IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Deep Learning-Based Facial Expression Recognition for Analysing Student Emotional Responses to Advertisements A CNN and Transfer Learning Approach with Multimodal Engagement

Darekar Pranav Pravin and Prof. Dr. Satish N. Gujar
Department of Computer Application
JSPM University, Pune
pranavdarekar00@gmail.com and sng.scos@jspmuni.edu.in

Abstract: *It is essential to study the emotions experienced by the audience in response to advertisements in order to analyze the effectiveness of advertising and engagement of the audience. The conventional way to measure these emotions relies on the use of surveys and feedback forms which may be cumbersome and may not always measure the emotions experienced by the audience. This research work seeks to employ deep learning based approach for the identification of emotions of the students based on their facial expressions while watching advertisements. The system will rely on the usage of CNNs and transfer learning techniques in order to detect emotions from the facial expressions. These emotions are further combined with other engagement parameters in order to compute the multimodal engagement scores indicating the level of interest and engagement of the students. The objective of the system is to classify facial expressions including happiness, sadness, anger, surprise, fear and neutral emotions using the deep learning models trained on the facial patterns from images and video frames. Abstract In today's digital age, the viewer's emotional response plays a vital role in measuring the effectiveness of advertisements. The existing methods to analyse viewer's response include questionnaires, interviews, and manual feedback; however, the methods cannot always determine real-time emotions of viewers. Emotional responses are usually reflected in one's facial expressions, which make Facial Expression Recognition (FER) an excellent method to automatically analyse viewer's emotions.*

Keywords: Facial Expression Recognition

I. INTRODUCTION

Nowadays, in the era of technologies, advertisements are considered to be an effective tool for company-client interactions. Due to the fast development of online platforms, social networks, and digital marketing, nowadays advertising does not only concern the provision of the information concerning some product or service but also tries to engage the viewer emotionally. Thus, now the process of researching audience reaction to advertisements becomes an important topic of study.

Human emotions are very important for understanding people's reactions to various advertisements. It goes without saying that positive advertisement more efficient, whereas negative emotions or inability to attract people's attention indicate failure. The main way of analyzing people's responses to advertisements used by advertisers and researchers was based on questionnaires, interviews, and feedback forms. However, there are some drawbacks in such kind of analysis since the response of people depends on their self-reporting skills and may not reflect their real emotions. The facial expression is a powerful tool which enables a person to convey his/her emotions to other people. Changes in facial features, such as eye movement, eyebrows and mouth positions can give much information concerning people's emotions.

DOI: 10.48175/IJARSCT-37161

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Currently, deep learning algorithms, in particular, Convolutional Neural Networks (CNNs), have been widely applied to solve emotion recognition tasks. CNNs can automatically learn the features from the images which are necessary for the task without any human intervention during the manual feature extraction stage. CNNs can find complicated patterns which relate to certain emotions, thus, are appropriate for analysing human expressions. However, to train a deep learning model from scratch, a huge number of labelled samples and computing power is required. To address this issue, transfer learning approach is employed, according to which, a model that has been previously trained on the general features of the images is adapted for a specific problem, like facial emotion recognition. One of the target audience groups for the analysis of advertisements is the students. Their preferences, behaviour and reaction to emotions affect future trends in consumers. process is needed. Thus, an automated solution that would be able to detect emotions from student faces during the ads viewing process is needed.

II. LITERATURE REVIEW

Facial Expression Recognition (FER) has become a key area of research in Artificial Intelligence, Computer Vision, and Human-Computer Interaction. The ability to automatically identify human emotions through facial expressions has many applications, such as in healthcare, education, marketing, security systems, and user behavior analysis. Researchers have developed various machine learning and deep learning methods to improve the accuracy of emotion recognition systems. Early methods for facial expression recognition mainly relied on traditional machine learning techniques. These approaches required manually extracting facial features and applying classification algorithms like Support Vector Machines (SVM), K-Nearest Neighbour (KNN), and Decision Trees. Common feature extraction techniques included Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Gabor filters. Although these methods performed reasonably well, their accuracy was limited because they depended heavily on manually selected features and struggled to capture complex facial patterns effectively. With the rise of deep learning, Convolutional Neural Networks (CNNs) became popular for facial expression recognition. CNN models automatically learn important features from facial images, removing the need for manual feature extraction. CNNs can detect low-level features like edges and textures, as well as high-level features like facial structures, making them very effective for emotion classification. Studies have shown that CNN-based approaches outperform traditional machine learning techniques. The Facial Expression Recognition 2013 dataset is really popular for training and testing models that recognize emotions. This dataset has a lot of grayscale pictures of faces. They are divided into seven groups based on the emotions they show: anger, disgust, fear, happiness, sadness, surprise and neutral. A lot of people have used the Facial Expression Recognition 2013 dataset to make new models using something called CNN architectures, which helps them get better at recognizing emotions.. There are still some problems, like the fact that the lighting is not always the same or the faces are at different angles or the pictures are not very clear and peoples faces are all different. To make things easier people have started using something called transfer learning. Transfer learning is when you take a model that has already been trained on a lot of pictures and you use it for a task. This way you do not have to start from scratch. Models like VGG16, ResNet50 InceptionV3 and MobileNet have been used for recognizing expressions and they work really well. These models are faster to train they need pictures and they are often better at recognizing emotions than traditional models. People like using models that are based on VGG because they're simple and they work well. The VGG model uses layers to look at the details of the faces. The ResNet model has something called connections, which helps with a problem called vanishing gradients and this allows us to make more powerful models. The MobileNet model is good for things that need to be fast and not use much power so it is good for recognizing emotions in real time.

III. RESEARCH METHODOLOGY

The proposed research method focuses on creating a framework that uses learning to understand how students feel when watching ads. The main goal of this research is to figure out how students feel by looking at their expressions and measuring how engaged they are with the ad content. The system uses computer vision, Convolutional Neural Networks (CNN) and techniques that help it learn from models to provide an automated way to understand how an audience

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

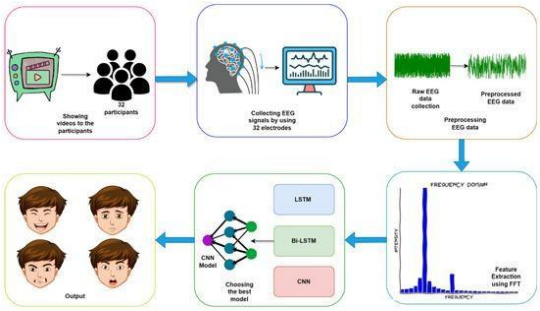
behaves. The method starts by collecting data on students facial expressions while they watch ad videos. This data is used to train and test the model that recognizes emotions. Facial expressions tell a lot about a persons state because emotions like happiness, sadness, anger, fear and surprise change how their face looks. In this research facial images and video frames are studied to capture these changes.

The table below summarizes key studies:

TABLE 1. SUMMARY OF KEY LITERATURE

Author	Year	Focus	Key Finding	Gap
LeCun et al.	2015	Deep learning techniques for image recogniton	Deep learning models, especially CNNs, provide effective automatic feature extraction for visual data analysis.	Requires large datasets and high computational resources for training.
Pantic and Rothkr antz	2000	Automati c facial expressio n analysis	Facial expressions provide important information for understandi ng human emotions and behaviour.	Traditional methods depend on manual feature extraction and have limited accuracy.
Lucey et al.	2010	CK+ facial expressio n dataset developm ent	Facial expression datasets improve the performanc e of emotion recognition models by providing labelled samples.	Limited diversity in real-world conditions such as lighting and backgroun d variations.

The first thing we do is collect a lot of data. Then we use some techniques to get the images ready for the deep learning model. We do this because the pictures of faces we have are not perfect. They have things in the background that we do not want. They are sizes. They have noise and the lighting is not always the same. All these things can make it hard for the model to work well. So the first thing we do is find the face in the picture. We call this face detection. The system looks at the picture. Finds the face. Then it separates the face from the rest of the picture. This way the model only looks at the face. Not the things we do not want.



After we find the face we make all the pictures the same size. This is so the model can look at them easily. Then we do something called image normalization. We change the numbers that make up the picture so they are all the same. This helps the model learn faster and work better. We also use something called data mentation. We take the pictures we have. Make new ones from them. We turn them around make them bigger or smaller and change the light. This helps the model learn about all the ways a face can look. It also helps the model not get too good at one thing. The model can look at a lot of faces and learn about them. Then it can look at faces and know what they are feeling. he next thing we do is use a kind of model called a Convolutional Neural Network. This model is really good at looking at pictures and understanding what is in them. It can find the things in a face like the eyes and the mouth. The model looks at things like how the eyesre moving, where the eyebrows are and what the mouth is doing. It uses all these things to figure out how someone is feeling.

DOI: 10.48175/IJARSCT-37161

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, June 2026 Impact Factor: 8.2

The model has a lot of layers. Each layer does something. The first layer looks at the picture. Finds the basic things, like edges and shapes. After we find the face we make all the pictures the same size. This is so the model can look at them easily. Then we do something called image normalization. We change the numbers that make up the picture so they are all the same. This helps the model learn faster and work better. We use something called transfer learning to make the model work better. This means we use a model that someone else has already trained. We do not have to start from scratch. This saves us a lot of time and work.We can use models like VGG16, ResNet50 and MobileNet. These models are all good at things. VGG16 is good at looking at pictures and finding all the details. ResNet50 is good at learning about things. MobileNet is fast and good, for things that need to happen in time. When we use transfer learning the model works better. It does not take as long to train. Once the model is trained it can look at a picture of a face

Tier 1: Data Collection Layer

This Data Collection Layer is where we get the facial expression data from students when they watch advertisements. The main thing this Data Collection Layer does is get real-time information, about how students feel. This information is used to train and test the deep learning model we are talking about.

Tier 2: Data Processing and Integration Layer

The facial data that we collect needs to be prepared before it is given to the deep learning model. The raw facial images that we have often have some things in them that we do not want, like the background and they can be sizes and the lighting can be different and there can be some noise, in them

Tier 3: Operational Performance Analysis

The CNN architecture is made up of different steps. The convolution layer is where it finds the things like edges and shapes of faces and eyes and how the mouth moves. This is where it gets the features that it needs. The activation layer is like a helper that makes the system better at learning by adding some complexity to the process.

Tier 4: Customer Insights and Business

This layer is about taking the responses and engagement information we found and turning them into something we can really understand. We want to know how students behave. If advertisements are really working. When the facial expression recognition model figures out how a student is feeling we look at the results to see if an advertisement is actually getting peoples attention and making them feel something.

V. JURISDICTION COMPARATIVE ANALYSIS

The study of how different places use facial expression recognition and emotion analysis is the focus of this research. It looks at how these technologiesre used in different situations and environments. The goal is to see how different areas and industries use systems that recognize emotions with the help of Artificial Intelligence to understand how people behave how customers react and how engaged they are. The development of facial expression recognition using learning has created new opportunities in areas like digital marketing, education, understanding customer behavior and how people interact with computers. Different places have levels of using these systems different rules about keeping data private different technology and different ideas about what is right and wrong that affect how facial emotion recognition systems are used. In places where a lot of progress has been made emotion recognition that uses Artificial Intelligence is used more and more to study how people react to advertisements. Companies use computer vision to look at the expressions of people watching advertisements and to see how emotionally engaged they are.

The old way of figuring out if an advertisement is good or not usually involves asking people what they think through surveys, interviews and feedback. However this way of doing things may not show how people really feel at the moment because people may not always say what they are feeling. Facial expression recognition that uses learning is a more

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 9, June 2026 Impact Factor: 8.2

objective way of doing things because it looks at how peoples faces really react. The system can tell if people are feeling emotions like happiness and excitement or bad emotions like confusion and dissatisfaction.

Facial expression recognition is used to understand how people react to things. It is used to see what people are feeling when they watch advertisements. Facial expression recognition can tell if people are happy or sad when they watch something. It is a way to know what people are feeling. The study of facial expression recognition is important to understand how people behave and how they feel. Facial expression recognition is used in areas, like digital marketing and education. It is a tool to understand peoples emotions and behaviors.

TABLE 2. COMPARATIVE JURISDICTION ANALYSIS

System / Approach	Data Sources Used	Key Features	Advantages
Tradition al Emotion Analysis System	Manual surveys, questionna ires, user feedback	Human-based emotion evaluation, subjective analysis, basic response measurement	Easy to implement, low cost, does not require complex technology
Machine Learning-Based Emotion Recogniti on	Facial image datasets, manually extracted facial features	Feature extraction using algorithms such as SVM and KNN,basic emotion classificati on	Faster than manual analysis, provides automated prediction
CNN-Based Facial Express ion Recogni tion	Facial expressio n datasets, video frames, image samples	Automati c feature extractio n, facial pattern recogniti on, emotion classification	Higher accuracy, reduces dependenc y on manual feature engineering
Adverti sement Engage ment Analysi s System	Advertise ment videos, facial expressio ns,	Measures emotiona l reaction, attention level, and audience engagement	Provides insights into advertisem ent effectivene ss

VI. ALGORITHMIC COMPLIANCE MECHANISMS

The proposed facial expression recognition system is about looking at peoples faces to see how students feel when they watch advertisements. This system uses Artificial Intelligence to figure out what people are feeling and how much they are engaged. It is really important that the system works correctly and is fair and safe. To make sure the system is working properly

we need to follow some rules. These rules help the system make decisions and use the data correctly. The system is made up of parts including collecting data getting the data ready teaching the system figuring out what people are feeling and seeing how much they are engaged. We want the system to give us information and not make too many mistakes. We also do not want the system to be unfair or hurt peoples privacy. The quality of the data is very important for the system to work well. The system has a way of making sure the data is good before it is used. We pick the data carefully. Make sure it is clean and ready to use. We throw away any pictures that're not clear or have too much noise. This helps the system make decisions.

If the data is not good the system might not work correctly. The system might make mistakes if the data does not have people with faces, lighting and cameras. The system might also make mistakes if people are not expressing their feelings in ways. The facial expression recognition system is looking at peoples faces to see how students feel when they watch advertisements.

The proposed facial expression recognition system has to be careful with the data it uses. The facial expression recognition system has to make sure the data is good fair.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

VII. ECONOMIC MODEL FOR ARTIST COMPENSATION

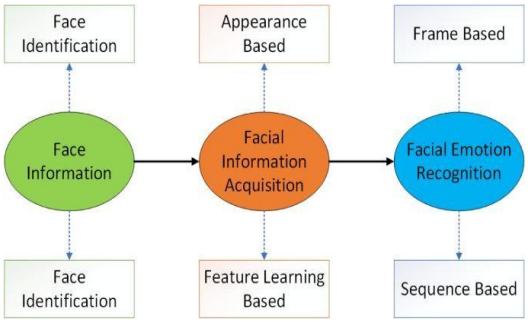
The way advertising and artificial intelligence work together has really changed how content creators, advertisers and audiences interact with each other. In the advertising world people like artists, designers, actors and video creators use their skills to make content that affects how people feel. This is because the people making this content want to influence the emotions of the people who see it.

However it has always been hard to measure the value of the work that creative people do.

This is because the way they get paid often depends on fixed amounts of money or contracts or how many people are thought to have seen their work.

The new research idea uses computer programs to recognize facial expressions and score how engaged people are. This gives us a way to really understand how audiences feel about the content they see.

By looking at how students engage with content how they react emotionally. How well advertisements work the system can help make sure that artists get paid in a way that is fair and based on how well they do.



VIII. ETHICAL CONSIDERATIONS

The goal of this study is to look at how people feel when they see advertisements and how they react to them. We do not want to find out who these people are or what they think about things.. We have to be careful because we are working with pictures of peoples faces and that is private information. If we are not careful it can cause problems with peoples privacy. It is not fair. So we have to think about what's right and wrong when we are doing this research. We have to think about peoples privacy when we are looking at their faces. Pictures and videos of peoples faces can tell us who they are and what they are like. In this study we are taking pictures of students faces to see how they feel when they watch advertisements. We have to make sure we are respecting their privacy when we do this. We need to tell the students why we are taking their pictures and what we are going to do with them. We should only take the pictures that we need to understand how they feel and not keep any information, about them that we do not need. We need to make sure we are doing it in a way that is fair and respectful. The students faces will be.

We will look at the recordings to see how they feel.. We have to make sure we are not keeping any secrets from them and that we are telling them everything they need to know.

IX. CONCLUSION

The proposed research "Deep Learning-Based Facial Expression Recognition for Analysing Student Emotional Responses to Advertisements: A CNN and Transfer Learning Approach with Multimodal Engagement Scoring" is about using Artificial Intelligence and Computer Vision to understand how people feel. This study is trying to make a system that can tell how students feel when they watch advertisements. It looks at their expressions and how they behave to see how engaged they are. Usually people use surveys and questionnaires to see if advertisements are working.. These methods may not be perfect because people may not always say how they really feel. This research is trying to do something. It uses Deep Learning to recognize expressions and tell how people feel. The system uses something called Convolutional Neural Networks to look at faces and find features like eye movements and mouth patterns.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 9, June 2026
Impact Factor: 8.2

Convolutional Neural Networks are good at finding patterns in pictures without needing help. The researchers also used something called transfer learning with models like VGG16, ResNet50 and MobileNet. This helps the system work better and understand emotions accurately. The proposed system is about Deep Learning-Based Facial Expression Recognition and it is trying to make advertisement effectiveness analysis better. The Deep Learning-Based Facial Expression Recognition system uses Computer Vision and Artificial Intelligence to make this happen.

REFERENCES

[1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521, 436–444.
[2] Goodfellow, I. et al. (2013). Challenges in Representation Learning: Facial Expression Recognition. ICML Workshop.
[3] Lucey, P. et al. (2010). The Extended Cohn-Kanade Dataset (CK+). IEEE CVPR Workshops.
[4] Mollahosseini, A. et al. (2019). AffectNet: Database for Facial Expression Recognition. IEEE Transactions on Affective Computing.
[5] Simonyan, K., & Zisserman, A. (2015). Very Deep CNNs for Large-Scale Image Recognition. ICLR.
[6] He, K. et al. (2016). Deep Residual Learning for Image Recognition. IEEE CVPR.
[7] Howard, A. et al. (2017). MobileNets: Efficient CNNs for Mobile Vision Applications. arXiv.