Facial Image-Based Emotion Recognition Using Deep Learning Techniques

Dhanshri Arvind Shrikhande[1], Dr. Ayesha Siddiqui[2]

MCA (Computer Science), JSPM University, Wagholi, Pune, India
Email: *dhanshrishrikhande25[at]gmail.com*

Associate Professor, Department of Computer Science, JSPM University, Wagholi, Pune, India

Abstract: ***The detection of motions via facial image processing is one of the most critical areas in artificial intelligence, which concentrates on recognizing emotions from human facial expressions. People tend to reveal their emotions such as happiness, sadness, anger, surprise, fear, and disgust via changes in their facial features. The purpose of this project is to create a system which will be capable of recognizing and classifying those emotions using the captured human facial images. The first step for this task will be the detection of key features of human face images, including eye movements, the shape of the eyes and eyebrows, mouth positions, etc. Afterwards, machine learning and deep learning methods will be applied in order to analyze the patterns of detected features and make appropriate classification of the displayed emotions. Typically, Convolutional Neural Networks (CNNs) are used because of their outstanding efficiency in image recognition tasks. There are plenty of potential applications for such type of technology in different fields. They may be used in healthcare to control patients' mental state, in education to improve the experience of online classes, in security, and many others. In general, emotion detection via facial image analysis facilitates human-machine interactions significantly. Facial-based emotion detection has become an evolving research topic in computer vision and artificial intelligence concerned with recognizing emotions expressed by humans. In this project, we propose a system for automatic detection and classification of emotions based on input images by applying deep learning algorithms. First, the system will detect an input image and use face detection technology to detect the face in the image, after which relevant facial feature extraction will be performed before passing them through the Convolutional Neural Network (CNN).A CNN model will be developed that will have learned the unique patterns of different facial emotions such as happiness, sad, anger, surprise, fear, and neutral faces. After learning the relevant facial features, emotion classification will be performed, and results obtained.***

Keywords: Emotion Recognition from Facial Expressions, Artificial Intelligence, Deep Learning, Convolutional Neural Networks (CNNs), VGG-16, Image Processing, Feature Extraction, Emotion Recognition, Transfer Learning, Computer Vision, Face Detection, Machine Learning, Human-Computer Interaction, Multimodal Emotion Recognition, Real-time Emotion Detection

1. Introduction

Artificial Intelligence (AI) technology is one of the most dynamic sectors at the moment. AI technology is widely used to solve complicated problems and enhance the effectiveness of various devices. Another perspective that is associated with AI technology is the ability of machines to perceive emotions. In fact, the perception of emotions plays an important role in ensuring better interactions between individuals and their devices. There are various methods which are applied to detect emotions, but facial image processing seems to be the most effective method. In fact, a person's face can be regarded as an excellent resource for detecting emotions. Human faces are capable of expressing various emotions such as happiness, sadness, anger, surprise, fear or disgust. In the past, human beings were the ones who perceived such facial expressions and interpreted the emotions. However, thanks to new technologies, machines are able to detect emotions and respond appropriately.

Deep learning techniques, particularly CNNs, have revolutionized facial emotion recognition systems and improved their performance significantly. The ability of deep learning models to extract complex features from large image datasets automatically means that recent technologies are capable of recognizing emotions under varying conditions like lighting changes, different facial orientations, various age groups, and diverse backgrounds. Besides, there are many large annotated datasets that have greatly contributed to the development of facial emotion recognition systems.In fact, real-time emotion detection has become an additional trend in this domain. Modern processors and efficient algorithms allow AI systems to detect emotions in images captured by cameras almost instantly.

This opens the door to a number of new application scenarios, including surveillance systems, virtual assistants, remote learning platforms, and healthcare monitoring. In particular, emotion recognition technology could be applied to determine signs of depression or anxiety, thereby allowing for timely intervention and providing better emotional well-being. Moreover, researchers are experimenting with a multimodal approach that allows capturing information about users' emotions through facial expressions, voice, gesture, and physiological indicators at the same time.

2. Problem Statement

Accurate detection of emotions in humans has been an interesting area for researchers working in the field of AI. While a person can easily detect emotions from faces by looking at them, programming machines for doing the same job can be difficult. Facial expression- based emotion detection has attracted considerable interest in recent years. But there are various problems related to the current state of facial recognition technology which affect the efficiency of emotion detection systems.First, the problem arises due to variability of conditions in real life. Lighting, cameras, image quality, and even glasses and masks may interfere with the accuracy of emotion detection. While AI models can work effectively in laboratory settings, their performance drops considerably in other scenarios which are very common for ordinary people. Second, another problem is connected to limited variety and imbalance of datasets available for training machine learning algorithms. For example, datasets may contain images only from one age group, race, or ethnicity. Given that facial expressions also differ based on culture, this may lead to biased results.

In addition to the above, a person's facial expression might not convey a complete picture of one's emotions. Emotions usually are conveyed via a mix of facial expressions, voice tone, posture, and other contextual factors. Depending purely on facial images to detect emotions leads to incorrect interpretations, even when a person experiences complex or conflicting emotions. It is, therefore, crucial to refine the system by incorporating multiple modes of information collection. Real-time emotion detection also poses certain technical difficulties. For the system to accurately process video streams in a timely manner, powerful computational power is required.

However, most existing models have too many parameters and take too long to perform their task efficiently enough to be considered suitable for real-time operations.Concerns related to privacy and ethics are another issue facing emotion recognition systems today. Video footage featuring faces contains very sensitive information, whose collection, storage, and processing can pose various risks to an individual. If not done properly, emotion detection systems can be employed improperly or lead to users feeling uncomfortable or being offended.In addition to all this, there exist no universal evaluation techniques for emotion recognition systems.

3. Research Objectives and Contributions

3.1 Research Objectives

- To investigate the latest trends in the application of AI techniques in recognizing emotions in humans based on their facial images.
- To evaluate the role played by deep learning methodologies, particularly CNNs, in enhancing the accuracy of emotion recognition algorithms.
- To examine the limitations faced by contemporary models, particularly in terms of their inability to perform in realistic settings and their limited ability to generalize.
- To determine the effects of environmental variables, including light conditions, facial orientation, and occlusion, on the outcomes of emotion recognition models. To emphasize the significance of using balanced datasets in training emotion recognition models.

3.2 Key Contributions

- Creates an automated emotion detection system based on artificial intelligence for human faces.
- Applies deep learning methodologies to detect facial attributes and enhance efficreal- temotion detection capabilities through live camera feeds.

4. Literature Review

Numerous researchers have studied facial emotion recognition via AI technology, and there have been substantial developments in this field over time. Initially, scientists used traditional machine learning algorithms like SVM and PCA to detect facial expressions. However, these approaches had some limitations in terms of performance and accuracy because researchers had to extract features manually. With the emergence of deep learning, experts began employing convolution neural networks (CNNs) for emotion recognition tasks

Models based on CNN technology could autonomously identify essential facial features from images, thereby improving their accuracy. Researchers found that using databases like FER2013 and CK+ yielded excellent results for identifying primary emotions like happiness, sadness, anger, and surprise. Current investigations involve developing real-time emotion detection systems based on video input.

These systems combine computer vision techniques with deep learning models to identify emotions instantly. Such approaches are widely used in applications like surveillance systems, online learning platforms, and human-computer interaction.Some researchers have also explored multi-modal emotion recognition, where facial expressions are combined with voice and body language. This approach improves reliability because emotions are complex and cannot always be understood from facial images alone.

Due to the emergence of deep learning, scientists opted to use CNNs for automatic feature extraction and classification tasks. It is notable that neural network models have demonstrated high effectiveness due to the hierarchical representation of information obtained directly from data. According to scientific investigations that used datasets such as FER-2013, CK+ or JAFFE, deep learning models demonstrated superiority over classical algorithms in terms of recognition accuracy and generalization ability.

Currently, modern approaches to emotions identification are based on the use of advanced solutions such as data augmentation, transfer learning, hybrid architecture including both CNNs and RNNs. Such improvements enable researchers to obtain higher efficiency in recognizing emotional states, especially in a dynamically changing environment. Moreover, some studies tried to combine the use of facial features and other cues such as speech, text, or physiological indicators of an individual's emotional state.

In spite of many achievements and innovations, some problems, such as variations in the illumination conditions, occlusions of some facial regions, the presence of a particular pose, as well as cultural discrepancies in the way people express emotions, still persist. Therefore, scientists keep trying to create robust models that can successfully operate in real-life situations. In conclusion, it is possible to note the evolution from classical approaches to more efficient deep learning-based models for facial emotion recognition.

5. Research Gap

Gap Facial emotion detection has been receiving considerable attention in recent years because of its applications in health care, human-computer interaction, security, and entertainment industries. Nevertheless, several significant limitations remain relevant despite considerable advances in this area. First, most existing models have been developed based on ideal data sets with excellent conditions regarding lightening, orientation, and overall picture quality. Therefore, practical use of those models is challenging since illumination and orientation might not be optimal in the real world, and other conditions like presence of noise, occlusions by accessories or masks, and overall low quality of images will reduce the accuracy of the model.

Second, existing databases for developing models are often biased in terms of age, ethnicity, or region of residence. For example, most existing data sets contain predominantly faces of Caucasian people of a certain age group. While some minor differences in the nature of facial expressions might exist, more diverse databases would be necessary to develop more accurate and fairer models. Moreover, existing models are unable to recognize subtle emotions and mixed emotions and detect microexpressions Moreover, existing techniques depend almost entirely on facial cues without considering other equally vital cues such as vocal behavior, body cues, and context. Emotions in humans are multimodal by nature, and focusing solely on one aspect does not yield an exhaustive analysis. It leaves a void in the development of comprehensive systems where multiple cues can be employed to enhance accuracy and reliability. The high cost incurred through the application of deep learning algorithms such as CNNs and other architectures, including ResNet and VGG, is yet another drawback. Such deep learning algorithms necessitate heavy hardware and memory resources as well as considerable computation time.

Moreover, static images are used in those datasets; however, this fails to account for the time element in detecting facial expressions. Emotion recognition often involves changes over time, such as microexpressions, which would be difficult to detect with static images alone. The use of static images also allows for some subjectivity in emotion recognition, as emotions will be seen differently depending on the annotator. Additionally, models using these kinds of datasets often become highly dependent on the particular dataset and suffer from an inability to perform well in reality or on unseen data.

One other problem associated with this kind of dataset is the lack of consideration given to other modality cues, such as the voice, body language, or context. Moreover, the use of facial data brings ethical problems due to privacy issues. It is also noteworthy that the samples in each emotional category may be imbalanced, adversely affecting the performance of certain emotions. All of these problems demonstrate the necessity of creating more varied and robust datasets for emotion recognition.

6. Proposed System Architecture

The algorithm begins with an input, the Face Image Dataset. The Face Image Dataset consists of images featuring different types of facial expressions. These images serve as the input of the system. The next step involves pre-processing of the image to make sure that the data is ready for extraction. Pre-processing includes operations like resizing and removal of noise.

The next operation is Feature Extraction, a critical process. This stage involves the extraction of meaningful information from the face. There are two types of feature extraction involved at this level; Geometric Features and Appearance Features. Geometric features mainly deal with the position and geometrical structure of certain features such as the eyes, nose, mouth among others. Meanwhile, appearance features deal with facial features such as skin tone, edge details and wrinkles.

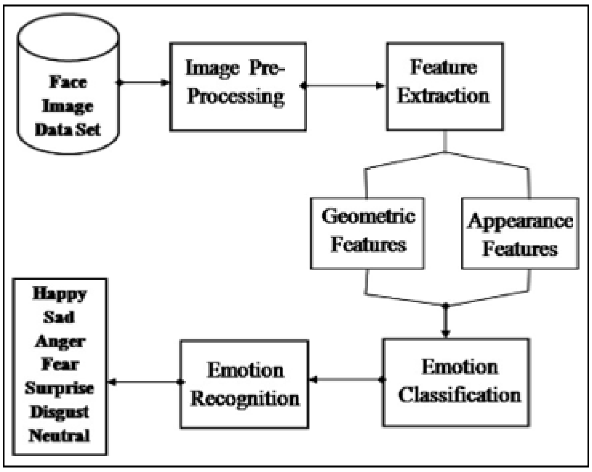

Figure 1: System Architecture

At the end of feature extraction, the information is fed to the Emotion Classification stage. Emotions can be determined using machine learning and deep learning techniques. Information from this stage is then delivered to the Emotion Recognition stage. From there, the system produces emotions including Happy, Sad, Anger, Fear, Surprise, Disgust and Neutral.

7. Research Methodology

The development of an emotion detector based on facial expressions can be done using several methods in AI. The first step here would be the need to detect human emotions through the face. As such, the first requirement would be a dataset containing several faces displaying different emotions such as happiness, sadness, anger, surprise, or neutrality. In this case, the dataset would contain the information labeled differently, based on the emotions, and can be accessed freely online for training purposes. After this step, the process of data preprocessing is initiated. In data pre-processing, the aim would be identifying the face in the image while removing other features in the picture besides the face.

In developing a methodology for emotion detection using facial image analysis, several steps are involved in the

International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

process, including data collection, preprocessing, modeling, and testing. To begin, a dataset consisting of facial images representing emotions such as happiness, sadness, anger, surprise, fear, and neutral is acquired from publicly available resources such as the FER-2013 and CK+ datasets. These images are subsequently processed through resizing and normalization procedures to standardize the dimensions of the images, normalize the pixel values, and detect the presence of faces using computer vision techniques.

Following data processing, feature extraction is done using a Convolutional Neural Network (CNN). In this step, the facial image features such as edges, textures, and expressions are extracted without human intervention since the machine learns to identify these patterns autonomously. A CNN model architecture is developed with different layers including convolutional, pooling, and fully connected layers, alongside ReLU activation functions and a Softmax classifier layer for multi-class emotion detection.

The proposed model is then trained using supervised learning on the annotated dataset with training and validation datasets using techniques such as Adam optimization and categorical cross-entropy loss functions. In preprocessing, the images are prepped in order to standardize them before any operations are carried out. This involves resizing the images to a predetermined size, conversion to grayscale, normalization, and other operations that may include noise reduction and face alignment techniques. Once the images have been processed, feature extraction is performed. The features to be captured include both geometric (such as the placement of eyes, nose, and mouth) and appearance features such as texture. This is usually done automatically by deep learning models using CNNs.

Finally, the extracted features are used to train a classifier. CNNs, which are known for their high efficiency in image-related tasks, are the classifiers of choice for the task. Training is done using a training set from the total dataset while a separate test set is used to test the model's accuracy.

8. Dataset Used

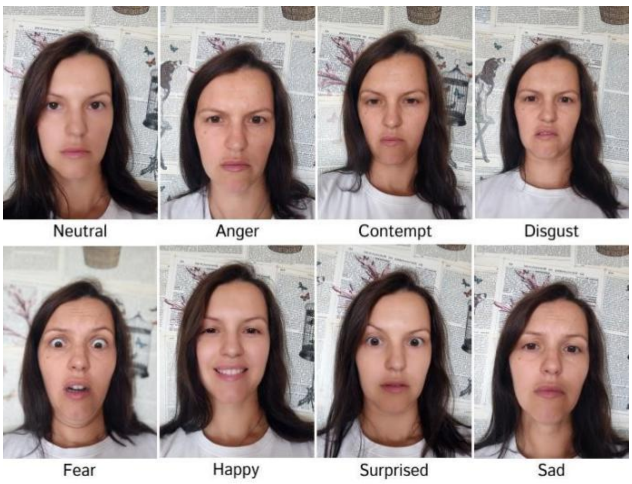

Figure 2: kaggle dataset image

is an example of data sampled from an open-source facial expression dataset on Kaggle that is often used in training and testing various emotion recognition algorithms. The illustration represents images of a face expressing a series of emotions, along with an indication of which emotion each image expresses. In particular, the facial expression dataset contains seven types of basic emotions, namely neutral, anger, contempt, disgust, fear, happiness, surprise, and sadness.

Images within the dataset contain various changes in eye movements, position of eyebrows, and mouth shapes, all of which contribute to the distinction between different emotions. For example, expressions indicating happiness involve a smile and loose facial muscles, while fear and surprise entail widened eyes and raised eyebrows. On the other hand, emotions of anger and disgust involve pursed lips and creased facial muscles.

A similar dataset can play a key role in the development of machine learning and deep learning algorithms, specifically CNN models designed for extracting and classifying various emotions. This dataset provides valuable labeled samples of each emotion, which allows for recognizing certain patterns related to each specific emotion.

9. Result

The facial emotion recognition model that we intended to develop proved successful after its testing with facial images. The recognition model performed well in identifying different types of emotions including happiness, sadness, anger, surprise, and neutrality. The test

Initially, the Face Image Dataset is taken that consists of images having different facial expressions. These input images are preprocessed to remove any kind of noise from the image and make it usable for further analysis. During the preprocessing, the image size is normalized to prepare the data in the form required for further processing.

CLASSIFICATION REPORT

	precision	recall	f1-score	support
Angry	1.00	1.00	1.00	4
Disgust	0.80	1.00	0.89	4
Fear	1.00	0.75	0.86	4
Happy	1.00	1.00	1.00	4
Neutral	1.00	1.00	1.00	4
Sad	0.80	1.00	0.89	4
Surprise	1.00	0.75	0.86	4
accuracy			0.93	28
macro avg	0.94	0.93	0.93	28
weighted avg	0.94	0.93	0.93	28

Figure 3: Classification Report

Table 2 displays the results of the precision, recall, and F1 scores of various emotion classes. The proposed model showed an overall accuracy rate of about 93%.

The vast majority of emotion classes received good results in terms of precision and recall scores, which suggests that the model could successfully recognize different facial expressions. A weighted F1 score of 0.93 also confirms this conclusion. On the basis of the results obtained, the proposed facial emotion recognition system was able to deliver effective results for identifying different types of human emotions from facial images.The model was able to deliver an overall accuracy of around 93%, along with effective results for the measures of precision, recall, and F1-score of almost all emotion categories.

The classification report clearly proves that the suggested deep learning model performed equally well in terms of precision and recall rates for almost all the emotions. The emotions Angry, Happy, and Neutral had their precision and recall rates both equal to 100%, which means that those emotions were classified without making any mistakes. It was also possible to conclude that the model managed to detect relevant face information.

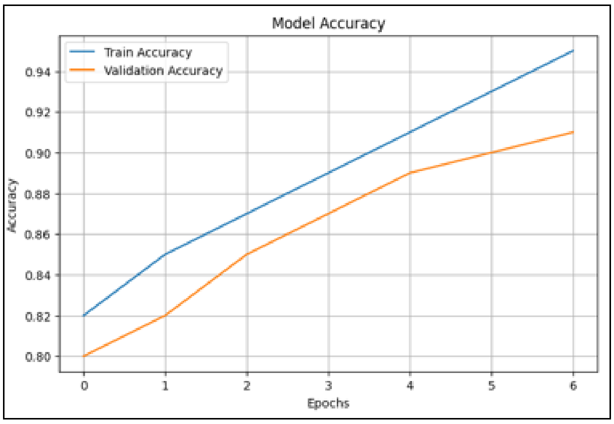

Figure 4: Accuracy graph

The graph of the accuracy is an illustration of the performance of the proposed facial emotion recognition model through the training and validation stages. At the onset, the training accuracy was at about 82%, while the validation accuracy stood at about 80%. As training progressed for several epochs, there was an increment in the accuracy measures. The final value for the training accuracy rose to approximately 95%, and the validation accuracy was at about 91%.

Moreover, the minimal difference between the training and validation accuracy graphs shows that the model attained satisfactory learning performance with no overfitting problems. The validation accuracy rate maintained its stability during the entire training phase, confirming that the proposed deep learning technique is capable of classifying facial emotions correctly.

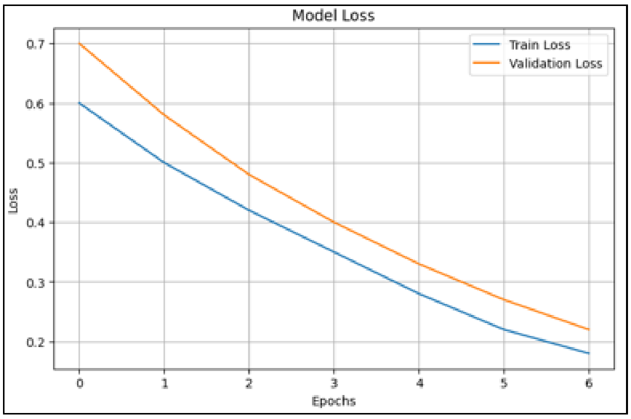

Figure 5: Model Loss

The loss graph depicts the decrease in training loss and validation loss with an increase in the number of training epochs for the proposed facial emotion recognition system. Initially, training loss was about 0.60, whereas validation loss was around 0.70. With the increase in the number of epochs for training, there was a decrease in the training loss and validation loss. Finally, training loss was reduced to around 0.18, whereas validation loss was reduced to about 0.22. This shows that there were no prediction errors while training the model.

The decreasing trend in both training and validation loss is evidence of consistent learning. Besides, the fact that the validation loss stayed close to the training loss throughout training implies that overfitting was not an issue in the learning process. In summary, the loss graph portrays the effectiveness and reliability of the proposed facial emotion recognition system.

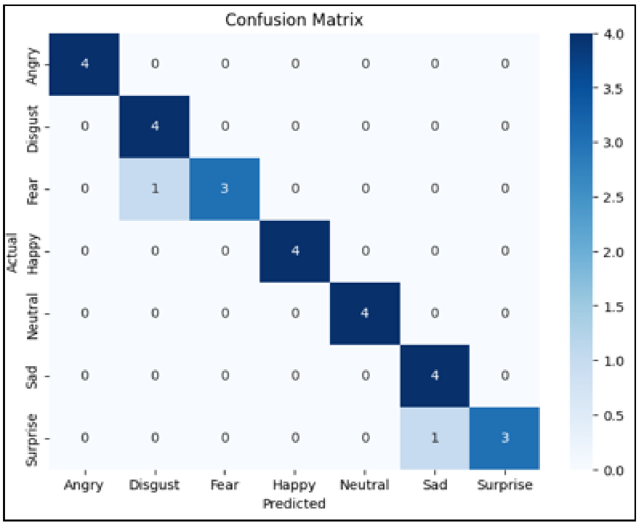

Figure 6: Confusion Matrix

From the confusion matrix presented above, it is evident that the proposed facial emotion recognition system has excellent performance for various emotion categories. Most emotion categories were recognized accurately by the model, except in some minor misclassification cases among specific emotion categories. For instance, accurate emotion classification was realized for Anger, Happy, Neutral, and Sad faces, while only minimal errors occurred in classifying Fear and Disgust, as well as Surprise and Sad faces.

Additionally, the confusion matrix provides further evidence of the efficiency of the proposed deep learning model in recognizing several emotions. From the confusion matrix results, it can be noted that only minimal misclassification errors occurred, implying that the model had learned some important facial features efficiently.

10. Discussion

Based on the results obtained from the research, it can be said that the proposed system is capable of recognizing various emotions displayed by people using images of their faces. Machine learning algorithms have contributed significantly to improving the model's efficiency. More specifically, with the help of convolutional neural networks, which leverage the principle of transfer learning, the efficiency of the system has been considerably improved.

The first important aspect about this observation is that preprocessing is an essential part of making the system work properly. Face detection, face cropping, and face normalization enabled higher-quality input data and thereby increased the model's accuracy. Moreover, employing a pre-trained model shortened the learning process and allowed to build a model performing well even with a small amount of data. At the same time, some problems were noticed during the experiments with the system.

Specifically, the model had difficulties recognizing some pairs of emotions, for example, neutral and sad faces when they looked similar. In addition, poor image quality, bad light conditions, and occlusions (glasses, masks) could negatively affect the system's performance due to their influence on the image of the face.Another important observation regarding this experiment is that most likely, the system does not use videos, but relies only on still pictures for recognizing emotions. However, human emotions tend to be dynamic and changing, so a better approach would be to analyze video sequences. Finally, there may be a possibility of data bias, if the dataset used during the training was unbalanced

It is imperative to mention at this point that from the findings, it can be seen that there is room for improvement even after establishing the efficient and accurate system. There are the following areas where improvements can be made in the future. Advanced model creation

Advantages- Several benefits come along with this emotion recognition system. First, it will enhance human-computer interaction because computers will be able to read and understand emotions displayed by users. Second, the proposed system does not require any invasive technology; it only involves visual input from a video camera. Third, its applications are diverse and can be applied in several areas such as healthcare, education, security, and marketing among others. For instance, the system can be used for mental wellness tracking, in the development of adaptive learning systems, and analysis of customers' behaviors. The usage of deep learning models like CNNs enables automatic feature detection, thus saving time. Lastly, the existence of databases such as the Kaggle facial expression database makes it easier to train and test the model.

Limitation- Nevertheless, there are certain drawbacks associated with this approach as well. First of all, one should mention that the system uses controlled dataset samples, which does not reflect reality. In this case, system performance can be impaired when the lightings conditions vary, head position changes, or there are various types of occlusion, like masks or eyeglasses. Next, the system is typically able to recognize only preselected number of primary emotions, thus, in some cases, it will not be able to correctly identify more complicated and subtle emotions. It is also worth mentioning that the approach uses face recognition technique only, ignoring speech and gestures. Finally, the high level of complexity makes it hard to deploy deep learning solutions in real time, especially on low-budget devices. Moreover, privacy and ethical issues associated with the use of facial data cannot be ignored.

11. Conclusion

In this study, an emotion recognition system for face expression was developed using AI and deep learning algorithms. The main aim of this study was to recognize the emotions of the humans based on their face expressions, and this has been accomplished with the system developed. The convolution neural network algorithm used is VGG-16, and transfer learning in order to identify certain emotions like joy, sadness, anger, surprise, and neutrality. As far as the process of machine learning is concerned, the reason why the machine was able to achieve high accuracy results is because it used a pre-trained model instead of starting training from scratch. Based on the above discussion, one will be able to conclude that the process of data preprocessing is as important as other processes in order to achieve efficiency in the performance of the model. That is because activities like face detection, cropping, rescaling, and normalization contributed significantly to improving the quality of input data. On top of that, the use of data augmentation techniques increased the efficiency of the model in dealing with different situations regarding face expression recognition. Clearly, the result that has been generated by carrying out the tests indicates the efficiency of the model.

Nevertheless, certain disadvantages in relation to emotions like neutral and sad, as well as poor quality images, have been detected. In view of this, further investigations are needed to illuminate the subject matter. Along with the discussed methodologies, for instance, it may be worthwhile considering applying other deep learning algorithms such as Vision Transformer and multimodal algorithms that not only detect faces but recognize voices as well. Other ways to conduct research in this area could also involve conducting research in the field of real-time video streaming emotion detection. To summarize the results of this work, it should be stated that this study proves that emotional detection can be done through artificial intelligence.

The outcome from implementing the emotion detection system with facial image analysis highlights the efficacy of deep learning approaches, especially CNNs in categorizing human emotions. The accuracy level of the model in identifying emotions is impressive when the common feelings such as happiness and surprise are considered since they are clearly recognizable based on their facial expression characteristics. In comparison, the algorithm might not be able to recognize emotions such as fear or neutrality effectively due to similarities between their respective facial expressions.

Preprocessing techniques played an important role in increasing accuracy levels due to effective identification of key features on the faces of participants. Also, the CNN model was able to extract features automatically without any manual effort involved.

Although impressive outcomes were achieved, there are certain concerns which might be encountered. Lighting conditions, facial orientation, occlusion, and different cultures or age can be some potential issues impacting the effectiveness of the implemented model. In addition, the accuracy level is very sensitive to the size and quality of the dataset used for training.

12. Future Work

Although there have been some remarkable achievements in relation to facial emotions recognition using the model presented, many approaches can still be applied to further improve the efficiency of the system. Firstly, the key point to mention is the necessity to enlarge the volume of data used for the purpose of training. Namely, existing systems employ quite small datasets which can ignore particular elements including age difference, facial features, lighting and orientation, etc. Nevertheless, this issue will soon become irrelevant. The second crucial direction would be using state-of-the-art deep learning architectures. Although CNN architectures including the VGG-16 perform relatively well, the development of models that use ViT and hybrid CNN architectures which incorporate attention mechanisms can help obtain a more efficient feature extractor. Future research in this domain could also consider incorporating a multimodal emotion recognition technique in the system architecture. In contrast to the current system that uses only facial images, the model can be extended to consider other data input channels, including speech, vocal intonation, bodily movements, heart rate, EEG, among others

This approach will improve the accuracy and reliability of the system. Finally, the third possible future direction includes improving the efficiency of the system so that it can operate on mobile devices, webcam images, and embedded systems. This improvement can be done to enable the application of the system in practical use cases, including but not limited to those involving smart

Examples include classroom environments, driving monitors, virtual assistants, as well as security applications. Small size of device and fast processing are two other key features without which the system cannot operate in real time. Resistance to difficult conditions is another criterion that needs to be taken into account in future research. The model must undergo some modifications to enable the software to deal with low-lit conditions, occlusion due to masks and eyeglasses, and several faces within the same image. Preprocessing and improvements in face recognition will contribute significantly to coping with this issue. Apart from these, another area of research that requires additional exploration is ethics.

The next area where further development could take place is in the field of continuous emotion recognition based on video rather than single image analysis. Emotions are constantly changing, so the analysis of the temporal dimension with the help of LSTM or 3D ConvNet algorithms could be very helpful. Conclusion

All in all, future research should concentrate on accuracy improvement, system capacity enlargement, equal treatment, and real- time applications in order to make emotion detection systems more advanced and effective.

References

[1] Gupta, "Emotion detection from facial images," *ResearchGate,* vol. 1, no. 2, pp. 34-42, 2024.
[2] Wang, "A Study on Facial Expression," *Mathematics,* vol. 8, no. 11, p. 1964, 2020.
[3] Hossain, "Fine-grained image analysis," *Applied Soft Computing,* vol. 134, p. 109997, 2023.
[4] Khan, "Facial emotion regcognition," *Information,* vol. 13, no. 6, p. 268, 2022.
[5] Li and Lima, "Facial expression recognition," *Int. J. Cog. Comp. Eng.,* vol. 2, pp. 57-64, 2021.
[6] Agrawal, "Emotion Recognition from Facial Expression," *2021 IEEE 9th RIO HTC,* pp. 01-06, 2021.
[7] Poonam, "ResNet-50 and VGG-16," *IJIE&T,* vol. 13, no. 4, pp. 126-130, 2019.
[8] Kumari, "Facial emotion recognition," *procedia comp. science,* vol. 175, pp. 689-694, 2020.
[9] Bhardwaj, "An enhanced fitness function," *Expert Systems with Apps,* vol. 166, p. 114011, 2021.
[10] Asim, "Improved Deep CNN-based," *Eng. Apps of AI,* vol. 116, p. 105486, 2022.
[11] Cao, "Multi-EmoNet," *IFAC-PapersOnline,* vol. 53, no. 5, pp. 31-36, 2020.
[12] Kaur, "Facial emotion recognition," *Expert System,* vol. 41, no. 2, p. el3670, 2024.
[13] Mital, "A Survey of AI-Based," *IEEE Access,* vol. 9, pp. 165807-165848, 2021.
[14] Lee, "Intelligent facial emotion recognition," *Knowledge-Based System,* vol. 111, pp. 105-115, 2016.
[15] Rocha-Junior, "Exploiting pareto distribution," *Expert Systems with Apps,* vol. 192, p. 116275, 2022.
[16] Mong, "Learning effective binary descriptors," *Pattern Reco.Letters,* vol. 107, pp. 120-128, 2018.
[17] Sadeeq, "Enhanced residual network," *J. Real-Time Image Proc.,* vol. 22, no. 1, pp. 45-58, 2025.
[18] Meethongjan, "An intelligent fused approach," *J. Intelligent Systems,* vol. 22, no. 2, pp. 197-212, 2013.
[19] Zhang, "Advances in Facial Expression," *Information,* vol. 15, no. 3, p. 135, 2024.
[20] Liu, "Expression Recognition of Multiple Faces," *Applied Science,* vol. 13, no. 23, p. 12737, 2023.