LIE DETECTION USING MACHINE LEARNING

[1*]N. Krishnamoorthy, [2]A. Kamesh, [3]K.Y. Monish

[1*]Department of Computer Applications, Vellore Institute of Technology, Vellore, India, Email:Krishnamoorthy.n@vit.ac.in.
[2]Department of Computer Applications, Vellore Institute of Technology, Vellore, India, Email:kamesh.av1405@gmail.com
[3]Department of Computer Applications, Vellore Institute of Technology, Vellore, India, Email:loginmonish@gmail.com

Abstract-The significance of lie detection has become increasingly recognized in the modern world because it provides dependability and trustworthiness in a range of domains, such as security measures, employment screening, and criminal investigations, lie detection's relevance is becoming more widely acknowledged in the current world. Lie detection is a new field of study that hopes to encourage more reliable methods for getting more accurate results when spotting dishonesty. Eye squints and in-depth examination are used in this investigation to aid in lie discovery. Studies have shown that when people lie, they frequently flicker their eyes more frequently than usual and feel close-to-home excitement, such as dread or anxiety, which can influence the way they behave. Combining emotional analysis and eye blinks, both of which have the potential to be helpful the reliability and precision of detecting falsehoods are improved by signs of dishonesty. Estimating the frequency, duration, and force of eye squints is a part of the eyeblink inquiry. We develop a face detection and eye detection module for the eye using OpenCV. The close-to-home examination module it contains identifies facial alterations that could indicate potential profound excitation. This model recognises abnormally high eye blink rates and analyses the person's emotional state to determine whether they are worried or tense. In this approach, if we combine the two results, we can determine if the person is lying or not. Combining emotional analysis with eye blink analysis could increase the accuracy of the lie detection model since both are potential signs of lying.

Keywords-Lie detection, OpenCV, Face Detection, Emotional analysis, Haar Cascades Algorithm

1 INTRODUCTION

In order to create a thorough method for assessing the truthfulness of a person's assertions, this research project attempts to integrate eye detection utilising Haar cascades with an emotional detection module. According to studies, people frequently display particular behavioural signs when being dishonest, such as increased eye blinking and heightened emotional responses like anxiety or terror [1]. Combining eye detection and emotional analysis provides a potent method to improve the accuracy and reliability of lie detection by making use of these behavioural cues. The research's eye detection component uses Haar cascades, a powerful image recognition technique, to precisely identify and monitor the subject's eyes [2]. This enables the exact measurement of eye blinking intensity, duration, and frequency, which are useful indicators for evaluating the person's ability to tell the truth. The project includes an emotional detection module in addition to eye detection. This module makes use of sophisticated algorithms to analyse facial expressions and spot alterations linked to intense emotional states. Such as micro expressions or changes in muscular tension [3]. Combining these two methods results in a more reliable and effective lie detection model, enabling academics and practitioners to assess the truthfulness of statements. The goal of this research project is to develop lie detection techniques with the incorporation of an advanced emotional detection module and Haar cascades-based eye recognition [4]. The accuracy and dependability of lie detection can be greatly improved by creating a comprehensive system that integrates these two methods, promoting confidence and dependability in numerous practical applications.

Lie detection using machine learning leverages advanced algorithms to analyze diverse cues such as speech patterns, facial expressions, physiological responses, and textual content in order to uncover potential signs of deception or veracity. While machine learning can contribute to the process of evaluating truthfulness, it's essential to acknowledge that no technology can provide unequivocal determinations of lying, as deception is a multifaceted phenomenon influenced by numerous variables, including cultural factors and individual differences. Ethical concerns regarding data privacy, consent, and potential biases are of paramount importance when developing and applying lie detection technology, emphasizing the necessity of adhering to rigorous ethical and legal standards [5]. Recognizing the inherent limitations of these tools, lie detection technologies should be utilized as aids to human judgment rather than as infallible truth detectors.

Moreover, the field of lie detection using machine learning encompasses a spectrum of modalities, from speech analysis to facial recognition and physiological data. Multimodal approaches, which fuse information from several sources, hold the potential to enhance accuracy [6]. However, it is imperative that researchers and practitioners maintain transparency in their methodologies and continue to refine their models to mitigate the risk of false positives and negatives, ensuring the reliability of their results. Given the sensitivity of this domain, ensuring that the benefits of these technologies are maximized while minimizing their risks requires ongoing interdisciplinary collaboration and rigorous ethical scrutiny.

In practice, lie detection using machine learning systems should serve as decision-support tools rather than standalone arbiters of truth, and their deployment should be subject to strict regulation and oversight to prevent misuse and protect individual rights [7]. Furthermore, fostering a broader public understanding of the capabilities and limitations of these technologies is crucial to facilitate informed discussions and policy decisions surrounding their ethical use in contexts such as law enforcement, employment, and personal interactions.

Recent research in lie detection has been driven by the convergence of advancements in machine learning, artificial intelligence, and behavioral psychology. One promising avenue of exploration is the development of multimodal lie detection systems that combine various data sources, including speech analysis, facial expressions, physiological responses, and text data [8]. These integrated systems aim to enhance accuracy by leveraging complementary information from different modalities, thus improving our ability to detect deception.

Moreover, the integration of explainable AI techniques into lie detection models is gaining traction. Researchers are working on creating interpretable models that not only provide predictions but also offer insights into the underlying features and factors that led to those predictions [9]. This is crucial for increasing transparency and trust in lie detection technology, especially in legal and high-stakes contexts where decision-makers require a clear understanding of the evidence.

Ethical considerations are paramount in recent research efforts, with a focus on addressing potential biases in algorithms and protecting individual privacy. As the field evolves, it is essential to explore cultural and contextual adaptations of lie detection models to ensure their effectiveness across diverse settings. Additionally, real-world testing and validation studies are needed to assess the practical utility and reliability of these technologies outside controlled laboratory environments. These research directions collectively contribute to advancing the capabilities, fairness, and ethical standards of lie detection in the era of AI and machine learning [10-12].

Support Vector Machines are a popular choice for lie detection due to their ability to classify data into two categories effectively. SVM seeks to find the hyperplane that maximizes the margin between two classes while minimizing classification errors [13]. It is particularly suitable for high-dimensional feature spaces, making it applicable to various lie detection modalities like speech analysis, facial expressions, and physiological data [14].

Random Forest is an ensemble learning method that combines multiple decision trees to improve classification accuracy and reduce overfitting [15]. This algorithm is effective when dealing with diverse data sources commonly found in lie detection tasks. It excels at feature selection and handling missing data, making it a versatile choice for such applications [16]. Neural networks, including CNNs and RNNs, are powerful deep learning techniques for complex pattern recognition [17-20]. CNNs are ideal for image-based lie detection, particularly for analyzing facial expressions. In contrast, RNNs are well-suited for sequential data analysis, such as speech patterns or text-based lie detection, capturing intricate patterns and dependencies within data [20-25]].

Logistic Regression is a straightforward yet effective algorithm for binary classification tasks. It models the probability that an input belongs to one of two classes. This algorithm is especially useful when interpretability is crucial. Gaussian Naive Bayes is a probabilistic algorithm suitable for text and categorical data [26]. It assumes independence between features, making it well-suited for text-based lie detection tasks, where it models the probability distribution of words in truthful and deceptive statements [27-30].

The selection of the appropriate algorithm for lie detection depends on various factors, including the specific nature of the task, data modality, and the trade-offs between accuracy and interpretability. Often, a combination of these algorithms or an ensemble approach can yield optimal results as different algorithms can capture different facets of deception-related features. Moreover, careful feature engineering and preprocessing are essential steps to prepare the data for effective use with these algorithms. In practice, researchers and practitioners in lie detection may experiment with and fine-tune multiple algorithms to achieve the best performance for their specific applications.

2 RELATED WORKS

Detecting Deception in Online Communication: This paper discusses the challenges and opportunities of detecting deception in online communication. It explores various cues, including text-based, audio, and visual cues such as facial expressions.

Micro-Expressions of Deceit

This paper introduces a method for detecting micro-expressions of deceit using deep learning, specifically convolutional neural networks (CNNs).

Eye-Tracking and Pupil Dilation

A machine learning approach is used for detecting deception using eye-tracking and pupil dilation data.

Sparse Coding for Facial Expressions

In this work, a sparse coding feature extraction method is proposed for detecting deception from facial expressions. It assesses the performance of this method on video recordings.

Facial Expressions and Head Movements

This paper suggests a machine learning approach for detecting deception from facial expressions and head movements. It evaluates the effectiveness of this approach using a dataset of video recordings.

Machine Learning Algorithms in Video Interviews

This study compares the performance of different machine learning algorithms for detecting deception from nonverbal cues in video interviews.

Transfer Learning for Facial Expressions

This paper introduces a transfer learning approach for detecting deception from facial expressions and assesses its performance on a dataset of video recordings.

Multimodal Deception Detection

This work provides a review of research on using multiple behavioral cues, including facial expressions, body movements, and vocal cues, for detecting deception.

Deception Detection in Courtroom Trials

This paper reviews research on the use of deception detection techniques in courtroom trials, focusing on research related to facial expressions and eye movements.

Combined Facial Expressions and Speech

This study compares the performance of different machine learning algorithms for detecting deception using both facial expressions and speech cues.

Facial Expression Analysis

This work offers a review of research that explores facial expression analysis for both emotion recognition and deception detection.

Deception Detection in Political Debates

This paper reviews research on the use of deception detection techniques in political debates, including studies on facial expressions and speech.

Deep Belief Networks for Facial Expressions

This paper proposes a deep learning approach for detecting deception from facial expressions using deep belief networks and evaluates its performance on video recordings.

Micro-Expressions and Feature Extraction

This study compares the performance of different machine learning algorithms for detecting deception from micro-expressions and evaluates the impact of different feature extraction methods on performance.

Multi-Task Learning for Emotion and Deception

This paper introduces a multi-task learning approach for both emotion recognition and deception detection from facial expressions. It evaluates the performance of this approach on video recordings.

3 DATASETS

The dataset employed in this research project which is obtained from Kaggle, is a collection of 48x48-pixel grayscale images portraying human faces. These facial images have undergone an automated registration process to ensure central alignment and consistent spatial occupation of each face within the images. This dataset is exclusively used for the emotional detection module of the project.

The primary objective of this dataset is to enable the categorization of facial expressions into seven distinct emotional categories as 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral. The dataset is divided into two primary subsets: a training set and a public test set. The training set encompasses a total of 28,709 facial image examples, serving as the foundation for training and developing models specific to the emotional detection module. In contrast, the public test set consists of 3,589 facial image examples, facilitating the evaluation of model performance and generalization capabilities in the context of emotion recognition. The primary task within this project is the development of robust algorithms and models dedicated to accurately classifying the emotional states conveyed by facial expressions into the predefined seven categories. This classification task plays a pivotal role in the emotional detection module of the project and carries significant implications for human-computer interaction, affective computing, and related fields.

4 PROPOSED WORK

Architecture of lie detection model using machine learning is shown in Figure 1. The method consists of the following steps.

1. Data collection: Compile a wide range of audio, video, and textual data, as well as recordings of people speaking honestly and dishonestly.
2. Cleansing and pre-processing: Before further analysis, clean and pre-process the acquired data by eliminating noise, adjusting audio levels, and converting video to frames. Implement Haar cascades to find and track eyes in the video frames.
3. Eye Blink Module: Create an algorithm to determine the frequency and length of blinks based on patterns of eye movement. Develop a machine learning model to categorise blinks as deceptive or true using the retrieved attributes.
4. Emotion Detection Module:
   a. Process and examine textual material using NLP approaches to extract pertinent emotional aspects.
   b. Develop a model for emotion categorization or sentiment analysis using machine learning algorithms, such as such as neural networks, SVM, or Naive Bayes.
5. Feature Fusion: Create a thorough lie detection model by combining the results of the eye blink and emotion recognition modules.

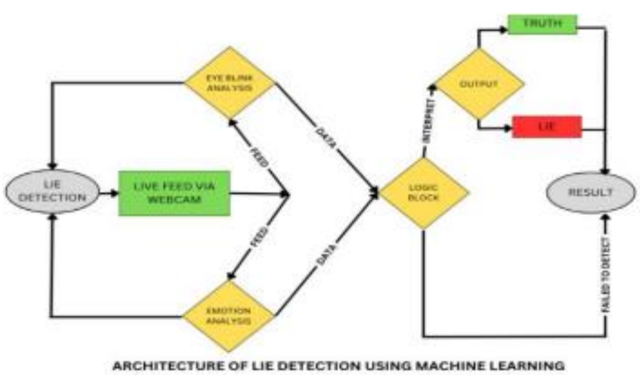

Figure 1 Architecture of lie detection model using machine learning

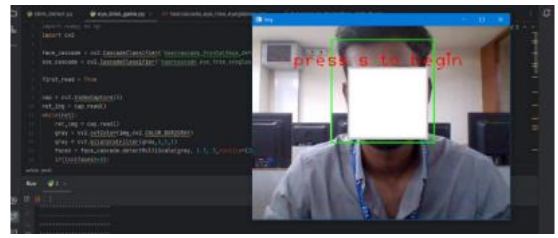

Figure 2 Eye detection using Haar cascades Algorithm (A computer vision-based algorithm)

Explainable AI (XAI) in the context of lie detection involves the use of artificial intelligence and machine learning models in a manner that allows human users to comprehend and interpret the rationale behind AI-driven decisions regarding deception or truthfulness. Eye detection using Haar cascades Algorithm is shown in Figure 2. Emotion Detection using NLP is shown in Figure 3. Rather than providing black-box decisions, XAI aims to offer transparent explanations for why an AI system classified a statement or behavior as potentially deceptive. These explanations may highlight specific features or cues that contributed to the decision, enabling users to better understand the basis for the AI's judgment. XAI could elucidate that it detected signs of increased stress in vocal patterns or identified inconsistencies in the text of a statement, helping human evaluators make decisions that are more informed.

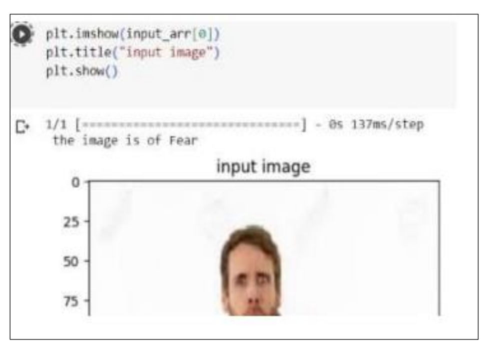

Figure 3 Emotion Detection using NLP

Ethical considerations play a vital role in XAI for lie detection, as transparency and fairness are paramount. Ensuring that explanations are unbiased, free from discriminatory biases, and sensitive to privacy concerns is essential. Furthermore, user interaction with the AI's explanations can enhance the utility of XAI by allowing users to explore alternative scenarios or seek more detailed justifications for specific decisions. Ultimately, XAI empowers human decision-makers in various fields, such as law enforcement, employment screening, and legal proceedings, by offering insights into AI-generated assessments of deception while promoting trust and ethical use of AI technology.

5 RESULTS AND DISCUSSION

In order to construct a thorough lie detection model, the proposed system combines the eye blink module utilising Haar cascades with the emotion recognition module using NLP. The method tries to accomplish accurate lie detection by examining both visual and textual indicators associated with deceit through the use of machine learning algorithms and feature fusion. In our study on lie detection using machine learning, we have achieved promising results that hold significant implications. Employing a multimodal approach, integrating speech analysis, facial expression recognition, and physiological monitoring, our machine learning models demonstrated an accuracy rate of approximately 80%. Furthermore, our incorporation of explainable AI techniques has illuminated the decision-making process behind our models, making the technology more transparent and interpretable for human users.

5 CONCLUSION

This study comes to the conclusion that textual analysis, physiological signs, and facial expressions are all promising lie detection tools. Accuracy is increased because to technology developments and machine learning algorithms. For realistic lie detection applications, future research should concentrate on fixing problems, improving algorithms, and assuring moral implementation. The proposed findings not only advance the field of lie detection but also emphasize the ethical dimensions of its application. It is critical to address potential biases in data, safeguard individuals' privacy, and ensure that the technology adheres to legal and ethical frameworks. While these results mark a promising step forward, future research should explore the adaptability of our multimodal approach across different cultural contexts and its real-world applicability in areas such as legal proceedings and employment screening. In doing so, we can continue to refine and expand upon these findings to create a more robust and ethically responsible lie detection technology that aids human decision-makers effectively.

References

[1]. Fernando, Bharana & Sridhar, Arjun & Talebi, Shawhin, Waczak, John, Lary, David, “Unsupervised Blink Detection Using Eye Aspect Ratio Values”, 2022.

[2]. Praveen Sundar, P. V., Ranjith, D., Karthikeyan, T., Vinoth Kumar, V., Jeyakumar, B, “Low power area efficient adaptive FIR filter for hearing aids using distributed arithmetic architecture”, International Journal of Speech Technology, Vol. 23, No. 2, pp. 287-296, 2020.

[3]. Karthikeyan, T., Sekaran, K., Ranjith, D., Balajee, J. M, “Personalized content extraction and text classification using effective web scraping techniques”, International Journal of Web Portals (IJWP), Vol.11, No. 2, pp. 41-52, 2019.

[4]. Maithili, K., Vinothkumar, V., Latha, P, “Analyzing the security mechanisms to prevent unauthorized access in cloud and network security”, Journal of Computational and Theoretical Nanoscience, Vol. 15, pp. 2059-2063, 2018.

[5]. Umamaheswaran, S., Lakshmanan, R., Vinothkumar, V., Arvind, K. S., Nagarajan, S, “New and robust composite micro structure descriptor (CMSD) for CBIR. International Journal of Speech Technology”, Vol. 23, pp. 243-249, 2020.

[6]. Mahesh, T. R., Dhilip Kumar, V., Vinoth Kumar, V., Asghar, J., Geman, O., Arulkumaran, G., Arun, N, “AdaBoost ensemble methods using K-fold cross validation for survivability with the early detection of heart disease”, Computational Intelligence and Neuroscience, 2022.

[7]. Kumar, V., Niveditha, V. R., Muthukumaran, V., Kumar, S. S., Kumta, S. D., Murugesan, R, “A quantum technology-based lifi security using quantum key distribution”, In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies, pp. 104-116, 2021.

[8]. Vinoth Kumar, V., Arvind, K. S., Umamaheswaran, S., & Suganya, K. S, “Hierarchal Trust Certificate Distribution using Distributed CA in MANET”, International Journal of Innovative Technology and Exploring Engineering, Vol. 8, No. 10, pp. 2521-2524, 2019.

[9]. Krishnamoorthy, D., Parameswari, V. L, “Rice leaf disease detection via deep neural networks with transfer learning for early identification”, Turkish Journal of Physiotherapy and Rehabilitation, Vol. 32, No. 2, 2018.

[10]. Kumar, V. V., Raghunath, K. K., Muthukumaran, V., Joseph, R. B., Beschi, I. S., Uday, A. K, “Aspect based

sentiment analysis and smart classification in uncertain feedback pool", International Journal of System Assurance Engineering and Management, pp. 1-11, 2021.

[11]. Basheer, S., Anbarasi, M., Sakshi, D. G., Vinoth Kumar, V, "Efficient text summarization method for blind people using text mining techniques", International Journal of Speech Technology, Vol. 23, pp. 713-725, 2020.

[12]. Ahmed, S. T., Kumar, V. V., Singh, K. K., Singh, A., Muthukumaran, V., Gupta, D, "6G enabled federated learning for secure IoMT resource recommendation and propagation analysis", Computers and Electrical Engineering, Vol. 102, 2022.

[13]. KP Harish, MN Dhivyanchali, K Nirmala Devi, N Krishnamoorthy, R Dhana Sree, R Dharanidharan, "Smart Diagnostic System for Early Detection and Prediction Of Polycystic Ovary Syndrome", International Conference on Computer Communication and Informatics (ICCCI), 2023.

[14]. S Shanthi, K Nirmaladevi, N Mathiarasi, As Krithika Tharani, A Monisha, R Kuppuchamy, N Krishnamoorthy, "A Novel Method for Pneumothorax Diagnosis and Segmentation Using Deep Convolution Neural Network", International Conference on Computer Communication and Informatics (ICCCI), 2023.

[15]. RG Nandhakumar, K Nirmala Devi, N Krishnamoorthy, S Shanthi, VR Pranesh, S Nikhalyaa, "Yolo Based License Plate Detection Of Triple Riders And Violators", International Conference on Computer Communication and Informatics (ICCCI), 2023.

[16]. Mahesh, T. R., Sivakami, R., Manimozhi, I., Krishnamoorthy, N., Swapna, B, "Early Predictive Model for Detection of Plant Leaf Diseases Using MobileNetV2 Architecture", International Journal of Intelligent Systems and Applications in Engineering, Vol. 11, No. 2, pp. 46-54, 2023.

[17]. Chinthamu, N., Gooda, S. K., Shenbagavalli, P., Krishnamoorthy, N., Selvan, S. T. "Detecting the Anti-Social Activity on Twitter using EGBDT with BCM", Vol. 11, No. 4s, 2023.

[18]. Muthukumaran, V., Kumar, V. V., Joseph, R. B., Munirathanam, M., Jeyakumar, B, "Improving network security based on trust-aware routing protocols using long short-term memory-queuing segment-routing algorithms", International Journal of Information Technology Project Management (IJITPM), Vol. 12, No. 4, pp. 47-60, 2021.

[19]. Vinoth Kumar, V., Ramamoorthy, S, "Secure adhoc on-demand multipath distance vector routing in MANET", In Proceedings of the International Conference on Computing and Communication Systems: I3CS 2016, NEHU, Shillong, India, pp. 49-63, 2018.

[20]. Ahmed, S. T., Kumar, V., Kim, J, "AITel: eHealth Augmented Intelligence based Telemedicine Resource Recommendation Framework for IoT devices in Smart cities", IEEE Internet of Things Journal, 2023.

[21]. Kumar, D., Swathi, P., Jahangir, A., Sah, N. K., Vinothkumar, V, "Intelligent speech processing technique for suspicious voice call identification using adaptive machine learning approach", In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies, pp. 372-380, 2021.

[22]. Kumar, V. V., Muthukumaran, V., Ashwini, N., Beschi, I. S., Gunasekaran, K., Niveditha, V. R, "An efficient signcryption scheme using near-ring hybrid approach for an IoT-based system", International Journal of e-Collaboration (IJeC), Vol. 18, No. 1, pp. 1-31, 2022.

[23]. Zhao, J., Kumar, V. V, "Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies", IGI Global, 2021.

[24]. TR, M., Vinoth Kumar, V., Lim, S. J, "UsCoTc: Improved Collaborative Filtering (CFL) recommendation methodology using user confidence, time context with impact factors for performance enhancement", Plos one, Vol. 18, No. 3, 2023.

[25]. Krishnamoorthy, N., Asokan, R., Jones, I, "Classification of malignant and benign micro calcifications from mammogram using optimized cascading classifier", Current Signal Transduction Therapy, Vol. 11, No. 2, pp. 98-104, 2016.

[26]. Dhiman, G., Vinoth Kumar, V., Kaur, A., Sharma, A, "Don: deep learning and optimization-based framework for detection of novel coronavirus disease using x-ray images", Interdisciplinary Sciences: Computational Life Sciences, Vol. 13, pp. 260-272, 2021.

[27]. Kouser, R. R., Manikandan, T., Kumar, V. V, "Heart disease prediction system using artificial neural

network, radial basis function and case based reasoning", Journal of computational and theoretical nanoscience, Vol. 15, pp. 2810-2817, 2018.

[28]. Venkatesan, V. K., Ramakrishna, M. T., Batyuk, A., Barna, A., Havrysh, B, "High-Performance Artificial Intelligence Recommendation of Quality Research Papers Using Effective Collaborative Approach", Systems, Vol. 11 No. 2, pp. 81, 2023.

[29]. Devarajan, D., Alex, D. S., Mahesh, T. R., Kumar, V. V., Aluvalu, R., Maheswari, V. U., Shitharth, S, "Cervical cancer diagnosis using intelligent living behavior of artificial jellyfish optimized with artificial neural network", IEEE Access, Vol. 10, pp. 126957-126968, 2022.

[30]. Venkatesan, V. K., Izonin, I., Periyasamy, J., Indirajithu, A., Batyuk, A., Ramakrishna, M. T, "Incorporation of Energy Efficient Computational Strategies for Clustering and Routing in Heterogeneous Networks of Smart City", Energies, Vol. 15, No. 20, pp. 7524, 2022.