NavaJyoti, International Journal of Multi-DisciplinaryResearch
Volume 7, Issue 2, August 2023

A COMPARATIVE ANALYSIS ON FACE EMOTION RECOGNITION TECHNIQUES

[1]Dr. Senthil Vadivu M, [2] Vaishnavi P

[1]Assistant Professor, Jyoti Nivas College Autonomous, Bangalore, India

[2] Student, Jyoti Nivas College Autonomous, Bangalore, India

Abstract: The field of emotion recognition aims to identify and analyze human emotions by utilizing different modalities, like facial expressions, speech intonations, and physiological indicators. The study uses the FER2013 dataset, which consists of grayscale images of faces labelled with seven basic emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral. The accuracy of each algorithm is evaluated, including the K-nearest neighbors' algorithm, Random Forest algorithm, and Convolutional neural network model, for classifying the seven different emotions. The project offers valuable insights into the strengths and limitations of each algorithm, helping to select the most appropriate one for emotion recognition applications.

Keywords: Facial Expressions, FER 2013

1.Introduction

The field of emotion recognition has garnered considerable interest due to its wide-ranging potential applications across domains such as psychology, human-computer interaction, security, marketing, and automotive industries. This research is important as it provides insights into the effectiveness of different algorithms for facial emotion recognition, informing the development of improved systems for various applications. By comparing the performance of CNN VGG-16, KNN, and Random Forest, it helps to identify the most appropriate algorithm for emotion recognition applications. This knowledge will play a vital role in improving the creation of more precise and efficient facial emotion recognition systems, with implications for psychology, human-computer interaction, security, marketing, and the broader field of machine learning. Through the analysis and comprehension of the comparative effectiveness between deep learning and traditional machine learning algorithms image classification tasks, this research aims to progress the emotion recognition field and enhance these systems in real-world scenarios.

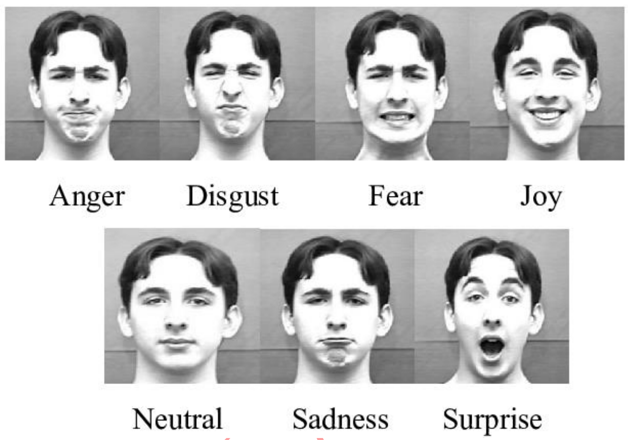


Fig 1: Different types of Human Facial Expressions

2. Literature survey

Abu et.al [1] conducted a study in 2021 focused on developing a real-time [13] emotion recognition application. They employed a Convolutional Neural Network (CNN) to analyze facial expressions. Their objective was to accurately recognize four specific emotions (happiness, sadness, surprise, and disgust) using a custom dataset and the MobileNet algorithm. The results showed an average accuracy of 92.50% and achieved a sensitivity of 85.00% and specificity of 95.00%.

Liliana et.al.[2] in 2018 focused on developing a facial expression recognition system which works on Convolutional Neural Network (CNN) architecture. The objective was to accurately classify eight different classes of facial expressions. The system was trained on the CK+ database using varying sizes of training data, and the results showed a decrease in mean square error as the number of training data increased [12]. The performance of the system reached an accuracy rate of 92.81%. Their study contributes to the advancement of emotion recognition technology and emphasizes the importance of training data size in improving system performance.

Kampel et.al.[3] conducted a study in 2019 that offers a comprehensive examination of facial emotion recognition utilizing Convolutional Neural Networks (CNNs). The paper provides a comprehensive overview of cutting-edge techniques and assesses their precision using well-known datasets, including FER-2013 and CK+. The highest accuracy achieved on the FER-2013 dataset was 75.2% using the Inception V3 architecture, while on the CK+ dataset, a combination of different architectures achieved an accuracy of 98.4%. Their results demonstrate the superiority of deep learning methods over traditional machine learning approaches in facial emotion recognition tasks.

Pandey et.al. [4] conducted a comprehensive review of facial emotion recognition techniques using deep learning, with a focus on Convolutional Neural Networks (CNNs). They reported notable achievements, such as the highest accuracy of 71.5% on FER-2013 using a deep CNN-based approach and 97.06% on CK+ using a CNN incorporating a spatial transformer network.

Ivan et.al. [5] proposed a hybrid approach combining Convolutional Neural Network (CNN) and local binary patterns (LBP) for facial expression recognition with masks during covid. The method achieves accuracies of 73.4% on the FER-2013 dataset and 95.4% on the CK+ dataset. They highlight the effectiveness of the hybrid CNN-LBP approach in accurately recognizing facial expressions.

Maryam et.al.[6] introduced a deep learning framework that combines Gabor filters with CNNs for efficient facial emotion recognition. Experimental results demonstrate high accuracy and reduced training time compared to traditional CNN methods. This approach shows promise for real-time applications requiring fast and accurate emotion recognition.

Venkata et.al.[7] focused on facial emotion recognition using NLPCA (Nonlinear Principal Component Analysis) and SVM (Support Vector Machine). The study utilizes Haar wavelet and Gabor wavelet for feature extraction, with Gabor wavelet for local features and Haar wavelet for global features. NLPCA is employed for dimensionality reduction, and weighted fusion combines the global and local features. The proposed method is evaluated on the Extended Cohn-Kanade database and demonstrates superior results compared to existing methods tested on the CK+ database. The approach achieves accurate recognition of six emotions: joy, surprise, fear, disgust, anger, and sadness.

Wafa et.al.[8] provides an overview of recent Facial Emotion Recognition (FER) research. It discusses different CNN and CNN-LSTM architectures proposed by researchers and emphasizes the need for larger databases and powerful deep learning models to recognize a wider range of emotions. The paper also highlights the shift towards multimodal analysis, combining audio, visual, and physiological data for more accurate emotion detection. Overall, the paper showcases advancements in FER and the importance of ongoing research in this field.

Kwon et al. [9] focuses on emotion recognition using speech signals. They incorporate velocity/acceleration of pitch and MFCCs to form feature streams. The study holds Quadratic Discriminant Analysis (QDA) and Support Vector Machine (SVM) as discriminative classifiers. Experimental results highlight the importance of pitch and energy in emotion recognition. The authors compare the performance of different classifiers on two databases and achieve high accuracy using Gaussian SVM. They emphasize the significance of pitch and energy in emotion recognition and suggest further improvements in feature representation and modeling of feature dynamics.

Hongli Zhang et.al.[10] proposed a facial expression recognition which worked on CNN model. The method effectively extracts facial features and achieves improved recognition rates in complex background scenarios. The authors emphasize the significance of weight initialization during training and suggest future work on exploring robust models for real-life conditions and recognizing dynamic expressions using 3D convolution technology.

3. Methodology

3.1 K-Nearest Neighbors (KNN) algorithm:

K-Nearest Neighbors (KNN) represents a simple yet effective classification algorithm. It operates by identifying the K nearest neighbors in the training dataset for a new instance and subsequently assigning the most frequent class among these neighbors as the predicted classification. Determining the ideal value for the hyperparameter K, which represents the number of neighbors, is required for achieving optimal performance.

3.2 Random Forest algorithm:

The Random Forest technique is an ensemble learning approach that enhances classification accuracy by amalgamating several decision trees. Within Random Forest, multiple decision trees are constructed using randomized subsets of the training data, and the ultimate predicted class is determined by the mode of the individual tree predictions. This approach aids in reducing overfitting issues while boosting the overall accuracy.

3.3 Convolutional Neural Networks (CNN):

Convolutional Neural Networks (CNNs) is a deep learning model which is extensively used for image classification duties. These networks utilize convolutional layers to extract pertinent features from input images, secondly followed by pooling layers to condense the dimensionality of the resultant feature maps. The outcome of the pooling layers is flattened and conveyed to fully connected layers that execute the ultimate classification process. The inherent strength of CNNs in image classification stems from their capacity to autonomously learn meaningful features from the input dataset.

3.3.1 VGG16 algorithm:

VGG16 is a specific CNN architecture that was designed to use in image classification. It consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. VGG16 has a simple and uniform architecture that makes it easy to understand and interpret. VGG16 has attained exceptional performance on various image classification benchmarks, notably including the ImageNet Large Scale Visual Recognition Challenge (ILSVRC).

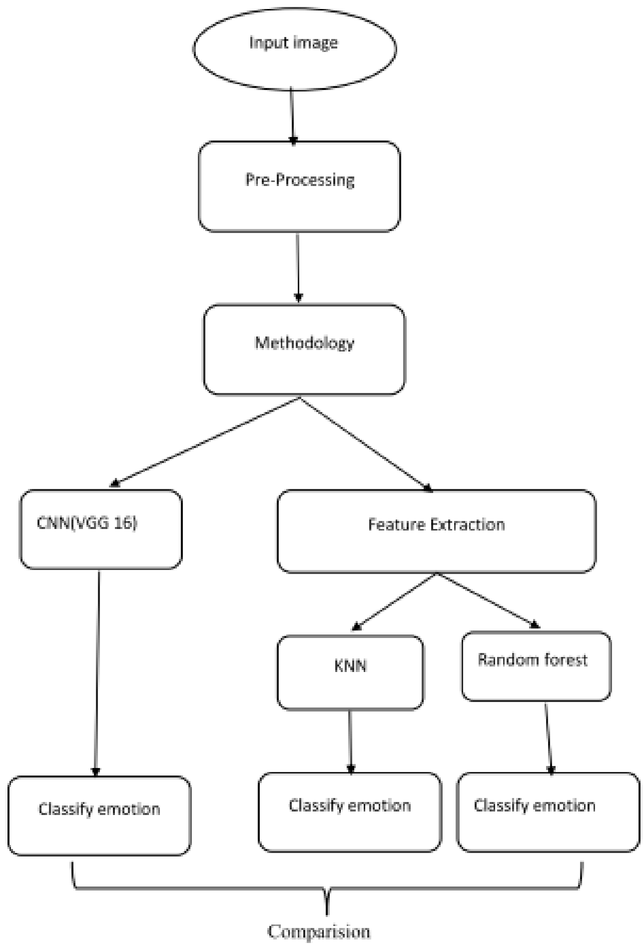


Fig 2: Methodology for comparing different techniques for face emotion recognition

3.4 Steps:

1. Input image: Load the image.
2. Pre-processing: Perform pre-processing steps such as resizing, normalization, and grayscale conversion to prepare the image for feature extraction.

3. Methodology: Select the machine learning methodology to be used for facial emotion recognition, such as KNN, CNN VGG16, or Random Forest.
4. Feature extraction: Extract relevant features from the pre-processed image using the chosen feature extraction method.
5. Classify emotion: Use the trained algorithm to classify the emotion present in the pre-processed image.
6. Compare: Compare the efficacy of different algorithms by considering accuracy and other metrics. Analyze the results and draw conclusions about the effectiveness of each algorithm and feature extraction method for facial emotion recognition.

3.5 Dataset:

FER2013 [11] is a widely used dataset for facial emotion recognition, consisting of 48x48-pixel grayscale images of faces with seven emotion labels. It contains over 35,000 images, divided into three subsets: training (28,709 images), public test (3,589 images), and private test (3,589 images) [11]. The dataset is given in CSV format, in which each row represents an image and its corresponding emotion label. The pixel values range from 0 to 255, representing the intensity of the pixel. The dataset has been used in various studies and benchmarks for facial emotion recognition, but it has some limitations, such as the imbalance of emotion classes, low resolution of the images, and noise in the emotion labels.

emotion	pixels	Usage
0	70 80 82 72 58 58 60 63 54 58 60 48	Training
0	151 150 147 155 148 133 111 140 1	Training
2	231 212 156 164 174 138 161 173 1	Training
4	24 32 36 30 32 23 19 20 30 41 21 22	Training
6	4 0 0 0 0 0 0 0 0 0 0 0 0 3 15 23 28 48	Training
2	55 55 55 55 55 54 60 68 54 85 151 1	Training
4	20 17 19 21 25 38 42 42 46 54 56 62	Training
3	77 78 79 79 78 75 60 55 47 48 58 73	Training
3	85 84 90 121 101 102 133 153 153 1	Training
2	255 254 255 254 254 179 122 107 9	Training
0	30 24 21 23 25 25 49 67 84 103 120	Training
6	39 75 78 58 58 45 49 48 103 156 81	Training
6	219 213 206 202 209 217 216 215 2	Training

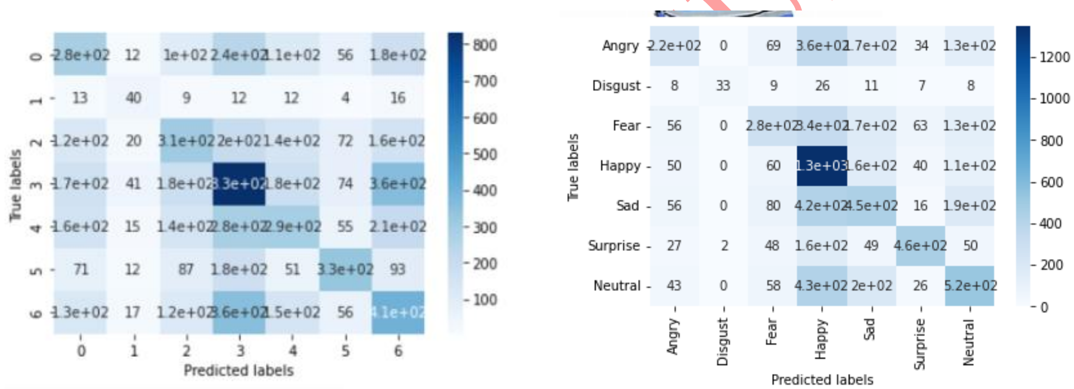
Fig 3: A View to the dataset

4. Result Analysis

The research project involves a comparative analysis of three machine learning algorithms, namely KNN, Random Forest, and CNN VGG16, for face emotion recognition. The objective is to address the need for accurate and reliable emotion recognition in applications such as security systems, human-computer interaction, and healthcare.

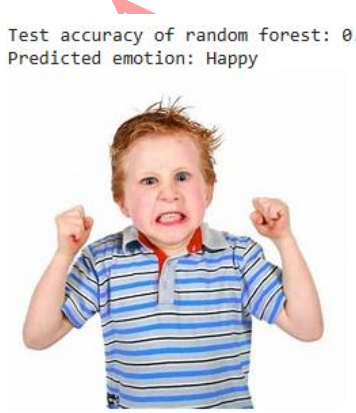
To begin, the dataset is pre-processed by resizing the images, converting them to grayscale, and normalizing pixel values. Feature extraction techniques, including Histogram of Oriented Gradients (HOG) [15], are applied to the pre-processed dataset.

Subsequently, the machine learning algorithms are trained on the pre-processed dataset. KNN is a distance-based algorithm that assigns labels to test data based on the nearest neighbors in the training set. On the other hand, Random Forest is an ensemble learning algorithm that combines decision trees to enhance accuracy and mitigate overfitting. CNN VGG16 is a deep learning algorithm that utilizes convolutional layers to extract features and learn patterns crucial for emotion recognition.



(a) KNN algorithm (b) Random Forest Algorithm

Fig 4: Confusion Matrix



(a)Random Forest (b) KNN (c) CNN(VGG16)

Fig 5: Results of the algorithm compared with same pictures

5. Performance Evaluation

The performance of the trained algorithms is evaluated on a separate test set of images using metrics such as accuracy. The results indicate that CNN VGG16 achieves the highest accuracy of 60%, followed by Random Forest with an accuracy of 46%, and KNN with an accuracy of 34.66%.

The results highlight the superior performance of deep learning algorithms, such as CNNs, compared to traditional approaches like KNN and Random Forest. The choice of algorithm should consider factors such as dataset characteristics, available computational resources, and desired accuracy levels.

Table 1: Results of Comparing the model

Sl.no	Algorithm	Accuracy
1	CNN (VGG 16)	0.60072
2	Random Forest Classifier	0.466
3	K Nearest Neighbors Classifier	0.3466

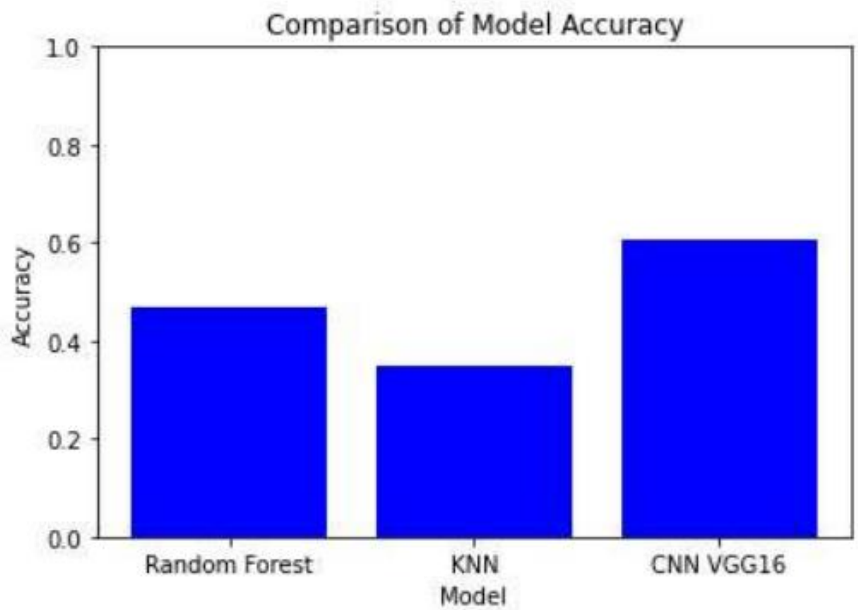


Fig 6: Comparison of Model Accuracy

6. Conclusion

The research aimed to compare three machine learning algorithms (KNN, Random Forest, CNN VGG16) for face emotion recognition. The CNN VGG16 got the highest accuracy of 60%, followed by Random Forest (46%) and KNN (34.66%). The CNN VGG16's deep architecture and feature extraction capabilities contributed to its superior performance. It performed well in predicting happiness and neutral emotions but struggled with sadness and disgust. The findings highlight the significance of algorithm selection in achieving accuracy. This research has implications for marketing, psychology, and healthcare by providing insights into consumer behavior, mental health, and patient care.

REFERENCES

[1] Nur Alia Syahirah Badrulhisham and Nur Nabilah Abu Mangshor 2021 J. Phys.: Conf. Ser. **1962** 012040**DOI** 10.1088/1742-6596/1962/1/012040

[2] D Y Liliana 2019 J. Phys.: Conf. Ser. **1193** 012004**DOI** 10.1088/1742-6596/1193/1/012004

[3] Pramerdorfer, Christopher and M. Kampel. "Facial Expression Recognition using Convolutional Neural Networks: State of the Art." *ArXiv* abs/1612.02903 (2016): n. pag.

[4] S. Pandey, S. Handoo and Yogesh, "Facial Emotion Recognition using Deep Learning," 2022 International Mobile and Embedded Technology Conference (MECON), *Noida, India*, *2022*, pp. 248-252, doi: 10.1109/MECON53876.2022.9752189.

[5] Gogić, I., Manhart, M., Pandžić, I.S. et al. Fast facial expression recognition using local binary features and shallow neural networks. Vis Comput **36**, 97–112 (2020).

[6]344190368_Fast_Facial_emotion_recognition_Using_Convolutional_Neural_Networks_and_Gabor_Filters

[7] Reddy C.V.R., Reddy U.S., Ki shore K.V.K. (2019). Facial emotion recognition using NLPCA and SVM, Traitement du Signal, Vol. 36, No. 1, pp. 13-22.

[8] Wafa Mellouk, Wahida Handouzi, Facial emotion recognition using deep learning: review and insights,Volume 175,

[9] Kwon, O. W. et al. "Emotion recognition by speech signals." 8th European Conference on Speech Communication and Technology (Eurospeech 2003) *(2003):* n. pag.

[10] H. Zhang, A. Jolfaei and M. Alazab, "A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing," in IEEE Access, vol. 7, pp. 159081-159089, 2019, doi: 10.1109/ACCESS.2019.2949741.

[11] fer2013 | Kaggle

[12] D Y Liliana. "Emotion recognition from facial expression using deep convolutional neural network", Journal of Physics: Conference Series, 2019

[13] Andrea F. Abate, Aniello Castiglione, Michele Nappi, Ignazio Passero. "DELEX", Proceedings of the International Conference on Advanced Visual Interfaces, 2020

[14] Proceedings of the International Conference on Paradigms of Computing, Communication and Data Sciences",Springer Science and Business Media LLC, 2023