Lie Detection System using Audio and Video Analysis

Kshitija Randive[1] Honey Pede[2] Gauri Wyawahare[3] Mehul Shahu[4] Prof.Chandrayani Rokade[5]
[1,2,3,4,5]Department of Computer Science and Engineering
[1,2,3,4,5]Government College of Engineering Nagpur, India

Abstract— The objective of this paper is to design and set up an efficient lie detector that can detect lies in spoken sentences in an object language. This work proposes an efficient lie detection technique based on the non-invasive method that involves analysing the subject's speech utterances through audio and video files. The goal of lie detection is to discover truth that is known to one person and which is concealed from others. Another aim of this lie detection system is to be an optimised replacement of the traditional Polygraph tests that are currently used for lie detection. The system is based on the theory that lying produces particular emotions, audio and verbal cues which produce corresponding measurable physiological responses.

Keywords: Lie Detection System, Audio and Video Analysis

I. INTRODUCTION

Human beings are the only species capable of masking their behaviour. The tone and expressions of a person can express many feelings such as emotions, anger, sadness, etc. Different aspects of a person's social position including physical states, psychic states, physical psychic states, experiences are unintentionally conveyed through the speech and facial behaviour. Lying could be a defence reaction to avoid hassle with the law, bosses, or authority figures. The occurrence of deception is common in daily life. Hiding information or misrepresenting it sometimes leads to huge problems. The importance of lie detection is increasing day by day. Lie Detection System is a website made for detecting lies whenever a subject is being deceitful in real-time and with recorded video and audio. The main objective is to create a deception detection software which is portable, accurate, and easy to use and could possibly be a healthy replacement for the traditional polygraph methodology which is quite a complex process to carry, in other words this provides an improved way to analyse a subject's truthfulness.. Our system uses facial and voice extraction features to generate outputs by giving the instances of a particular behaviour or image portrayed by the person under examination. The system mainly focuses on felony detection by the vocal and expressions of the person and can have many wide scale applications like police investigations, court cases and many more. In other words this provides an improved way to analyse a subject's truthfulness.

Introducing an easy and simplified way to detect lies that could have wide scale applications and an increased public usage could be very helpful and could even be used in day to day dispute solving. . Also one of the benefits of our system is that one need not have any prior experience or knowledge about the working of the deception system i.e even a newbie is eligible and would be comfortable to use it rather than the traditional methods which require prior knowledge and specifics for its use.

II. RELATED WORK

Most of the deception researchers agree that lying involves processes or factors such as arousal and felt emotion (Zuckerman et al., 1981). Therefore, emotional facial expressions can be valid behavioural cues to deception. Meanwhile, there are involuntary aspects of emotional expression. As noted by Darwin, some actions of facial muscles were the most difficult to be voluntarily controlled and were the hardest to be inhibited (the so-called Inhibition Hypothesis (see also Ekman, 2003). When a strongly felt genuine emotion is present, the related facial expressions cannot be suppressed (Baker et al., 2016). Hurley and Frank (2011) provided evidence for Darwin's hypothesis and found that deceivers could not control some particular elements of their facial expression, such as eyebrow movements. The liars would feel fear, duping delight, disgust, or appear tense while lying, and would attempt to suppress these emotions by neutralising, masking, or simulating (Porter and Ten Brinke, 2008). However, the liars could not inhibit them completely and the felt emotion would be "leaked" out in the form of microexpressions, especially under high-stake situations (Ekman and Friesen, 1969).

Wright Whelan et al. (2014) considered a few cues that had successfully told liars and truth-tellers, including gaze aversion and head shakes. They combined the information from each cue to classify individual cases and achieved an accuracy rate as high as 78%. Meanwhile, Wright Whelan et al. (2015) found non-police and police observers could reach an accuracy of 68 and 72%, respectively, when required to detect deception in high-stake, real-life situations. Matsumoto and Hwang (2018) found that facial expressions of negative emotions that occurred for <0.40 and 0.50 s could differentiate truth-tellers and liars. These studies all suggested that leaked facial expressions could help human beings detect liars successfully.A meta-analysis by Bond and DePaulo (2006) showed that human observers only achieved a slightly-better-than-chance accuracy when detecting liars. Su and Levine (2016) found that emotional facial expressions (including microexpressions) could be effective cues for machine learning to detect high-stake lies, in which the accuracy was much higher than those reported in previous studies.Bartlett et al. (2014) showed that computer vision could differentiate deceptive pain facial signals from genuine pain facial signals at 85% accuracy. Barathi (2016) developed a system that detected a liar based on facial microexpressions, body language, and speech analysis. They found that the efficiency of the facial microexpression detector was 82%. Similarly, the automated deception detection system developed by Wu et al. (2018) showed that predictions of microexpressions could be used as features for deception detection, and the system obtained an area under the precision-recall curve (AUC) of 0.877 while using various classifiers.

III. IMPLEMENTATION

The system has simple working principles involved and does not require much of a hardware aid, a camera and a microphone does the work. It takes the audio and video file as an input and the pre-trained or pre-installed datasets are checked and compared for deceptions and the output is shown in percentage format with different emotions and behaviours.

A. Dataset collection

The datasets are collected from earlier projects done or by real time court cases or debates that would help easily distinguish between a truth and a lie. These Datasets contain audio files with both truth and lies audio files. The dataset initially has audio files with one defendant answering to the cross-examiner. This audio file will be later preprocessed by removing the voice of the cross examiner and removing the background noises in it. This is done using PyAudioAnalysis Library

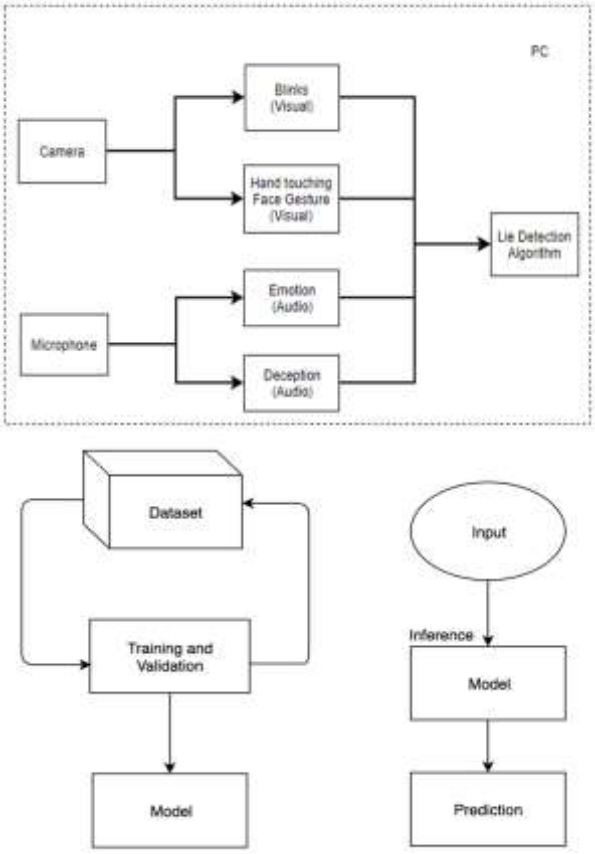
B. Model Training

For audio, we use PyAudioAnalysis to train data sets using the audio from real life court cases. This teaches the machine the differences between the two data sets (Truthful datasets and Untruthful datasets) and leads to the machine being able to differentiate between truths and lies. This is followed by using Paura2 to record audio in real-time and split the audio whenever there is a pause in the audio. These audio recordings are then used by PyAudioAnalysis to detect deception. Deception detection using audio is achieved for real-time as well as recorded audio. For video, we use OpenCV to detect the faces in recordings as well as real-time. From research, we found that blinking the eyes and touching the face too frequently was a sign of lying so we focused on these two parts. The two major components of visual deception detection were eye blink detection and hand tracking for the face. This is achieved using imutils, dlib, OpenCV, and tensorflow.

C. Feature Extraction

Feature extraction refers to extracting emotions of the defendant using his voice and facial behaviours. This is done using a module named Audio_Feature_Extraction present in PyAudioAnalysis Library. Facial region is detected and feature extraction is done to reduce the dimensionality of the input space. Principal Component Analysis is used to recognize statistical patterns for data reduction. These extracted features have a great role in distinguishing input patternsThe interrogation of the subject is recorded and the recorded speech is converted to text. The converted text is analysed for verbal cues that are exhibited by a liar and the emotional state of the subject is recognized using the dictionary approach. The process of extracting relevant features from data is critical to micro-expression recognition. To recognize micro-expressions accurately and effectively, features (both static and temporal) were extracted from both static and temporal micro-expression samples. These features were extracted using a holistic approach whereby spatio-temporal features are extracted from whole facial images instead of blocks of facial regions.

IV. LIE DETECTION ALGORITHM



V. RESULTS AND DISCUSSION

The proposed system would have an easy and simple to use interface that would be helpful for the not-so technology equipped people to use. Also one of the benefits of our system is that one need not have any prior experience or knowledge about the working of the deception system i.e. even a newbie is eligible and would be comfortable to use it rather than the traditional methods which require prior knowledge and specifics for its use. The primary purpose of this project is to present a simple and easy to use lie detection system that has a comparative result producing capacity as the traditional Polygraphic method but instead has much more wide scale usage due its operation and working. Our aim is to give the public and the system an approach that one can use to detect lies and felony in the person and this can have wide scale applications with many authorities and bodies under the usage of the same as the traditional polygraph method has many drawbacks ranging from requirement of specific skills for its use to a complex machine setup and equipments. Our system provides the counter to all these with a downright accuracy and most important wide range usage and application.

VI. CONCLUSION

This paper proposes an efficient technique to detect lies by using the method of audio analysis that involves recording the subject's speech utterances. The model in this research is an ensemble learning classifier which is constructed from a majority voting system, based on the outputs from the

Gradient Boosting Classifier (GBC).Using this classifier, the maximum accuracy achieved is 55.8%, which presents an improvement over the baseline chance accuracy of 50% and human accuracy of 48%. The results suggest that there is the possibility of detecting lies through the speech of the individual using audio analysis to the problems presented at the beginning of this paper. The main contribution of this research paper was to relate the stress level of the voice as an inducer to detect lies

REFERENCES

[1] A. Rosebrock, "Facial landmarks with dlib, OpenCV, and Python," PyImageSearch, 03-Apr-2017.
[2] C. DeAmicis, "How to build your design system in Figma - Figma Design - Medium," Figma Design, 30-Nov-2017.
[3] D. Pandey, "Eye Aspect Ratio(EAR) and Drowsiness detector using dlib," Analytics Vidhya, 21-Apr-2021.

