COMPUTER VISION-BASED MOBILE APPLICATION FOR VISUALLY IMPAIRED

Project Reference No.: *47S_BE_1964*

College *: East Point College of Engineering and Technology, Bengaluru*
Branch *: Department of Computer Science and Engineering*
Guide(S) *: Mrs. Divya U. H.*
Student(S) *: Mr. Kalluri Saiteja*
Mr. Krishna Chaitanya S.
Mr. Hari Balaji V.
Mr. Kancham Hariprasad

Keywords:
TensorFlow Lite, Indoor Object Detection and Recognition (IODR), Convolutional Neural Network (CNN), Neural Architecture Search (NAS), Image Segmentation Algorithms, Deep Convolution Neural Network (DCNN).

Background:

The field of computer vision, a subset of artificial intelligence, has made remarkable strides in recent years. It is dedicated to equipping machines with the ability to interpret and comprehend visual data from their environment. By analyzing digital images or videos, computer vision systems can extract valuable insights, identify objects, and discern patterns. This progress has been fueled by breakthroughs in deep learning algorithms, augmented computational capabilities, and the availability of extensive datasets.

Concurrently, the global prevalence of visual impairment poses profound challenges for individuals in their daily lives. With approximately 285 million people worldwide affected by visual impairment, including 39 million who are blind, there is a critical need to empower these individuals with tools that enhance their independence and quality of life. Recognizing this urgency, advancements in technology have ushered in innovative solutions aimed at addressing the unique needs of the visually impaired community.

This project is a proactive response to the challenges faced by visually impaired individuals in accessing information and navigating their environments. Leveraging the capabilities of modern technology, particularly in computer vision and speech

recognition, this project endeavors to provide a comprehensive solution tailored to the unique needs of the visually impaired community.

The genesis of this project stems from a deep understanding of the daily obstacles encountered by visually impaired individuals. Traditional methods of assistance, such as guide dogs and canes, offer valuable support but are limited in their scope. With the rapid advancement of technology, there arises a tremendous opportunity to harness its potential in addressing the specific needs of this demographic.

Modern smartphones, equipped with powerful processors and a myriad of sensors, serve as an ideal platform for developing innovative solutions. By integrating computer vision algorithms, these devices can analyze visual information in real-time, opening up a world of possibilities for assisting visually impaired individuals in object recognition and environmental navigation.

Additionally, advancements in speech recognition and synthesis technologies further enrich the user experience by enabling natural interaction with the mobile application. By seamlessly integrating these technologies, the project aims to create a user-friendly interface that empowers visually impaired individuals to interact with their surroundings more independently.

Objectives of the Project:

1 Develop a user-friendly mobile application interface that integrates advanced computer vision algorithms, facilitating real-time object recognition within the camera feed, thus empowering visually impaired users to efficiently identify and locate objects in their surroundings.

2 Implement an intuitive object location grid system within the application to provide precise guidance to visually impaired individuals, allowing them to navigate their environment with confidence by pinpointing the location of identified objects.

3 Integrate cutting-edge speech recognition technology into the application, enabling seamless interaction through voice commands, thereby enhancing accessibility and usability for visually impaired users who may face challenges with traditional input methods.

4 Utilize state-of-the-art text-to-speech synthesis to provide dynamic auditory feedback.

Methodology:

The proposed system consists of following listed modules, which collectively form the essential components of the mobile application, utilizing specific algorithms, libraries, and functionalities to deliver a comprehensive and effective solution for assisting visually impaired individuals.

1. **Camera Input Module:** This module manages the integration of the Smartphone's camera functionality into the application for capturing real-time images or video frames, Android CameraX API or Camera2 API used for camera integration.

2. **Object Detection Module:** Handles the identification and localization of objects within captured images or video frames using computer vision techniques, TensorFlow Lite with pre-trained models like MobileNet, YOLO (You Only Look Once) used for object detection.

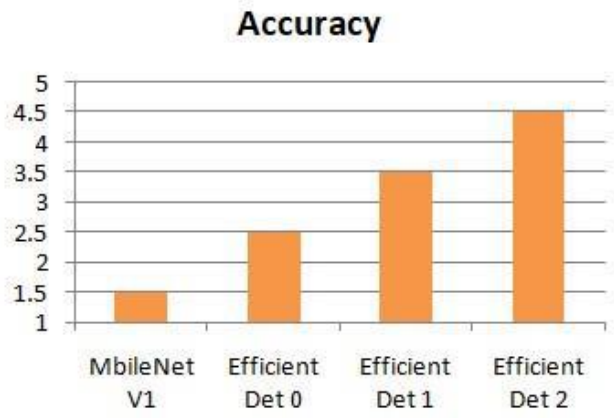
3. **Voice Feedback Module:** Converts text or information into audible output to provide real-time feedback to users based on detected objects or surroundings. Text-to-speech (TTS) libraries like Google Text-to-Speech API or Android's Text-To-Speech class for converting text into speech.

4. **User Interface Module:** Manages the graphical interface of the application, ensuring ease of use and accessibility for visually impaired users. Android's User Interface (UI) components, possibly utilizing accessibility features like high contrast, large text, or voice commands.

5. **Testing and Validation Module:** Conducts rigorous testing and validation procedures to ensure the functionality, accuracy, and reliability of the application. Testing frameworks like Espresso or JUnit used for functional testing, along with manual validation methods for user experience and accessibility.

Result:

The accuracy graph illustrates the performance of each model variant, including MobileNet, EfficientNet-B0, EfficientNet-B1, and EfficientNet-B2, as they progress through the given dataset. The graph demonstrates a consistent upward trend in accuracy for each model variant, indicating an improvement in recognition capabilities over time.



This upward trajectory underscores the effectiveness of the models in accurately identifying objects within the dataset, with each successive variant showing incremental enhancements in accuracy. Such a visualization is invaluable for understanding the comparative performance of different model architectures and guiding decisions regarding model selection and optimization strategies.

Conclusion:

This project focused on the development of a computer vision-based mobile application tailored for visually impaired individuals, aiming to enhance their accessibility and independence in navigating the world around them. Through the

integration of advanced technologies such as object detection, speech-to-text, and text-to-speech, the application successfully provided real-time support and assistance to users in identifying objects, interpreting commands, and receiving auditory feedback. The main findings underscored the effectiveness of the application in facilitating seamless interaction for visually impaired individuals, fostering inclusivity and empowerment within their daily lives.

Furthermore, the project's outcomes demonstrated significant contributions to the field of assistive technology, showcasing the potential of computer vision and deep learning algorithms in addressing the unique needs of visually impaired individuals. However, it's important to acknowledge certain limitations that impacted the project's outcome, such as constraints in real-world testing environments and the need for further refinement to improve the accuracy and efficiency of the application. Despite these limitations, the project holds substantial significance in advocating for accessibility and inclusivity, paving the way for future advancements in assistive technology aimed at enhancing the quality of life for individuals with visual impairments.

Description of the innovation in the project: -

The innovation in this project lies in the integration of computer vision, deep learning, and OCR technology into a specialized mobile application designed for individuals with visual impairments. By leveraging the smartphone's camera, the app identifies and describes objects in real-time, providing users with a better understanding of their surroundings. Additionally, the OCR feature converts printed text into audio, expanding access to printed materials. Overall, this innovation aims to improve accessibility and support for individuals with visual impairments in various aspects of their daily lives.

Future work scope: -

Several potential enhancements could be implemented to further improve the computer vision-based mobile application for visually impaired individuals. Firstly, enhancing the accuracy and efficiency of object detection algorithms through continuous refinement and training on diverse datasets could significantly improve the application's performance in identifying various objects in different environments. Additionally, integrating machine learning models to personalize user experiences by learning from user interactions and preferences could enhance the application's adaptability and responsiveness to individual needs.

Furthermore, expanding the application's capabilities to include features such as scene recognition, facial recognition, and currency identification could further enhance its utility for users in navigating their surroundings and accessing essential information. Moreover, incorporating augmented reality (AR) technology to provide spatial audio cues and immersive auditory feedback could offer users a more intuitive and immersive experience in interacting with the environment.

In terms of accessibility, implementing compatibility with braille displays and refreshable braille output devices could cater to users with different preferences and needs. Moreover, localization and language support for a wider range of languages could broaden the application's accessibility and usability for users worldwide.

Lastly, fostering collaboration with relevant stakeholders such as accessibility organizations, researchers, and end-users through user-centred design methodologies and participatory design workshops could ensure that future enhancements align closely with the evolving needs and preferences of visually impaired individuals, ultimately leading to a more inclusive and empowering user experience.