International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2393 - 2400
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0354

An AI-Powered Intermediate Accessibility App for Visually Impaired Users with Real-Time Voice and Vibration Support

Aiswariya R P[1], Dharshini A[2], Gayathri R[3], Vigneswari D[4], Dhivakar B [5], Sabarish M[6], Anita Dorothy T[7]
[1,2,3,4,5,6]UG – Artificial Intelligence and Data Science, Saranathan College of Engineering, Trichy, India.
[7]AP, Artificial Intelligence and Data Science, Saranathan College of Engineering, Trichy, India.
Emails: *aiswariya0308@gmail.com[1], dharshiniannadurai0910@gmail.com[2], gayathrirajendran3967@gmail.com[3], vigneswari11173@gmail.com[4], dhivakarb1807@gmail.com[5], msabarish1210@gmail.com[6], anitadorothy7086@saranathan.ac.in[7]*

Abstract
This paper presents the design and implementation of an AI-powered, multilingual smart assistant application aimed at enhancing smartphone accessibility for visually impaired individuals. Acting as an intermediate control layer, the application enables complete hands-free operation of essential mobile applications including WhatsApp, YouTube, Contacts, and Maps through intuitive voice commands. The system integrates a suite of assistive technologies to improve user autonomy and safety: voice-based app navigation; real-time emotion recognition; object and product identification using OCR and barcode/QR scanning; and fall detection with emergency alerts. A key innovation lies in the integration of a compact wearable camera attached to the user's clothing and connected via Bluetooth to the smartphone which captures images of the surrounding environment. These images are processed on the mobile device, enabling real-time feedback through speech synthesis. The application supports multilingual interaction in English, Tamil, and Hindi, enhancing accessibility for diverse user groups. Developed using the Flutter framework, the system ensures cross-platform compatibility and is optimized for devices with limited hardware capabilities. This work contributes a scalable and inclusive solution that leverages artificial intelligence and wearable technology to empower visually impaired users. By bridging the gap between standard mobile interfaces and assistive needs, the proposed system promotes independence, mobility, and digital equity.

Keywords: *Assistive technology, voice user interface, Bluetooth camera, visually impaired, accessibility, emotion recognition, OCR, wearable devices.*

1. Introduction
The widespread adoption of smartphones has significantly transformed the way individuals communicate, navigate, and access information. However, for the estimated 39 million people who are blind and the additional 285 million with moderate to severe visual impairments, as reported by the World Health Organization (WHO), interacting with mobile technology presents a considerable challenge. Despite the availability of built-in accessibility tools such as screen readers and voice assistants, many of these solutions offer limited functionality, lack contextual awareness, or require significant user training factors that often hinder complete independence. To address these limitations, recent advancements in artificial intelligence (AI), computer vision, natural language processing (NLP), and wearable technology offer promising opportunities to create more intuitive and responsive accessibility solutions [1]. This paper introduces an AI-powered, voice-controlled smart assistant application designed to serve as an intermediate accessibility layer for visually impaired smartphone users [2]. The system provides seamless voice-based interaction with key mobile applications such as WhatsApp, YouTube, Contacts, and Google Maps allowing for complete hands-free operation. The proposed solution

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2393 - 2400
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0354

incorporates several intelligent features that extend beyond conventional accessibility tools [3]. These include real-time voice-based app control, emotion recognition through facial expression analysis, object and product identification using optical character recognition (OCR) and barcode/QR code scanning, and voice-guided navigation. Additionally, the application includes automatic fall detection and emergency alert capabilities, enhancing personal safety. A distinctive component of this system is the integration of a compact, wearable camera affixed to the user's clothing and connected to the smartphone via Bluetooth. The camera captures environmental images on command, which are then processed locally on the smartphone to extract relevant information. This design provides a more flexible and context-aware user experience, enabling visually impaired individuals to understand and interact with their surroundings more effectively. The application supports multilingual interaction in English, Tamil, and Hindi, making it accessible to a diverse population. Built using the Flutter framework, the system is designed to operate efficiently on low- to mid-range devices, ensuring broad accessibility without requiring high-end hardware [4]. In summary, this paper presents a scalable and intelligent accessibility solution that leverages AI and wearable technology to improve digital inclusion, autonomy, and safety for visually impaired users. The following sections describe the system architecture, implementation details, and evaluation results that demonstrate the effectiveness of the proposed approach [5].

2. Problem Statement

Despite the proliferation of smartphones and advancements in mobile technology, visually impaired individuals continue to face significant barriers in accessing and interacting with mobile applications [6]. Existing accessibility tools, such as screen readers and basic voice assistants, provide only limited functionality and often require complex navigation procedures, which can be unintuitive and cumbersome for users with severe visual impairments. Furthermore, these tools typically lack contextual awareness, emotional feedback, and real-time environmental interpretation, which are critical for promoting true independence. Current systems also fall short in providing integrated safety features such as fall detection and emergency alerting, which are essential for supporting visually impaired users in dynamic, real-world environments [7]. Moreover, while some assistive applications support voice commands, few allow seamless control over third-party apps like WhatsApp, YouTube, and Maps. Additionally, many solutions are language-restrictive, excluding non-English-speaking users from fully benefiting from accessibility technology. Another critical limitation is the absence of wearable technology integration. The lack of a portable, real-time visual input mechanism prevents users from gaining situational awareness about their surroundings, such as identifying objects, recognizing emotions, or navigating unfamiliar spaces. Thus, there is a clear need for a comprehensive, AI-driven, voice-controlled, and multilingual mobile application that not only offers full smartphone control but also integrates environmental sensing through wearable hardware, enabling real-time interaction, navigation, and safety for visually impaired users. This paper addresses this gap by proposing a scalable solution that bridges modern mobile functionality with advanced assistive support tailored specifically for the visually impaired [8].

3. Literature Review

Over the past decade, researchers and developers have explored various technologies to enhance accessibility for visually impaired individuals. Traditional solutions such as screen readers, including TalkBack for Android and Voice Over for iOS, provide basic auditory feedback by reading out on-screen content. However, these tools rely heavily on gesture-based input and offer limited functionality when navigating complex or third-party applications. Voice assistants like Google Assistant, Siri, and Alexa have introduced hands-free interaction with smartphones, but their capabilities are often restricted to predefined commands and lack deep integration with third-party apps. Additionally, they offer minimal support for multilingual users, which is a barrier in linguistically diverse regions such as India. Several AI-based mobile applications have emerged to address object recognition and text reading challenges. Apps like Seeing AI by Microsoft and Lookout by Google use computer vision to describe scenes, read documents, and recognize currency. While these apps offer significant improvements, they typically require users to manually switch modes or interact with the screen, limiting true hands-free usability. For real-time navigation, wearable technologies such as the OrCam MyEye and Aira Smart Glasses provide auditory guidance and object identification. However, these devices are often expensive, hardware-dependent, and not widely accessible in low-resource settings. Furthermore, these solutions do not offer full smartphone control or emotional context recognition. Fall detection and emergency alert systems have also been explored in prior works. Research by Nguyen et al. utilized smartphone accelerometers for detecting abnormal motion patterns, while other studies proposed wearable sensor-based methods. Nonetheless, integration of such features into a unified mobile application remains limited. The integration of Bluetooth-connected wearable cameras for environmental understanding remains relatively underexplored in the field of mobile accessibility. Such cameras, when combined with AI-driven analysis on smartphones, can significantly enhance situational awareness by identifying objects, interpreting scenes, and even recognizing emotions.

While each of these technologies addresses specific challenges faced by visually impaired individuals, few offer an integrated, voice-controlled, real-time system capable of handling multiple accessibility tasks-including app navigation, object and emotion recognition, and safety alerts-in multiple languages. This paper addresses these limitations by proposing a comprehensive, Flutter-based mobile solution that combines AI, voice recognition, wearable input, and multilingual support to enhance digital independence for visually impaired users.

4. Proposed Solution

To address the challenges faced by visually impaired individuals in interacting with smartphones and navigating their environments, this paper proposes a comprehensive, AI-powered, voice-controlled smart assistant application. The system is designed to serve as an intermediate accessibility layer, enabling seamless interaction with both native and third-party applications through natural language voice commands, without requiring visual or tactile input [9].

4.1 System Overview

The proposed solution integrates multiple assistive features into a single mobile application built using the Flutter framework for cross-platform compatibility. The core functionality is centered around voice interaction, where users can initiate and control various smartphone operations such as opening apps, reading and replying to messages, placing calls and consuming media content [10]. The application is enhanced with a Bluetooth-connected wearable camera, affixed to the user's clothing, which captures real-time images of the environment. These images are transmitted to the mobile device where AI-powered models perform tasks such as object detection, text recognition (OCR), and facial expression analysis to recognize the emotional state of people in the camera's view. The resulting information is conveyed to the user through speech synthesis in their preferred language [11].

4.2 Key Features

Voice-Based App Navigation: Users can open and control applications like WhatsApp, YouTube, Contacts, and Maps using simple voice commands, enabling full device usability without screen interaction. Emotion Recognition: Through facial analysis of individuals in the camera's field of view, the system detects emotional expressions (e.g., happy, sad, angry) and informs the user audibly. Object and Product Identification: Leveraging OCR and barcode/QR scanning, the application identifies objects, products, and printed text, and provides the user with details such as product name, price, expiry date, and more [12].

4.3 Gesture-Based Input Support

For users with speech impairments or in environments where speaking is impractical, gesture-based input using the smartphone camera or wearable sensors could be incorporated as an alternative interaction mode. Fall Detection and Emergency Alerts: Using the smartphone's accelerometer and gyroscope sensors, the system detects falls and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2393 - 2400
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0354

immediately sends location-based alerts to predefined emergency contacts. Multilingual Interaction: Supports natural language processing in English, Tamil, and Hindi, allowing users to communicate with the system in their native language for improved usability. Figure 1 shows App Launching Page.

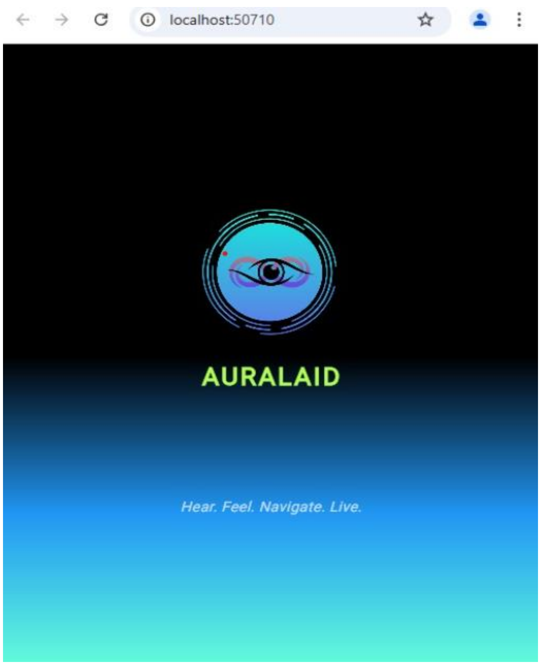

Figure 1 App Launching Page

4.4 Wearable Camera Integration

A key innovation in the system is the incorporation of a compact, Bluetooth-enabled wearable camera. This device acts as a vision input tool, extending the smartphone's ability to "see" the environment on behalf of the user. The captured visual data is processed using on-device machine learning models, ensuring privacy and reducing latency. This design provides users with real-time insights about their surroundings and the people within them [13].

4.5 Accessibility and Affordability

The application is optimized for devices with limited hardware resources, ensuring that it can run effectively on entry-level smartphones. By using a cross-platform development framework and offloading intensive processing to lightweight AI models, the system maintains performance while remaining cost-effective. Figure 2 shows App Main Page.

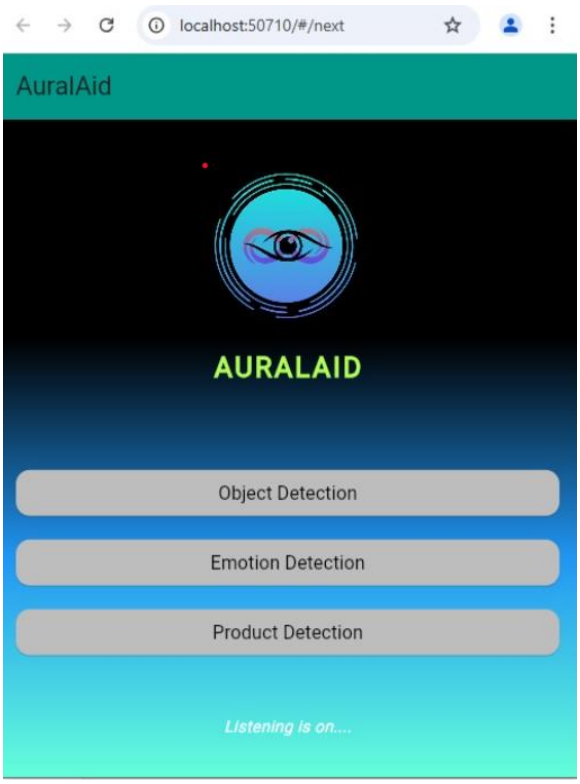

Figure 2 App Main Page

5. System Architecture and Data Flow

5.1 System Architecture Overview

The proposed system architecture is composed of three main components:

- **User Interface Layer:** Provides the voice-based interaction interface using speech-to-text (STT) and text-to-speech (TTS) modules. Facilitates multilingual communication and interprets voice commands in English, Tamil, and Hindi.

Processing & Intelligence Layer: Core logic of the application that handles:

- Natural language processing (NLP)
- Object and text recognition
- Emotion detection
- Fall detection logic

Integrates with on-device AI models and external APIs as needed. Controls flow between commands, wearable device input, and app functions.

- **Hardware & Sensor Layer:** Consists of the smartphone sensors (accelerometer, gyroscope, GPS) and a Bluetooth-connected wearable camera. Captures environmental visuals for

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2393 - 2400
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0354

processing. Collects motion and location data for fall detection and emergency response. Figure 3 shows Object Detection.

Figure 3 **Object Detection**

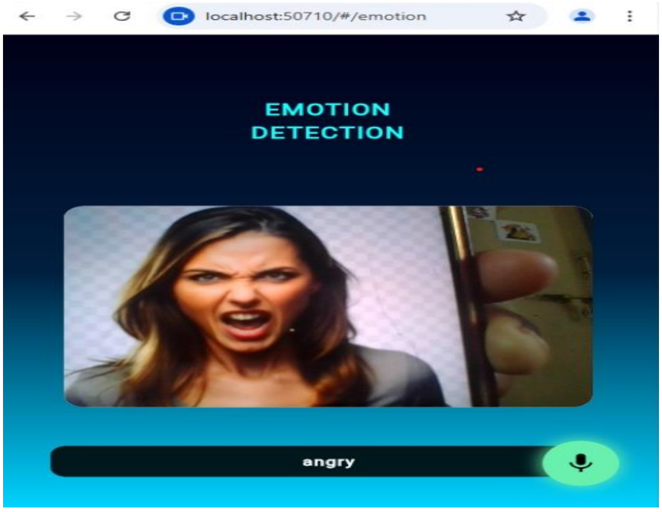

Figure 4 **Emotion Detection**

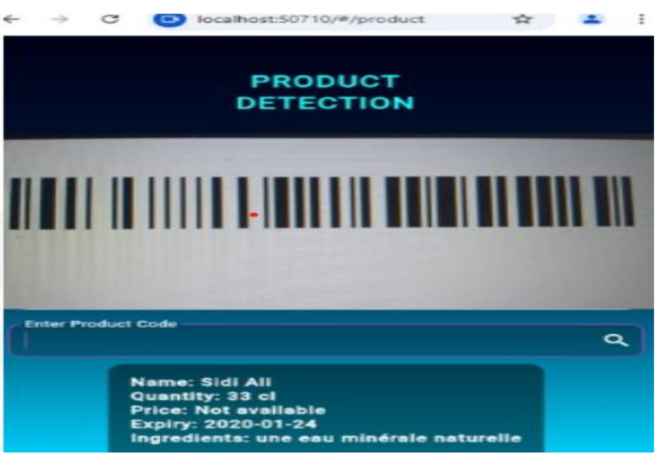

Figure 5 **Product Detection**

5.2 Data Flow Diagram

Here's the step-by-step flow of data through the system:

- **Voice Input:** The user gives a voice command (e.g., "Read my messages," "Where am I?"). The microphone captures the input and sends it to the STT engine. Figure 4 shows Emotion Detection.
- **Command Processing:** The STT engine converts the voice input into text. The NLP engine interprets the intent and selects the corresponding action (e.g., open WhatsApp, start navigation).
- **Wearable Camera Activation (If Required):** For tasks such as object recognition or emotion detection, the mobile app triggers the Bluetooth-connected wearable camera. The camera captures images and transmits them to the smartphone.
- **AI Processing:** On-device AI models perform: Object detection and OCR. Emotion recognition using facial expression analysis. Barcode/QR decoding. Results are formatted into understandable speech output.
- **Fall Detection (Parallel Process):** The motion sensors continuously monitor user movement. If a fall is detected, an alert with GPS location is automatically sent to emergency contacts.
- **Response Output:** The TTS engine converts results into spoken feedback (e.g., "You are at Main Street. The person in front of you looks happy."). Optional haptic feedback can be provided via vibration for alerts. Figure 5 shows Product Detection.

6. Advantages of The Proposed Solution

The proposed AI-powered intermediate accessibility application offers several key advantages over existing solutions, especially in terms of usability, accessibility, safety, and affordability for visually impaired users:

- **Complete Hands-Free Operation:** The application supports full voice-based control, allowing visually impaired users to operate smartphones without the need for screen interaction. This enables independent access to third-party apps such as WhatsApp, YouTube, Contacts, and Maps, greatly enhancing digital inclusion [14].

- **Real-Time Environmental Awareness:** The integration of a Bluetooth-enabled wearable camera provides continuous visual input. Combined with on-device AI models for object recognition, text reading (OCR), and emotion detection, the system offers real-time awareness of the user's surroundings and social context.
- **Enhanced Safety Features:** The fall detection module continuously monitors user movement using accelerometer and gyroscope data. In the event of a fall, the system triggers immediate alerts to emergency contacts with GPS location, improving user safety and response time.
- **Multilingual Support:** Unlike many existing applications that are limited to English, this solution supports multiple regional languages including Tamil and Hindi. This enhances accessibility for a broader user base, particularly in linguistically diverse regions.
- **Affordable and Scalable:** Built using Flutter, the application is cross-platform and optimized for low-resource devices, making it accessible even on entry-level smartphones. The use of cost-effective wearable hardware ensures scalability and affordability for deployment in rural and urban communities alike.
- **Seamless Integration with Existing Apps:** The intermediate accessibility layer enables the user to interact with existing mobile applications through voice commands, eliminating the need for dedicated accessible versions of each app.
- **Privacy-Preserving and Offline Capability:** Processing is done locally on the device for tasks such as image recognition and NLP, ensuring user data privacy and enabling offline functionality in low-connectivity environments.

7. Future Work and Enhancement

While the proposed solution presents a robust and feature-rich accessibility layer for visually impaired users, there are several areas identified for future development and improvement. These enhancements aim to increase system intelligence, adaptability, and integration with emerging technologies:

- **Integration of AI-Powered Voice Assistant with LLMs:** Future versions of the application could integrate advanced Large Language Models (LLMs) to enable more natural and intelligent conversational interactions. This would allow the system to better understand complex user commands and context, and provide more human-like, personalized assistance.
- **Indoor Navigation Using Beacon Technology:** While current navigation relies on GPS, it may be limited indoors. Future enhancements could include the use of Bluetooth beacons or Wi-Fi fingerprinting to enable precise indoor navigation, helping users move safely within large buildings such as malls, airports, and hospitals.
- **Cloud-Based Processing for Advanced Recognition Tasks:** Though the current version performs processing locally to support offline functionality, future iterations could offer optional cloud-based processing for more computationally intensive tasks such as scene description, multi-object detection, and facial recognition at scale.
- **AI-Powered Learning and Personalization:** Machine learning algorithms could be employed to learn user preferences over time and automatically adapt responses, feedback speed, language usage, and frequently used commands. This would create a more personalized user experience.
- **Smart Wearable Integration:** Future updates may explore integration with smart glasses or haptic feedback wearables to provide more immersive feedback. For example, vibration cues could help with object localization or indicate directional changes during navigation [15].
- **Community and Caregiver Dashboard:** A web-based or companion mobile dashboard

for caregivers or family members could be developed to monitor user location, receive fall alerts, and assist remotely with navigation or communication when needed.

Conclusion

This paper presents a comprehensive, AI-powered accessibility solution designed to empower visually impaired individuals through real-time voice, vision, and vibration support. By integrating a multilingual, voice-controlled mobile application with a Bluetooth-enabled wearable camera, the system provides users with hands-free access to smartphone functionalities, enhanced environmental awareness, and critical safety features such as fall detection and emergency alerts. The application not only enables seamless interaction with third-party apps using voice commands but also leverages on-device AI for object recognition, text reading, and emotion detection. These features collectively contribute to a higher degree of independence and digital inclusion for visually impaired users. Built using the Flutter framework, the system ensures cross-platform compatibility and affordability, making it accessible to a wide range of users, including those in low-resource settings. Its multilingual support further broadens its usability across diverse linguistic backgrounds. The proposed solution represents a significant step toward inclusive technology design, transforming smartphones into intelligent assistive devices. With continued enhancements and integration of emerging technologies, this solution has the potential to evolve into a fully adaptive assistive ecosystem, bridging the digital divide and promoting autonomy and safety for the visually impaired community.

Acknowledgements

We would like to express our sincere gratitude to Mrs. Anita Dorothy T, Assistant Professor, Department of Artificial Intelligence and Data Science, Saranathan College of Engineering, Trichy, for her invaluable guidance, continuous support, and encouragement throughout the development of this project. Her expertise and insights greatly contributed to the successful completion of our research work. We also extend our heartfelt thanks to the management and faculty members of Saranathan College of Engineering for providing us with the necessary resources and a conducive environment to carry out this work. Finally, we are deeply thankful to our families and friends for their constant motivation and support during the course of this project.

References

[1]. World Health Organization (WHO). 2014. Visual Impairment and Blindness. [online] available: http://www.who.int/mediacentre/factsheets/fs282/en/ [accessed 31 May 2017]

[2]. Wordpress.com. 2015. How does technology help people who are blind or visually impaired? [online] available: https://sandysview1.wordpress.com/2015/05/07/how-does-technology-help-people-who-are-blind-or-visually-impaired/ [accessed 31 May 2017]

[3]. knfbreader.com. Introduction the KNFB Reader, [online] available: http://www.knfbreader.com/ [accessed 31 May 2017]

[4]. bookShare: Read2Go, [online] available: https://www.bookshare.org/cms/help-center/reading-tools/read2go [accessed 31 May 2017]

[5]. areTec - Products for the Blind and Visually Impaired. Colorino Talking Color Identifier and Light Probe. [online]available: http://www.caretec.at/ColorTest_Colorino.32.0.html [accessed 31 May 2017]

[6]. D. Kendrick, "PenFriend and Touch Memo: A Comparison of Labeling Tools", Access World Magazine, Vol. 12, No. 9, 2011.

[7]. Polish Association of the Blind. Seeing Assistant home. [online], available: http://seeingassistant.tt.com.pl/home/ [accessed 31 May 2017]

[8]. Tap See. Blind and Visually Impaired Assistive Technology. [online], available: httphttp://taptapseeapp.com/ [accessed 31 May 2017]

[9]. OpenCV 2.4.13.2 documentation. OpenCV4Android SDK, [online] available: http://docs.opencv.org/2.4/doc/tutorials/intro

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2393 - 2400
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0354

duction/android_binary_package/O4A_SDK.html [accessed 31 May 2017]
[10]. GitHub: ENCP/CNNdroid. [online] available: https://github.com/ENCP/CNNdroid [accessed 31 May 2017]
[11]. Cat choom: CraftAR Pro SDK. [online] available: https://catchoom.com/product/craftar/craftar-pro-sdk/ [accessed 31 May 2017]
[12]. Android Developers: Android Studio the Official IDE for Android, [online] available: https://developer.android.com/studio/index.html [accessed 31 May 2017]
[13]. w3schools: PHP Create MySQL Database. [online] available: https://www.w3schools.com/php/php_mysql_create.asp [accessed 31 May 2017]
[14]. MathWorks: MATLAB. [online] available: https://www.mathworks.com/products/matlab.html [accessed 31 May 2017]
[15]. K. J. Thakkar, "What is OpenCV? OpenCV vs. MATLAB An insight" (Blog), 2012.available: https://karanjthakkar.wordpress.com/2012/11/21/what-is-opencv-opencv-vs-matlab/ [accessed 31 May 2017]