Facial Emotion Detection In Bot Interviews Using Deep Learning Algorithms

/¨K V S Sravanth kumar[1]   K S S Jagannadh[2]*
Department of Computer science and Technology
Sathyabama Institute of Science and Technology, Chennai
@gmail.com[1]
@gmail.com[2]"/*

Abstract---Human communication relies heavily on facial expressions to convey emotions and ideas. Numerous industries, including safety, bot interviews, healthcare, and financial security, use emotion detection. These days, not every business visits campuses to conduct interviews; rather, a lot of businesses adopt a hybrid method in which internet chatbots perform interviews. During face-to-face interviews, the panelists establish a human connection to help the candidate feel at ease. Online bots, however, are unable to develop human interaction or comprehend a candidate's feelings.Our goal is to identify and comprehend a candidate's emotions throughout the interview. We will send a notification to the candidate to help them feel at ease and allow for human contact between the bot and the individual. We propose a deep learning strategy for facial emotion recognition based on convolutional neural networks (CNNs) and recurrent neural networks (RNNs). On multiple benchmark datasets, the proposed method demonstrated the effectiveness of deep learning-based face emotion recognition techniques. As well as analyzing the advantages and disadvantages of the proposed approach, we also discuss its advantages and disadvantages. A number of contexts, such as human-machine interactions, health, safety, and customer satisfaction, have been addressed by automatic facial expression recognition. The goal of computer prediction researchers in this field is to decipher and extract facial expressions in order to improve computer predictions.

Keywords: Detect emotion, Support vector machine, Convolutional neural network, Feature extraction, OpenCv, Restricted boltzmann machine.

I. INTRODUCTION

The identification of facial emotions has emerged as a noteworthy issue across multiple domains, such as psychology, human-computer interface, marketing, and others. Since emotions are a basic component of human behavior, precisely and instantly identifying them has a multitude of uses, from boosting consumer satisfaction to improving mental health diagnosis and treatment. Recent developments in deep learning have led to a rise in the popularity of deep learning-based methods for detecting emotions on faces. These methods achieve state-of-the-art performance on the FER-2013 dataset and are capable of learning intricate patterns and representations from big datasets.

People use a variety of techniques to communicate their feelings, including body language, texting, gestures, and facial expressions. The other person's reaction to the speaker during a conversation frequently depends on their facial expression. Emotions are important because they are a quiet means of communication in human life.In this work, we use a two-step face recognition approach to identify emotions from facial expressions. When someone approaches the camera, the first stage is to identify their face, and the second is to identify their expression.Emotion detection is useful in a variety of contexts, including commercial promotions where achieving increased profits depends heavily on client pleasure. AI systems are able to detect emotions in real time by taking a picture of a person's face and analyzing it to determine whether or not the individual is pleased. In order to facilitate human contact and identify an individual's emotions during an interview, we are integrating this application into bot interviews. The different categories of human emotions are fear, neutral, disgust, anger, surprise, sad, and joyful. Since human emotions are frequently subtle, it can be difficult to distinguish between them.We employ a dataset with over 60 million human facial expressions to tackle this problem. Anger, for example, is characterized by lowered and furrowed eyebrows, a fixed stare, and a lifted chin. lifted brows, wide eyes, and a lowered jaw are the indications of surprise, whereas lifted corners of the mouth indicate happiness.

.. Furrowed brows, gaping mouth, and wide eyes are the facial expressions associated with fear, while lip corner depressors and furrowed brows are associated with grief. Lastly, the expressions we choose for neutral are those of a typical face. Convolutional neural networks are used to identify and categorize these emotions (CNNs).When engaging in social interactions, emotions are exhibited. In our lives, the ability to identify emotions is crucial. People use a variety of expressions, such as facial expressions, gestures, voice, and text, to convey their feelings. Here, we'll use their facial expressions to gauge their emotional state. Human emotion detection is used in various contexts where further security or personal data is needed.It can be viewed as a follow-up to face detection in which we might need to install an additional security layer that detects emotions in addition to faces. To ensure that the person in front of the camera is more than just a two-dimensional image, this can be helpful.

Business promotions are a significant area in which emotion detection is crucial. The majority of businesses rely heavily on consumer reactions to all of their offerings and products. An artificial intelligence system is capable of making a choice if it can recognize and record emotions in real time from a user's picture or video. The following categories apply to human emotions: fear, neutral, disgust, anger, surprise, sadness, and happiness. These feelings are quite nuanced. Because even little variations produce distinct expressions, it can be difficult to discern even the smallest variations in facial muscular contortions. Additionally, because emotions are so context-dependent, different people—or even the same people—may exhibit the same feeling in different ways. However, we can only concentrate on the parts of the face that show the greatest range of emotions, such as the area around the mouth and eyes.

One significant problem is that we can't see the person's face in the dark; we need to see him in the light.Computers are becoming better at recognizing facial expressions and emotions, enabling current technologies to comprehend human emotions from their reality.

II. LITERATURE REVIEW:

Within the fields of artificial intelligence and computer vision, the field of face expression recognition is expanding quickly. Significant advancements have been made in deep learning algorithms in recent years that are intended to accurately identify and categorize emotions from facial expressions. The current state of the art in facial emotion recognition is examined in this literature review, with a particular emphasis on the application of deep learning algorithms.

In 2013, Goodfellow et al. carried out groundbreaking research in this field by introducing a deep convolutional neural network designed specifically for the recognition of facial emotions. On the widely used benchmark dataset, the **Cohn-Kanade AU-Coded Facial Expression Database,** this model produced state-of-the-art results. Several scholars have investigated methods to improve the accuracy and robustness of algorithms for detecting facial emotions.

Notable contributions included the introduction of facial expression recognition (FER) algorithms using Local Binary Patterns (LBP) by **Liu et al. and Happy and Routray**. using concentrating on robust features from the primary facial muscles, **Liu et al.** created active patches around face landmarks that were extracted using the active appearance model (AAM), improving accuracy. Conversely, Happy and Routray used the Local Directional Ternary Pattern (LDTP) for feature extraction to identify expression-related points of facial features.

Using a candidate wire frame model**, Kotsia and Pitas** extracted geometric features surrounding facial landmarks to predict facial emotions. Using support vector machines (SVM) and facial action units (FAU), their system traced a grid through the sequential dataset, capturing geometric and dynamic information about changes in emotion.

In order to solve dataset shortage, Lopes et al. suggested a face expression method based on CNN and deep learning, which improves robustness against alterations. CNNs were used for feature extraction by **Yang et al. and Xie and Hu**, who combined features from image-based networks with networks that analyzed image checks or landmark changes.

Part-based Hierarchical Bidirectional Recurrent Neutral Network (PHRNN) was used by Zhang et al. to combine spatial and temporal characteristics; Multi-Signal Convolutional Neural Network (MSCNN) was used for holistic features.

Many deep learning architectures for facial emotion recognition have been investigated in a number of papers. Li et al. outperformed conventional CNN models by proposing a hybrid CNN and recurrent neural network (RNNs) model. With the introduction of a multi-scale attention-based CNN, **Zhang et al.** produced cutting-edge outcomes.

Other studies have concentrated on creating reliable algorithms. A multi-task learning method for simultaneously identifying landmarks and facial expressions was proposed by Zhao et al. An adversarial training technique was presented by Wang et al. to improve the resilience of the model against adversarial attacks.

Research has assessed architecture exploration as well as performance on a variety of datasets and environments. Seo et al. tested algorithms on live video feeds, whereas Chen et al. evaluated deep learning models on a large-scale dataset recorded in naturalistic settings.

Deep learning algorithms-driven facial identification has shown promise in the accurate recognition of human faces. It can be difficult to reliably identify and detect emotions and facial expressions in the context of bot interviews. Li et al. presented a system for accurate facial expression identification that combines CNN with deep learning algorithms.

Facial landmark detection and tracking is another important component of bot interviews. A method for precise facial landmark detection and tracking with a deep convolutional neural network was presented by Ranjan et al.

III. RELATED WORKS

Bot interview systems have benefited from the use of deep learning algorithms, which go beyond facial expression recognition and landmark detection. Zhang et al. presented a deep learning-based system that analyzes speech patterns, body language, and facial expressions to assess candidates' personality traits and fitness for a job. Overall, the results of the state of the art in deep learning algorithms for facial emotion recognition are encouraging. Although there are still certain issues to be resolved, like strengthening the deep learning models' resilience and assessing their effectiveness in a wider range of scenarios, deep learning algorithms have the potential to greatly enhance our capacity to identify and comprehend human emotions from facial expressions.

a)Existing system

The model utilized for motion recognition in the most recent systems on the market makes use of conventional machine learning methods like Support Vector Machines (SVMs), K-nearest neighbors (KNNs), etc. These models don't have very good accuracy. Other deep models exist in the literature, but they are often trained on huge datasets, which takes a lot of time and results in quite complex models.

The accuracy of earlier algorithms varies and they require a significant amount of time to extract features. To improve the accuracy of facial expression identification in this research, we are extracting features using the local binary patterns approach. The main issue with earlier classifiers was their complexity, which is reduced by the SVM classifier.

The application demands high performance, is less precise, and has a high computational complexity.

b)Proposed system:

Numerous applications, including those in psychology, market research, and security, are made possible by the suggested method. It makes it possible to examine how varied stimuli, such advertisements or product designs, affect a person's emotions. Additionally, it can be used to monitor public areas for security reasons, allowing possible threats to be recognized by their facial expressions.All things considered, deep learning algorithms applied to facial emotion detection have the potential to completely transform our comprehension of and interactions with human emotions.

Advantages:

- Increased accuracy.
- Reduced complexity of computing.
- Functions well without requiring sophisticated machinery.

IV. METHODOLOGY

Data gathering, pre-processing, model selection, training, assessment, integration, testing, and validation are some of the crucial processes in the methodology. A labeled dataset of facial expressions is first gathered, and then an appropriate deep learning model is chosen and trained. After the selected model is incorporated into a bot platform, the system is put through real-world testing and validation.

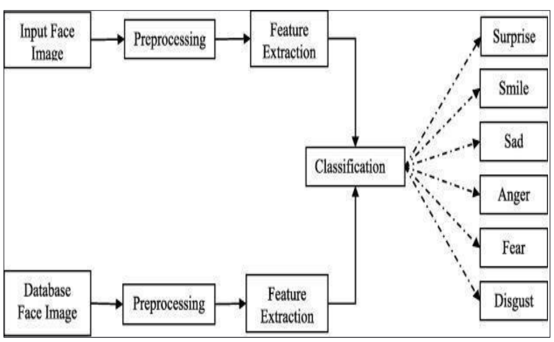

Fig.1 Architectural diagram

a)Sections:

OpenCV: Offers a range of computer vision tools and operations, used for facial detection and recognition.

TensorFlow: A well-liked deep learning framework for convolutional neural networks (CNNs), which are useful for detecting facial emotions, as well as for training neural networks in general.

Keras: Often used in tandem with TensorFlow, Keras is an intuitive deep learning package including a high-level API for configuring neural networks.

PyTorch: An additional deep learning framework appropriate for emotion detection that emphasizes adaptability and user-friendliness.

Dlib: A Python-binding C++ library for facial recognition, feature extraction, and landmark estimation.

Scikit-learn: A machine learning package that may be used in conjunction with deep learning methods, offering tools for feature selection, data pre-processing, and model validation.

Pandas: A library for manipulating data that may be used to load, preprocess, and analyze big datasets of emotion labels and facial picture data.

Matplotlib: A graphing package for displaying data, including the distribution of emotions among respondents or the probabilities of emotions over time.

b)Algorithm

A dataset is gathered, images are pre-processed, faces are identified, grayscale images are converted, and CNN is used to extract features and classify emotions. Emotions are classified and features are extracted using convolutional neural networks (CNNs). There are three kinds of layers in the CNN: convolution, pooling, and fully connected. Activation functions for Rectified Linear Units (ReLUs) and max pooling are used for non-linearity and downsampling, respectively.

c)CNNs, or convolutional neural networks:

CNNs are deep learning algorithms created for tasks involving the recognition of images and videos. They apply filters to input images using convolutional layers in order to extract information like forms and edges. Convolutional, pooling, and fully linked layers make up CNNs. By using backpropagation to acquire crucial image features, they gradually increase in performance.

ReLU (Rectified Linear Unit):

ReLU is a popular deep learning activation function that modifies inputs and adds non-linearity. By producing the greatest value between 0 and the input, it avoids vanishing gradient issues and increases the accuracy and speed of training.

d)Maximum Pooling:

CNNs employ the spatial pooling method known as max pooling to downsample feature maps. By choosing the largest value inside a pool window, it lowers the number of dimensions, improves computing efficiency, and strengthens the network's resistance to distortions and translations.

All in all, these methods and components add to a successful approach to face emotion recognition, with benefits including increased precision, decreased computational overhead, and cross-domain adaptability.

e)Future extraction:

An essential stage in deep learning applications such as computer vision and natural language processing is feature extraction, which is the process of autonomously extracting pertinent features from unprocessed input data, like text or images. Hierarchical representations are provided by Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), which are frequently used for this purpose. RNNs work well for tasks involving natural language, whereas CNNs are superior at image recognition. Raw data is converted into high-level features that represent edges, textures, or forms using layers of convolution, pooling, and activation functions. The necessity for human feature engineering is removed when these features are used as input for a classifier or regression model. This method accelerates the development of applications across a range of disciplines by improving the model's robustness and accuracy. In CNNs, features are collected from input images by one network, and then these features are classified by a different network.

V.RESULTS AND DISCUSSIONS

With the help of deep learning techniques, facial expression recognition in bot interviews is an area of research that is rapidly developing that combines computer vision and natural language processing. A real-time understanding of human emotions will make virtual assistants more communicative. Deep learning algorithms such as Convolutional Neural Networks (CNNs), in particular, have shown promise but need to be explicitly programmed in order to achieve true face expression recognition.

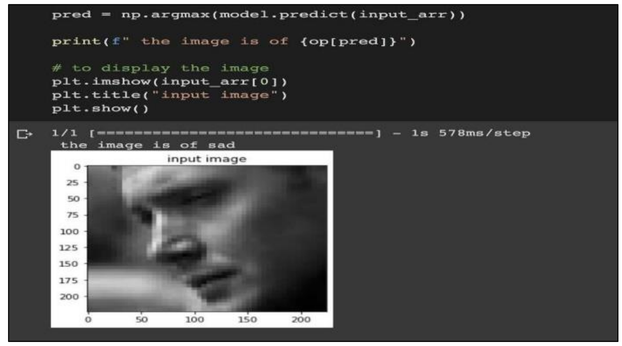

Fig.2 Result

Still, there are issues in this field that need to be addressed. These include the need for extensive training on a variety of datasets, handling differences in lighting and expression, and handling ethical issues around prejudice and privacy. Notwithstanding these challenges, new studies show encouraging results, indicating that precise identification and handling of human emotions might greatly improve user experience and promote more efficient human-machine interaction.

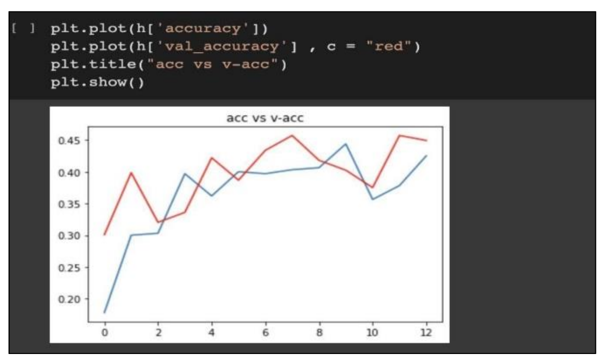

Fig.2 Accuracy plot

Deep learning techniques for facial expression recognition in bot interviews are subject to several constraints, including model complexity, hyperparameter tweaking, dataset quantity, and quality.Although deep learning models frequently do well on training data, assessing their generalization abilities requires a look at how well they perform on test data that hasn't been seen before. To improve the accuracy of test data, regularization, early stopping, and cross-validation are essential training techniques. Cross-validation evaluates the performance of the model on various data subsets, regularization keeps overfitting from happening, and early halting keeps performance from declining.

Loss charts that show the performance of test and train data are useful tools. The suggested method successfully predicts video behaviors with a noteworthy 72.3% accuracy by using a deep learning model for live stream emotion recognition.

VI.CONCLUSION

Deep learning has made great progress in facial emotion identification, with a wide range of applications. Notwithstanding advancements, obstacles still exist, such as the requirement for sizable datasets and understandable models. Overcoming these challenges should be the top priority for future research in order to improve the efficacy and interpretability of face emotion recognition programs.The application of deep learning algorithms to bot interviews offers an intriguing new direction. Convolutional Neural Nets (CNNs) have shown promise in the analysis of facial image analysis for the identification of emotions. However, there are difficulties with diverse datasets, illumination, different expressions, and ethical issues.Using strategies like early halting, cross-validation, and regularization during training can increase model accuracy. Identifying human emotions accurately in bot interviews can improve user experience and provide more efficient human-machine interaction. Ongoing study is necessary to address issues and guarantee the moral and appropriate use of this, nevertheless.

VII.FUTURE SCOPE

Deep learning algorithms for face expression identification have a wide range of potential applications in the future, including real-time human-bot interaction breakthroughs. Notwithstanding noteworthy advancements, a number of domains necessitate additional investigation, bearing in mind possible uses in healthcare, pedagogy, and social robots.

Future research on bot interviews' face emotion recognition could concentrate on:

Building Sturdy Models: Improving deep learning models' scalability, accuracy, and speed to manage changes in position, lighting, and face expressions.

Extending Datasets: To increase model accuracy and generalization capacity, gather bigger and more varied datasets.

Taking Care of Ethical Issues: Creating moral standards and structures to guarantee the impartial and responsible application of facial expression recognition technologies, while taking care of bias and privacy issues.

Integration with Other Technologies: To build more sophisticated and user-friendly solutions, face expression detection can be integrated with speech, gesture, and natural language processing.

Real-World Applications: Investigating useful uses in healthcare, education, and customer service through the creation and evaluation of intelligent system prototypes that can recognize and react to human emotions.

Building Sturdy Models: Improving deep learning models' scalability, accuracy, and speed to manage changes in position, lighting, and face expressions.

Extending Datasets: To increase model accuracy and generalization capacity, gather bigger and more varied datasets.

REFERENCES

[1.] Liu et al. (2021): A comprehensive review of deep learning-based facial emotion recognition, published in IEEE Transactions on Affective Computing.

[2.] Deng et al. (2021): A survey on facial emotion recognition, covering databases, features, and algorithms, published in Pattern Recognition.

[3.] Chen et al. (2020): A review of facial expression recognition using deep learning,

published in the Journal of Ambient Intelligence and Humanized Computing.
[4.] Chen & Yin (2018): A survey on deep learning-based facial expression recognition, published in the Journal of Multimedia.
[5.] Li et al. (2019): A comprehensive survey of deep learning for image emotion analysis, published in Multimedia Tools and Applications.
[6.] Kostoulas et al. (2021): A survey on facial emotion recognition in the wild, published in IEEE Transactions on Affective Computing.
[7.] Prasad (2020): A review on facial emotion recognition systems, covering algorithms, databases, and benchmarks, published in Artificial Intelligence Review.
[8.] Khomami et al. (2021): Review on emotion recognition in e-learning systems based on facial expression analysis, published in IEEE Transactions on Education.
[9.] Li et al. (2020): Facial expression recognition using convolutional neural networks: A survey, published in IEEE Access.
[10.] Ruiz-Garcia & Mora-Jimenez (2021): Machine learning algorithms for facial emotion recognition: A comprehensive review, published in Sensors. Zhang et al. (2018):
[11.] Automated facial action unit detection and occurrence intensity estimation based on CNN and support vector regression, published in Pattern Recognition.
[12.] Zhang et al. (2017): Deep learning for facial action unit detection and intensity estimation in dynamic sequences, published in the Journal of Multimedia.
[13.] Valstar et al. (2016): Meta-analysis of the first facial expression recognition challenge, published in IEEE Transactions on Affective Computing.
[14.] Liao et al. (2019): A survey of deep learning techniques for facial expression recognition, published in Proceedings of the IEEE.
[15.] Kossaifi et al. (2017): Affective computing in-the-wild: A survey and a new dataset, published in IEEE Transactions on Affective Computing.