Lie Detection through Facial Recognition Using Machine Learning Techniques

Manuel Cortez Muñoz

December 3, 2024

Abstract

This article addresses lie detection through facial recognition using machine learning techniques and image analysis. Traditionally, lie detection has relied on devices like the polygraph, but significant advances have been made in detecting deception through facial expression analysis. Ethical and practical methods are explored for use in psychological assessments, identity fraud detection, and deception prevention in recruitment processes. The research focuses on using Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), particularly LSTM networks, to analyze testimony videos and determine if a person is lying. This approach holds the potential to enhance lie detection accuracy in various contexts.

1 Introduction

Lie detection is a crucial aspect of human communication, especially in situations where it is vital to distinguish between truth and falsehood. These situations include police investigations, judicial processes, and security measures at airports. Traditionally, lie detection has been performed using polygraphs, which measure physiological responses such as blood pressure, heart rate, respiration, and skin conductivity.

However, the polygraph has its limitations and is not infallible. In recent years, significant advances have been made in lie detection using machine learning techniques and image analysis. New methods have emerged to detect lies by analyzing facial expressions, utilizing machine learning algorithms to analyze behavioral patterns and micro-expressions that may indicate deception. The ethical and practical issues these methods address include:

- **Psychological Assessment:** Evaluating whether a person is attempting to evade the law through false statements.
- **Identity Fraud:** Detecting if someone is using another person's data to misrepresent themselves.

2 Objectives

The main objectives of this research are:

- To analyze the effectiveness of machine learning techniques in lie detection.
- To implement a facial recognition system that classifies facial expressions into "truthful" and "lie" categories.
- To evaluate the model's performance using various metrics and improve its accuracy through iterative training and testing.

3 Literature Review

Research on lie detection has evolved significantly over the years. Traditional methods, such as polygraph tests, have limitations in accuracy and reliability. Recent studies have shown that facial expressions can provide valuable insights into a person's emotional state, which can be indicative of deception. The application of machine learning techniques has shown promising results in improving the accuracy of lie detection systems.

4 Methodology

The methodology adopted in this research consists of several key steps:

- **Data Collection:** A dataset of videos containing deceptive and truthful statements was compiled.
- **Data Pre-Processing:** The collected videos were processed to extract frames and perform image classification.
- **Modeling Techniques:** CNNs and LSTMs were used to build the classification model.
- **Evaluation:** The model's performance was evaluated using metrics such as accuracy, precision, recall, and the F1 score.

4.1 Data Collection

I compiled a dataset of 121 videos: 61 videos of deception attempts and 60 videos of truthful statements. This dataset served as the foundation for training and testing the machine learning models.

4.2 Data Pre-Processing

Data pre-processing converts raw data into a clean dataset. It involves cleaning and transforming the data into an appropriate format for analysis. Key steps include:

- **Data Cleaning:** Handling missing values, outliers, and incorrect data.
- **Data Transformation:** Scaling, normalizing, or encoding categorical variables to ensure compatibility with machine learning algorithms.

4.3 Modeling Techniques

To build my model, I used both CNNs and LSTMs to achieve optimal accuracy. The model's performance was evaluated using a validation dataset, and hyperparameters were tuned accordingly.

5 Work Performed

5.1 Project Analysis

I began by analyzing the thesis titled "Lie Detection through Facial Recognition Using Machine Learning," presented by Engineer Sebastián Torres Mojica at Universidad de los Andes, Colombia. His study processed 121 videos: 61 focused on deception attempts and 60 on truthful statements.

5.2 Code Implementation

For the code implementation, I used OpenCV for Python, a tool specialized in image and video processing. I employed the Haar Cascade algorithm from OpenCV, which uses machine learning to train a cascade function from positive and negative images. This function was then applied to detect faces in other images.

5.3 Video Processing

I processed each video in the dataset by defining the frame rate and segment duration, reading each video frame by frame, converting them to grayscale, and detecting faces using the algorithm mentioned. Each detected face was cropped and stored in the output directory as an image.

5.4 Image Classification

For image classification into "truthful" and "lie" categories, I used the Keras library. The pre-trained InceptionV3 model was selected as the base, with a new classification layer added. Images were resized to 299x299 pixels, converted into arrays, and preprocessed using Keras functions. I split them into training and test sets using the scikit-learn train-test split function.

5.5 Model Training

The model was compiled using the Adam optimizer, binary cross-entropy loss, and accuracy metric. After 10 epochs of training, the model reached 96.90% training accuracy and 97.6% validation accuracy.

5.6 Training Data Adjustment

Due to computational constraints, I reduced the dataset to 14 videos (7 truthful, 7 deception attempts). Despite this, the model achieved high accuracy in both training and validation sets.

5.7 Evaluation on Test Set

When evaluated on the test set, the model achieved an accuracy of 97.57%, effectively classifying images into "truthful" and "lie" categories.

5.8 Real-Life Application

I implemented a deep learning model for real-life lie detection, loading a pre-trained model to extract image features. A function was defined to preprocess images from directories, and images were split into training and test sets for model evaluation.

5.9 LSTM Model Configuration

I configured an LSTM model to classify images based on extracted features. The model was trained with 10 epochs, achieving a training accuracy of 98.71% and a validation accuracy of 97.66%.

5.10 LSTM Model Results

After evaluating the LSTM model on the test set, an accuracy of 97.01% was obtained, demonstrating its ability to classify "truthful" and "lie" images with high precision.

6 Conclusions

Throughout this project, I made decisions based on my analytical capabilities at the time. The results are promising, but further improvements can be made, especially in expanding the dataset and enhancing computational resources. My knowledge was limited at the start of the project, which made it challenging to understand all the processes that could improve the AI model. However, the results were positive, and I believe this project can be further developed in the future.

References

[1] Pascanu, R., Gulcehre, C., Cho, K., & Bengio, Y. (2013). How to construct deep recurrent neural networks. *arXiv preprint arXiv:1312.6026*.

[2] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.

[3] Carissimi, N., Beyan, C., & Murino, V. (2018). A Multi-View Learning Approach to Deception Detection. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (pp. 599-606). doi: 10.1109/FG.2018.00095.

[4] Anta, J. Á. (2012). Detection of deception: polygraph vs verbo-corporal analysis. *Quadernos de criminología: revista de criminología y ciencias forenses*, (19), 36-46.

[5] Torres Mojica, S. (2021). Lie detection through facial recognition using Machine Learning.

[6] Torres, J. (2019). Recurrent Neural Networks. *Jordi Torres.AI*.

[7] Shynu, T., Rajest, S. S., & Regin, R. (2023). A Review on Using Machine Learning to Conduct Facial Analysis in Real Time for Real-Time Profiling. *International Journal of Human Computing Studies*, 5(2), 18-37.

[8] Mapkar, S., Kapade, I., Panvel, N. M., Lawand, S., & More, S. Lie Detection System using Machine Learning. *PCE JCE*, 8(1), 32.

[9] Ekman, P. (2005). *Telling lies: Clues to deception in the marketplace, politics, and everyday life*. New York: W. W. Norton & Company.

[10] Zhang, J., Wang, X., & Liu, J. (2022). A deep learning approach for deception detection based on facial expressions. *IEEE Access*, 10, 48134-48144.