Emotion prediction using deep Learning algorithms

## 

## Abstract

Understanding and predicting human emotions using different deep learning algorithms and models has become one of the important domains to research as emotion recognition helps in understanding human behaviors. Recognizing emotions by analyzing the different facial expressions is the main aim of this research. This paper investigates the application of different machine learning algorithms especially deep learning algorithms to understand human emotion based on a photographic image of a human face and predict the emotion where the dataset used to train the model is collected using open web source Kaggle and is an image-based dataset, as the model input will be a different human face. The dataset is annotated with different emotions such as ‘angry’, ‘happy’, ‘sad’, ‘surprise’ etc. The primary focus during the entire research is understanding and exploring different Convolutional Neural Networks (CNN) which will be building blocks of the deep learning model as they help in the extraction of different spatial features from the image in a more efficient manner and different evaluation metrics such as accuracy, precision and f1 score will be used to understanding how the model is behaving.Recurrent neural network models with long short-term memory (LSTM) are useful for sequential data. They use memory cells and gates to regulate data flow between layers, allowing the outcome of one layer to be repeatedly transferred back to the previous layer. Different Optimization techniques such as Adam and RMSprop are used to fine-tune and optimize the deep learning model and these optimization techniques are based on the classification algorithm. This research work ends with an overview of the findings, inferences from the study, and recommendations for future enhancements for face emotion recognition which enhances the field of face expression recognition.

Keywords - Deep learning algorithms, Convolutional Neural Network (CNN), evaluation metrics,Long short-term memory (LSTM), Adam Optimizer, RMSprout

## Introduction

In the growth of artificial intelligence and human-computer interaction, where computers are taught how to behave differently based on human input, and in this particular research the recognition of human emotion is performed with the help of different deep learning algorithms. Understanding human emotion has witnessed a surge area of interest and has a wide range of applications in different fields such as Virtual Reality, Facial Recognition devices, and health care where using the ability of computer knowledge and algorithms emotion is detected, and based on it the computer or device reacts.

The dataset [1] chosen to understand and explore the deep learning model is collected from open source and can be used for research and learning purposes in which there are images of 7 different human emotions which are ‘sad’, ‘happy’, ‘angry’, ‘surprise’, ‘disgust’, ‘fear’ and ‘neutral’ where each image are present in grayscale format and each emotion has around 3000 to 4000 image which is sufficient to understand the pattern for deep learning model but to increase the number of training and testing image different image related processing will be applied to remove extra noise and add some sampling image using data augmentation [2], which will increase the final accuracy of the model. Recurrent neural network models with long short-term memory (LSTM) are useful for sequential data. They use memory cells and gates to regulate data flow between layers, allowing the outcome of one layer to be repeatedly transferred back to the previous layer. This feature also helps when there is insufficient data for model training.

The deep learning model will be fine-tuned and optimized using different Optimizer functions [3] such as Adam and RMSprop and will be based upon a classification algorithm as there are 7 different classes present for each of the images and the probability of each class will be predicted and the class which has the highest probability will be chosen as the final output for the given image.

### 1.1 Background

The past few decades have seen a significant rise in the usage of machinery and automation in society. These days, a wide range of sectors use automated tools and models in order to get a task done or to predict or identify any factor. Machine vision develops when a machine can recognize and understand its environment. People utilize their senses to learn about the world around them. These days, devices may record the condition of their surroundings using a variety of cameras and sensors. Therefore, machine vision can be produced by combining this data with the appropriate algorithms. The application of Deep Learning algorithms has shown to be quite effective in this area in recent years. According to the research findings, emotion recognition is essential for machines to perform their tasks more effectively.

The application of deep learning techniques could enable automated systems to recognize the mood of the other person if they are able to acquire a series of photos of the facial expressions. Within this framework, deep learning holds promise for fostering improved human-machine connection while giving machines a measure of self-analysis of their human counterparts and how to enhance communications and interaction through artificial and natural intelligence.

### 1.2 Problem Statement

The proble statement of this research is how could deep learning methods be enhanced to recognise human emotions and understand those emotions more accurately and robustly in a variety of environmental settings and population groups which has various applications like in the field of health monitoring, security and surveillance, human-computer interactions, gaming, customer feedbacks, etc. Recognizing emotions through facial expressions can be simplified and automated using many deep-learning techniques and algorithms. To start with, a variety of elements, including illumination, body posture, expressions and individual variances in face structure might make it difficult for current facial emotion detection systems to handle changes in emotional features. Reduced precision and accuracy in recognizing emotions can result from these complicated feature extraction and process of classification.

### 1.3 Aim and Objectives

The purpose of this research work is to establish a powerful, precise and intelligent system which can easily and automatically detect facial features and recognize the emotions of humans using deep learning techniques. From a technical point of view, the project's objective is to use labelled photos of different facial expressions which can be used to train a deep neural network model. In this research, the main goal is to create and apply various deep learning techniques and algorithms which can efficiently extract the facial features and characteristics which are necessary to distinguish between a variety of expressions and characterise them into different types of emotions.

The overall purpose of this system is to achieve a cutting-edge machine or software which makes it possible for various applications like healthcare, human and machine interaction, identification of human emotion and further reacting to it accordingly.

### 1.4 Solution Approach

Building an effective system which can easily detect emotions by just looking at the facial expressions and other characteristics can be very challenging as the model needs to be trained well with a data set which contains a variety of expressions. The approach to implementing this system can be divided into two parts. In the first part, a labelled dataset of different facial expressions is used which is picked from an online source Kaggle. This data set is divided into two parts - training and validation. Both are categorized into images of a variety of emotions like angry, disgust, fear, happy, neutral, sad and surprise which are used as training data sets to train the deep learning model. The amount of training and testing images will be increased, and various image-related processing techniques will be used to add some sampling images via data augmentation and eliminate excess noise, increasing the model's ultimate accuracy. Long short-term memory (LSTM) is a type of recurrent neural network model which is helpful in sequential data and use memory cells and gates to control the data flow between different layer where output of one layer can be transfer back to previous layer multiple times which also help in case enough data is not present for model training.

For the second part, the deep learning model will be based on a classification algorithm and will be adjusted and optimized using various optimiser functions like Adam and RMSprop. Each image contains seven different classes, and the class with the highest probability will be selected as the final output for that particular image. The probability of each class will be predicted.

## Literature Review

Deep learning techniques for facial expression analysis have garnered a lot of interest lately because of their broad use in fields including marketing, emotional computing, the interaction between humans and machines, and wellness. This review provides an overview of the main research findings, approaches, challenges, and developments in this area.

Numerous studies have demonstrated the effectiveness of deep learning architectures in identifying emotions on faces. The most popular method for obtaining features from facial images is Convolutional Neural Networks (CNNs).

Regardless of the advancements, there are still a number of obstacles to overcome in utilizing deep learning to recognize facial emotions. One of the biggest challenges is how well the models can tolerate changes in facial emotions brought forth by lighting, head orientation, shadows, and unique facial features.

The lack of diverse and well-annotated datasets, which are necessary for deep learning model evaluation and training, is another major obstacle. Due to the partiality in most available datasets towards particular demographics or emotional states, the models created are not well-suited to be applied to a wide range of groups and real-world situations. Several studies have developed methods to synthesize facial expression data in order to improve upon pre-existing datasets or produce fresh labelled data samples in order to address this particular limitation. Furthermore, there is still much to be concerned about when it comes to the interpretability of deep learning models for facial emotion recognition, especially in situations where understanding model predictions is essential to building transparency and confidence. In an effort to improve interpretability, researchers have taken a number of approaches, including afterwards analysis tools to clarify predictive models and methods for visualization to emphasize important areas within facial images.

As a result of developments in model designs, data augmentation strategies, transfer learning techniques, and understanding advancements, facial emotion recognition through deep learning methodology has advanced remarkably in recent years. However, issues including handling differences in facial expressions, resolving biases in datasets, and improving the interpretability of models remain, highlighting the need for ongoing research targeted at creating more accurate, consistent, and comprehensible face emotion identification systems.

### 2.1 Related Work

Using deep learning algorithms to recognize facial emotions has become a crucial field with numerous possibilities in several fields. A thorough analysis of the body of research in this field offers significant perspectives into current modern techniques, difficulties, and developments. The effectiveness of deep learning techniques, particularly Convolutional Neural Networks (CNNs), in precisely identifying and deciphering emotions from facial expressions has been demonstrated in a number of studies. Taking an example of a research paper done in 2016, it had built a system to recognize emotions by analyzing facial expressions using the Convolutional Neural Network (CNN) technology which produced effective results. In order to achieve outstanding results and lower computing needs, another research work used transfer learning to refine already presented CNN models for facial emotion identification tasks. These results highlight CNNs' supremacy as the principal framework for extracting features in facial emotion recognition projects. However, there are still issues with guaranteeing resilience to changes in facial emotions brought forth by things like illumination, body orientation, closures, and unique facial features. This issue was resolved by another researcher [5] who tackled this issue by introducing a multi-task learning framework, which simultaneously addressed facial expression recognition and facial action unit detection, resulting in enhanced adaptability to diverse facial expressions. Furthermore, the scarcity of diverse and well-annotated datasets presents a significant impediment to the training and evaluation of deep learning models. In order to improve predictive accuracy across a range of demographics and situations in reality, a different researcher [6] tackled this restriction by putting forth methods for producing artificial facial expression data to augment pre-existing datasets or create unique annotated examples. Assuring the comprehensibility of deep learning models for face emotion identification is still a major challenge despite recent developments, especially for applications where understanding model predictions is essential for building clarity and confidence. To improve the understanding of the model, one of the researchers, another researcher [7] experimented with a number of strategies, such as various analysis techniques to clarify the model's forecasts and visual aids to emphasize important areas in facial photos. These initiatives highlight the critical need to create facial expression detection systems that are easier to understand and interact with, in order to increase user adoption and confidence. All things considered, the corpus of research on deep learning algorithms for face emotion recognition demonstrates notable advancements in model structures, data enhancement tactics, transfer learning procedures, and comprehension of the models. To create more precise, dependable, and understandable facial emotion detection systems, research must be conducted indefinitely due to enduring issues such as assuring resistance to alterations in facial expressions, resolving flaws in datasets, and improving model understanding.

### 2.2 Proposed System

Deep learning algorithms implementation of face expression identification is a broad and competitive field of study, indicating the growing interest and advancement in this field. The problem of precisely recognizing and deciphering human emotions from facial expressions in a variety of contexts is addressed by facial emotion recognition utilizing deep learning algorithms. Due to obstacles, illumination, head posture, unique variations in face anatomy, and other factors, existing facial emotion identification systems sometimes have trouble processing changes in facial emotions. Reduced recognition precision and dependability result from these variations' complicated extraction of features and categorization processes. Two components make up the implementation strategy for this system. An internet source called Kaggle is used to select a tagged dataset of various facial expressions for the first section. The training and validation portions of this data set are separated. The deep learning model is trained using a training data set consisting of images representing a range of emotions, including anger, disgust, fear, happiness, neutrality, sadness, and surprise. The number of training and testing photos will be expanded, and different image-related processing methods will be applied to remove unwanted information and add some sample images via data enhancement, ultimately improving the reliability of the model. In the second section, the deep learning model will be tuned and made more efficient using several optimizer functions such as Adam and RMSprop. It will be based on a classification method. Every image comprises seven distinct classes, and the class with the highest likelihood will be chosen as the image's final output.

The study of human emotion has gained a lot of attention lately and has many uses in the healthcare industry, augmented and virtual reality, and facial recognition technology, among other areas. These applications rely on computers and algorithms to detect emotion and act accordingly. Since emotion detection aids in the understanding of how people behave, the understanding and prediction of human emotions using various computational algorithms and models has emerged as one of the key areas of study. Scholars and experts in the business can design facial expression detection systems with high accuracy by utilizing the valuable data provided by these tactics and algorithms.

In conclusion, the complexities of model interpretability and transparency have also been studied by academics, who have tried to develop methods for visualization and analysis that promote user confidence and understanding. To create more robust and flexible facial expression identification systems, current research has also attempted to address real-life problems such as changing lighting, a variety of head positions, and obstructions. All things considered, the literature review highlights the complexity of facial emotion recognition using deep learning algorithms and emphasizes the need for continued study to progress this subject and overcome new challenges. Recent developments in model structures, data enhancement tactics, transfer learning methods, and model comprehension have brought about tremendous progress in the field of facial emotion recognition utilizing deep learning approaches. It is still necessary to conduct research to create facial emotion detection systems that are more precise, dependable, and interpretable because of issues including an understanding of the models, dataset presumptions, and resistance to changes in facial expressions.

## 

## 

## 

## 

## References

[1] www.kaggle.com. (n.d.). Face expression recognition dataset. [online] Available at: <https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset>.

[2] Takimoglu, A. (2021). What is Data Augmentation? Techniques, Benefit and Examples. [online] research.aimultiple.com. Available at: <https://research.aimultiple.com/data-augmentation/>.

[3] Zaheer, R. and Shaziya, H. (2019). A Study of the Optimization Algorithms in Deep Learning. *2019 Third International Conference on Inventive Systems and Control (ICISC)*. doi:<https://doi.org/10.1109/icisc44355.2019.9036442>.

[4] Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M. and Pollak, S.D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, [online] 20(1), pp.1–68. doi:<https://doi.org/10.1177/1529100619832930>.

[5] Küntzler, T., Höfling, T.T.A. and Alpers, G.W. (2021). Automatic Facial Expression Recognition in Standardized and Non-standardized Emotional Expressions. *Frontiers in Psychology*, 12. doi:<https://doi.org/10.3389/fpsyg.2021.627561>.

[6] Mellouk, W. and Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. *Procedia Computer Science*, 175, pp.689–694. doi:<https://doi.org/10.1016/j.procs.2020.07.101>.

[7] Khan, A.R. (2022). Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges. *Information*, 13(6), p.268. doi:<https://doi.org/10.3390/info13060268>.

‌

‌

‌