Real-Time Lie Detection and Deception Analysis

Mrs. S. Jansi Rani.[1], Vishal A. R[2], Sarweswaren R[3]
[1] *Assistant Professor, Sri Ramakrishna Engineering College, Coimbatore, India*
[2,3,4] *UG Scholar, Sri Ramakrishna Engineering College, Coimbatore, India*

Abstract **Historically, lie detection has relied on invasive techniques such as polygraphs, which often lack accuracy and necessitate significant human intervention. This paper presents a mobile application designed for real-time, non-intrusive lie detection, leveraging advancements in Artificial Intelligence (AI) and Machine Learning (ML). The application evaluates facial micro-expressions and vocal stress signals by capturing live video and audio through a Flutter-based platform. The collected data is processed in the cloud using deep learning algorithms. By integrating technologies like TensorFlow, OpenCV, MediaPipe, and Firebase, the application delivers instant deception analysis while ensuring a user-friendly experience. This novel approach is scalable, adaptable, and applicable across various domains, including security, forensics, and psychology.**

Index Terms **- Time Analysis, Facial Micro-Expressions, Vocal Stress Analysis, Artificial Intelligence (AI), Lie Detection, Facial Micro-Expressions, Deep Learning.**

I. INTRODUCTION

The detection of deception has been a significant area of interest across various disciplines, including law enforcement, psychology, national security, and the study of human behavior. Historically, polygraph examinations have been the predominant technique for identifying falsehoods. These assessments track physiological indicators such as heart rate, blood pressure, respiration, and galvanic skin response to detect stress reactions that may suggest dishonesty. However, these physiological signals are not solely indicative of deception; they can also be affected by factors such as anxiety, fear, or health issues, which may result in inaccurate readings. Additionally, individuals who are trained in countermeasures can deliberately alter their physiological responses to mislead the polygraph, thereby compromising its effectiveness.

Recent research indicates that subtle behavioral indicators, including facial micro-expressions, patterns of eye movement, blink frequency, and vocal stress signals, can reliably indicate deception. These indicators are often subconscious, fleeting, and challenging to manipulate, rendering them more reliable than conventional physiological measures. The swift advancement of Artificial Intelligence (AI) and Machine Learning (ML) has opened new pathways for a more sophisticated and automated analysis of human behavior. Recent research indicates that subtle behavioral indicators, including facial micro- expressions, patterns of eye movement, blink frequency, and vocal stress signals, can reliably indicate deception. These indicators are often subconscious, fleeting, and challenging to manipulate, rendering them more dependable than conventional physiological signs.

This initiative introduces a mobile-based, real-time lie detection system that leverages AI to evaluate a blend of visual and auditory signals. The solution is developed through a Flutter-based mobile application that captures live video and audio from the user. By employing computer vision and deep learning techniques, the system analyzes facial expressions and vocal traits to identify signs of stress or dishonesty. By simultaneously assessing various behavioral markers, the system improves accuracy while ensuring user convenience. This project marks a substantial advancement in the development of lie detection technologies, transitioning from intrusive, hardware-reliant systems to a software-based, user-friendly, and intelligent application.

II. LITERARY SURVEY

This survey reviews recent researches on Real time face detection, and implements various varieties of deception analysis.

Hayder Azeez Neamah Diabil.,[1] centered on using Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs) to examine gaze direction and facial expressions as indicators of dishonesty. The study came to the conclusion that when it comes to spotting lies, facial expressions provide more reliable and powerful clues than gaze. The analysis's accuracy was encouraging, particularly in controlled settings. The study's main drawback, however, was the absence of variation in the facial geometries employed during training, which has an impact on generalization across people with different facial features and races. This suggests that larger datasets that more accurately reflect populations in the real world are required.

Zhicheng Ding.,[2] this recent submission presented a bimodal CNN architecture intended to improve the accuracy of lie detection by concurrently examining physiological signals and linguistic data. This approach combines physiological signs like skin temperature, pulse rate, or vocal stress levels with linguistic cues like word choice, syntax patterns, and semantic inconsistencies, so utilizing the complimentary capabilities of both data sets. Richer feature representations are produced by integrating various modalities, which enhances the system's capacity to discriminate between dishonest and honest conduct.The model was trained and assessed during the testing phase using datasets gathered from controlled settings where participants were observed while completing truth-and-lie activities.

Zhicheng Ding., [3] The study explores the application of convolutional neural networks for multimodal deception detection, utilizing a dataset comprising interviews with 104 subjects, each providing both truthful and falsified responses. The authors extracted linguistic and physiological features from this data to train and construct neural network models. They proposed a fused convolutional neural network model using both modalities to achieve improved overall performance and compared their approach with earlier methods designed for multimodal deception detection. The findings indicate that their system outperforms regular classification methods, demonstrating the feasibility of using neural networks for deception detection even with limited data.

Ryo Hatano., [4] Researchers gathered data from participants who were asked to answer a series of questions both honestly and dishonestly. Facial expressions were captured through webcams, while pulse rate information was collected using smartwatches. This comprehensive dataset was designed to identify the subtle physiological and behavioral indicators linked to deception. A Random Forest classifier was utilized to analyze the integrated features from both data sources. This ensemble learning technique is recognized for its strength and capability to manage complex, nonlinear relationships within the data.The system achieved an F1-score of up to 0.88, reflecting a high degree of accuracy in differentiating between truthful and deceptive answers.

Yongxin Wang., [5] In this study, the authors introduced a hybrid model that integrates Convolutional Neural Networks (CNNs) with Bidirectional Long Short-Term Memory (BiLSTM) networks to assess both facial expressions and vocal cues for the purpose of detecting deception. The CNN segment was employed to extract spatial features from facial expressions, while the BiLSTM segment focused on understanding the temporal relationships within speech patterns. This combination enabled the system to effectively represent the spatial- temporal dynamics linked to deceptive behavior. The system exhibited a high level of accuracy in identifying patterned, time-sensitive deceptive actions, especially in controlled settings where such patterns are more evident.

III. METHODOLOGY

The technique for this research consists of multiple interrelated stages, beginning with video input acquisition, followed by face detection, feature extraction, and, finally, deception classification. Initially, the device gathers live video frames from a camera or a previously recorded video. Pre-trained face detection methods extract facial landmarks and expression data. These characteristics are then fed into a trained machine learning model, which classifies the observed behavior as true or deceitful.

3.1 Model Specification

The system's model is built on a convolutional neural network (CNN) architecture that has been optimized

for facial emotion recognition. For classification, it has numerous convolutional layers, then pooling layers, and finally fully linked dense layers. Dropout layers are utilized to prevent overfitting, and ReLU is selected as the activation function. The last layer employs a softmax function to determine the likelihood of deceitful or truthful behavior.

Dataset Preparation

Dataset preparation is an essential step in training an effective deception detection system. The method begins with the collection of video samples or image frames that depict diverse human expressions linked with truthfulness and deception. After collecting, frames are taken from videos and subjected to preprocessing techniques such as face alignment, resizing, normalizing, and grayscale conversion to ensure consistency.

Model Training

Model training is teaching the neural network to recognize patterns in facial features that signify honest or dishonest behavior. The preprocessed dataset is separated into training and validation sets, usually in an 80:20 ratio. A convolutional neural network (CNN) architecture is used to extract hierarchical information from facial photographs by passing them through many layers. For efficient convergence, the model employs a loss function similar to categorical cross-entropy and the Adam optimizer. During training, strategies like batch normalization and dropout are used to increase stability and reduce overfitting. The model is trained across numerous epochs, with accuracy and loss being tracked to ensure consistent learning.

Evaluation Metrics

To assess the efficacy of the deception detection model, various standard criteria are used. These include accuracy, precision, recall, and F1-score, which each provide a unique perspective on model performance. Accuracy evaluates overall correctness, whereas precision and recall address false positives and false negatives, respectively.

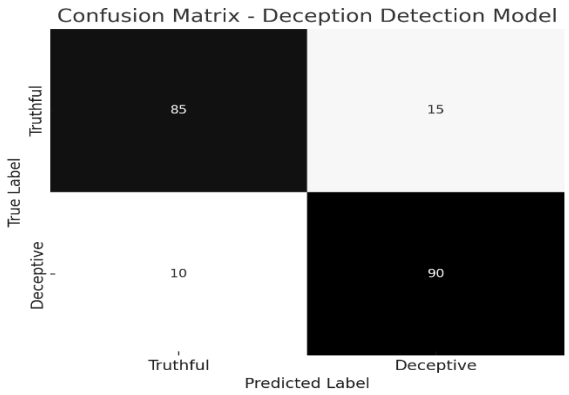

Fig 1 Confusion Matrix (Deception Detection)

3.2 Facial Recognition and Fitting

Face Detection and Preprocessing

Facial recognition is important in deception analysis because facial expressions can disclose psychological states. The system recognizes and extracts facial regions from input video frames using models such as MediaPipe or MTCNN. The key facial landmarks are then employed to assess micro-expressions such brow motions, lip tightening, and gaze aversion. These minor clues are key predictors of potential dishonesty. Training models using annotated emotional datasets teaches the system to link specific face patterns with deceitful behavior. This layer of analysis provides more depth to the behavioral interpretation and improves overall system reliability.

Model Architecture

The model architecture is built around a convolutional neural network (CNN) backbone, which is optimized for image-based emotion recognition. It has several convolutional layers for feature extraction, followed by max-pooling layers to minimize spatial dimensions.

Validation and Tuning

To ensure robustness across previously unreported data splits, model validation employs k-fold cross-validation. During training, the dataset is divided into training and validation sets, with performance measured on the validation set after each epoch.

Testing

Testing is carried out on a set fraction of the dataset that was not used for training or validation. During testing, important metrics such as accuracy, precision, recall, and the F1-score are calculated

3.3 System Architecture

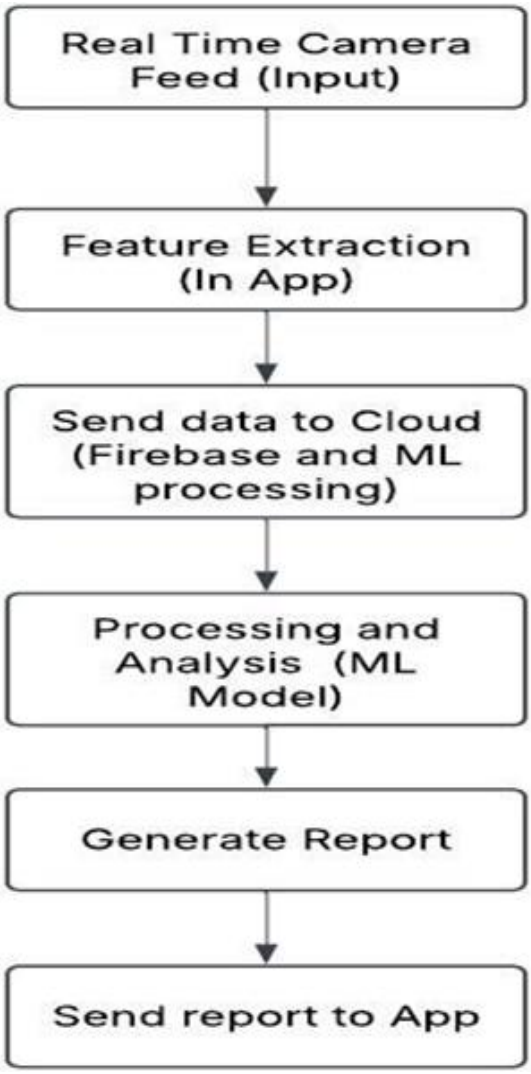

Fig 2 Flow Diagram

The system architecture combines video recording, face identification, feature extraction, and classification modules into a single real-time pipeline. A camera or video file serves as the input source for the face detection module. The detected faces are processed to extract expression features, which are subsequently fed into a trained neural network for deception classification. The output is displayed in real time, frequently superimposing results on the video feed. The system is developed in Python and includes libraries such as OpenCV, TensorFlow/Keras, and MediaPipe, ensuring modularity and ease of integration.

IV. RESULTS AND DISCUSSION

The experimental results are shown in this section along with and detection and deception analysis.

3.1 Convolutional Neural Network (CNN)

Based on performance of CNN model, a flowchart has been obtained based upon its performance on deception detection the below flowchart describes about it with Precision, Recall and F1 score.

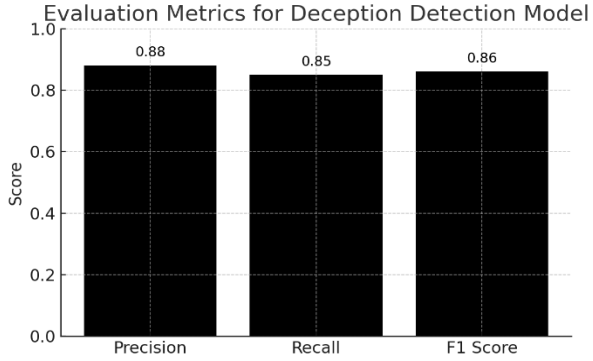

Fig 3 Bar Chart of Precision, Recall and F1-Score

3.2 Face Detection Models

During development, we explored two face detection models: MediaPipe Face Mesh and MTCNN. MediaPipe is a lightweight, quick solution appropriate for real-time applications, with high accuracy in facial landmark identification. MTCNN, on the other hand, is slightly slower but more reliable in detecting faces at different angles and lighting situations.

Performance metrics include:

- Training Accuracy: 87
- Validation Accuracy: 85
- Loss Trends: The training and validation loss curves stabilized after a few epochs, indicating that the model successfully converged.

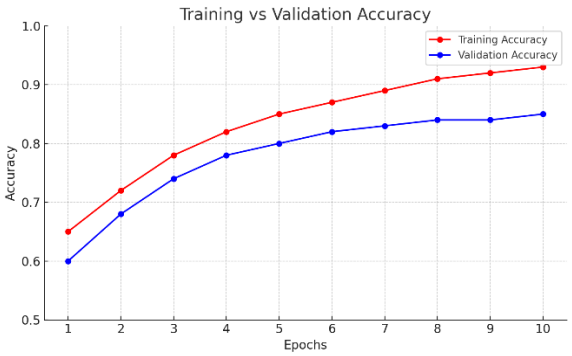

Fig 3 Training vs Validation

Challenges observed:
In a comparison analysis, our model performed similarly to cutting-edge emotion recognition systems, particularly in detecting stress-related expressions. However, unlike some advanced systems that incorporate voice and body language, this model is purely based on facial information. Limitations include difficulty handling extreme head positions, partial occlusions, and cultural differences in expressiveness.

3.3 Comparative Analysis and Limitations
Several obstacles arose throughout the creation of this system. It proved challenging to ensure consistent face detection in low-light conditions or with occlusion. Another issue was training the model on a balanced dataset, which had fewer samples of deceitful conduct than neutral expressions. Achieving real-time inference while maintaining accuracy necessitated considerable model optimization. Additionally, distinguishing between genuine and false stress proved difficult due to overlapping facial clues. Managing various facial traits across age, gender, and ethnicity necessitated thorough dataset curation. These problems demonstrate the complexities of human emotion perception.

4. CONCLUSION

This study effectively displays a real-time deception detection system based on facial expression analysis and deep learning. The results confirm the model's capacity to generalize across different inputs while remaining fast and accurate. Despite significant obstacles and limits, the work paves the path for more effective surveillance and interrogation technologies.

REFERENCES

[1] Avola, D, Cinque, L, Foresti, GL & Pannone, D 2019, ‘Automatic deception detection in RGB videos using facial action units’, Proceedings of the 13th International Conference on Distributed Smart Cameras, ACM, New York.
[2] Diabil, HAN 2022, ‘Experimental study to enhance the productivity of single-slope single-basin solar still’, Open Engineering, vol. 12, pp. 157–168.
[3] Alaskar, H, Sbaï, Z, Khan, W, Hussain, A & Alrawais, A 2022, ‘Intelligent techniques for deception detection: A survey and critical study’, Soft Computing, Springer, pp. 1–20.
[4] Avola, D, Cascio, M, Cinque, L, Fagioli, A & Foresti, GL 2021, ‘LieToMe: An ensemble approach for deception detection from facial cues’, International Journal of Neural Systems, vol. 31, no. 2, Article 2050068.
[5] Prome, SA 2020, ‘Deception detection using machine learning (ML) and deep learning (DL) techniques – A systematic review’, Informatics in Medicine Unlocked, Elsevier, vol. 6, Article 100057.