INTELLIGENT DECEPTION DETECTION FOR ONLINE INTERVIEWS

M.L.M. Fernando
209323R

Master of Science (MSc) in Computer Science
Specialized in Data Science

Department of Computer Science & Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

October 2022

INTELLIGENT DECEPTION DETECTION FOR ONLINE INTERVIEWS

M.L.M. Fernando
209323R

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Science (MSc) in Computer Science

Dept. of Computer Science & Engineering

Faculty of Engineering

University of Moratuwa
Sri Lanka

October 2022

DECLARATION

Unless otherwise noted, I hereby declare that the content of this dissertation is entirely original work by myself, and that no part of it has been previously submitted for a master's degree, a degree, or a diploma at another university or institution of higher learning without acknowledgement. The University of Moratuwa has the nonexclusive right to reproduce and distribute my dissertation in print, electronic, or any other medium without my permission. All or part of this content may be incorporated into future works of mine.

Student Name	**Registration Number**	**Signature**
M.L.M. Fernando	209323R	Lochana Fernando

Under my supervision, the aforementioned candidate is conducting research for his master's dissertation.

………………………………… ……………………………

Signature of the supervisor
(Dr. Chathura De Silva)

Date

ABSTRACT

When it comes to human communication, lying is a common practice. Recently, the detection of lies has become an important focus of judiciary, law enforcement, and security, interviews, etc.[1] Due COVID-19 the pandemic of interviews being conducted online; this is a main problem where a person may give false information specially in the visa applying process. Nonverbal behavior is constantly being transmitted by humans in opposition to spoken language. where visual and auditory cues like facial expressions, postures, gestures, and nonverbal vocal sounds can be used to detect deception intelligently. These human signals are known as deception indicators, and they are primarily associated with deceptive communication. The hiring of unskilled workers can eventually lead to a company's demise if an online interviewer exaggerates or fabricates his or her abilities.

ACKNOWLEDGMENT

This project Many people contributed to the success of this endeavor. Prior to anything else, my heartfelt gratitude goes to my project supervisor, Dr. Chathura De Silva for his support in helping me to drive the project towards the path of success by providing valuable feedback about my work and providing with the necessary guidance for the success of the project with his constant support and also, I gratefully acknowledge our Lecture in- charge, Head of the Research Groups Dr. Charith Chitraranjan and Head - Dept. of Computer Science & Engineering, Faculty of Engineering Prof. Indika Perera for his guidance through the course.

Thanks to all of my colleagues and friends for their assistance, support, interest, as well as valuable advice. Last but not least, I would like to thank everyone else who has helped in some way and encouraged me to make this a great success.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Term	Definition
EEG	Long-term electroencephalographic
IOT	Internet of Things
WHO	World Health Organization
ANN	Artificial Neural Networks
SVM	Support Vector Machines
CMC	Computer Mediated Communication

Table 1 - List of abbreviations

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Backgrounds

Using new cognitive psychology-based approaches, Deception Detection offers practical advice on spotting deception in the field.[2] Analyzes current issues in the field, such as counter interrogation, networks of lying, deception in cross-cultural, and the distinction between correct and incorrect intention. The use of evidence strategically, the imposition of cognitive load, response times, and covert lie detection can all be improved by new approaches based on cognitive psychology [3]. In accordance with the standard global classification of emotions, the visual function is classified into four main expressions such as Fear, Disgust, Contempt, Sadness, Surprise, Happiness, Anger and neutral [4]. when a person is trying to hide or suppress their emotions, they will use facial micro expressions and vocal expressions., which is a one technique for detecting lies. [5]. The manual calculation of micro-expressions and features is difficult, time consuming, and inaccurate.

The silent talker has an NVB conceptual model that assumes that different mental states are deceptive. The silent talker's main feature since it is a machine learning system is that the system takes some features from the interviewee as its input and determines whether the inputs are true or lies.[6][7] This system has a higher accuracy level since human interviewers may misunderstand the interviewee's actions. The accuracy is also achieved since the system uses complex interaction over time between many micro gesture channels to determine whether the behavior is truthful or deceptive.[8] One of the limitations of the current proposed solutions are identifying facial expressions only on pre-trained models where no error correcting algorithm is implemented to fine tune the system and lack of identifying vocal expression shows major drawbacks in proposed solutions [9][10].

1.1.1 Literature review on existing systems

This sub- from the literature review carried on deception detection in online interviews, found that the solutions for interviews have been addressed individually and lack of portability which makes unable full fill requirements.

1. A Comparison of Features for Automatic Deception Detection in Synchronous Computer-Mediated Communication [11]
    - The goal of this study is to compare the performance of automatic deception detection with and without structural features in synchronous (CMC). Furthermore, the findings suggest that structural features can be useful in detecting deception, and that combining structural and linguistic features can improve deception detection performance. According to their methodology, the use of structural features extracted from online social networks to detect deception in text-based synchronous communication is examined CMC (Structural Features and Content Features).

      Following machine learning classifiers were used Nave Bayes, NN, IBK, JRip, SVM, J48 and AdaBoost.

        i. Structural Features
            1. In-/Out- Degree Centrality
            2. Weighted In/Out Centrality
            3. Betweenness Centrality
            4. Closeness Centrality
            5. Clustering Coefficient

      Their findings show that structural characteristic features (80%) and combined features (92%) have higher accuracy rates, linguistic characteristic features have a low accuracy rate (51 ~ 73.7%).

2. A Portable Cost-efficient Lie Detector [12]
    - Recent advances in low-cost, small-scale EEG scanners have paved the way for a slew of new ideas in disease detection, social sciences, cyber security, and other fields. They look into the possibility of using EEG signals to detect deception or lies in this paper. MUSE, a new EEG scanner, is also being tested to see if it can provide the necessary information to detect potential lies. Studies have shown that the MUSE can provide the necessary components for a low-cost lie detector of moderate quality with a low degree of complexity. Integrating machine-learning techniques into the presented results would result in far more accurate detection, opening the door to a plethora of new possibilities.

      Volunteered participants' EEG waves were used as a detection tool to distinguish between Truth and Lie states. The Deception Test was broken down into three stages: Data collection, data filtering and processing, and analysis of the results Data acquisition is the first phase; data filtering and processing is the second phase.

3. Bag-of-Lies: A Multimodal Dataset for Deception Detection [13]
    - According to this publication present multimodal approach is presented which, contains data for identifying deception detection using a variety of different models such as video, audio, EEG, and gaze data. The dataset analyzes on cognitive aspect of deception by combining it with vision to investigate the cognitive aspect of deception. Using a realistic scenario, the dataset presented here was compiled from 35 unique subjects who collectively contributed with an even distribution of truth and deception throughout. On proposed dataset, the advantages of incorporating multiple modalities for fusion are also investigated. As a result, they believe that better deception detection algorithms will be developed as a result of this dataset's availability.

      EEG has been studied for deception detection and has shown promising results in a small pilot study. There is a new dataset called Bag-of-Lies. The people who participated are allowed to choose whether to be genuine or deceived in these situations. It combines video, audio, EEG, and eye gaze with other modalities. The EEG data was collected using a 14 channel Emotive EPOC EEG and gaze data was collected using a Gaze point GP3 Eye Tracker. In addition, for the experiment, 21 distinct, content-heavy, and descriptive images were collected. For analyzing the performance of different modalities, different datasets used. (tried it with video, audio, and EGG).
        i. Video
        ii. Audio
        iii. EEG
        iv. Gaze - During the recording.

      It has been found that using gaze data provides the best accuracy of 62 percent and 56 percent.

4. Deception detection using artificial neural network and support vector machine [14]
    - The subject is put to the test with a series of 15 questions, some of which are relevant and some of which are not. At the same time, the subject's physiological and linguistic characteristics are documented. The process of feature extraction is carried out in a solitary setting. In order to get the best results, the subject is instructed to keep their hands and heads still during the procedure. Standard wired sensors, are used to extract physiological features. Mice are used to extract

speech signals in a noise-free environment, and various algorithms are used to extract their features. Classification of Physiological and Speech Characteristics: Classification is done using artificial neural networks (ANNs) and support vector machines (SVMs).[15]

1. Physiological Features
2. Speech Features

There are two methods for classifying data: neural networks and support vector machines (SVM). For each subject, the feature extraction process generates a feature vector that is unique. Each subject's feature-level fusion is done by concatenating the features from the two modalities. A decision classifier is trained using the concatenated feature vectors to detect deceptive situations. Error correction validation and the multilayer perceptron neural network concept help achieve a better classification result.

5. Detection of Deception Using Facial Expressions Based on Different Classification Algorithms [16]
   - In this paper, In this work, the Facial Action Coding System (FACS) is used to extract facial features. The main idea of FACS is to use Action Units (AUs) to describe all facial actions. All of the AUs are linked to some facial muscle movement. Facial pattern vectors are made up of eight AUs that work together. These Data is collected for 43 people (20 males, 23 females). Most of them are between the ages of 18 and 25. Each of these algorithms is used in its own. MLP, VG-RAM, KNN, and SVM are some of the algorithms used. Individuals who classify things with the help of VGRAM and KNN achieve the best results. The main accomplishments of this work are the development of new DDS classification techniques, the creation of a real database that can be used to assess how well any DDS works with facial expressions, and the improvement of facial features.
     1. Video Capturing and Pre-processing
     2. Features Extraction
     3. Decision Maker

     Capturing video and pre-processing, extracting features, and making a decision There are two classes of extracted features in any DDS: the lie class and the true class. The FACS-based system uses these extracted features as AUs. As an individual, you can choose from four different classification methods for this task: VGARAM WNN.

     The accuracy rates for MLP and SVM are 83 percent and 84 percent, respectively, when applied to all datasets, while the accuracy rates for MLP and SVM are 86 percent and 89 percent, respectively, when applied to male datasets. Each of the results for Multilayer perceptron and Support Vector Machine was less than K-Nearest Neighbour and VG- RAM's.

6. Evaluation of Voice Stress Analysis Technology [17]
    - For use by the military and law enforcement. Marketed as computer-based and able to measure stress and deception in some systems, this technology has been made available commercially. These devices are described as less invasive, less constrained, and easier to use than standard polygraphs. According to the findings of this study, VSA technology can detect stress more accurately than chance alone, with results that are on par with those of today's polygraph systems. The technology, however, is still in its infancy and cannot be used in a court of law. They also discovered that more experienced and trained individuals were more accurate than those who were less trained. When combined with polygraph technology, this technology could be an effective interrogation tool.
        1. VSA-2000
        2. Computerized Voice Stress Analyzer
        3. Digital Voice Stress Analyzer
        4. TiPi 6.40
        5. Diogenes Lantern

      Polygraph testing is similar to the VSA's method of detecting deception. Pre-test interviews are a tool used by VSA examiners to get to know their subjects before the actual test. For the purpose of this process, the subject's behavior will be observed in order to develop test questions and break down any internalized barriers the subject may have to making admissions, and the test procedures will be established as credible (Pre-Test, Test, Post-Test).[18]

7. Lie Detection Using Image Processing [19]
    - Many human behaviors can be identified using the eye blink pattern. Image processing techniques can be used to detect lies and determine the truthfulness of an alibi, according to this paper. The mechanism for detecting lies is based on the experimentally proven hypothesis that liars are more cognitively demanding than truthtellers. As a result, after the lie is told, there is a significant drop in the number of eye blinks, followed by an increase in the number of eye blinks. The HAAR Cascade algorithm is used to detect blinks in MATLAB. The skin detection algorithm is applied to detected eye images to determine whether the eyes are open or closed.

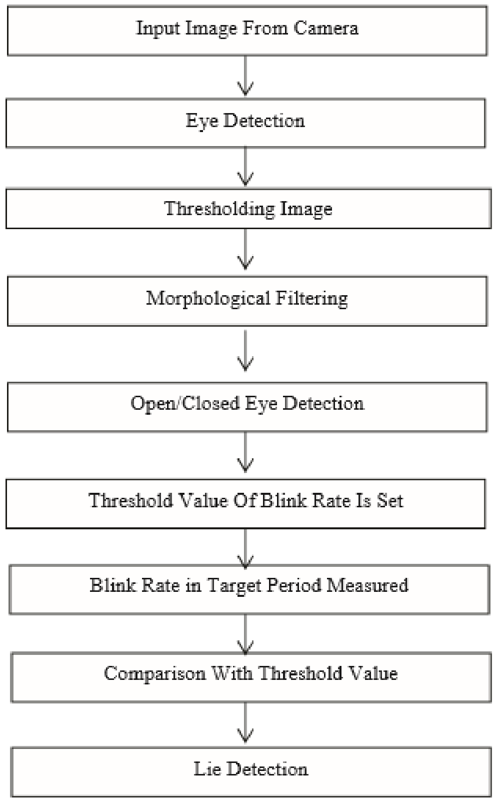


Figure. 1: Lie Detection Using Image Processing

The eyes are identified using Paul Viola's HAAR cascade method [4] by separating individual frame images from the video input. The eyes must be open or closed in order to proceed. YUV images of the detected eyes are used to obtain the histogram back projection, which is then used to calculate the results. Using the back projections, each pixel's probability of having skin present is calculated. If necessary, the outcome can be fine-tuned using morphological operations such as erosion, dilation, and so on.

8. Real Time Deception Detection for Criminal Investigation [20]
    - The MUSE 2 Headband was used to collect EG brain signals for analysis. In the humane cerebrum there are four parts the frontal, parietal, occipital, and temporal. The frontal lobe is responsible for vision, while the parietal lobe is responsible for hearing, Higher executive functions, such as emotional regulation, planning, reasoning, and problem solving, are typically located in the frontal lobe.
        i. Pre-Processing Noises and artifacts are common in the EEG signals, making them difficult to analyze. These artifacts must be reduced or eliminated prior to pre-processing the EEG, in order to improve the signal quality. Interference from electronic amplifiers, power lines, and other sources all contribute to the noise. Artifacts are signals that originate in

the non-cerebral part of the brain. Among the most common EEG artifacts are eye blinking, jogging of the eyelids, and movement of the head, shoulders, legs, and fingers. To obtain information pertinent to the five rhythms, the data will be subjected to a bandpass filter operating 0.5-50 Hz frequency range.

ii. Extraction of useful features in accordance with the scenario is the second step in feature extraction. Alpha, beta, and gamma frequencies of brain waves are all within the range of 8–12 Hz, but theta, 1–4 Hz, and 12–30 Hz are the most commonly observed (above25Hz)

iii. In order to detect lies and divert attention, the binary classification and the multiclass classification were used. [30].

iv. Evaluation Metrics – Accuracy, F-measures, Recall.

9. The Design and Development of a Lie Detection System using Facial Micro-Expressions [21]

    - In several fields, such as national security, government officals investigations, and counterterrorism, it is critical to detect lies. When people are trying to keep their emotions hidden or suppressed, they display facial micro-expressions, short, involuntary expressions that appear on their faces. Micro expressions are difficult to measure manually because they are so small. Facial Micro-Expressions are used in this study to develop a Lie Detection System. LabVIEW was used to create and implement an automated vision system. The interview is recorded using an EVS (Embedded Vision System). As a result, the video is broken down into a series of frames and processed in four sequential stages by a LabVIEW application. Color conversion and filtration are the focus of the first two steps. In the third stage, key characteristics of the facial structure are specified using geometric-based dynamic templates applied to each frame. Face micro-expression detection is done in the fourth stage by extracting the necessary measures. It has been proven that this system is capable of interpreting eight facial expressions. It generates precise results where can be applied to a variety of different types of research, including psychological testing. There is a hardware and software component to the proposed deception detection system that uses facial expressions. A high-speed camera is utilized to capture the face, which is subsequently segmented into various parts. A new dataset of facial micro-expressions is developed and manually marked as a ground value truth for evaluating this approach to facial expression recognition. 85% of the phrases in a database of five are correctly recognized using the template models that were developed.

10. Toward End-to-End Deception Detection in Videos [22]

When it comes to real-world scenarios, deception detection can be a game-changer in everything from commercial videos to airport screenings to courtroom proceedings to job interviews. As a result, there is a huge demand for video deception detection. Because of the inherent complexity of video and the lack of detective labels present in many real-world applications, traditional deception detection methods are severely hindered. We investigate video deception detection in this paper. Deceptive Videos can be detected automatically, even with limited training data, using a framework known as DEV, where it's based on a ethical approach to capturing information into a model. Experiments on real-world videos show that the proposed framework is effective.

Detecting deception in videos presents a number of challenges, including (a) the complexity of video data and the deception itself; (b) complex temporal dependencies in a video; and (c) typically limited labeled data. An automated feature extraction component, a visual component based on an attention mechanism for tackling the challenge (b), and finally a classification component based on a metric learning approach make up this model's three major components. Automated

i. Feature Extraction
ii. Visual Interpretability

Cross-validation is used to fine-tune the hyperparameters of the Random Forest classifier. On the test set, baseline methods and our model were run 10 times each to see how accurate they were. The detection accuracy results are shown in Table I. To begin, because the test set contains an equal number of deceptive and truthful videos, a random guess of 50% is made.

11. Visual Cues of Facial Behavior in Deception Detection [23]

Surveillance researchers are increasingly interested in the facial expressions of deception. Deception is studied by analyzing the facial behavior of deception. They give a brief overview of facial deception cues. Finally, run an experimental one-on-one interview in a controlled and experimental setting to see if you can establish rapport with your subjects. The experiment has been designed to include three stages in accordance with standard procedure. Stage one: a session of facilitation in which the starting points for each participant are established. During stage two, participants are asked to answer questions about two different topics, one of which must be true and the other fabricated. It's now time for participants to self-report on what they learned in stages two and three. It was then analyzed and compared with facial behavior cues identified in the literature after the data was collected. A non-parametric sign-test was used for statistical analysis. According to the findings, facial behavior can be used as a cue to predict deception. To wrap things up, we'll talk about how we can use subtler methods of facial action analysis in future work. Throughout the trial period, both during the introduction session and during the questions, participants' facial behavior was monitored. Facial expressions were also measured during both questions and answers because many initial responses were likely to be elicited while the Examiner was still talking. The majority of the questions were designed to elicit responses of 2 to 10 seconds. That's what we're hoping will be enough to accurately represent the full range of facial expressions during the question period. A FACS coder manually measured facial behavior. The FACS coder was kept in the dark about the coding conditions and the meaning of the cues in order to avoid bias. The Examiner was able to correctly identify 80 percent of truth tellers, but only 58.3 percent of liars based on verbal and nonverbal communication. It was easy for the Examiner to see things that an untrained eye might miss.

12. Verbal and Nonverbal Clues for Real-life Deception Detection [24]

They will be able to analyze both verbal and nonverbal behaviors in relation to deception with the help of this research multi-modal collection of real deception occurrences. A major challenge they faced in this task was the lack of available data, and most research so far has been based on acted or artificially collected data. " Thus, the generated deception models lack evidence from the real world. For the purpose of detecting deception, they investigate the use of multimodal real-world data. They build deception tools based on verbal and nonverbal features in a new dataset of real-life scenarios. Using a model that extracts and fuses features from the linguistic and visual modalities, they achieve classification accuracy of 77% to 82%. The results are superior to what humans are capable of in terms of spotting deception. Based on this, they decided to concentrate on linguistic and gestural elements.

i. Features of Verbal and Non-verbal Behaviors
    1. Verbal Features
        a. Unigrams
        b. Psycholinguistic Features.
        c. Syntactic Complexity.

2. Non-verbal Features
    a. Facial Displays
    b. Hand Gestures

They concluded that in order to develop a fully automated deception system, they plan to focus on the use of automated gesture and facial expression recognition and automated speech transcription.

13. Thermal Facial Analysis for Deception Detection [25]

Temperature readings from a person's face can be used to determine how stressed they are. Using thermal imaging, the authors investigate whether the periorbital region's thermal variations can provide a distinguishing signature for spotting deception. More questions per subject were scored, a person-centered approach to learning from data was emphasized, a framework for validating decision-making was proposed, and an accurate evaluation of generalization performance was evaluated in this paper. Classifying the thermal responses was done using a k-nearest neighbor classifier, which uses different strategies for data representation. Predicting the lie/truth responses with an 87 percent accuracy is possible thanks to fivefold cross validation. Modeling deception using a between-person approach does not generalize well across the training data, according to the results as well.

In addition to deception, other factors, such as facial expressions, body metabolism, changes in the underlying Musculo-thermal activities, thermal emissions from the surrounding environment, and illness, can affect skin surface temperature according to their conclusion Because of this, it is critical to account for the fact that each person's initial baseline temperature differs. Using the generic anxiety level as a measure of deception can produce high misclassification rates.

14. A Non-Contact Lie Detector using Radar Vital Signs Monitor (RVSM) Technology [26]

Subjects in traditional lie detector tests are tethered to various sensors. Where a huge number of people enter at a higher rate, necessitate the use of other methods, making this impractical. Security persons at checkpoints currently have to quickly determine whether a person is being deceptive and whether or not they need to be searched. Using a non-contact sensor, the RDD concept allows the checkpoint officer to gather physiological data and use it to make an informed decision. An unobtrusive and non-contact method of monitoring the body's physiological signals must be employed in order to detect deception. In the GHz frequency range, RVSM technology uses electromagnetic waves to detect human respiration and heartbeats from a distance and without physical contact. Long-term, non-thermal effects and RF safety have become more important concerns for society. Despite the lack of proof of non-thermal effects, the general public should be exposed to as little radiation as possible. The FCC regulations on radio frequency exposure must be met for any device that could potentially be used on the general population.

15. Facial Emotion Recognition in Continuous Video [27]

Video games, medicine, and affective computing all benefit from facial emotion recognition, which is the detection of emotional states from video of facial expressions. In spite of significant progress, a method that performs well on the nontrivial Audio/Visual Emotion Challenge 2011 data set has yet to be discovered. The majority of approaches still rely on single frame classification or temporally aggregated features to classify images. Emotional video should be classified based on the change in features rather than simply combining them, according to our research. They use histogram differencing and Gaussian derivatives to calculate the feature derivatives and a hidden Markov model to model the changes. When it comes to derivatives, we are the first to incorporate time information. The approach's effectiveness is evaluated using the challenging AVEC2011 dataset, which shows an increase in classification rates of up to 13%. Detection,
Alignment and Features

i. Modeling Temporal Changes
ii. Local Derivatives with DoG
iii. Global Derivatives with HD
iv. Observation Quantization

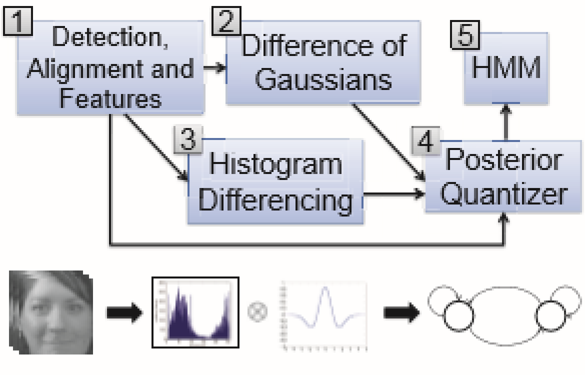


Figure. 2: Facial Emotion Recognition in Continuous

1.1.2 Proposed system

The procedure which is going to be followed can be divided into two stages. The first stage to identify the facial expressions, eye tracking whether a person takes eyes out of camera, question to answer response rate, and voice stress level are identified through the sequence of video frames(Calculate the amount of transformation needed to tell which emotions are strong and weakest in terms of their geometric shift from neutral) and output as quantifiable data for each section. An LSTM network efficiently propagates information across a sequence of inputs and is capable of detecting temporal correlations in sequential data. Emotional intensity will be quantified using a metric of deviation from the neutral face, which will be used to create synthetic face stimuli.

The second part using Deep Recurrent Convolutional Neural Network will be created by the currently available identified data from the dataset and gives the grading to quantifiable output from the first stage to which extent the person is lying. At the end of the interview the interviewer's judgement will be considered as feedback to the system which acts as an error correcting algorithm where if the results differ the result will be considered to fine tune the model. From this approach the model will adapt according to the questions that were asked based on the area the system is working on (ex - immigration domain, security domain).

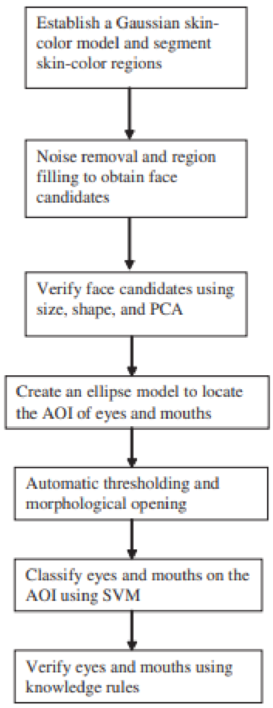
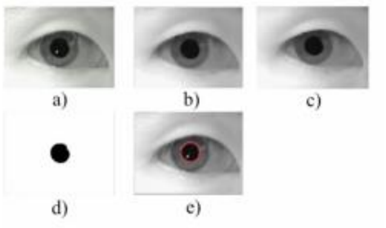
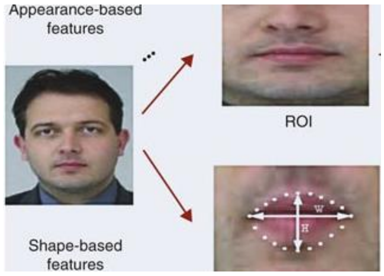


Figure. 3: Feature Identification

In order to achieve above methodology extract features from several feature detection methods/algorithms to detecting facial features. It is necessary to first segment a color image into skin and non-skin regions using a gaussian skin-color model and mathematical morphology and region filling techniques for noise reduction and hole filling before using a pulse coupled neural network (PCNN) for facial feature detection. The PCNN-generated 3D binary map series (BMS) effectively describes image feature information such as edges and regional distribution, or by creating an ellipse model to locate eyes and mouths and applying the support vector machine in the general area (SVM) facial features can be detected using a similar method. [32].





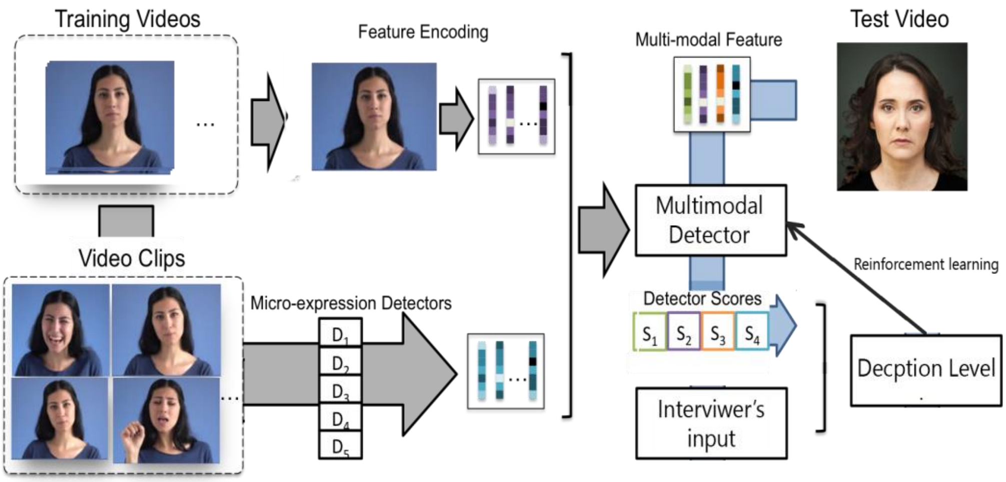


Figure. 4: Methodology

After detecting facial features these identified characteristic converted in to facial actions. (Lip contour, eye corners, chin lowest point, eye pupil dilation, eye blinking) then these feature variations will to classify them after overtime collection features and the feature variation will converted into high level deceptive actions. Then these identified deceptive actions will cross check with interviewers' feedback and give output as percentage of the deception level.

1.2 Research Gap

Since deception detection in online interviews has been a problem for a long time there are systems with advanced technology in the market but they are not focused on image processing rather they based fully on machine based. Most of these solutions are using a machine sensor based approach to get the information and the detection is occurs onsite.

Table 2: Research gap

	Computerized Polygraph LX5000-S	Computerized Polygraph LX6-S	Other	Proposed Solution
Blink count calculation	✗	✗	✗	✓
Face direction	✗	✗	✗	✓
Emotion's identification	✓	✗	✓	✓
Machine based	✓	✓	✓	✗

1.3 Research Objectives

1.3.1 Main Objective

There several feature detection methods/algorithms to detecting facial features (principal component analysis and independent component analysis, linear discriminate analysis, Neural Networks in face recognition, Gabor wavelet-based solutions, 3D-based face recognition). After detecting facial features these identified characteristic converted in to facial actions. (Lip contour, eye corners, chin lowest point, eye pupil dilation) then these feature variations will to classify them after overtime collection features and the feature variation will converted into high level deceptive actions. Then these identified deceptive actions will cross check with interviewers' feedback and give output as percentage of the deception level.

- Identify the facial expressions & vocal expressions
- Converting identified expressions into quantifiable output.
- Training dataset to predict which percentage the input data is true.
- Creating the system to get the interviewers response as judgement as feedback to the system (acting as error correcting algorithm which will fine tune the model

1. METHODOLOGY

2.1 Methodology

During literature review to study various approaches to get a solution came across. After going through those stuff it has been figured out that to get the facial attributes in order to perform desirous actions following steps will be involved:

- Face Detection
- Facial Emotion Classification
- Face Angle Calculations
- Calculate Blink Count
- Analyzing all Stats
- Decision Making

Face detection is the main step in order to reach out the solution. After getting the face from the image the various techniques of images processing are supposed to be applied. After that feature matrix is made. Then on the base of the feature matrix the step of classification of truthiness in the sentence will be calculated.

2.1.1 System Overview

Face Detection

The first page in the development of the solution is accurately detecting the face region. During the literature review several face detection methods have been studied. The details of all methods to detect the face have been presented below.

- **Haar Cascade Classifier**

The most effective way to detect faces in the images is Haar cascade classifiers. There are three main key steps (PAUL VIOLA, 2003):

- The first step is introducing new image representation, which allows features to compute quickly, which are used by our face detector.
- The second step is simple classifier which uses AdaBoost algorithm to select small numbers of visual features.
- The third step is combining classifier in a cascade which focuses on face like regions by discarding background.

That method has been found out by Viola and Jones. That method is also included in computer vision library named OPENCV.

- **Face Detection Using Oriented Matching**

Real time face detection on the gray scale images can be performed using edge-oriented information. Edge oriented approach is powerful tool to detect face like objects in the image. That technique is performed by template matching and object modeling in order to obtain a face model through edge orientation.

- **Artificial Neural Network**

Artificial Neural Networks are statistical learning models which have amazing abilities to get useful information from raw and indefinite data. Therefore, the neural network can be used to figure out various trends from input data, which are quite complicated for humans. Artificial neural networks can also be used to detect the faces in the image efficiently**.**

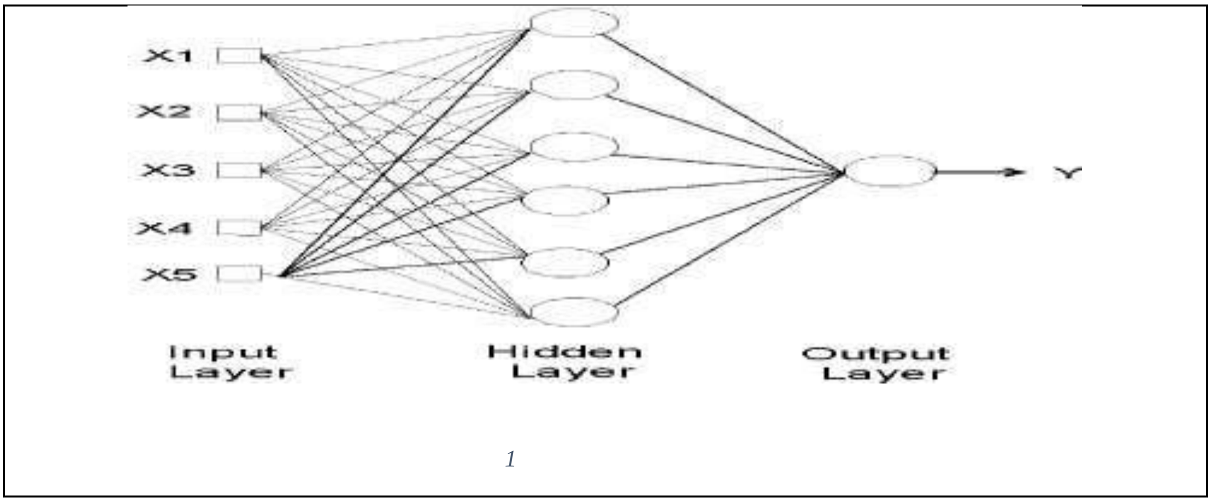


Figure. 5: Neural Network

A well-connected neural network examines the image and decides that whether there is face or not. A bootstrap algorithm can be used to make false detections in order to eliminate non face areas in the image.

Apart from above mentioned techniques to find out faces in the image, there are some efficient libraries are also available to detect faces from the image. Some of the wildly used libraries and API are following:

- **OpenCV**

OPENCV is one of the most efficient libraries of computer visions which contains more than 2500 algorithms. OPENCV is open-source library that was developed to provide common infrastructure to computer vision applications

- **OpenFace**

OPENFACE is also an open-source library for face recognition and detection. OPENFACE library is based on the deep neural networks, which intensify its performance amazingly.

- **FaceMark**

In terms of facial feature detection, FaceMark is a powerful API. There are 68 points in front and 35 in profile for a frontal face. It is widely used in the various applications of computer vision.

- **Face++**

FACE++ is very strong API to detect facial landmarks, attributes, gender, and age of the Person.

- **Opencv Caffe Model**

Opencv also provided the CNN based face detection model. The Caffe based CNN model is a deep learning-based model, which is the state of the art face detection model. That model can also perform face detection in real time.

Each Caffe model mandates the shape of the input image, usually also image preprocessing is needed like mean subtraction to eliminate the effect of illumination changes. In a typical case the mean values will be calculated from the training dataset images (Those are the mean RGB values across all images in the ImageNet dataset.)

In the implementation of that module the Caffe Based CNN model has been used. The reason for using that model is its highest accuracy.

Facial Emotions Classification

After detecting the face region, it is one of the most important steps to extract features to determine various characteristics of facial attributes. Then based on those features a deep learning-based model has been trained to determine the facial emotions.

For example, CNNs can be used for image segmentation, recognition, and classification. They are one of the most common and used deep learning algorithms. Deep learning algorithms are one of the most common ways to figure out how people feel (e.g., happy, surprise, and neutral). We had made an algorithm that used deep learning called Mini-Xception. The main goal of the proposed model is to be able to automatically detect emotions and predict how people will feel with high accuracy. In this method, we used facial expression images from the FER dataset to look at the

results of our experiments. We also used some open-source images from Google Images to train the model. These images are fed into the model, which then uses them to refine its behavior. Afterwards, the proposed model identifies which facial expression is being displayed. In that model a data classification approach has been followed to classify the facial expression on the detected face.
In the dataset there were 2800 images of each expression. In the training of the model 80% of the data has used in the training and 20% of the data has been used in the testing of the model.

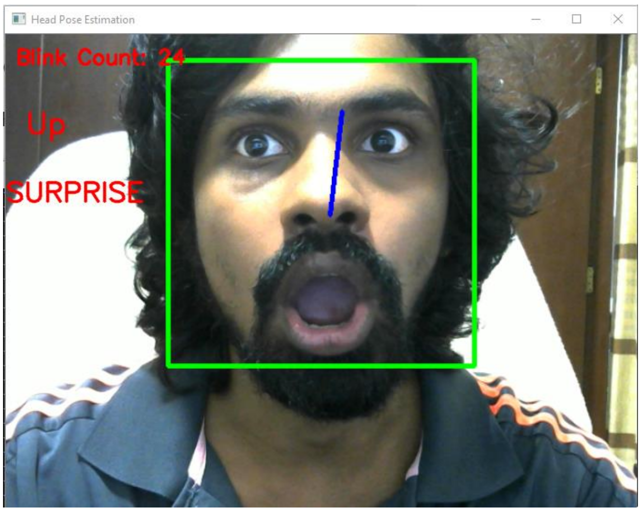


Figure. 6: Surprise emotion

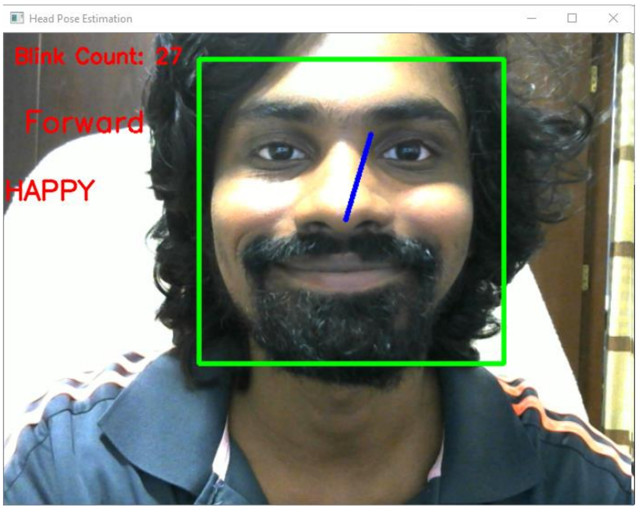


Figure. 7: Happy Emotion

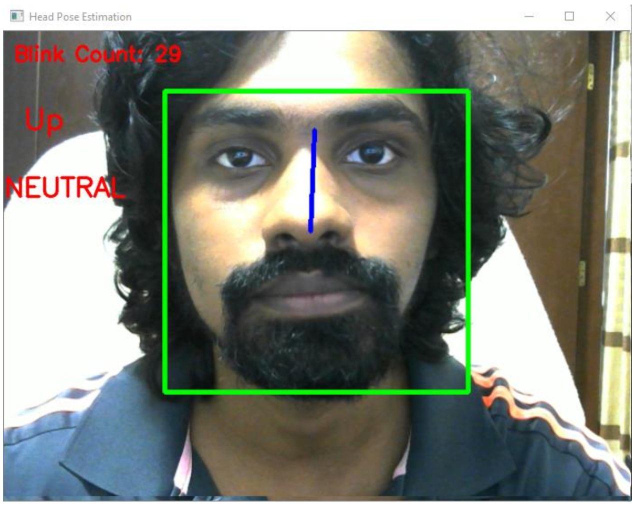


Figure. 8: Neutral Emotion

Figure. 9: Emotions Trained

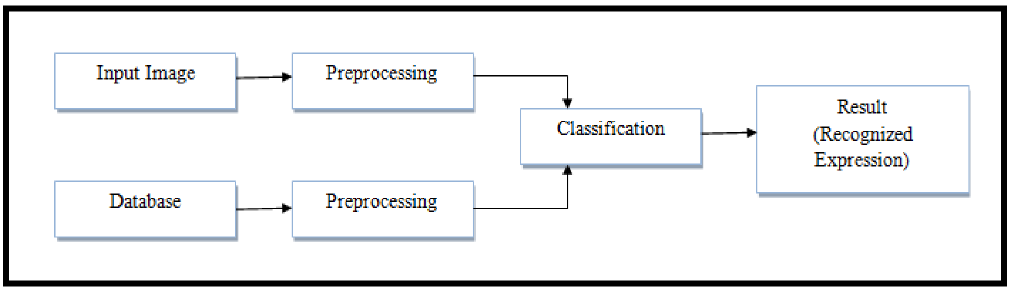


Figure. 10: System flow of the facial expression recognition system

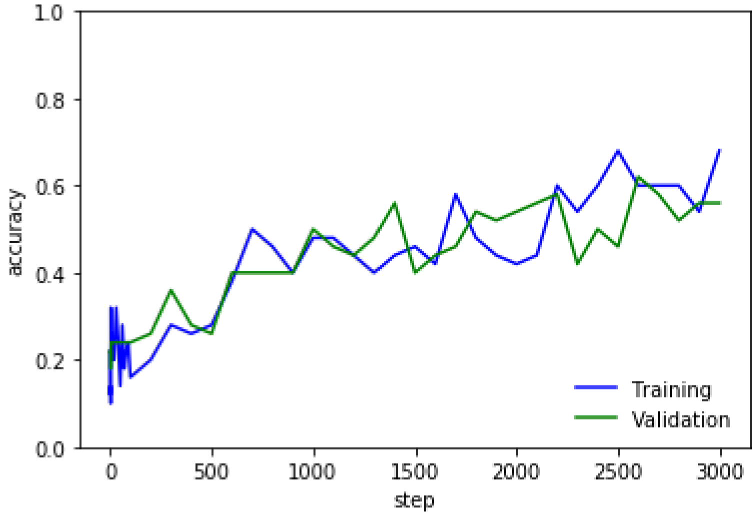


Figure. 11: Result Analysis Visualization

Even tough model was trained for 7 emotions , the accuracy with higher level was only found for surprise , happy and neutral emotions. Therefore other emotions were not taken for consideration when predicting the final out come off the system.

Table 3: Accuracy for emotions

Emotion	Accuracy
Surprise	89%
Sad	63%
Disgust	45%
Fear	32%
Anger	43%
Happy	85%
Neutral	88%

Face Angle Calculations

In calculating the facial angle that either person is looking at left, right or center, it is mandatory to detect the facial landmarks of the face.
For that purpose, there were two solutions under considerations, which are following:

- Dlib facial landmarks
- Mediapipe FaceMesh

The dlib facial landmarks solution gives 68 facial landmarks. The Mediapipe gives 468 facial landmarks. We have the option to use both of them but the dlib based facial landmark has some accuracy issues, as it works only on the frontal face but we are supposed to detect the left, right and front face direction as well.
Due to that constraint the Mediapipe solution has been used which gives 468 facial landmarks. It is one the most accurate and best solutions which has been offered by

Google. The pictorial view of the facial landmark detected by Mediapipe can be seen below.

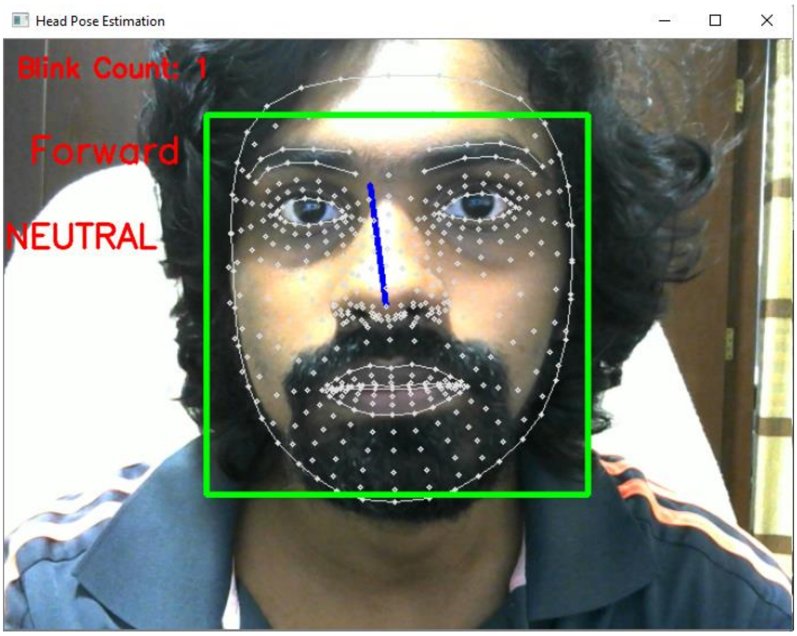


Figure. 12: Facial Landmarks

After detecting the facial landmarks, it is mandatory to make a pivot point from where we can determine the projection angle.
For that purpose, the nose has been chosen to calculate the angle of projection to determine the direction of the face in real time. The pictorial view can be seen below:

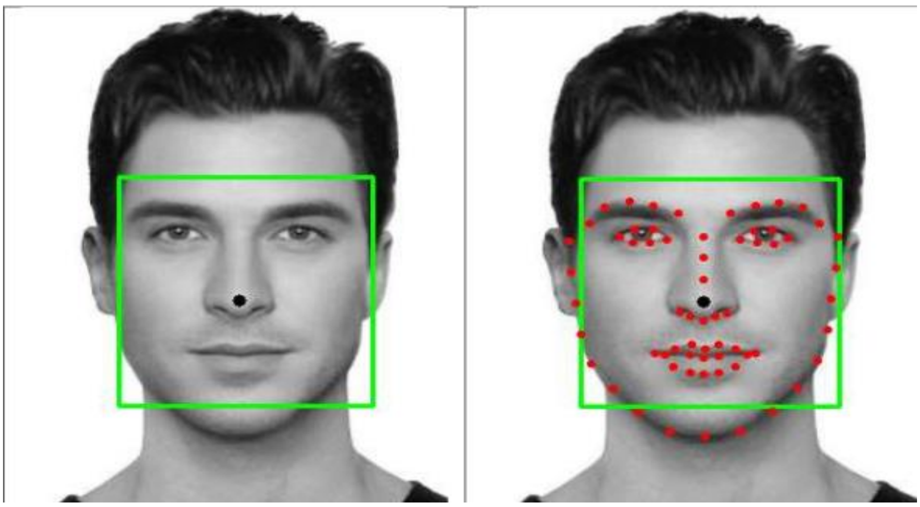


Figure. 13: Facial Landmarks 2

By deciding the nose as pivot point the direction of the face has been calculated. The pictorial view of angle calculation can be seen below:

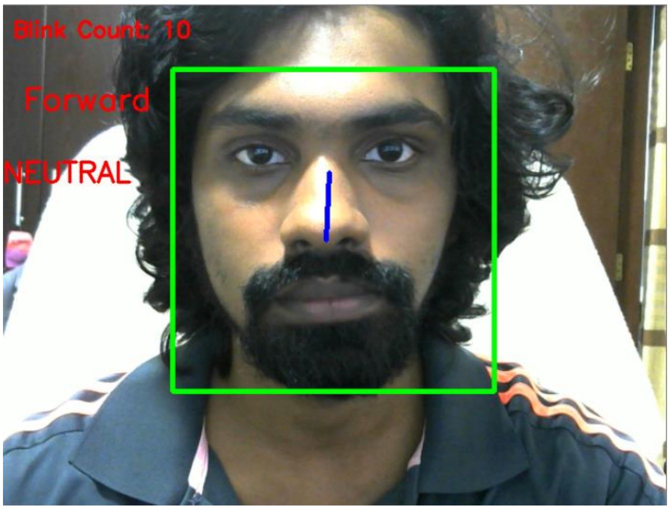


Figure. 14: Forward Direction

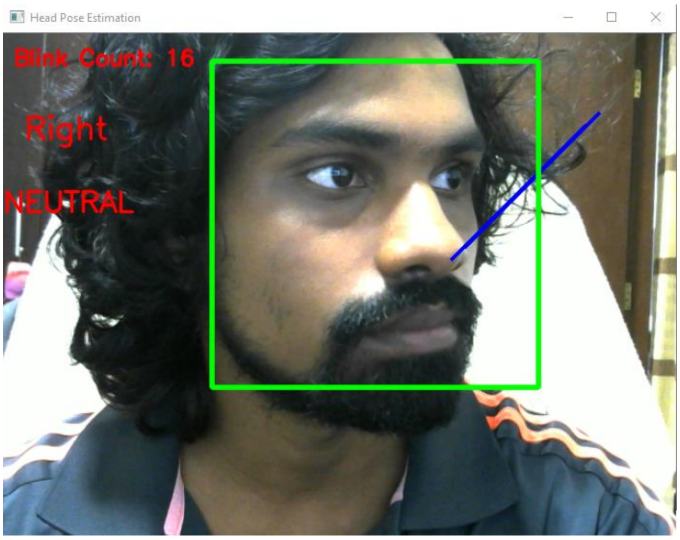


Figure. 15: Right Direction

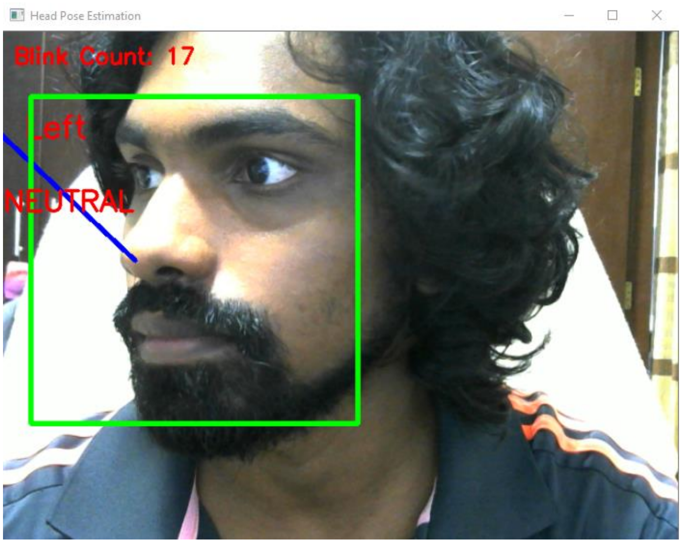


Figure. 16: Left Direction

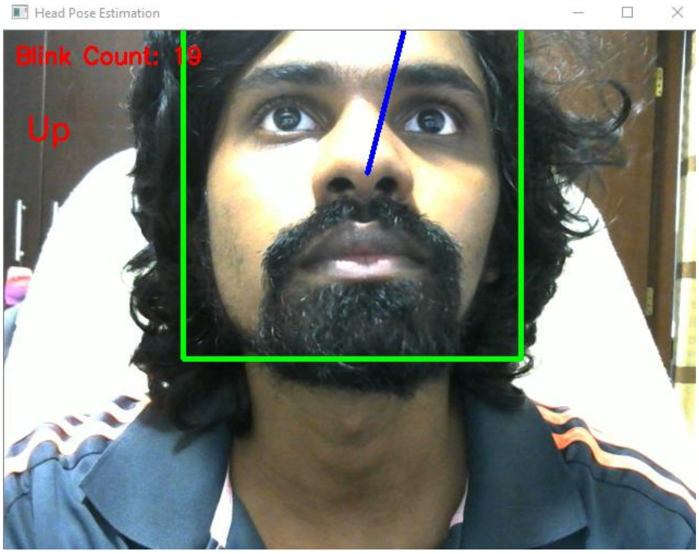


Figure. 17 : Up Direction

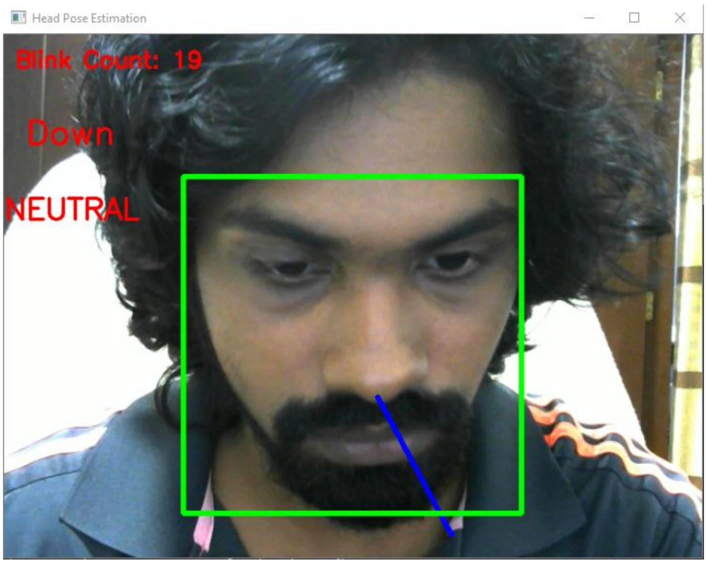


Figure. 18: Down Direction

In that way it can be easily estimated that either person is looking at right or left or at front from an angle of projection on the face.

Calculate Blink Count

The last parameter in determining that either person is telling a lie or truth is counting the blink count of the person. It is common practice that if the person is confused about something or wants to manipulate something the count of his/her blink count increases. In that implementation we covered the blink count from facial landmark detection of the eyes.
The pictorial view of approach of calculating the blink count can be seen below:

Figure. 19: Facial Landmarks 3

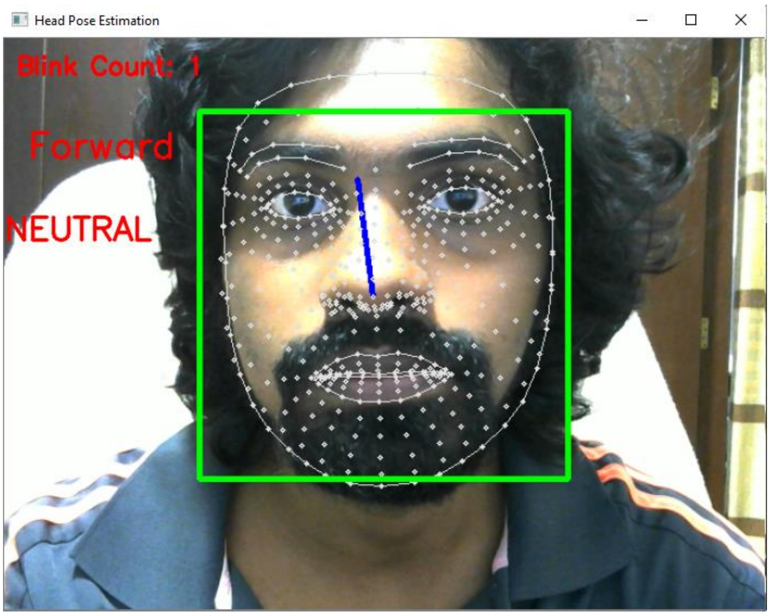


Figure. 20: Facial Landmarks 4

After detecting the landmarks of the eyes, it has been calculated how the euclidean distance between the upper and lower parts of the eyes change. Then based on that euclidean distance the status of blink has been determined. The pictorial view of calculating that distance can be seen blow:

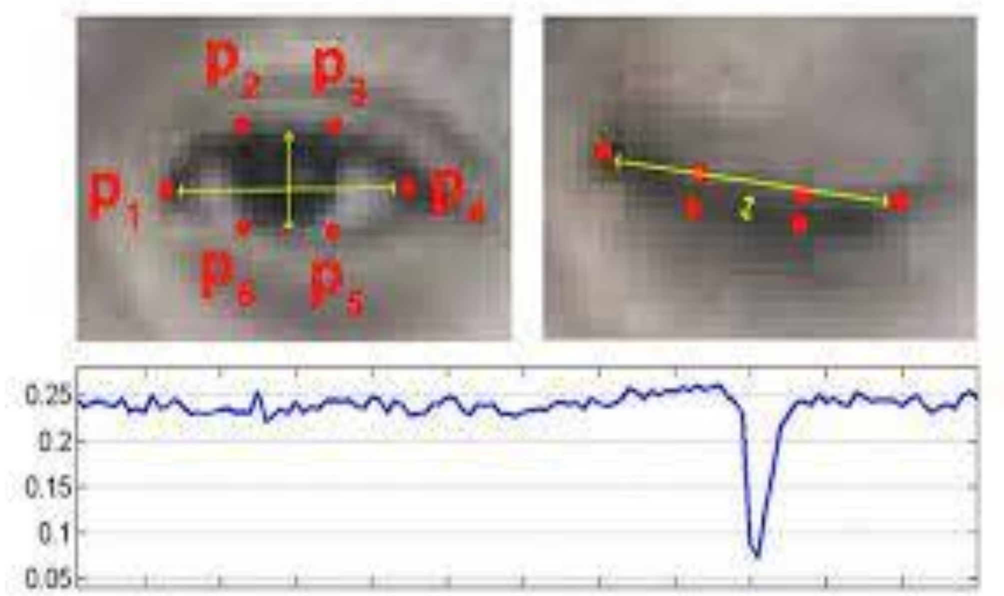


Figure. 21: Eye Blink

In that way the blink count from the detected face of the person has been calculated. In that way, all parameters of facial emotion, face angle and blink count have been analyzed to determine the speech status in order to determine that percentage of truthiness has been calculated from the speech signals using facial features.

Calculate the Results and Making Decision

In the current solution I am making the threshold of 5 seconds. After each 5 seconds the average of each attribute is calculated. So, after each 5 seconds, we will be getting the average emotion (happy, surprise and neutral), average face direction (left, right, up and down) and average of blink count in the time frame of 5 seconds. Based on these stats the case base reasoning strategy has been utilized to get the percentage of the truthiness and the percentage of the deception in the speech sentence has been calculated. Because we have not been provided with a truth table to classify the sentence so based on case base reasoning, we made assumptions to find the truthiness in the sentence.

2.1.2 Testing

Software testing is an investigation that is done to give people who are interested in the quality of the output product or service being tested more information. Software testing can also give a business an objective, independent look at the software so that they can see and understand the risks of putting new software in place. There are many different ways to test things, but they don't have to be limited to running a program or application to look for software bugs.

Table 4: Test cases

Test Case ID	Test case Description	Test Procedure	Test Inputs	Expected Outputs	Actual Outputs and decision	Test Result
1	Identifying the face	Focus the camera on a face(person)	Camera feed	Person's face detected with a percentage (92%)	person face detected with a percentage (92%)	Pass
2	Identifying blink count	Focus the camera to person	Mark the frame with detected eyes and count of blink	Blink count for the particular period	Variation with actual blink counts and identified	Pass
3	Identifying emotions	Focus the camera to person	Identify the emotions of the person	Output correct emotions (89%)	Output correct emotions (89%)	Pass
4	Calculate the percentage of lying	Get the results from identifications	Identified values	Output percentage of deception	Output percentage of deception	Pass

2.1.3 Technologies Used

Software interfaces

- OpenCV
- Pyttsx
- Tensor Flow
- Anaconda Navigator

Communication interfaces

- Camera Adapter

2. RESULTS AND DISCUSSIONS

3.1 Results and discussions

3.1.1 Results

The result of this research project is a identify deception in online interviews. This solution will be created to overcome the problems that have identified in the literature review phase.

The result of testing the system with interview

Table 5: Results for head position identification

N = 15	Predicted Yes	Predicted No	
Actual Yes	9	1	10
Actual No	1	4	5
	10	5	15

We test our suggested solution for different emotions. Below chart is the table of results that we got.

Table 6: Result for emotion identification

N = 15	Predicted Yes	Predicted No	
Actual Yes	7	3	10
Actual No	3	2	5
	10	5	15

```
o enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
he Average Expression of 5 sec is NEUTRAL
he Average direction of 5 sec is Forward
he Average blink count of 5 sec is 1
he approximate percentage of Lying is 29 %
he approximate percentage of Truth is 76 %
he Average Expression of 5 sec is NEUTRAL
he Average direction of 5 sec is Forward
he Average blink count of 5 sec is 0
he approximate percentage of Lying is 29 %
he approximate percentage of Truth is 71 %
he Average Expression of 5 sec is NEUTRAL
he Average direction of 5 sec is Right
he Average blink count of 5 sec is 0
he approximate percentage of Lying is 47 %
he approximate percentage of Truth is 40 %
he Average Expression of 5 sec is NEUTRAL
he Average direction of 5 sec is Right
he Average blink count of 5 sec is 0
he approximate percentage of Lying is 48 %
he approximate percentage of Truth is 50 %
```

Figure. 22: Results

3.2 Discussion

This section is mainly focused on the problems faced during the project design and implementation and how to overcome issues and how we solved them. Furthermore, it describes how the successful solutions implemented are gained.

I have allocated more time for analysis since the entire project depends on understanding the research problem. To find information about this problem I have went through articles, research papers and got the information. According to that literature survey, I have identified that there were no similar products available in but there were some products which had some functionality but they were sensor-based solutions and not directly address the problems and some systems just solved one scenario. In order to provide a solution, I have divided the solution's overall content into several functionalities. As a result, the architectural diagram was created and the technical aspects that needed to be considered were discussed. Once understood the problem and designed a solution to come up with the optimal solution through the knowledge gathered.

4. FUTURE WORK

- Lie detection through speech
- Creating the system to get the interviewer's response as judgment as feedback to the system.
- Model to give interviewer feedback and compare the output of the system

5. CONCLUSION

In this research I have focused on detecting deception on online interviews. This research divided into four main parts identifying blink count, identifying emotions, identifying face direction and end results the experimental results show that method achieves accurate and robust detection, as can be seen in the previous result section. Even though the system shows close results the accuracy is less than having results with connected to a machine.

In Further work, addition of deception detection by analyzing video to the machine output can get us higher accuracy than the implemented solutions.

6. REFERENCES

[1] I. Lakshan, L. Wickramasinghe, S. Disala, S. Chandrasegar and P. S. Haddela, "Real Time Deception Detection for Criminal Investigation," 2019 National Information Technology Conference (NITC), Colombo, Sri Lanka, 2019, pp. 90-96, doi: 10.1109/NITC48475.2019.9114422.

[2] Martinez, Aleix & Du, Shichuan. (2012). A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives. Journal of machine learning research : JMLR. 13. 1589-1608.

[3] Blandon-Gitlin, Iris & Fenn, Elise & Masip, Jaume & Yoo, Aspen. (2014). Cognitive-load approaches to detect deception: Searching for cognitive mechanisms. Trends in Cognitive Sciences. 18. 441–444. 10.1016/j.tics.2014.05.004.

[4] Li, Ming & Xu, Hao & Huang, Xingchang & Song, Zhanmei & Liu, Xiaolin & Li, Xin. (2018). Facial Expression Recognition with Identity and Emotion Joint Learning. IEEE Transactions on Affective Computing. PP. 1-1. 10.1109/TAFFC.2018.2880201.

[5] H. Karimi, J. Tang and Y. Li, "Toward End-to-End Deception Detection in Videos," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 1278-1283, doi: 10.1109/BigData.2018.8621909.

[6] M. Owayjan, A. Kashour, N. Al Haddad, M. Fadel and G. Al Souki, "The design and development of a Lie Detection System using facial micro-expressions," 2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Beirut, 2012, pp. 33-38, doi: 10.1109/ICTEA.2012.6462897.

[7] N. Baghel, D. Singh, M. K. Dutta, R. Burget and V. Myska, "Truth Identification from EEG Signal by using Convolution neural network: Lie Detection," 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), Milan, Italy, 2020, pp. 550-553, doi: 10.1109/TSP49548.2020.9163497.

[8] D. Barsever, S. Singh and E. Neftci, "Building a Better Lie Detector with BERT: The Difference Between Truth and Lies," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, United Kingdom, 2020, pp. 1-7, doi: 10.1109/IJCNN48605.2020.9206937.

[9] N. Srivastava and S. Dubey, "Deception detection using artificial neural network and support vector machine," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2018, pp. 1205-1208, doi: 10.1109/ICECA.2018.8474706.

[10] Z. Labibah, M. Nasrun and C. Setianingsih, "Lie Detector With The Analysis Of The Change Of Diameter Pupil and The Eye Movement Use Method Gabor Wavelet Transform and Decision Tree," 2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), Bali, 2018, pp. 214-220, doi: 10.1109/IOTAIS.2018.8600918.

[11] J. Pak and L. Zhou, "A comparison of features for automatic deception detection in synchronous computer-mediated communication," 2015 IEEE International Conference on Intelligence and Security Informatics (ISI), Baltimore, MD, USA, 2015, pp. 141-143, doi: 10.1109/ISI.2015.7165955.

[12] E. Youssef, H. T. Ouda and M. Azab, "MUSE: A Portable Cost-efficient Lie Detector," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2018, pp. 242-246, doi: 10.1109/IEMCON.2018.8614795.

[13] Gupta, Viresh & Agarwal, Mohit & Arora, Manik & Chakraborty, Tanmoy & Singh, Richa & Vatsa, Mayank. (2019). Bag-of-Lies: A Multimodal Dataset for Deception Detection. 83-90. 10.1109/CVPRW.2019.00016.

[14] Zhiyu, Wang. (2010). Based on physiology parameters to design lie detector. ICCASM 2010 - 2010 International Conference on Computer Application and System Modeling, Proceedings. 8. 10.1109/ICCASM.2010.5619088.

[15] N. Srivastava and S. Dubey, "Deception detection using artificial neural network and support vector machine," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2018, pp. 1205-1208, doi: 10.1109/ICECA.2018.8474706.

[16] H. H. Thannoon, W. H. Ali and I. A. Hashim, "Detection of Deception Using Facial Expressions Based on Different Classification Algorithms," 2018 Third Scientific Conference of Electrical Engineering (SCEE), Baghdad, Iraq, 2018, pp. 51-56, doi: 10.1109/SCEE.2018.8684170.

[17] Hopkins, C.S. & Ratley, R.J. & Benincasa, D.S. & Grieco, J.J.. (2005). Evaluation of Voice Stress Analysis Technology. 20b- 20b. 10.1109/HICSS.2005.254.

[18] J. O’Shea, K. Crockett, W. Khan, P. Kindynis, A. Antoniades and G. Boultadakis, "Intelligent Deception Detection through Machine Based Interviewing," 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018, pp. 1-8, doi: 10.1109/IJCNN.2018.8489392.

[19] Singh, Birender & Rajiv, Pooshkar & Chandra, Mahesh. (2015). Lie detection using image processing. 1-5. 10.1109/ICACCS.2015.7324092.

[20] Lakshan, L. Wickramasinghe, S. Disala, S. Chandrasegar and P. S. Haddela, "Real Time Deception Detection for Criminal Investigation," 2019 National Information Technology Conference (NITC), Colombo, Sri Lanka, 2019, pp. 90-96, doi: 10.1109/NITC48475.2019.9114422.

[21] M. Owayjan, A. Kashour, N. Al Haddad, M. Fadel and G. Al Souki, "The design and development of a Lie Detection System using facial micro-expressions," 2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Beirut, Lebanon, 2012, pp. 33-38, doi: 10.1109/ICTEA.2012.6462897.

[22] Karimi, Hamid & Tang, Jiliang & Li, Yanen. (2018). Toward End-to-End Deception Detection in Videos. 1278-1283. 10.1109/BigData.2018.8621909.

[23] Yap, Moi Hoon & Rajoub, Bashar & Ugail, Hassan & Zwiggelaar, Reyer. (2011). Visual Cues for Facial Behaviour in Deception Detection. ICCAIE 2011 - 2011 IEEE Conference on Computer Applications and Industrial Electronics. 10.1109/ICCAIE.2011.6162148.

[24] Pérez-Rosas, Verónica & Abouelenien, Mohamed & Mihalcea, Rada & Xiao, Yao & Linton, CJ & Burzo, Mihai. (2015). Verbal and Nonverbal Clues for Real-life Deception Detection. 2336-2346. 10.18653/v1/D15-1281.

[25] Rajoub, Bashar & Zwiggelaar, Reyer. (2014). Thermal Facial Analysis for Deception Detection. Information Forensics and Security, IEEE Transactions on. 9. 1015-1023. 10.1109/TIFS.2014.2317309.

[26] J. Geisheimer and E. F. Greneker, "A non-contact lie detector using radar vital signs monitor (RVSM) technology," in IEEE Aerospace and Electronic Systems Magazine, vol. 16, no. 8, pp. 10-14, Aug. 2001, doi: 10.1109/62.942213.

[27] G. Sapijaszko, T. Alobaidi and W. B. Mikhael, "Adaptive Feature Extraction Algorithm using Mixed Transforms for Facial Recognition," 2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS), Windsor, ON, Canada, 2018, pp. 226-229, doi: 10.1109/MWSCAS.2018.8623830.