IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Fake Smile Classifier

[1]Divya H N, [2]Y Sandeep, [3]Y Tejeesh Raghav, [4]Suchita, [5]Shashidhar Firangi
[1]Associate Professor, [2] Student, [3]Student, [4]Student, [5]Student
[1]Computer Science and Engineering,
[1] Dayananda Sagar Academy of Technology & Management, Bengaluru, India

Abstract: This research presents an innovative AI-driven framework for detecting and classifying facial expressions, focusing on distinguishing fake smiles, genuine smiles, and identifying emotional states like anxiety and depression. By harnessing the power of deep learning, specifically through the ResNet-50 architecture and Convolutional Neural Networks (CNNs), the system is capable of analyzing subtle facial features with high accuracy. The model's real-time analysis capability offers transformative potential in mental health diagnostics, providing a non-invasive and efficient method for early detection of psychological conditions. This project aims to bridge the gap between traditional psychological assessments and cutting-edge AI technologies, enabling scalable and personalized mental health monitoring. The results demonstrate significant advancements in the field of emotion recognition, emphasizing the role of AI in enhancing mental health care through timely and precise interventions.

Keywords- *Artificial Intelligence, ResNet-50, Convolutional Neural Networks, Fake Smile Detection, Real-Time Emotion Analysis, AI in Healthcare, Psychological Well-being.*

I. INTRODUCTION

The Fake Smile Classifier project is an important confluence of AI and deep learning techniques, aiming to analyze and classify facial expressions as either real or fake smiles. This system makes use of the advanced neural networks (NN), such as ResNet-50 and Convolutional Neural Networks (CNNs), in order to detect minute variations in facial features that are not easily detectable by the human eye. The core objective of the project is the differentiation between genuine and fake smiles, but the scope of detection also includes different emotional states such as anxiety and depression. Since these emotions are very subtle, they are sometimes difficult to be recognized, which makes the AI and deep learning a very useful tool for proper identification.

The importance of this project lies in its use for self-monitoring and assessing mental health issues. Finding a person's emotional states through facial expressions is the most crucial aspect wherein information about one's mental health can be extracted before the manifestation of symptoms that relate to anxiety, depression, and other psychological illnesses. These abilities promise much to psychologists, therapists, and medical professionals, because they can assist in the earlier diagnosis and design of treatment programs for patients. Real-time applications of emotion recognition on mobile apps or wearable devices would make it possible to track a patient's emotional health over time and help them maintain an active stance regarding their mental health.

Furthermore, the Fake Smile Classifier fills in a critical gap in mental health diagnostics by incorporating smile detection and analysis of other emotional states, such as anxiety and depression. As AI continues to advance, its role in improving emotional well-being through non-invasive methods is becoming highly important. The project, therefore, provides a comprehensive approach to assessing emotions, and this may even revolutionize how diagnoses are approached and therapy delivered. The contribution of the Fake Smile Classifier to the ongoing development of mental health is its ability to provide an accessible, efficient, and reliable tool for emotion detection, thus opening up new possibilities both in clinical applications and personal well-being management.

II. LITERATURE REVIEW

The detection of fake smile using deep learning has emerged as one of the important research areas due to their extensive applications in psychology, healthcare, and social robotics. Fake smiles are often called "non-Duchenne smiles," which differ from genuine smiles by the absence of activation of the orbicularis oculi muscle around the eyes. With the invention of CNNs and ResNet, this accuracy and reliability have improved for such subtle variations in faces, greatly changing the realm of emotion recognition [1][7].

2.1 Facial Emotion Recognition Techniques

Facial Emotion Recognition, a crucial constituent of human-computer interaction, emphasizes the examination of facial expressions for the better interpretation of emotions. Modern FER techniques focus on pre-processing methods like detecting facial features (e.g., eyes, eyebrows, nose, and mouth) and extracting fine-grained characteristics to improve the quality of input data and enhance emotion classification accuracy. Classifiers such as CNNs, SVMs, and RNNs are trained on labeled datasets to generalize and generate emotion labels for unseen data [8].

FER is a multi-stage process that involves facial feature detection, feature extraction, and training of the classifier. Preprocessing standardizes video frames or images so that the data quality is improved for emotion classification. Fine-grained feature extraction identifies subtle movements in facial muscles, which are important for the detection of complex emotions such as anxiety or happiness. Modern FER systems use large-scale datasets such as AffectNet or FER2013 to enhance their generalizability. Despite the progress made, issues such as dataset imbalance, spontaneous expressions, and cultural diversity in facial expressions remain critical issues to be addressed [3][8].

2.2 Fake Smile Detection

Eddie Wu et al. proposed a Model for Fake Smile Detection using CNNs, showing the efficiency of machine learning in detecting fake smiles based on minute facial changes and changes in muscle movement [1]. Nagaraj et al. again moved a step forward within the domain with the help of CNNs to detect fake faces and contributed significantly to the idea of feature extraction techniques within the domain of facial expression classification [7].

Fake smile detection systems mainly rely on facial dynamics, paying particular attention to distinguishing between real and false smiles in terms of muscle activation patterns. For example, a real smile requires the coordination of both the zygomatic major muscle and the orbicularis oculi muscle, which is not present in most cases of fake smiles. Deep learning models, such as ResNet-50, improve the capability to extract hierarchical patterns from images, allowing better accuracy in distinguishing true from false smiles [2][3]. Moreover, by including anxiety and depression detection, these systems extend their application towards mental health diagnostics, providing a non-invasive approach for psychological assessment. Real-time emotion classification improves its practical usability in therapy and behavioral studies further [4][7].

2.3. Multimodal Emotion Recognition

The multimodal data integration has proved to be very promising for enhancing the accuracy of emotion detection. Facial expressions along with audio cues are combined for better understanding of emotional states. Rohan Agrawal et al. have shown that voice-to-text systems can be feasible in emotion recognition, where visual and auditory data can be utilized for further analysis. Multimodal systems, which integrate modalities such as speech tone, facial expressions, and head gestures, are designed to overcome the limitations of single-modal approaches, which are often sensitive to environmental noise or occlusion [6][11].

For instance, when a video conferencing is conducted, audio cues combined with face expressions are more likely to give more accurate insight into the participant's emotional state, even with poor lighting or low resolution video. The combination of these modalities ensures robustness; therefore, multimodal emotion recognition becomes a key focus for future research work [11].

2.4 Challenges with Real-World Applications

Real-world applications of smile detection face several challenges, including pose variations, complex backgrounds, and diverse lighting conditions. Junkai Chen et al. explored smile detection in uncontrolled environments, revealing that while deep CNNs exhibit robustness to pose changes, handling intricate and variable backgrounds remains a significant limitation. These findings emphasize the need for models that can generalize effectively across diverse real-world scenarios [7].

Addressing such difficulties includes methods of data augmentation: artificially augment the diversity of datasets by incorporating various pose, brightness, and backgrounds. Further adaptation techniques on data from different domains might be implemented in order to allow models to better adapt to real situations. Robust extraction of features might be used using attention mechanisms applied in neural networks to tackle difficulties caused by a complex background of images [5][7][10].

2.5 Detection of Deception and Emotional Intelligence

A very important application of CNNs has been in the field of deception detection. Laslo Dinges et al. proposed a system based on facial cues such as gaze, head pose, and expressions for automated deception detection. Their work indicates the need for scenario-specific training data and suggests that combining several modalities like visual and auditory cues would lead to better accuracy in detection. The aim of the present Fake Smile Classifier is to match these objectives through detecting nonverbal emotional markers-including anxiety, depression, or other forms-simultaneous with fake smiling [7][9].

The deception-detecting devices will try to catch the disparity between the contradictory facial expressions as well as speech against non-verbal indicators of lie, such as altered gaze patterns, micro expressions, or other variations in pitch. These systems would yield higher performance on tasks tied to high-stakes scenarios, such as testimonies in a courtroom or security screenings, by being trained on the said datasets. Furthermore, inclusions of AU analysis enable models to detect fine-grained facial muscle movements, allowing more insights on emotional verisimilitude [5][7][9].

2.6 Improvements Through Advanced Architectures

Recent advancements in AI architectures, such as visualizing and understanding convolutional layers through techniques proposed by Zeiler and Fergus, have further refined the interpretability of deep learning models [2]. The adoption of architectures like ResNet-50 and their modifications has streamlined the process of feature extraction, making these systems highly efficient for large-scale image recognition tasks [3][4]. Techniques such as Ant Colony Optimization (ACO) have also been explored to optimize image processing pipelines, adding robustness to segmentation and preservation processes in emotion recognition [9][10].

2.7 Mental Health and Well-Being Monitoring Application

The Fake Smile Classifier project integrates AI with deep learning methodologies to analyze facial expressions and classify them as either real or fake smiles. It uses powerful neural networks, such as ResNet-50 and CNNs, to capture minute variations in facial features that are often imperceptible to the human eye. Although the system is primarily designed to differentiate between real and fake smiles, it can be extended to detect other emotional states like anxiety and depression. These emotions are usually subtle and can be detected with high accuracy.

The importance of this project is that it can be applied in monitoring mental health and well-being. Face-emotion recognition detects the emotional state of a person, which reflects the mental condition of the person. This technique can be applied to detect symptoms of anxiety and depression at their early stages. This is the most significant significance for psychologists, therapists, and healthcare professionals in terms of detecting mental health issues at an early stage and customized treatment. Furthermore, real-time emotion recognition could be used to monitor emotional health over time using daily-use applications such as mobile apps or wearable devices [12].

By integrating smile detection with the analysis of anxiety and depression, the Fake Smile Classifier fills an important gap in mental health diagnostics. As technology advances, the role of AI in enhancing emotional well-being through non-invasive methods is increasingly crucial. This project offers a holistic approach to understanding emotional states, improving both diagnostic and therapeutic processes. Its real-time capabilities and reliability make it a transformative tool for emotional assessment, fostering scalable and accessible mental health solutions [12].

In a nutshell, the works above highlight the transformative ability of deep learning techniques, especially CNN and ResNet architectures, toward improving the fields of emotion recognition and fake smile detection. Thereby, these systems address existing challenges, utilize multimodal data integration, and exploit advanced architectures to improve accuracy and robustness, opening the doors for impactful applications in healthcare, psychology, and beyond [12].

III. WORKFLOW

This project pipeline was carefully designed to classify facial emotions, with particular attention paid to the detection of fake smiles, real smiles, anxiety, and depression. The complete pipeline ensures that data handling and model training are done systematically and efficiently, making it possible to detect emotional states in real time through facial expressions. Here is a step-by-step explanation of each stage in the pipeline:

Data Collection and Preprocessing: The very first step in the data processing pipeline is the collection of a dataset that consists of labeled facial images representing various emotional states, such as fake smiles, real smiles, anxiety, and depression. In most cases, public datasets like FER-2013 and AffectNet are used. These datasets are enormous collections of annotated facial images and are, therefore, perfect for training and testing the model. They come as various images from other different demographic populations. This implies that the model will be solid and generalization for different sets of populations will not be challenging.

The following preprocessing steps were done on these images in preparation for the data in the system input:

i. **Image Resizing**: The images in the dataset are all resized to the same dimension that is 224x224, ensuring uniform and compatibility with a pretrained ResNet-50, which requires particular input dimensions of images. Image resizing is of paramount importance because that is what enables the model to process the input images.

ii. **Image Normalization**: This is a standard mean and a standard deviation used in normalizing pixel values of images, as predefined in datasets such as ImageNet, such as mean=[0.485, 0.456, 0.406] and std=[0.229, 0.224, 0.225]. This technique standardizes the data for images to make them more usable in deep learning algorithms. The pixel values for all images will be scaled in the same way because of normalization. Because it makes things go much faster and will train with much greater accuracy, it makes things go much better.

iii. **Data Augmentation:** The idea behind data augmentation is to enhance the capabilities of the model so that it performs better with reduced training datasets. The above set of techniques used includes random rotations, flips, zooms, and cropping of images. Through these artificial augmentations to the dataset, the model is less likely to overfit with respect to training data and can tolerate more variability in facial expressions while on deployment. Augmentation also makes sure that the model learns to recognize facial expressions under different orientations and lighting conditions, which is a requirement for real-world applicability.

iv. **Feature Extraction with Pretrained ResNet-50:** The next stage of the pipeline involves feature extraction, which plays a crucial role in extracting the necessary visual cues from facial expressions. For this task, the project uses the ResNet-50 model, a widely known deep convolutional neural network (CNN) that has been pre-trained on a large-scale dataset like ImageNet. ResNet-50 is specifically very good at extracting hierarchical features from images, such as complex patterns, like facial structures and expressions, which are important in identifying various emotional states.

The pretrained ResNet-50 model learns to identify a vast range of visual features, from edges and textures to shapes and other subtle facial cues. All these features are highly relevant to the detection of real versus fake smiles and even to the recognition of emotional states like anxiety and depression. The pretrained layers of ResNet-50 are maintained in the pipeline because these have been optimized to capture those complex visual patterns across different domains. Using such pretrained layers enables the model to draw upon knowledge obtained from a large-scale dataset, ImageNet, which reduces training by a lot and improves the overall efficiency of the model.

v. **Model Training and Fine-Tuning:** The training and fine-tuning process is the heart of the pipeline's effectiveness. The last fully connected (fc) layer of ResNet-50 is replaced to suit the number of classes in the classification task at hand, namely, fake smile, real smile, anxiety, and depression. This is because the original ResNet-50 model was trained for a general image classification task, and its last layer needs to be adapted to suit the specific classification problem at hand.

This pipeline uses transfer learning, an approach that allows the model to retain the valuable feature extraction abilities learned from ImageNet while adapting to the new task with minimal computational cost. The pretrained convolutional layers are frozen during the fine-tuning process, meaning they are not updated during training, and only the new fc layer is trained. This method significantly reduced the time required for training while using fewer computational resources, thus more efficient and just as accurate in its predictions.

Fine-tune the fc layer to enable it to learn on the specific relation between the obtained features and emotions, so there is a good discriminative capability distinguishing between slight conditions such as false smiles and anxious states.

vi. **Classification of Emotion:** Once the model has been trained, it is ready to classify new input images. When a new image is passed through the model, it is processed by the pretrained convolutional layers of ResNet-50, which extract the relevant features from the face. These features are then passed through the newly trained fc layers, which map the extracted features to the emotional states defined in the classification task.

The model outputs a probability distribution over the possible emotional states real smile, fake smile, anxiety, or depression based on the extracted facial features. The class with the highest probability is selected as the final prediction. This classification step is crucial for identifying the presence of specific emotional states from the facial expressions.

vii. **Real-Time Emotion Detection:** The deployed trained model then enables real-time emotion recognition in dynamic environments such as video feeds or live camera streams. In such a real-time application, the system will continuously process individual frames, applying the same preprocessing steps followed by feature extraction and emotion classification for each frame. This allows the system to monitor facial expressions in real-time and accurately detect emotions such as fake smiles, anxiety, and depression during activities like therapy sessions or research studies. Real-time emotion detection opens up significant possibilities for continuous emotional assessment in various contexts, offering timely insights into an individual's emotional well-being.

This real-time analysis ensures that there is a direct tracking of changes in emotions; this is most important in the clinical environment, and it offers timely intervention and continuous feedback regarding the emotional condition of a patient during treatment or therapy sessions.

After classification, the system may apply post-processing techniques to refine the results and give meaningful feedback. For example, the system may show confidence scores for each detected emotion, providing a level of certainty about the predictions made. Moreover, aggregated results may be used to track emotional trends over time, which can help in giving a comprehensive overview of an individual's emotional state.

If the presence of anxiety or a forced smile is sensed by the system, it may produce an alert and feedback to psychiatric professionals to properly assess the emotions of the patient. Such understandings can facilitate further evaluation and personalized treatment towards better therapeutic outcomes.

Besides this, the post-processing steps of this system ensure it is not only accurate in terms of emotion detection but also actionable for healthcare professionals. The confidence level of its predictions and aggregation of data from the system add to its usability for real-world applications in mental health monitoring and assessment of emotional well-being.

The integration of these stages within the processing pipeline creates a potent tool for the detection of emotion, one which can provide immediate feedback for use in both the clinical and personal realms. Such an approach seeks to combine innovative deep learning approaches with practical use, ensuring accuracy and efficiency while identifying and interpreting subtle facial expressions related to mental health.

IV. METHODOLOGY

4.1 Data Gathering and Preparation

The success of any deep learning project largely depends on the dataset quality and diversity. For this project, data has been gathered from known sources like FER-2013 and AffectNet along with proprietary data which ensures the variability in age, gender, and ethnicity. Data includes labelled images of facial expression of varied states such as artificial smiles, actual smiles, anxious, and depressive. This diverse dataset allows the model to generalize effectively across various populations and emotional expressions.

Raw image data is prepared in this preprocessing step. All images are resized into 224x224 pixels size, which is the input size for the pre-trained model, ResNet-50. The uniform resizing within the dataset enables fast training of models. Pixel values also get normalized to standard mean and standard deviation from ImageNet so that the images are standardized, promoting faster convergence during the training process.

Further, data augmentation techniques are utilized to make the model generalize more. These data augmentation techniques like random flipping, rotation, and brightness adjustments artificially increase the size of the dataset by introducing different variations in the facial expressions and orientations. It is particularly helpful in making the model more robust in real conditions where facial expressions may vary with different environmental factors or lighting conditions.

4.2 Model Architecture and Customization

The model is built on the architecture of ResNet-50, which is a well-known feature extractor that effectively extracts hierarchical features from images. ResNet-50 is built using multiple convolutional layers that can capture important visual features such as facial landmarks, edges, and textures. A pre-trained ResNet-50 model provides a rich set of learned features, which are generally applicable to many tasks.

In the case of adapting the ResNet-50 model for emotion classification, the last fully connected (fc) layer is replaced by a custom dense layer that will classify images into one of four classes: real smile, fake smile, anxiety, and depression. Softmax activation is applied to this layer in order to get the probability distribution over the four classes. It utilizes transfer learning where the convolutional layers pre-trained are frozen and the dense layer is customized to fine-tune for the task of emotion classification. This enables faster training with better accuracy because the model will be concentrating on the exact features that matter for the recognition of subtle emotional differences.

4.3 Training and Optimization

Training is done in two stages:. In the first stage, only the custom dense layer is trained while the ResNet-50 layers are frozen. This allows the model to learn the classification-specific features related to emotions. In the second stage, selective layers of ResNet-50 are unfrozen, allowing the model to fine-tune its deeper features to better capture emotional cues. The Adam optimizer is used with adaptive learning rates to ensure efficient convergence. Utilized is categorical cross-entropy loss function, where learning rate schedulers dynamically changed the learning rate to avoid overfitting.

4.4 Real-Time Implementation and Evaluation

This model is finally deployed in the real-time application using OpenCV, which is responsible for capturing video streams in live and processing the frames to get the facial expressions. The instantaneous emotion classification from the system could detect the existence of smile markers, anxiety, and depression. With respect to evaluating a model's performance, a plethora of metrics include accuracy, precision, recall, and F1-score. These specific misclassifications are diagnosed from a confusion matrix and inform subsequent refinements. Real-time feedback guarantees applicability in practical, real-life scenarios for clinicians and researchers alike.

This method can be understood as an extensive procedure in constructing a deep-learning facial emotion classifier powered by artificial intelligence.

V. CONCLUSION

This project uses state-of-the-art AI techniques in classifying facial emotions, targeting the differentiation of fake smiles and real smiles as well as indicating anxiety and depression. The system uses a pretrained ResNet-50 model in combination with transfer learning to evaluate subtle facial expressions for accurate classification of emotional states. This method ensures precise classification, making it highly effective at recognizing nuanced facial cues that denote emotional health.

The integration of both smile detection and emotional health indicators enhances the system's ability to offer a comprehensive tool for mental health assessment. With the ability to detect a range of emotions in real-time, the project is well positioned to be a valuable asset in clinical and research settings where it can be used for early detection and continuous monitoring of mental health conditions. Its potential to aid psychologists, therapists, and other healthcare professionals in diagnosing and understanding emotional states opens new doors for personalized and proactive care.

To ensure the model's high accuracy and robustness, techniques such as data augmentation and model fine-tuning are used, allowing the system to generalize effectively across diverse datasets. These processes enable the model to adapt to variations in facial expressions and environmental factors, further increasing its reliability in real-world scenarios.

This project is a great example of how AI transforms the support of mental health, by providing real-time feedback on emotional states and improving the understanding of non-verbal cues. The system's ability to assess emotional well-being can pave the way for AI-driven tools that monitor and promote emotional health, potentially improving quality of life and providing valuable insights into mental health management.

VI. REFERENCES

[1] Eddie Wu, Yixuan Li, Kuan-Cheng Chiu, (2018) "Fake-Smile-Detection-Model" eddiewu01/Fake-Smile-Detection-Model: Fake Smile Detection Model using CNNs in Python

[2] Zeiler M.D., Fergus R. (2014) Visualizing and Under-standing Convolutional Networks. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer, Cham

[3] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[4] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp.1097-1105).

[5] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[6] Raj Kumar, Smitha, Altaf Raja Anuhya,Dr C Nandini" Controlling Multiple Applications with Hand Gesture Using Convolution Neural Network", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.10, Issue 4, pp.e73-e78, April 2022

[7] Nagaraj M Lutimath, Sneha Reddy M V, Shravani N, Yasaswitha Reddy S, Pavan Sai C, "Identification of Fake Faces Using Convolutional Neural Network", 9 International Journal of Research and Analytical Reviews (IJRAR), Vol 11, Issue 1, Jan 2024, pp. 53-56.

[8] Manjunath, D. R., Selva Kumar S., Sai Shiva Sumanth Reddy, and Lohith J. J. "Enhancing Personalized Learning Based on AI-Driven Lesson Plan Generator Using Mistral-7B for Efficient Content Extraction and Summarization." *Gradiva Journal*, June 2024, https://gradiva.it/june-2024/

[9] P. Nandihal, P. K. Pareek, V. H. C. De Albuquerque, M. R. B, A. Khanna and V. S. Kumar, "Ant Colony Optimization based Medical Image Preservation and Segmentation," *2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE)*, Bangalore, India, 2022, pp. 1-7, doi: 10.1109/ICATIECE56365.2022.10047584.

[10] S A, K. ., Nandihal , P. ., K, S., D R , M. ., & Liyakathunisa. (2022). PRIOR DETECTION OF ALZHEIMER'S DISEASE WITH THE AID OF MRI IMAGES AND DEEP NEURAL NETWORKS . *Malaysian Journal of Computer Science*, 16–28. https://doi.org/10.22452/mjcs.sp2022no2.2

[11] Rohan Agrawal, Rohan Rajesh Purundhar, Siddharth Mehta, Yash Bhalla, Dr. C. Nandini, "Lyrebird: Voice-To-Text Note Making Automated Software with Speech Recognition," International Journal of Research Publication and Reviews, Vol(2) Issue (7) (2021) Page 272-276.

[12] Manasa Sandeep, C. Nandini, "An Extensive Survey on 3D Face Reconstruction based on Passive Method" in International Research Journal of Engineering and Technology (IRJET), Volume 8 Issue 12 December 2