive learning environment and help students achieve their full potential.

- **Ethical Considerations**

The proposed system has the potential to address ethical considerations surrounding privacy and consent by ensuring transparent data collection and analysis processes. Educators can implement the system with strict adherence to privacy regulations, protecting students'

rights and confidentiality.

Overall, the proposed system offers several advantages, including enhanced emotional understanding, improved facial expression recognition, real-time feedback for educators, potential for personalized education, and ethical considerations, making it a promising tool for analyzing students' emotional responses in educational settings.

Methodology

The breakdown of the proposed research work is given in Table 1 for three modules, each with specific objective and functions:

Table 1: Proposed work module

Module	Objective	Functions
Data Gathering	Collect relevant data from various sources, including user information, content metadata, and video files from users who upload content to Social Media Networks (SMNs).	Collect user profiles, activity history, and other metrics for determining user trustworthiness. Capture data from external sources like community reports and moderation feedback to create a comprehensive dataset.
Data Preprocessing & Enhancement	Clean, process, and enhance the collected data to prepare it for further analysis.	Remove redundant or irrelevant information and address inconsistencies. Perform video transcoding for compatibility across platforms and conduct quality checks to identify anomalies or security risks. Add metadata, extract key features, and perform preliminary analysis to aid subsequent trust evaluation processes.
Data Upload	Upload preprocessed and enhanced data to the SMN platform, ensuring only trusted and verified content is uploaded.	Integrate with the SMN's content management system to manage the upload process.Implement security measures like digital watermarking and integrity checks to ensure content remains secure during delivery.Interact with automated content moderation systems to categorize content, determine trust levels, and flag suspicious activities for review. This modular approach ensures efficient and organized handling of data throughout the process, from collection to upload, contributing to a safer and more reliable SMN environment.

Architecture

The proposed research work architecture is given in Fig.1.

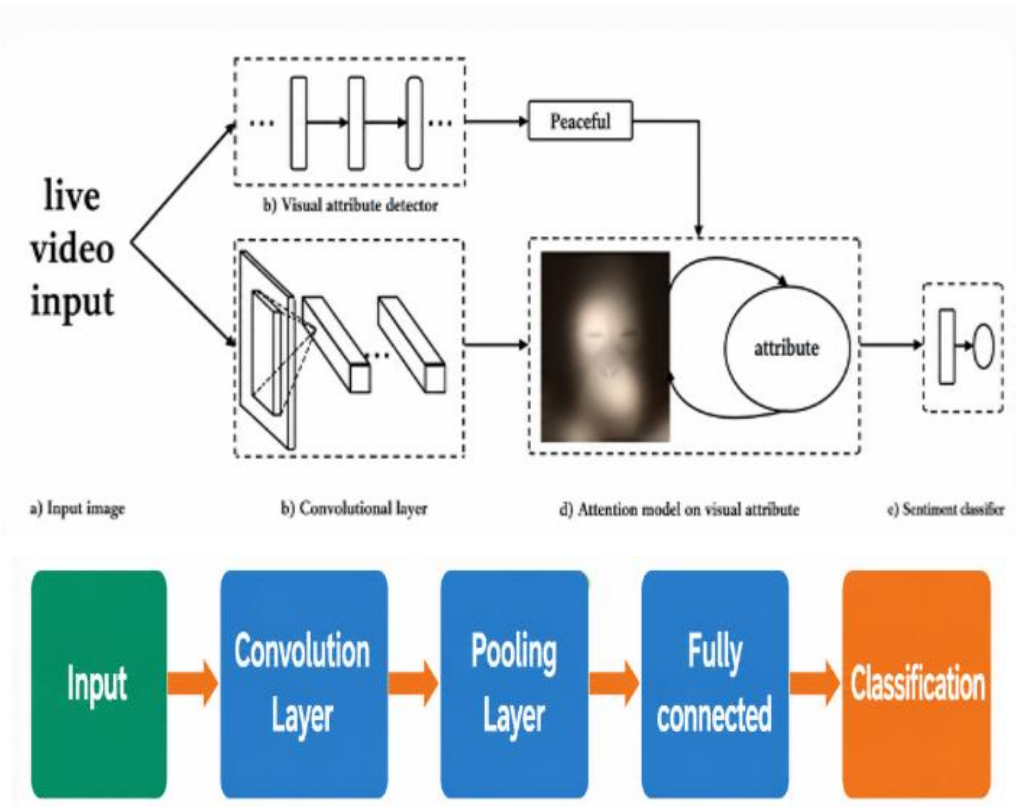

Fig 1: Proposed Research Work Architecture

Result and discussion

The system evaluates the detected face against seven emotions given as follows:

Angry, Disgust, Scared, Happy, Sad, Surprised, and Neutral

The sample real-time inputs are given to the systems to different set experiments performed on facial expression and predicts the emotional state using CNN techniques is shown in following Fig.2-5. The Facial Emotion Recognition system works into two main parts discussed for “Happy Face” as follows:

Face Detection

The left side shows the input image captured through a camera.

The system detects the person's face and highlights it with a red bounding box.

The predicted emotion, “happy,” is displayed above in Fig2. as detected “Happy Face”.

Emotion Classification

The right side displays the probability distribution for different emotional categories.

The system evaluates the detected face against seven emotions considered in the experiment.

Anusandhanvallari
Vol 2024, No.1
December 2024
ISSN 2229-3388

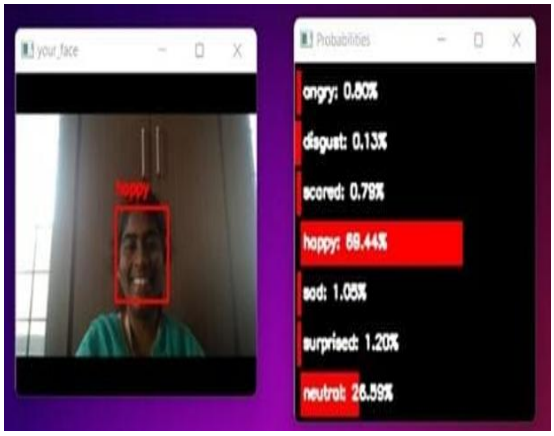

Fig.2: Happy face detection

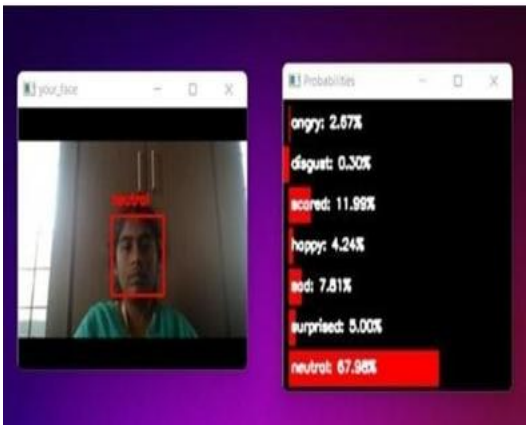

Fig.3: Neutral face detection

Fig.4: Angry face detection

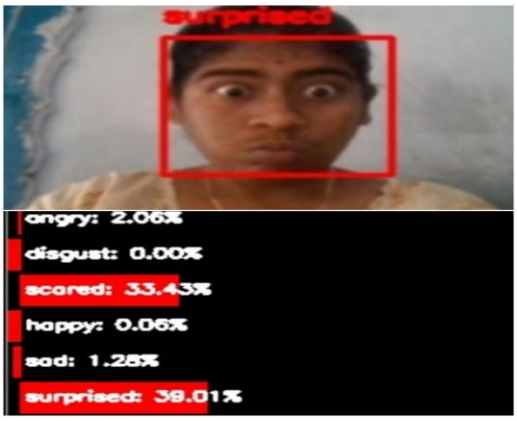

Fig.5: Scared face detection

The system is detecting correctly with the average accuracy for the specified real-time emotions given as input. The above results demonstrated for the located face and then estimate the person's emotional state from facial features such as the eyes, eyebrows, mouth, and overall facial configuration. The Table 2 shows the probability distributions for different face detection and emotional features considered for the experimentation.

Table 2: Probability distributions

Type of Emotion input	Probability Distributions for detected face against seven emotions features						
	Angry	Disgust	Scared	Happy	Sad	Surprised	Neutral
Happy	0.80	0.13	0.78	**69.44**	1.05	1.20	26.59
Neutral	2.67	0.30	11.99	4.24	7.81	5.00	**67.98**
Angry	**31.1**	27.56	14.22	9.73	6.50	0.69	10.12
Scared	2.06	0.00	**33.43**	0.06	1.28	**39.01**	23.98

The results also indicate a degree of similarity between the facial features associated with the “scared” and “surprised” emotions. Both emotions can produce similar facial characteristics, such as raised eyebrows, widened

eyes, and an increased degree of facial openness.

Conclusion

In conclusion, visual sentiment analysis presents a compelling avenue for understanding human emotional responses to visual stimuli such as images and videos. The success of current models in this field owes much to the development of robust algorithms rooted in computer vision. While existing approaches primarily focus on extracting features from the entire image or video, our project highlights the significance of local areas in eliciting human emotional responses. By CNN algorithms, we proposed a novel approach to visual sentiment analysis that specifically targets facial expressions. The applied method involved a detecting face within images, identifying emotional states associated with each face, and displaying the corresponding emotion with an emoji. This approach not only enhances the interpretability of emotional responses but also provides a more nuanced understanding of the visual content.

Moving forward, further research and development in this area could explore refining the accuracy and efficiency of emotion recognition algorithms, as well as incorporating contextual information to better capture the complexities of human emotions. Additionally, the application of image recognition techniques to analyze sentiments and emotions in real-world scenarios holds promise for diverse fields such as marketing, healthcare, and human-computer interaction. Overall, our project contributes to the advancement of visual sentiment analysis by offering a focused and accessible framework for understanding human emotional responses in visual content.

References:

[1] R. G. Harper, A. N. Wiens, and J. D. Matarazzo, Nonverbal communication: the state of the art. New York: Wiley, 1978.

[2] P. Ekman and W. V. Friesen, “Constants across cultures in the face and emotion,” Journal of Personality and Social sychology, vol. 17, no 2, p. 124 129, 1971.

[3] C. Tang, P. Xu, Z. Luo, G. Zhao, and T. Zou, “Automatic Facial Expression Analysis of Students in Teaching Environments,” in Biometric Recognition, vol. 9428, J. Yang, J. Yang, Z. Sun, S. Shan, W. Zheng, et J. Feng, Éd. Cham: Springer International Publishing, 2015, p. 439 447.

[4] Savva, V. Stylianou, K. Kyriacou, and F. Domenach, “Recognizing student facial expressions: A web application,” in 2018 IEEE Global Engineering Education Conference (EDUCON), Tenerife, 2018, p. 1459 1462.

[5] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, “The Faces of Engagement: Automatic Recognition of Student Engagementfrom Facial Expressions,” IEEE Transactions on Affective Computing, vol. 5, no 1, p. 86 98, janv. 2014.

[6] N. Bosch, S. D'Mello, R. Baker, J. Ocumpaugh, V. Shute, M. Ventura, L. Wang and W. Zhao, “Automatic Detection of Learning-Centered Affective States in the Wild,” in Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI ’15, Atlanta, Georgia, USA, 2015, p. 379 388.

[7] Krithika L.B and Lakshmi Priya GG, “Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric,” Procedia Computer Science, vol. 85, p. 767 776, 2016.

[8] U. Ayvaz, H. Gürüler, and M. O. Devrim, “USE OF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS,” Information Technologies and Learning Tools, vol. 60, no 4, p. 95, sept. 2017

[9] Y. Kim, T. Soyata, and R. F. Behnagh, “Towards Emotionally Aware AI Smart Classroom: Current Issues and Directions for Engineering and Education,” IEEE Access, vol. 6, p. 5308 5331, 2018