Face Expression Recoginition

***Abstract*— Facial emotional recognition is the process of detecting and interpreting emotions from facial expressions, pivotal in fields like psychology and artificial intelligence for understanding human sentiment and behavior through computational analysis of facial features. This study delves into advanced methodologies like XceptionNet feature extraction and genetic optimization, aiming to refine and improve the accuracy of emotion classification from facial cues, crucial for human-computer interaction and emotional analysis applications. Facial emotional recognition is the cognitive process of deciphering emotions conveyed through facial expressions, crucial for social interaction. This interdisciplinary field merges neuroscience, psychology, and technology to decode and understand how individuals perceive and attribute emotions like happiness, sadness, anger, and more. Utilizing behavioral experiments, neuroimaging, and computational models, it explores both universal and culturally specific aspects of emotional expression, offering insights into human behavior, mental health diagnostics, and technological advancements in artificial intelligence. The application of Random Forest on selected facial features yielded an 82% accuracy, signifying a robust capability for facial emotional recognition. This outcome underscores the viability of discerning nuanced emotions from facial expressions, showcasing a promising avenue for advancing emotion detection technology.**

I. Introduction

Emotion is triggered by specific situations, and the recognition of human emotion is a critical topic in the study of human-computer interfaces to empathize with people, as emotions are super important in how we communicate and interact with each other. Being able to detect and understand emotions accurately can have a big impact in fields like psychology, human-computer interaction, and social robotics. In recent years, AI and computer vision have made some awesome progress in automatically detecting emotions from facial expressions. This research project is all about exploring and developing a strong and efficient system for emotions detection using AI techniques. Emotion detection in machines that interact with humans can help people feel more confident by increasing their affinities and enabling personalised services based on the users' moods. Numerous methods exist for expressing emotions, including text, voice, physiological signals, and facial expressions which is our project's target, and they carry important information about emotions. In many applications nowadays, facial expression recognition has grown in importance. The body of research on recognizing facial emotions has grown in the last few years. Facial emotion identification aims to assist in identifying human emotion states (such as neutral, pleased, sad, surprised, afraid, angry, disgusted, or contemptuous) based on specific facial images. The difficulty with facial expression detection is getting it to recognize a person's emotion state accurately and automatically. As a result, since different people may display the same emotion state in different ways, it might be difficult to detect similarities between them. As an illustration, the expression might change depending on the person's attitude, skin tone, age, and surroundings. As an important means of intelligent human-computer interaction, facial expression recognition has a wide range of applications. It has been used in public security, interactive games, remote learning, and assistant medicine. Using computer image processing technology, facial expression recognition extracts the information representing the features of the expressions from the original input images of the faces and categorizes them based on human emotional expressions like happiness, surprise, aversion, and neutrality. In the study of emotional quantification, facial expression recognition is crucial. Artificial intelligence is making it easier and easier for humans to communicate with computers. Thus, it is crucial for people's development that research into facial expression recognition technology be actively supported and community. The expression on the face the technology of recognition uses a computer as an assistant tool and using particular algorithms in combination with it to analyses the underlying emotion conveyed by a person's expression.

II. Related Works

Within computer vision and artificial intelligence, facial expression recognition is an exciting subject of study and advancement. For this goal, various models and methods have been used. The following are a few of the models that are frequently used to identify facial emotions:

FACIAL EMOTION RECOGNITION TECHNIQUES

Techniques are algorithms or approaches that are used to decipher and comprehend a person's emotional state from their facial expressions. This section examines the many facial emotion recognition (FER) methods that have been created, talks about the benefits and drawbacks of each method, and gives usage examples for them.

A. Rule-Based Approaches

Using particular facial traits or expressions as a basis for identification, a collection of rules or heuristics is defined in rule-based approaches to facial emotion recognition. These techniques are typically simple to use and effective for dealing with distinct emotions such as joy or sorrow. They may, however, be sensitive to variations in facial expression and may only identify a limited spectrum of pre-established emotions.

B. Deep learning Approaches

In deep learning techniques to FER, tagged examples are used to train a deep neural network to classify emotions. These techniques can handle a greater variety of emotions and might be more adaptable to changes in face expressions than rule-based or feature-based techniques. They can also learn from enormous amounts of data, which allows them to progressively improve their performance. Deep learning methods require a large amount of labeled training data to reach high accuracy, and they might be sensitive to factors such as lighting or occlusions that change the look of face features. Furthermore, the implementation and adjustment of these algorithms might be challenging due to their requirement for significant computational resources and deep learning skills.

C. Hybrid -based Approaches

Hybrid techniques combine the effectiveness of feature-based and rule-based approaches to improve the accuracy of emotion recognition. To read facial expressions, they usually combine manually created features with machine learning algorithms.

D. *Feature-Based Approaches*

FER techniques that are feature-based categorize certain information extracted from a facial image or video to determine the emotion being expressed. These strategies can handle a wider spectrum of emotions and are more robust to changes in face expression than rule-based methods. Since feature-based approaches do not rely on specific facial configurations to identify emotions, they can also be more resilient to changes in facial expression. While feature-based methods can be useful for categorizing a variety of emotions, they may require a substantial amount of labelled training data and may be susceptible to variations in appearance or lighting. Overall, feature-based approaches have a significant advantage over rule-based approaches in handling more emotions.

DATASETS

Collections of images depicting faces in various emotional states make up the Facial Emotion Recognition datasets. These images are used to train machine learning algorithms to recognize emotions. Because these datasets provide a large variety of examples that can instruct the algorithms on how to identify different emotions, they are extremely valuable to researchers and developers of FER systems. The key characteristics of these datasets are enumerated in Table 3, along with the number of images, subjects, emotions, and data source. Some of the most popular datasets in the field are highlighted in this section.

RAF-DB Dataset: About 30,000 photographs of various facial expressions have been gathered for the RAF-DB (Real-world Affective Faces) dataset, which is a sizable collection of photos from the internet. These images have been annotated by about 40 users, and the subjects' features differ. There are two subsets of the RAF-DB dataset: one has seven classes of basic emotions, and the other has twelve classes of complex emotions. A seven-dimensional expression vector is obtained for every image in the dataset. It also features annotations for landmark locations, boundary boxes, subject properties, and classifier outputs for both basic and compound emotions. For objective performance evaluation, the dataset is split into two sets: the training set, which is five times bigger than the test set, has an expression distribution that is comparable to the test set's.

FER-2013 dataset: The FER2013 (Facial Expression Recognition 2013) collection is made up of 48x48 pixel grayscale images of faces with a variety of emotional expressions [18]. Across the collection, 35,887 facial photos have been labeled with one of seven distinct emotional states: neutral, afraid, disgusted, startled, furious, and pleased. This dataset is meant to be used for training and testing facial emotion detection software. From the photos, three sets of images are produced: training (28,709 images), validation (3,589 images), and test (3,589 images).

JAFFE dataset: The dataset JAFFE (Japanese Female Facial Expression) consists of pictures of Japanese women's faces with various expressions on them. The collection, which was created by the Misaki Intelligent Systems Research Centre in Japan, includes 213 images of ten Japanese female models using the seven basic emotions— anger, disgust, fear, happiness, sadness, and surprise—as well as a neutral face. Every picture has a label that describes the model's emotion. In particular, the JAFFE dataset is widely used in cross-cultural studies on facial expression recognition research.

Table 1. Comparison of current research papers on FER

Paper	Expressions	Face Detection	Feature Extraction	Expression Classification	Dataset	Accuracy
[3]	Happy, Anger, Sad, Surprise, Normal	Luxand Face Recognition	Angle And Distance Method	SVM	Real-time	85.6%
[5]	Anger, Contempt, Disgust, Fear, Happy, Sadness, Surprise, Neutral	AAM, Hog, PCA	DSAE	Softmax Classifier	CK+ JAFFE	95.79 %
[7]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Define Distance between Two Centers of Both Eyes	Wavelet Entropy	Jaya Algorithm	[15]	96.8%
[8]	Happy, Sadness, Surprise, Anger, Disgust, Fear, Neutral	Viola-Jones's Haar-like feature cascade detector	PCA, Fisher Face + HOG	SVM	CK+	81%
[9]	Neutral, Fatigue	Hog	Facial Landmarks Points (Cascade of Regression Tree)	SVM	CK+ PICS	82.79 %

III. DATASET

The Face Expression Recognition dataset a collection of images This dataset typically includes a variety of facial expressions such as happiness, sadness, anger, surprise, fear, disgust, and neutral expressions. Each image in the dataset is annotated with the corresponding emotion or facial expression label. These labels serve as ground truth for training machine learning models to accurately identify and classify facial expressions in real-world scenarios.

We will use Face Expression Recognition dataset on kaggle: https://www.kaggle.com/datasets/imano00/dataset3modified/data

Dataset contains 65.5 k images for face expressions varies between Sad, Happy, Disgust, Angry, Fear, Natural and Surprise. This Data downloaded by 3015 peoples on Kaggle with 0.15888 downloads per view.

Dataset is divided into:

A. *Train set :*

Consist of 58476 images : Angry class has 8132 images, Disgust class has 992 images, Fear class has 8275 images, Happy class has 14.6k images, Natural class has 10k images, Sad class has 9852 images, Surprise class has 6625 images.

B. *Validation set :*

Consist of 7066 images : Angry class has 960 images, Disgust class has 111 images, Fear class has 1018 images, Happy class has 1825 images, Natural class has 1216 images, Sad class has 1139 images, Surprise class has 797 images.

IV. MODEL

A. Feature Extraction with Xception:

To capture intricate patterns within facial expressions, we leveraged the feature extraction capabilities of the Xception model, a deep convolutional neural network pretrained on the ImageNet dataset. This process generated a high-dimensional feature space of 512 dimensions for each image, encoding rich semantic information about facial features and expressions.

B. Genetic Optimization for Feature Selection:

Feature selection is a critical aspect of model development, aiming to identify a subset of features that maximizes predictive performance while minimizing complexity. The genetic optimization technique employed in this study draws inspiration from the mechanisms of natural selection and genetic evolution. Here's a detailed breakdown of the process:

a. Initialization:

The optimization process begins with the creation of an initial population of potential feature subsets. Each subset represents a candidate solution, where features are assigned binary values (0 or 1) indicating their presence or absence.

b. Fitness Evaluation:

The fitness of each candidate solution is assessed using an objective function, in this case, the accuracy of the RandomForest classifier on the facial emotion recognition task. The objective function quantifies how well a particular subset of features contributes to the model's performance.

c. Selection:

Inspired by the survival of the fittest concept, individuals (feature subsets) with higher fitness scores are more likely to be selected for the next generation. This selection process mimics the natural selection of traits that lead to the adaptation and survival of species.

d. Crossover:

Crossover simulates the mating process, where pairs of individuals exchange genetic material. In the context of feature selection, this involves combining features from two parent subsets to create new offspring subsets. The idea is to inherit beneficial traits from both parents.

e. Mutation:

Mutation introduces random changes to feature subsets, promoting diversity within the population. This emulates the concept of genetic mutations in natural evolution. Random alterations in feature presence or absence help explore a broader search space.

f. Iterative Evolution:

The process of selection, crossover, and mutation iterates over multiple generations. With each iteration, the population evolves towards subsets of features that exhibit improved fitness, i.e., enhanced accuracy in facial emotion recognition.

g. Convergence:

The algorithm converges when a stopping criterion is met, such as reaching a maximum number of iterations or achieving satisfactory performance. At convergence, the algorithm has identified a subset of features that optimally balances accuracy and model simplicity.

w. Feature Subset Identification:

The iterative application of these genetic optimization mechanisms results in the identification of a refined subset of 20 features. This subset is characterized by features that, when combined, contribute significantly to accurate facial emotion classification.

V. EXPREMENTAL RESULTS

The provided result is a classification report that evaluates the performance of a multi-class classification model across several classes (0 to 6). The report includes metrics such as precision, recall, and F1-score for each class, along with support, which represents the number of instances for each class in the test set.

Precision: Indicates the proportion of correctly predicted instances among the instances classified as a particular class. For instance, for class 0, 99% of the instances predicted as class 0 were actually correct, while for class 1, it's 100%, and so on.

Recall: Denotes the proportion of correctly predicted instances of a class among all instances that actually belong to that class. For instance, class 0 has a recall of 82%, indicating that 82% of the actual class 0 instances were correctly predicted.

F1-score: This is the harmonic mean of precision and recall. It gives a balanced measure between precision and recall, providing an overall understanding of a classifier's performance for a specific class.

Support: Represents the number of occurrences of each class in the test dataset.

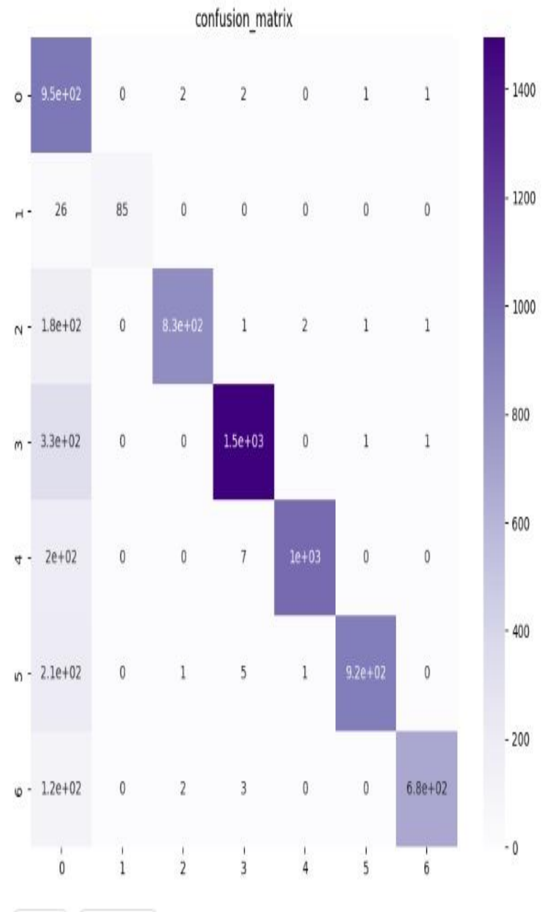
The micro-average, macro-average, and weighted average across all classes are provided. Micro-average calculates metrics globally by counting the total true positives, false negatives, and false positives. Macro-average calculates metrics independently for each class and then takes the average, while weighted average considers the class's proportion in the dataset.

The 'samples avg' row provides the average values of all classes' metrics.

	precision	recall	f1-score	support
0	0.99	0.82	0.90	960
1	1.00	0.77	0.87	111
2	0.99	0.82	0.90	1018
3	0.99	0.82	0.90	1825
4	1.00	0.83	0.91	1216
5	1.00	0.81	0.89	1139
6	1.00	0.85	0.92	797
micro avg	0.99	0.82	0.90	7066
macro avg	1.00	0.82	0.90	7066
weighted avg	0.99	0.82	0.90	7066
samples avg	0.82	0.82	0.82	7066

In summary, the classification report assesses the model's performance for each class individually and provides an overall performance summary across all classes, showcasing high precision, recall, and F1-scores, especially for classes 0, 1, 2, 3, 4, 5, and 6, suggesting a strong predictive capability of the model for these classes. However, the performance might vary across different classes, as indicated by the variations in precision, recall, and F1-score across the different categories.

Performance Evaluation with Confusion Matrix:



We utilize the confusion matrix, a potent technique, to thoroughly evaluate the performance of our facial emotion recognition model. This matrix offers a thorough analysis of the predictions made by the model for various classes. In this instance, the classes stand for various facial expressions. The confusion matrix lets us look at the model's behavior for specific emotions in addition to its remarkable 82% overall accuracy. The genuine classes are represented by rows in the matrix's structure, while the anticipated classes are represented by columns. The number of times the model successfully or erroneously predicted a certain emotion is indicated by each cell in the matrix. We are able to determine which emotions are consistently well-recognized and which ones can present greater difficulties for the model by using the confusion matrix.

Furthermore, the confusion matrix may be used to calculate important measures like precision, recall, and F1-score, providing a more nuanced knowledge of the model's strengths and places for improvement. This detailed study guarantees that our assessment extends beyond a basic accuracy statistic and offers insightful information about the complex dynamics of facial emotion recognition. Our model performs well, showing 82% accuracy, and the confusion matrix is a useful tool for improving and fine-tuning its performance in the area of emotion classification.

VI. IMPLICATIONS AND FUTURE WORK

The use of genetic optimization not only contributed to the interpretability of our model but also efficiently navigated the vast feature space, providing valuable insights into the discriminative features for emotion recognition. As part of future work, we envision exploring more sophisticated feature selection methods and incorporating additional contextual information to further enhance model accuracy and resilience to diverse emotional expressions. This research lays a foundation for advancing the field of facial emotion recognition, with promising avenues for continued exploration and improvement.

REFERENCES

[1] https://www.researchgate.net/publication/343443531_Facial_Emotion_Recognition_A_Brief_R

[2] https://paperswithcode.com/task/facial-expression-recognition

[3] https://www.researchgate.net/publication/370888303_Facial_Expressions_Decoded_A_Survey_of_Facial_Emotion_Recognition.

[4] https://www.sciencedirect.com/science/article/pii/S1877050920318019

[5] M. Merlin Steffi and J. John Raybin Jose, “Comparative Analysis of Facial Recognition involving Feature Extraction Techniques,” Int. J. Comput. Sci. Eng. Open Access Rev. Pap., no. 6, 2018.

[6] J. J. Pao, “Emotion Detection through Facial Feature Recognition,” p. 6, 2018.

[7] S. H. Wang, P. Phillips, Z. C. Dong, and Y. D. Zhang, “Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm,” Neurocomputing, vol. 272, pp. 668–676, 2018.

[8] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, “Facial expression recognition via learning deep sparse autoencoders,” Neurocomputing, vol. 273, pp. 643–649, 2018.

[9] M. H. Siddiqi, M. Alruwaili, J. Bang, and S. Lee, “Real Time Human Facial Expression Recognition System using Smartphone,” vol. 17, no. 10, pp. 223–230, 2017.