International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 10 | October - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

VisiAssist: "Android-based Visual Aid and Health Predictor"

Pranjali Patil
Department of Computer Engineering
Smt. Kashibai Navale college of Engineering

Vivek Kakade
Department of Computer Engineering
Smt. Kashibai Navale college of Engineering

Vaishnavi Dhakane
Department of Computer Engineering
Smt. Kashibai Navale college of Engineering

Prof. Supriya Jagtap
Department of Computer Engineering
Smt. Kashibai Navale college of Engineering Vadgaon (SPPU- Pune)
supriyajagtap_skncoe@sinhgad.edu

Abstract:- This innovative application is built for the visually challenged people. By using algorithms to analyze the camera feed and identify objects within the user's environment. The application be able to identify a diverse range of objects commonly found in everyday scenarios, such as furniture, obstacles, vehicles, people, and more. The detected objects are then be translated into auditory feedback, which the user can interpret and understand without needing to rely solely on visual cues.

Keywords— Android application; YOLOv3 algorithm, object detection

I. INTRODUCTION

Object detection has been known as the core of computer vision and attracted much research attention in recent years especially because of its close relationship with video analysis and image understanding. According to the abundant research on object detection, many traditional object detection methods have been proposed. This paper introduces some famous traditional methods, which are based on SIFT, HOG, SURF, and ORB.

However, due to the characteristics of large amount of computation and simple training structure, the traditional detection method has low detection speed. With the fast rise of deeper learning, stronger devices are implemented to address the problems that exist in conventional architectures. In the architecture of the network, training and optimization functions etc., these models are special.

Through this project, we seek to revolutionize the field of personality assessment, offering a more efficient and insightful method for understanding and utilizing personality insights in various domains.

In this paper, we review the frameworks for object detection based on deep learning. We begin our review with the methods based on Convolutional Neural Networks. Then typical methods of object detection and some helpful modification to improve detection performance are introduced.

Moreover, the methods based on YOLO and SSD are introduced. Vision is one of the most significant human senses among all the human senses present, and it assumes a crucial job in the understanding of the surrounding environment. Visually impaired people find it difficult to move around outdoors without any supervision.

Mobile devices are evolving into convergent platforms for communication, computing, and personal sensing. The inclusion of these features in mobile devices would improve communication. more productive and efficient. Android is a technology platform for the smartphone application, which is among mobile technologies, the largest market share. The initial stage in image processing is identifying objects in a given image.

Particularly in complicated settings, when numerous targets must be processed instantly and automatically. Particularly crucial are extraction and recognition. As computer technology has advanced and the extensive usage of computer vision techniques, the to track targets in using computer image processing Real-time is becoming more and more common.

The goal of object detection is to find and identify all the prominent objects in the entire image that have received a lot of attention. Real-time responsiveness and precision are crucial features of the whole system .

As per the survey, there are the following learning techniques for data classification.

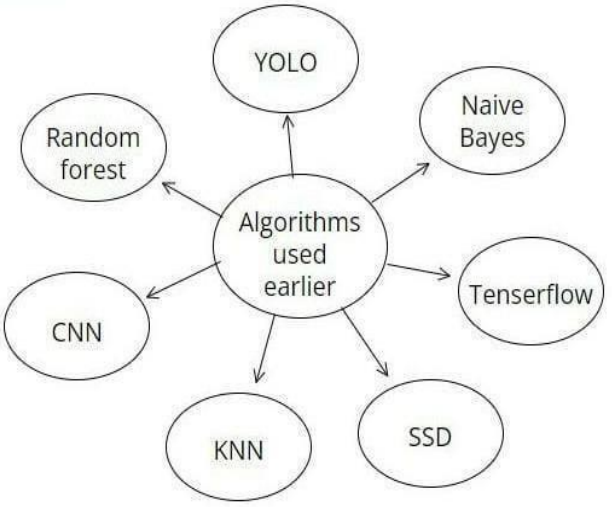

Fig. 1. Algorithms used earlier

You Only Look Once (YOLO)
Extracting the features automatically from the large datasetand get the proper result.
It has excellent processing speed where it can process 45 frames per second can be applied to classify users' textdata into personality traits by learning patterns andassociations between wordsor phrases and specific traits. It's effective when dealing with high-dimensional data,making it suitable for text classification tasks.

Random Forest
An ensemble learning technique for problems like regression and classification is called Random Forest.
During training, it builds a large number of decision trees and outputs the class that represents the mean forecast of each individual tree or the mode of classes.The algorithm doesn't need a lot of data and is resistant to outliers. It can handle problems involving both regression and classification, and it is very resilient to noisy data.
By combining predictions from several decision trees, Random Forest may be utilized to capture intricate correlations between text elements and personality qualities

Single Shot MultiBox Detector(SSD)
This detector operates in a stage, which implies that it predicts both the bounding boxes and class probabilities of objects in an image simultaneously. To accomplish this SSD employs a network (CNN) to extract image features. Subsequently it utilizes layers to estimate the bounding boxes and class probabilities, for the objects, in the image. The speed and accuracy of SSD make it well suited for real time object detection tasks.

Tensorflow object detection API:
The Tensorflow object detection API offers a range of, to use models that can be applied to tasks including object detection, instance segmentation pose estimation and motion tracking. Additionally it allows users to develop and train their custom models by providing the API with a dataset of images and associated labels..

Convolutional Neural Network (CNN)
CNNs are primarily designed for image data but can also be adapted for text analysis using techniques like text convolution.
Text convolution involves applying filters to sequences ofwords to capture local patterns.
CNNs could be used to capture complex text patterns, such as specific word combinations or phrases, related to personality traits.

K-Nearest Neighbors (KNN)
KNN is a basic instance-based classification method thatis frequently applied to applications involving regressionand classification .Because it is a non-parametric approach, no assumptions are made about the underlyingdata
The training set's k points that are closest to the new datapoint are found by the algorithm.

Naive Bayes Classifier
The probabilistic classification method Naive Bayes is basedon the Bayes theorem.
It makes the simplistic but useful premise that characteristicsare independent of one another in text categorization. NaiveBayes uses user text data to analyze word frequencies and correlations in order to predict personality characteristics.

RELATED WORK

A literature survey consists of different learning techniques research data as follows:
[1] Ankush Singh, Ashish Yadav, Saloni Shah, and Renuka Nagpure conducted a study that was published in 2022. A vast amount of patient data is produced by the medical sector, and this with the aid of machine learning. Many of the current systems have lower accuracy and can only predict one disease at a time. Inaccurate results can gravely endanger a patient's health.

[2] Akhilesh Salunke, Manthan, Shayantam Santra, and Mrs. Sumedha Bhagwat conducted a study that was published in a paper in 2021. This essay outlines a straightforward Android app that could aid those who are blind in comprehending their environment. It was difficult to detect objects in real time and turn them into an audio output. enabled the creation of numerous applications for real-time object detection.

[3] Anil Thapa, Akash Bongane, Madhuri Rokade, and Nilesh Wani conducted a study that was published in 2021. People who are blind or visually impaired must overcome many challenges in their daily lives. The purpose of the proposed project is to develop an Android application for blind or visually impaired individuals. The suggested framework uses Google's TensorFlow Lite object detection API in conjunction with machine learning..

[4] In a study led by Dr. Ketaki B. Naik, Sayali Wanewe, Rutuja Nikam, and Tejal Adep, which was published in 2021. Blind people deal with this issue on a daily basis. They require assistance to even walk. They frequently rely on assistance from others. Numerous technologies have been developed to aid individuals who are visually impaired. Because they are accessible and affordable, computer vision-based solutions are looking like one of the most promising options among the various technologies being used to assist the blind. A system for those who are visually impaired is proposed in this paper.

[5] In a study by Shoji Yachida, Keiko Yokoyama, and Shigeaki Namiki that was published in 2021. Detecting and identifying objects moving at high speeds can be accomplished with high-speed camera imaging (e.g., 1,000 fps) because the temporally dense images produced by a high-speed camera can typically capture the optimal moment for object detection and recognition. We suggest a unique framework that combines high-speed camera imaging with real-time object recognition. The essential procedures of data ensemble.

[6] In a 2020 paper, Sunit Vaidya, Naisha Shah, and Niti Shah conducted a study. Vision is one of the most vital senses for survival. Millions of people worldwide suffer from some form of visual impairment. These individuals struggle to get around independently and safely; they also have trouble communicating and getting information. The proposed work aims to inform blind individuals about objects in their path, thereby transforming the visual world into an auditory one. By using the real-time object detection system, this will enable visually impaired people to navigate independently without the need for outside assistance.

[7] In a study conducted by Pooja Maid, Omkar Thorat and Sarita Deshpande,paper published in 2020. The Smart Blind Navigation is fill gap, providing accurate and contextually rich information about the environment around the user current location, and simplifying the navigation and increasing the overall accuracy of the system. Preventing the user from dangerous locations. They have very little information on self-velocity objects, direction which is essential for travel.In research conducted by Peng Wang, Yun Yan, Xiangping Zhan, Gancheng Wu, Jun Wang, and Yunsheng Pan, publishedin IEEE in May 2020, Support Vector Machine (SVM), K- Nearest Neighbors (KNN), Naive Bayes (NB), and Logistic Regression (LR) were applied for personality analysis. The researchers emphasized the use of SVM and LR as linear algorithms, but specific results were not mentioned.

[8]Reya Sharma, Baijnath Kaushik, and Naveen Gondhi, in their study published in IEEE in March 2020, explored various machine learning algorithms, including Support Vector Machine, Artificial Neural Network, Naive Bayes, and Convolutional Neural Network. They conducted data acquisition, preprocessing, segmentation, feature extraction, and classification. The overall model accuracy was reported as61%, and the advantage noted was the ability of multiple algorithms to maintain accuracy when one failed at classification.

[9] Sayali Ambekar and Rashmi Phalnikar conducted a study that was published in a paper in 2018. Handling a lot of data in the healthcare industry requires the use of data analysis. Rather than relying on prediction, earlier medical research handled and assimilated vast amounts of hospital data. The biomedical and healthcare industries are experiencing a massive data growth, which makes accurate medical data analysis advantageous for early disease detection and patient care

[10] Xiaoyin Xu and Eric L. Miller conducted a study that was published in 2000. We tackle the challenge of identifying objects or buried mines from multichannel sequentially collected ground penetrating radar (GPR) data for uses in archaeology, utility line mapping, and landmine and unexploded ordnance remediation. Typically, it is unknown what the target signal's precise form is. To address this issue, we create and evaluate a low-complexity, sequential detection strategy based on decision theory.

TABLE I. SUMMARY OF RELATED WORK / GAP ANALYSIS

Ref No	Algorithm and Parameters	Highlights	Limitations and Future Work
1	You Only Look Once (YOLO)	- Faster than traditional algorithms. - Can detect objects in real-time video streams. - YOLO achieves high accuracy and precision in object detection, even in cluttered scenes.	- Limited Accuracy for Small Objects. - Difficulty in Detecting Overlapping Objects. - Can't handle objects well that vary significantly in size across different images.
2	Random Forest Algorithm	- Random Forest typically provides high accuracy in both classification and regression tasks. - It is resistant to overfitting, especially when compared to decision trees.	- Random Forests can be computationally intensive, especially when dealing with a large number of trees. - Can still overfit noisy data
3	Convolutional Neural Network (CNN)	- CNNs automatically learn hierarchical features from raw data. - CNNs are capable of recognizing patterns in different positions within the image.	- CNNs require significant computational resources, especially for training large networks on large datasets. - CNNs perform best when trained on large datasets.
4	K-Nearest Neighbour (KNN)	- KNN is easy to understand and implement. - Unlike many other algorithms, KNN does not require a training	- KNN can be computationally expensive, especially when dealing with large datasets. - KNN stores all the training data, which can be memory-intensive, particularly for large datasets.

5	Single Shot Detection (SSD)	- SSD is incredibly fast compared to many other object detection algorithms. - Despite its speed, SSD maintains a high level of accuracy in object detection.	- SSD may not be as accurate as some two-stage detection methods like Faster R-CNN, especially for small or densely packed objects. - SSD processes the entire image as a single shot, which means it might lack the context provided by region proposal methods in two-stage detectors.
6	Tenserflow	- Its flexibility allows developers to implement various machine learning models and algorithms efficiently. - Suitable for both small-scale and large-scale machine learning tasks.	- Its complex API and extensive documentation can be overwhelming for newcomers. - TensorFlow code can be verbose, requiring developers to write more lines of code compared to other frameworks.
7	Naïve Bayes Algorithm	- Naive Bayes is easy to understand and implement. - Naive Bayes algorithms train quickly, especially on large datasets, because they assume independence between features.	- Naive Bayes assumes that all features are independent, which is often not the case in real-world scenarios. - It is highly sensitive to the quality of the input data.

II. SYSTEM ARCHITECTURE

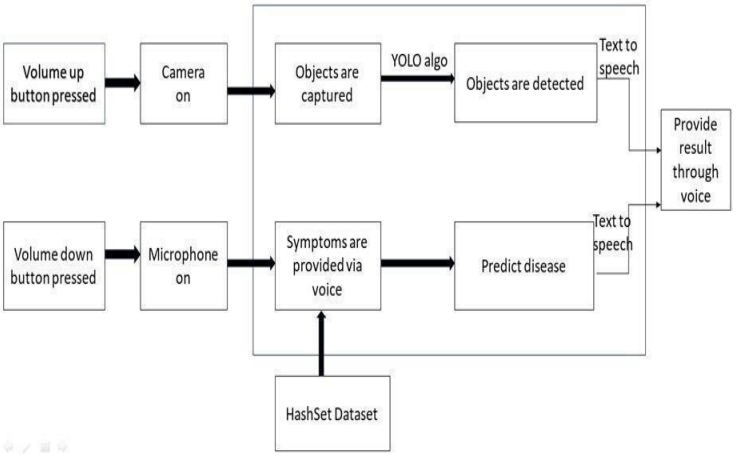

Fig. 2. System Architecture

I. *Image Object Detection Module:*
This innovative system combines advancedtechnologies to create a seamless user experience. Using a connected camera, the system captures real- time images, and with the power of YOLOv3 algorithm, it identifies and labels objects within theseimages accurately and swiftly.

II. *Speech Output Module:*
The detected objects are then translated into spoken words using natural language processing, providing users with an auditory understanding of their surroundings. This feature not only assists individuals with visual impairments but also offers a convenient hands-free interaction method.

III. *Voice Input and Disease Recognition Module:* Moreover, the system incorporates a microphone to accept voice inputs from users. Through automatic speech recognition, spoken words are converted into text. A sophisticated disease recognition module analyzes these textual inputs, identifying specific symptoms mentioned by the user. Leveraging machine learning algorithms and medical databases, the system then predicts potential diseases based on the provided symptoms.

IV. *Disease Prediction and Audio Output:*
The results are communicated back to the user in an audio format, offering insights into the predicted diseases and related information. This interactive voice interface not only enables real-time object recognition but also empowers users to describe their symptoms verbally, aiding in potential self-diagnosis and raising awareness about their health conditions.

VI. Testing and Validation:
Conduct user testing and validation to ensure that the systemprovides fair and accurate personality assessments. Collect feedback from users and make necessary improvements to the system.

VII. Deployment and Integration:
Deploy the text-based personality analysis system for use byindividuals, organizations, and researchers.
Explore integration options with existing platforms orservicesto enhance accessibility.

VIII. Testing and Validation:
Conduct extensive testing and validation of the system's predictions and advice.
Solicit user feedback and fine-tune the system to enhance accuracy and user satisfaction.

Ethical Considerations:
Prioritize user privacy and data protection throughout the project.
Address potential biases in the dataset and the algorithm to ensure fairness in disease prediction.

Deployment and Impact:
Deploy the object detection system for widespread use, considering integration with existing platforms and services.
Monitor and measure the impact of the system in assisting personal and professional development.

Continuous Improvement:
Continuously update the system with new data and refinethe YOLO algorithm to adapt to evolving user needs.
Stay informed about advancements in YOLO and machine learning for potential enhancements.

Software Requirements Specification:

Vision impairment affects millions of people, making everyday activities challenging. For the blind, mobility is particularly difficult, requiring constant assistance. Our project aims to empower visually impaired individuals by leveraging computer vision technology on the Android platform. By using real- time object detection and disease prediction, we intendto enhance their independence and safety. This system utilizes YOLOV3, a powerful object detection algorithm, and Convolutional Neural Networks (CNN) for accurate predictions. The Android platform was chosen due to its popularity and user-friendly interface. The project's goal is to provide immediate, accurate disease predictions based on user-entered symptoms, thereby aiding in timely healthcare decisions.

3.1.1 Project Scope

Our system focuses on object detection and disease prediction. Through audio output, users receive information about their surroundings and disease predictions based on symptoms input.

3.1.2 User Classes and Characteristics

The system caters to visually impaired individuals, offering object identification and disease predictionfunctionalities.

3.1.3 Assumptions and Dependencies

The system relies on YOLOV3 for object detectionand CNN for disease prediction. Kotlin programming language and Android Visual Studio are used for development, ensuring efficiency and seamless integration.

3.2 Functional Requirements

The system includes understanding the problem statement, defining hardware/software requirements, system understanding, planning, designing, programming, and testing.

3.3 External Interface Requirements

3.3.1 User Interface

The application is designed for Android devices, requiring a minimum of 8GB RAM, an Intel i5 Processor, and 500GB hard disk space.

3.3.2 Software Interfaces

The system operates on Windows 10, utilizing Android Visual Studio and Kotlin programming language.

3.4 Non-Functional Requirements

3.4.1 Performance Requirements

The system must perform efficiently, particularly in data encryption and virtual environment provision.

3.4.2 Safety Requirement

Modular design allows for easy error detection and fixes, ensuring safe installation and updates.

3.4.3 Security Requirements

Robust security measures authenticate users, ensuring reliable

protection against unauthorized access.

3.4.4 Software Quality Attributes

The software is adaptable, available to all, easily maintainable, reliable, user-friendly, and secure. It is thoroughly tested to ensure quality.

3.5 System Requirements

3.5.1 Database Requirements

Firebase Realtime Database is utilized for real-time data synchronization, allowing collaborative usage across devices.

3.5.2 Software Requirements

Kotlin, a cross-platform programming language, and Android Visual Studio are used for development.

3.5.3 Hardware Requirements

The system requires a minimum of 8GB RAM, an Intel i5 Processor, and 500GB hard disk space.

3.6 Analysis Models: SDLC Model

The software development life cycle involves stages such as Requirement Analysis, System Design, Implementation, Testing, Deployment, and Maintenance. Each phase ensures a systematic approach to software development, allowing for efficient and error-free deployment.

CONCLUSION AND FUTURE SCOPE:

The application features a user-friendly interface specifically designed to cater to visually impaired users. Upon launching the application, the camera initiates real-time video capture. When the user activates a button, the backend server algorithm commences processing. The output is then conveyed to the user via audio notifications. The YOLOv3-tiny algorithm is employed to detect objects and their positions around blind individuals. Users can halt the algorithm by pressing the same button again.

The future scope of this project involves enhancing the system's capability to recognize multiple items in a display with greater accuracy and speed. By improving the frame rate, the system can expand its functionality to identify specific entity types effectively. The text-to-speech module has been developed with a forward-looking approach, ensuring adaptability for future advancements.

REFERENCES

[1] Archana Singh ,Rakesh Kumar, “Heart Disease Prediction Using Machine Learning Algorithms”, 2020 IEEE, International Conference on Electrical and Electronics Engineering (ICE3).

[2] S. Vaidya, N. Shah, N. Shah and R. Shankarmani, "Real- Time Object Detection for Visually Challenged People," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 311-316,doi:10.1109/ICICCS48265.2020.9121085.

[3] Priyanka Sonar, Prof. K. JayaMalini,” DIABETES PREDICTION USING DIFFERENT MACHINE LEARNING APPROACHES”, 2019 IEEE ,3rd International Conference on Computing Methodologies and Communication (ICCMC).

[4] Nasreen, W. Arif, A. A. Shaikh, Y. Muhammad and M. Abdullah, "Object Detection and Narrator for Visually Impaired People," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS),Kuala Lumpur, Malaysia, 2019, pp. 1- 4,doi:10.1109/ICETAS48360.2019.9117405.

[5] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, “Disease prediction by machine learning over big data from healthcare communities,” IEEE Access, vol. 5, no. 1, pp.8869–8879, 2017.

[6] Meghajit Mazumdar, Dr. Sarasvathi V, Akshay Kumar “Object Recognition in Videos by Sequential Frame Extraction using Convolutional Neural Networks and Fully Connected Neural Networks” International Conference on Energy, Communication, Data Analytics and Soft Computing 2017.

[7] “You Only Look once : Unified, Real-Time Object Detection .” J Redmon, S Divvala, R Girshick, A Farhadi,IEEE transactions ,May 2016.

[8] SSD: Single Shot MultiBox Detector Wei Liu.” , Dragomir Anguelov , Dumitru Erhan , Christian Szegedy, Scott Reed , ChengYang Fu , Alexander C. Berg, IEEE transactions ,Jan 2016

[9] Yide Ma, Dong Hwan Kim, and Sung-Kee Park “Region- Based Object Recognition by Color Segmentation Using a Simplified PCNN” IEEE transactions on neural network and learning system, Vol, 26 No. 8 Aug 2015

[10] Jizhong Xiao, Kevin Ramdath, Manor Iosilevish, Dharmdeo Sigh, and Anastasis Tsakas. A low cost outdoor assistive navigation system for blind people. In Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on, pages 828–833. IEEE, 2013.

[11] Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(9):1627