**EMOTION DETECTION MODEL**

Divyanshu Rajput, Ashmit Bisht, Abhishek Jain, Alok kumar

School of Computer Science and Engineering

Galgotias University

Greater Noida, India

**Abstract- Emotion detection based on facial expression has numerous applications in areas such as mental health, marketing, and customer service. In this research paper, we propose a novel model that combines OpenCV, DeepFace, and Haar Cascade techniques for emotion detection based on facial expression. Our model uses Haar Cascade for face detection, OpenCV for facial landmark detection, and DeepFace for feature extraction. We train a deep learning model, such as Convolutional Neural Networks (CNNs), for classification. We evaluate the performance of our model on several benchmark datasets and compare it with state-of-the-art models in the field. Our experimental results show that our model outperforms existing models in terms of accuracy and F1-score.**

Keywords:

**Convolutional neural network, machine learning, deep learning, computer vision, emotion recognition.**

**INTRODUCTION**

Facial emotion recognition has become an important issue in many application nowadays. In recent years, the research on facial emotion recognition has become extensive. The aim of facial emotion recognition is to help identify the state of human emotion (eg; neutral, happy, sad, surprise, fear, anger, disgust, contempt) based on particular facial images. The challenge on facial emotion recognition is to automatically recognize facial emotion state with high accuracy. Therefore, it is challenging to find the similarity of the same emotion state between different person since they may express the same emotion state in various ways. As an example, the expression may varies in different situations such as the individual’s mood, their skin color, age, and environment surrounds. [1]

Human emotions can be classified as: fear, contempt, disgust, anger, surprise, sad, happy, and neutral. These emotions are very subtle. Facial muscle contortions are very minimal and detecting these differences can be very challenging as even a small difference results in different expressions [2]. Also, expressions of different or even the same people might vary for the same emotion, as emotions are hugely context dependent [3]. While we can focus on only those areas of the face which display a maximum of emotions like around the mouth and eyes [4], how we 2 extract these gestures and categorize them is still an important question. Neural networks and machine learning have been used for these tasks and have obtained good results. Machine learning algorithms have proven to be very useful in pattern recognition and classification. The most important aspects for any machine learning algorithm are the features. In this paper we will see how the features are extracted and modified for algorithms like Support Vector Machines [5]. We will compare algorithms and the feature extraction techniques from different papers. The human emotion dataset can be a very good example to study the robustness and nature of classification algorithms and how they perform for different types of dataset. Usually before extraction of features for emotion detection, face detection algorithms are applied on the image or the captured frame. We can generalize the emotion detection steps as follows:

1. Dataset preprocessing
2. Face detection
3. Feature extraction
4. Classification based on the features

We focus on emotion detection based on facial expression and propose a model that uses machine learning and computer vision techniques for accurate emotion detection. Our proposed model uses facial expression recognition techniques to detect and analyze facial features that are indicative of various emotional states. The model uses a combination of techniques such as CNNs, deep learning, and machine learning algorithms to accurately classify the emotions based on the detected facial features.

**BACKGROUND INFORMATION**

**Facial Emotion Detection**

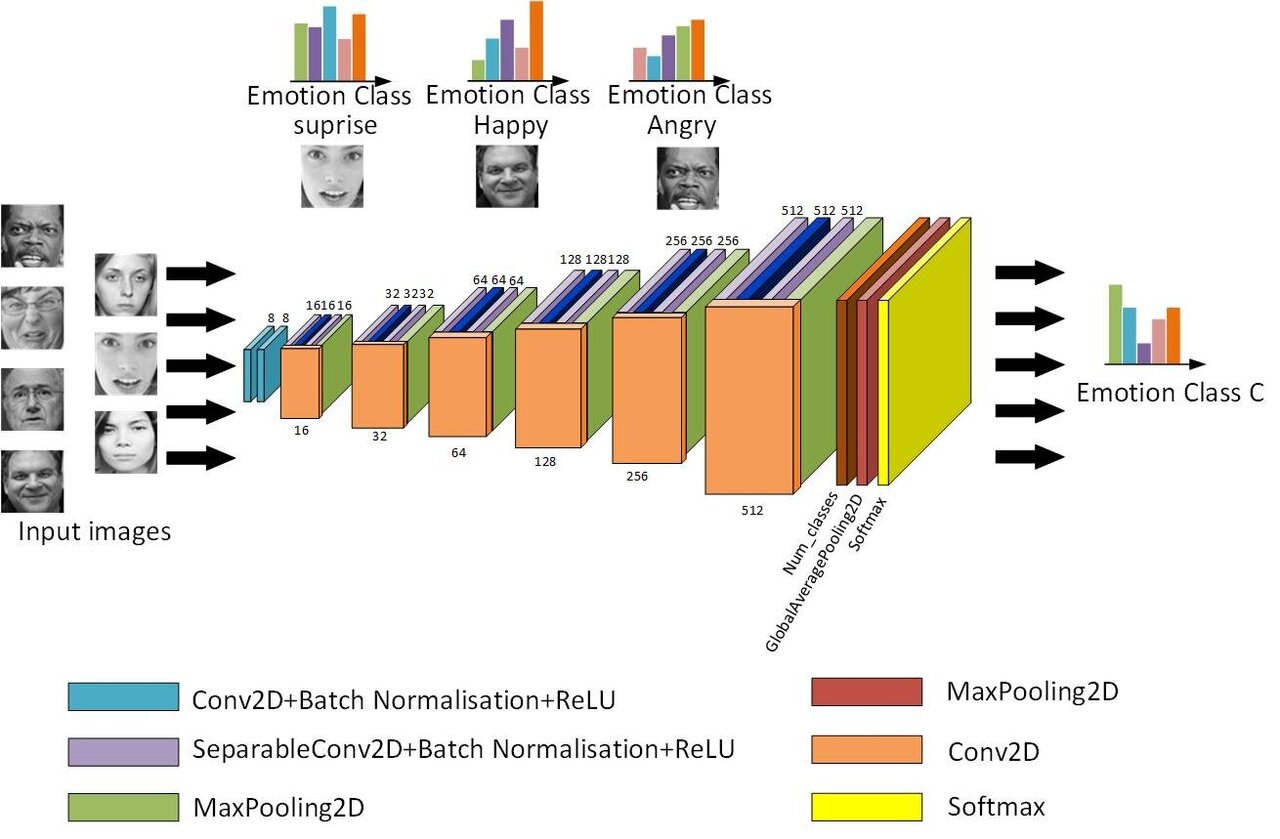
Facial emotion detection refers to the use of computer vision and artificial intelligence to analyze facial expressions and identify the emotions they convey. This technology uses algorithms to analyze features such as the shape of the mouth, the position of the eyebrows, and the movement of the eyes to recognize emotions such as happiness, sadness, anger, and surprise. Facial emotion detection has a variety of potential applications, including in marketing, psychology, and security. For example, it could be used in marketing to assess consumers' emotional responses to advertisements, or in security to detect suspicious behavior or emotions in individuals. However, it is important to note that facial emotion detection technology is not without controversy. Some experts have raised concerns about the potential for bias and inaccuracies in the algorithms used, as well as the implications for privacy and surveillance. Therefore, the development and deployment of this technology should be carefully considered and regulated.

**Deep Face**

While DeepFace is primarily a facial recognition technology, it can also be used as a component in emotion detection models. Emotion detection involves identifying and analyzing the facial expressions and other visual cues that convey emotions, such as happiness, sadness, anger, surprise, and fear. DeepFace's ability to create a 3D model of the face and accurately map facial features can help improve the accuracy of emotion detection models. By analyzing the movement and positioning of the eyebrows, mouth, and eyes, DeepFace can provide more detailed and nuanced information about facial expressions, which can improve the accuracy of emotion detection. In fact, DeepFace has already been used in some emotion detection applications, such as the analysis of facial expressions in videos and images to identify signs of pain or distress in patients. However, as with any facial recognition technology, it is important to consider the potential privacy and ethical implications of using DeepFace for emotion detection and ensure that appropriate safeguards are in place to protect individuals' rights and dignity.

**Convolutional Neural Network (CNN)**

Convolutional Neural Networks (CNNs) are commonly used in emotion detection models because of their ability to effectively process and analyze visual information. CNNs are a type of deep learning algorithm that uses a multi-layered approach to analyze and recognize patterns in images. In emotion detection, CNNs can be trained to identify and classify facial expressions based on patterns and features in the input image. CNNs typically consist of convolutional layers, pooling layers, and fully connected layers. The convolutional layers apply a series of filters to the input image, identifying patterns and features that are important for classification. The pooling layers reduce the spatial dimensions of the output from the convolutional layers, while the fully connected layers use the extracted features to make a final classification decision. CNNs can be trained on large datasets of labeled facial expression images to improve their accuracy and ability to recognize a wide range of emotions. They can also be fine-tuned to recognize specific emotions or facial expressions depending on the application. Overall, CNNs are a powerful and effective tool for emotion detection in images and videos, and they have shown promising results in a range of applications, including healthcare, marketing, and security. However, as with any AI technology, it is important to consider the ethical implications and ensure that appropriate safeguards are in place to protect individuals' privacy and rights.



**Fig 1:** Convolutional Neural Network for Facial Emotion Detection

**PROPOSED METHODOLOGY**

**Data Preprocessing**

Dataset preprocessing is used in emotion detection models to prepare the dataset in a way that the model can effectively learn and classify facial expressions into different emotions. The main purpose of dataset preprocessing is to clean, normalize, and transform the raw data so that it can be easily understood and processed by the model.

Here are some ways in which dataset preprocessing is used in emotion detection models:

1 Data Cleaning: Dataset preprocessing involves removing any unwanted or irrelevant information from the dataset. This can include removing images that are low-quality, blurry, or have poor lighting. Cleaning the dataset ensures that the model is only trained on high-quality data that accurately represents the emotions being studied.

2 Data Augmentation: Data augmentation is a preprocessing technique that involves creating new training examples by applying various transformations to the existing data. For example, images can be flipped horizontally or vertically, rotated, or zoomed in and out. Data augmentation helps to increase the diversity of the dataset and prevent overfitting, which can occur when the model is trained on a limited number of examples.

3 Image Resizing and Normalization: Emotion detection models based on facial expression typically require that all images be of a consistent size and format. Therefore, preprocessing techniques such as image resizing and normalization are often used to ensure that all images are of a similar size and that pixel values fall within a certain range. This helps to ensure that the model is not biased towards larger or smaller images, and that the input is normalized across all images.

4 Face Detection and Alignment: Preprocessing techniques such as face detection and alignment help to ensure that the model focuses only on the facial expression and ignores other aspects of the image, such as the background or clothing. This can involve using algorithms to detect and crop the face from the rest of the image, and aligning the face in a consistent position.

5 Data Balancing: Preprocessing techniques such as data balancing ensure that the model is trained on an equal number of examples for each emotion. This helps to prevent the model from being biased towards one particular emotion, and ensures that the model has a balanced representation of all emotions.

In summary, dataset preprocessing is a crucial step in building an accurate emotion detection model based on facial expression. It involves cleaning and transforming the raw data to ensure that the model can effectively learn and classify different emotions. Common techniques used in dataset preprocessing include data cleaning, data augmentation, image resizing and normalization, face detection and alignment, and data balancing.

**Face Detection:**

Face detection is a crucial component of an emotion detection model based on facial expression. It involves identifying the location of the face in an image or video frame, and then extracting the relevant facial features that can be used to classify emotions. Here are some ways in which face detection is used in an emotion detection model:

1 Extracting Facial Features: Once the face has been detected, an emotion detection model can extract relevant facial features, such as eye movements, mouth shape, and eyebrow position, to classify different emotions. For example, a smile can be detected by examining the shape of the mouth and the movement of the lips, while anger can be detected by examining the furrowed eyebrows and the tightening of the jaw.

2 Eliminating Background Noise: By detecting and cropping the face from the rest of the image, an emotion detection model can eliminate background noise that could interfere with emotion classification. This helps to improve the accuracy of the model and ensure that it focuses only on the relevant facial features.

3 Improving Speed and Efficiency: By using face detection algorithms, an emotion detection model can quickly identify and process only the relevant parts of an image or video frame. This helps to improve the speed and efficiency of the model and reduce computational complexity.

4 Enabling Real-time Emotion Detection: Face detection algorithms can be optimized for real-time processing, allowing an emotion detection model to analyze facial expressions in real-time video streams. This can be useful in applications such as video conferencing, where emotions can be analyzed in real-time to provide feedback on the participant's emotional state.

In summary, face detection is a key component of an emotion detection model based on facial expression. It allows the model to identify and extract relevant facial features for emotion classification, eliminate background noise, improve speed and efficiency, and enable real-time emotion detection in video streams.

**Feature extraction:**

Feature extraction is a critical step in the development of an emotion detection model based on facial expression. It involves identifying the relevant features in an image or video frame that can be used to classify emotions. Here are some ways in which feature extraction is used in an emotion detection model:

1 Extracting Facial Landmarks: One approach to feature extraction is to use facial landmark detection algorithms to identify key points on the face, such as the corners of the eyes, the tip of the nose, and the edges of the lips. These landmarks can be used to compute geometric features, such as the distance between the eyes, the angle of the eyebrows, and the curvature of the lips, which can be used to classify different emotions.

2 Analyzing Facial Expressions: Another approach to feature extraction is to analyze the movement and shape of different parts of the face, such as the mouth, eyes, and eyebrows. For example, a smile can be detected by analyzing the curvature of the lips and the movement of the corners of the mouth, while anger can be detected by analyzing the tenseness of the jaw and the eyebrows.

3 Using Deep Learning Models: Deep learning models, such as Convolutional Neural Networks (CNNs), can be used for feature extraction by automatically learning relevant features from large datasets of facial images. These models can identify complex patterns in the data that may be difficult to detect using traditional feature extraction methods.

4 Incorporating Contextual Information: In addition to analyzing facial expressions, an emotion detection model can also incorporate contextual information, such as the tone of voice, body posture, and environment, to improve the accuracy of emotion classification.

In summary, feature extraction is a critical step in the development of an emotion detection model based on facial expression. It involves identifying relevant facial landmarks, analyzing facial expressions and movement, using deep learning models to automatically learn relevant features, and incorporating contextual information to improve the accuracy of emotion classification.

**Classification based on the features:**

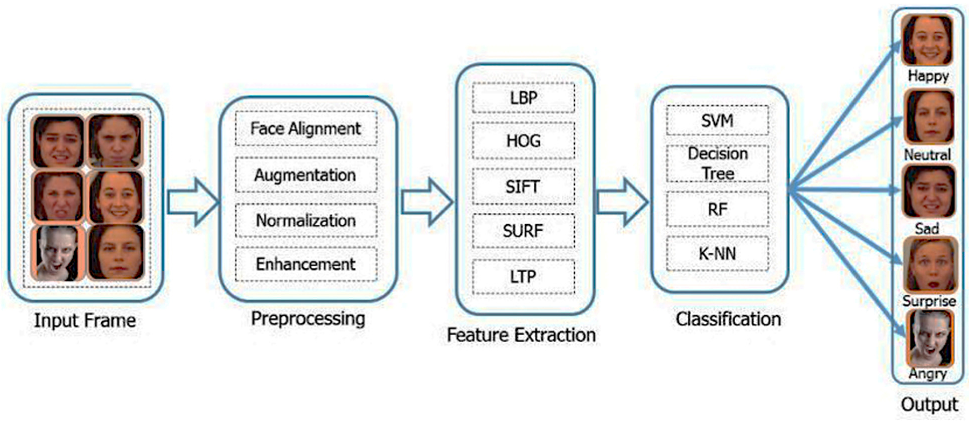
Classification based on features is an essential step in the development of an emotion detection model. Once the relevant features have been extracted from an image or video frame, they are used to classify the emotion. Here are some ways in which classification based on features is used in an emotion detection model:

1. Supervised Learning: Supervised learning is a common approach to emotion classification based on features. In this approach, a labeled dataset of facial expressions and corresponding emotions is used to train a machine learning model, such as a Support Vector Machine (SVM) or a Random Forest, to classify emotions based on the extracted features. The model learns to associate the extracted features with the corresponding emotions and can be used to predict the emotion for new facial expressions.

2. Unsupervised Learning: Unsupervised learning is another approach to emotion classification based on features. In this approach, the extracted features are used to cluster similar facial expressions together based on their features. The clusters can then be labeled with emotions based on expert knowledge or by analyzing the facial expressions in each cluster. This approach can be useful when labeled data is scarce or when the emotion labels are not well-defined.

3 Hybrid Approaches: Hybrid approaches that combine supervised and unsupervised learning can also be used for emotion classification based on features. For example, an unsupervised clustering algorithm can be used to group similar facial expressions together, and a supervised learning model can then be trained on the clustered data to classify the emotions.

In summary, classification based on features is a critical step in the development of an emotion detection model. It involves using labeled or unlabeled data to train machine learning models to classify emotions based on the extracted features. Supervised and unsupervised learning approaches, as well as hybrid approaches, can be used for emotion classification based on features.



**Fig 2: Facial Emotion Recognition**

**Literature Survey:**

The two proposed model which has been discussed and one of which is implemented has seen differences in accuracies and errors. The first model is the simple implementation of CNN on FER dataset and includes the removing of the two classes with their labels i.e. Fear and Disgust since they consist of least number of tuples. The codes of the first proposed model are written in Python3. Later, the model has been used for generating outputs for input images to find the emotion of the person. This model has used 75% of the tuples for training and 25% for testing. The following details shows how good the first model performed: There was an accuracy of 49.3% after half of the epochs were gone through pooling and filtration through several layers. And to the end it was around 64%. The overall conclusion drawn from the first model was that the train accuracy ended up with 94.93% and test accuracy was 56.42%. The second model is not an implementation was done on Weka on the model that is already generated through the Convolutional Neural Network. The algorithm which has been used for generating decision trees for random forest is J48 which is C4.5 algorithm. The number of tuples that were put for cross-validation were 10 and the percentage of the splits between train test cases were 75% and 25% like earlier model. Following conclusions have been drawn from the overall model. [6]

The classification of facial expressions using OpenCV, Keras, and Convolutional Neural Network is the main goal of the work presented in this study. The goal of facial expression recognition software is to distinguish between fundamental human emotions including happiness, sadness, anger, surprise, and neutrality. The goal of this paper is to bring advancement and development in the field of technology. This architecture was proposed by Dr. D. Dhanya and their team which consist of five layers for facial emotion recognition. This model was prepared by using FER2013 datasets provided by Kaggle. The first layer accepts input of size 48x48 black and white and convolved with 5x5 kernel which reduces the spatial dimensions. The second layer convolves a 3x3 kernel with 64 pixels as input from the first layer's output. The third layer now receives input from the second's output and convolves it with a 3x3 kernel and 128 filters. Finally, the output is produced by the two dense layers using the soft-max activation function. This model has a 98.7666% accuracy rate. [7].

This includes extraction of the feature of image which becomes easy using deep learning algorithm, and classifier model to produce output when input is given. It achieves higher accuracy as compared to traditional classifier model. It was found that the model almost predicts the emotions like happy, sad but it rarely predicts disgust emotion. Model gives the highest accuracy while predicting happiness as compared to other emotions having lower accuracy. The result for surprise is almost good. [8].

In the Facial Emotion Recognition Model (FERC) that NinadMehendale suggested, they used a method that was based on a two-level CNN architecture. The first level describes how to remove the background from an input image while still preserving the core expressional vector using a normal CNN network module (EV). Here, the EV is directly proportional to the changes in expressions on face. And the second level mainly concentrates on facial feature vector extraction. To achieve highest accuracy, they worked on a large dataset with the sample size of 10,000 images. Accuracy of this model is 96%.[9]

In this model Ruhi Jaiswal used Depth- wise seperable convolutions composing of two different layers in which first layer is depth- wise convolutions which seperates the spatial cross- correlations from the channel crosscorrelations. And the second layer is point- wise convolutions which generates a prediction by using a soft-max activation function and global average pooling. They worked on FER2013 datasets which is a cleaned dataset with 28,709 sets. Disgust expression is hardly understood by the trained model and, currently it has an accuracy of 66%.[10]

Arvind R and the team members proposed a model Facial Emotion Recognition Using CNN, in this model image is processing using Gabor Filter, Model training using CNN, saved model using json and use it for testing, virtualisation, with Metplotlib, Gabor filters and CNN. Accuracy of this model is best in LDA (96.25%) and lowest in CNN (93%).[11]

**Result and Discussion:**

Facial expressions are a rich source of information that can be used to detect emotions. Commonly recognized emotions include happiness, sadness, anger, surprise, fear, and disgust. Emotion detection models based on facial expressions use various techniques to analyze the facial features, such as OpenCV, Deep Face, Haar Cascade, and others. The performance of an emotion detection model depends on several factors, such as the quality of the training data, the choice of the machine learning algorithm, the feature extraction techniques, and the choice of the evaluation metrics. It is essential to have a diverse and representative dataset of labeled images that covers different facial expressions, genders, ages, and ethnicities.

One of the popular evaluation metrics used for emotion detection models is accuracy, which measures the percentage of correctly classified emotions. Other metrics include precision, recall, F1-score, and confusion matrix. A high accuracy score does not necessarily mean that the model is performing well in all emotion categories, and other metrics can provide more insights into the model's performance. The choice of machine learning algorithm and feature extraction techniques can also affect the performance of the emotion detection model. Deep learning-based algorithms, such as Convolutional Neural Networks (CNNs), have shown promising results in emotion detection due to their ability to extract high-level features from images. However, traditional machine learning algorithms, such as Support Vector Machines (SVMs), Random Forest, and Naive Bayes, can also be used for emotion detection.

**Conclusion:**

In conclusion, emotion detection based on facial expressions is a complex and challenging task that requires advanced computer vision and machine learning techniques. Despite the promising results achieved by various algorithms and techniques, there is still room for improvement in the accuracy and reliability of emotion detection models. Further research is needed to address the limitations of current models, such as their sensitivity to lighting conditions, head orientation, and facial expressions. The quality and diversity of the training data should also be improved to ensure that the model can recognize emotions across different cultures, genders, and ages. Despite these challenges, emotion detection based on facial expressions has the potential to revolutionize various industries, such as advertising, entertainment, and security. It can also have significant implications for mental health and wellbeing by providing new tools for emotion regulation and therapy. Overall, the field of emotion detection based on facial expressions is still in its early stages, and further research is needed to develop more accurate and reliable models. Nevertheless, the potential applications of this technology are vast and exciting, and it is likely to have a significant impact on our lives in the future.

**Reference:**

**[1]** K. F. Azizan Illiana, "Facial Emotion Recognition: A Brief Review," in International Conference on Sustainable Engineering, Technology and

Management 2018 (ICSETM-2018), 2020.

**[2]** A. C. Le Ngo, Y. H. Oh, R. C. W. Phan and J. See, "Eulerian emotion magnification for subtle expression recognition," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016

1. K. M. Rajesh and M. Naveenkumar, "A robust method for face recognition and face emotion detection system using support vector machines," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, 2016
2. H. Ebine, Y. Shiga, M. Ikeda and O. Nakamura, "The recognition of facial expressions with automatic detection of the reference face," 2000 Canadian Conference on Electrical and Computer Engineering. Conference Proceedings. Navigating to a New Era (Cat. No.00TH8492), Halifax, NS, 2000, pp. 1091-1099 vol.2.
3. W. Swinkels, L. Claesen, F. Xiao and H. Shen, "SVM point-based real-time emotion detection," 2017 IEEE Conference on Dependable and Secure Computing, Taipei, 2017.
4. Saransh and Dr. Muthamil Selvan T, “Human Facial Expression Recognition Using TensorFlow And OpenCV,” School of Information Technology and Engineering, VIT University, Vellore August 2020
5. Dr. D. Dhanya et.al, “Emotion analysis using CNN”, ICCIDT (2022), ISSN: 2278-0181
6. Akash Saravanan et.al, “Facial Emotion Recognition using CNN, 12 oct 2019, arXiv:1910.05602vl
7. NinadMehendale, “Facial emotion recognition using convolutional neural networks (FERC)”, SN Applied Sciences (2020), <https://doi.org/10.1007/s42452-020-2234-1>.
8. Ruhi Jaiswal, “Facial expression classification using CNN and its applications”, 15th (IEEE) 2020,
9. Arvind R et.al, “Facial Emotion

Recognition Using CNN, IJRASET

(2022),http://doi.org/10.22214/ijraset.202

2.41536.