MJIT 2024 **Malaysian Journal of Industrial Technology**

SUPPORTING PATIENTS OF AGE-RELATED MACULAR DEGENERATION IN MALAYSIA USING MOBILE APPLICATION ENABLED WITH OBJECT AND TEXT RECOGNITION

H. P. Fung.
Technology Consulting, Hewlett Packard Enterprise, Kuala Lumpur, Malaysia.
han-ping.fung@hpe.com

*Corresponding author's email: han-ping.fung@hpe.com

ARTICLE INFO

Handling Editor: Rahimah Mahat

Article History:
Received 21 May 2024
Received in revised form 2 August 2024
Accepted 25 August 2024
Available online 16 September 2024

Keywords:
Age-related macular degeneration; Mobile application; Object recognition; Text recognition; Speech generation.

ABSTRACT

When people are diagnosed with Age-related Macular Degeneration (AMD) disease, object or text recognition through their eyesight pose a challenging task as they are not blind but lose the central vision. Despite various assistive technologies available to help AMD patients to see better in advanced or western countries but there is lack of research how lightweight mobile application with object and text recognition capability to help Malaysian AMD patients especially in situations without Internet connection. The research objective is to understand Malaysian AMD patient's requirement then develop a simple but useful mobile application with object and text recognition capability to generate audio speech for patients to perceive and understand their surrounding environment better. After literature reviewed, appropriate machine learning models are selected and trained for object and text recognition functionality. Then the mobile application is developed using Kotlin programming language and Android Studio before series of testing are conducted. Three AMD patients also are selected to test the mobile application. The developed mobile application yields mean Average Precision (mAP) of 42% for object recognition and 83% for text recognition. Despite lower mAP for object recognition however through Unit and Usability Testing conducted on AMD patients, the test results met their requirements and the mobile application did support them to have a better grasp of their surroundings. This research contributed a small step in relieving Malaysian AMD patients to see better and live more independently. Limitations and future improvements were also provided.

1.0 Introduction

Eye-related disease named Age-related Macular Degeneration (AMD) is one of the leading causes of vision loss and it is expected to affect 288 million people by 2040 [1]. AMD disease has caused loss of central vision of many aging patients whereby they themselves are not fully blind [2]. The following two pictures in Figure 1 depict a vision of a normal person and the vision of an AMD patient.

Figure 1. Normal Person and AMD Patient Visions

Due to AMD, patients need to adjust their head or sometimes even their body movement so that they can see or walk better. This vision impairment, awkward physical body adjustment and loss of independence have significant impact on AMD patients' self-reliance and self-confidence [3], [4]. In order to support AMD patients to see better, some assistive tools or technologies have been invented like low vision devices e.g. magnifying glass, artificial intelligence, machine learning, virtual reality headset, objects and images recognition software like Microsoft Seeing AI, Google Lens et al. [5], [6], [7], [8].

There is research on how the above tools and technologies can help vision impaired and AMD patients in developed or western countries to see better [9], [10], [3], [4], [6], [11], [12]. However, problem statement is still lack of empirical study how a mobile application, in the absence of Internet with object and text recognition functionalities capable to help AMD patients to see better especially in developing countries like Malaysia. The mobile application will be developed to run on Android phone because more than 65% of Malaysian are Android phone users [13]. Although this research is conducted in English but most of the Malaysians are multi-linguals whereby English generally is not their first language especially for those above 60 years old. Hence, it is interesting to evaluate how AMD patients in Malaysia are reacting towards the mobile application developed and administered in English. In addition, AMD patients above age of 60 in Malaysia are generally at retirement age whereby their financial independence and socioeconomic status are slightly different from their counterparts in developed or western countries who are still under employment or below the retirement age. Moreover, this mobile application should be light-weight, user friendly and helpful so that AMD patients still can perceive properly despite unavailability of Internet connection or telecommunication signal in certain environments e.g. in sheltered rooms, buildings, car parks, remote locations et al.

The objective of this applied research is to understand and develop a mobile application based on object and text recognition functionalities in helping AMD patients to perceive better their surrounding environment. Research questions for this study includes:

- RQ1: Generally, how useful is the mobile application based on object and text recognition capabilities in helping AMD patients to perceive the surrounding environment better?
- RQ2: Are the provided features or functions in the mobile application are ease of use?

Above research questions will be answered after the study is completed.

Following Figure 2 illustrates an overview of the mobile application architecture and design.

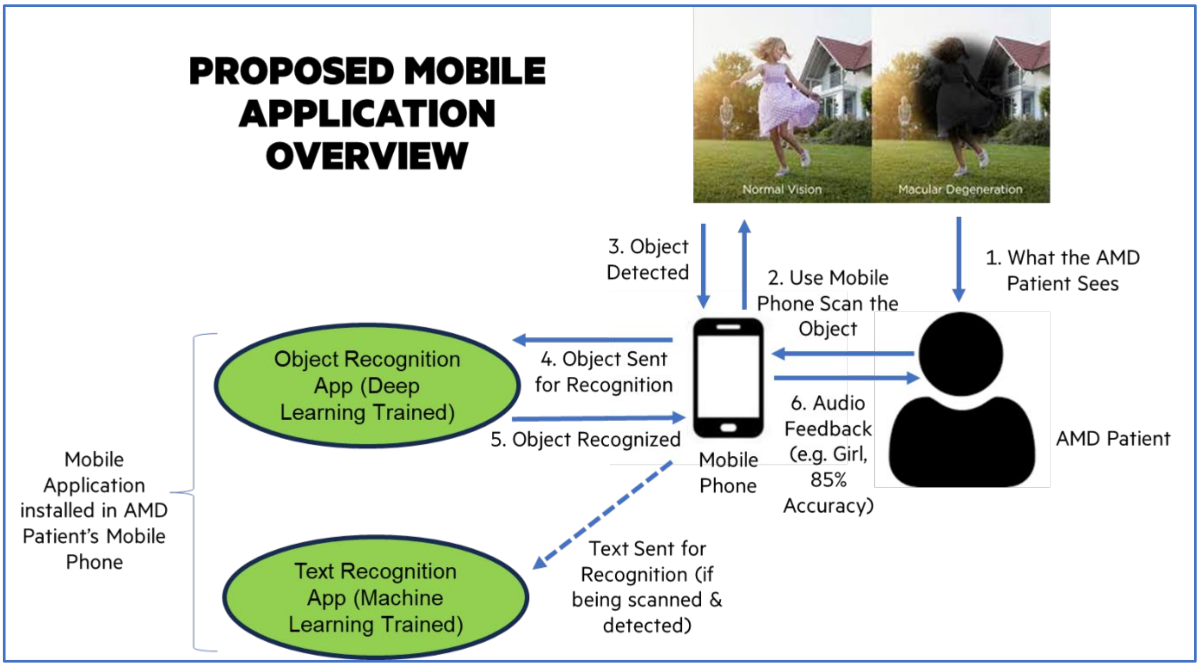

Figure 2. Proposed Mobile Application Overview

In order to compensate AMD patients for what they can't see properly, they can use their mobile phone to scan an object or text before letting the mobile application to process and respond with audio feedback on what they are seeing as illustrated from Step1 to 6 of the above Figure 2. Likewise, similar audio feedback will be generated when text is scanned by the AMD patient's mobile phone.

2.0 Literature Review

As we aged, some of us above 60 years old can suffer from Age-related Macular Degeneration (AMD) disease whereby the central part of our vision is lost [2]. The loss of central vision can continue to spread until total blindness covers both eyes. According to [2] besides aging, other risk factors like genetics, hypertension, overweight, smoking and imbalance dietary lacking carotenoids intake also contributing to AMD. Patients suffering from AMD are not immediately blind but rather losing the central part of the vision to black or grey spots whereby they need to adjust their head or body movement to see better. This vision impairment, awkward physical actions or positions and lack of independence have profound impact on AMD patients' self-reliance and self-confidence in terms of working, walking, reading, writing et al. as well as socioeconomic consequences. The loss of productivity and cost incurred to rehabilitate the impaired vision amounted to significant economic burden for the AMD patient, their family and society [3], [4]. As of now, there is no cure or way to reverse the damage already caused by AMD disease.

While there are different ways of AMD treatments to lessen the impact and preventive measures available from medical field; parallelly there are several assistive technologies being researched and deployed to support AMD patients to see or understand better their surrounding environment. From literature reviewed, these assistive technologies are divided into 3 categories: 1) Low vision devices, 2) Entity recognition, 3) Navigation systems. Low vision devices including magnifying glass or electronic visual enhancements et al. that can manipulate the image formed in the eye's retina by adjusting contrast, color correction, luminance, remapping and spatial frequency filtering [5]. Entity recognition includes object, text, color, image and people recognition and then converting them into audio or speech by leveraging Deep Learning or Machine Learning [6], [7], [8], [12]. Some of the simpler entity recognitions e.g. object and text recognition can operate without Internet whereas more complex entity recognitions like image, people, facial et al. require larger model size or even Internet connection. Some complex entity recognition running as mobile application includes Microsoft Seeing AI, Google Lens, Be My Eyes, virtual reality headset [14], [12]. Despite Deep Learning or Machine Learning is adopted in entity recognition, they are also being used to develop learning models to train, predict and diagnose AMD disease [15], [16]. In such use cases, they are not considered part of these assistive technologies for AMD patients to perceive their surrounding environment. Navigation systems sometimes are also being coined as traveling aid systems or walking assistants [12]. Regardless of the terminology, assistive navigation systems provide directions to users to navigate from one location to another. Navigation systems are complex as they encompass technologies that monitor the surrounding environment and rely on Global Positioning System (GPS) as well as require various methods of user interaction. Within the assistive technologies mentioned earlier, entity recognition specifically object and text recognitions are the focus of this applied research.

Estimated 6.8 billion users worldwide are using mobile phones in 2023 whereby more than 50% of them are actively using this tool for various activities in their daily lives [17]. One of the contributors towards the wide adoption of mobile phone is the proliferation of various applications development that produce different mobile applications installed into the mobile phone. Mobile phone as an indispensable tool to mankind is continuously improving throughout different generations which encompass improved capabilities in Short Messaging Service (SMS), Multi-media Messaging Service (MMS), video calling, Internet access, various application functionalities as well as increased data bandwidth [18]. Hence, mobile phones are suitable tools to install the object and text recognition application to better support AMD patients.

This applied research is to leverage object recognition through deep learning and text recognition through machine learning to develop an application that can run in AMD patients' mobile phones so that they can see better. When an AMD patient points his or her mobile phone with camera turned-on, any detected and recognized object or text will generate a speech to augment the AMD patient's handicapped vision. This research is also underpinning on theoretical framework named Technology Acceptance Model (TAM) whereby a user behavior is determined by the behavioral intention and in turn, the latter will be influenced by perceived usefulness and perceived ease of use [19]. Similarly, the perceived ease of use and usefulness of the mobile application for object and text recognition also can influence the AMD patient's behavioral intention and then their actual behavior.

In order to develop the mobile application that is capable of object recognition, TensorFlow Lite Machine Learning (ML) model is adopted because it is Open Source-based, lightweight and efficient ML that the application is easier to run in Android-based mobile phone compares to TensorFlow models [20]. The Android phone is selected instead of other mobile phone Operating System (OS) e.g. iOS, Windows et al. due to its comparatively lower cost of development and mobile device acquisition. In-built TensorFlow visualization tool i.e.

TensorBoard is used to support ML experimentation by visualizing and tracking accuracy and loss metrics. Common models adopted for object recognition include Convolutional Neural Networks (CNN), Single Shot Detector (SSD), You Only Look Once (YOLO) and EfficientDet [21], [22], [23], [24], [25]. Following Table 1 summarized their differences [26].

Table 1 – Key Differences among Different ML Models for Object Recognition

Feature	CNN	SSD	YOLO	EfficientDet
Architecture	Convolutional neural network	Single-shot multibox detector	One-stage	Compound scaling
Accuracy	Low to medium	Good	Good	High
Speed	High	High	Very high	High
Complexity	Low to medium	Medium	Low	High
Scalability	Low to medium	Good	Good	High

Based on the above table, EfficientDet is selected as the model used in this research due to its overall effectiveness compared to other models' capabilities in terms of accuracy, speed and scalability. Despite EfficientDet model is complex but it is still worth the effort to deploy in exchanging for achieving other beneficial outcomes to support AMD patient.

To recognize texts detected before converting them into speech in Android application, following two ML libraries are being considered in which there are differences in terms of architecture, integration and feature set:

1. Play Services ML Kit Text Recognition [27]
2. Firebase ML Kit Text Recognition [28]

Advantages of Play Services ML Kit include smaller application size that yields faster performance, and the application can be installed and run on mobile device without Internet connection [27]. On the other hand, Firebase ML Kit overcomes some of the disadvantages of Play Services ML Kit which include no cloud-based documentation recognition features and manual efforts are required to download or update the model into the mobile phone. However, Firebase ML Kit has its own disadvantages which include larger application size that require larger storage space in mobile phone, reliance on Internet connectivity as well as limited text recognition capability in offline mode [28]. Based on the above comparisons, Play Services ML Kit is a better option for text recognition deployment in mobile phone due to its smaller library size and work offline without Internet connection.

Camera2Api is Android low level software used to streamline the linking of application and mobile phone hardware's camera so that to successfully capture the AMD patient's surrounding in real time. LabelImg as Open Source and lightweight image annotation tool is used to label the boundary box of detected object in image [29]. Java-based programming language Kotlin and Python are used for application development and ML training due to their mobile phone OS flexibility, robustness, security rich features, built-in mobile development function, straightforward syntax and lower cost option [30], [31].

3.0 Research Methodology

3.1 Requirements

Based on interaction with some AMD patients that are also using mobile phone in their daily lives, requirements for the mobile application include: various object classes recognition, user's camera clearly detect one object or text at a time within 2 meters away from the user, capable to work without Internet connection, displaying boundary boxes around object detected, displaying confidence score and object name as well as discarding incorrectly detected recognition results lower than 70% of confidence score.

3.2 Object Recognition Model Preparation

Since EfficientDet model is selected as explained in literature review, this research will consider following four EfficientDet Lite models of different sizes that have already undergone sufficient training by the TensorFlow team.

1. Model1: lite-model_efficientdet_lite2_detection_metadata_1.tflite (size 7.4MB)
2. Model2: lite-model_efficientdet_lite3_detection_metadata_1.tflite (size 11.7MB)
3. Model3: lite-model_efficientdet_lite3x_detection_metadata_1.tflite (size 13.6MB)
4. Model4: lite-model_efficientdet_lite4_detection_metadata_2.tflite (size 20.4MB)

Each file from the above list is a merger of their respective TensorFlow Lite (TFLite) Model and its metadata file combined to enable seamless deployment and management of the TFLite models. These models are used to develop a mobile application that can assist AMD patients to listen and understand what is being detected and recognized by the application through their own mobile phones.

The following Figure 3 illustrates how the data format is being transformed into a Deep Learning Model that can be used in object recognition.

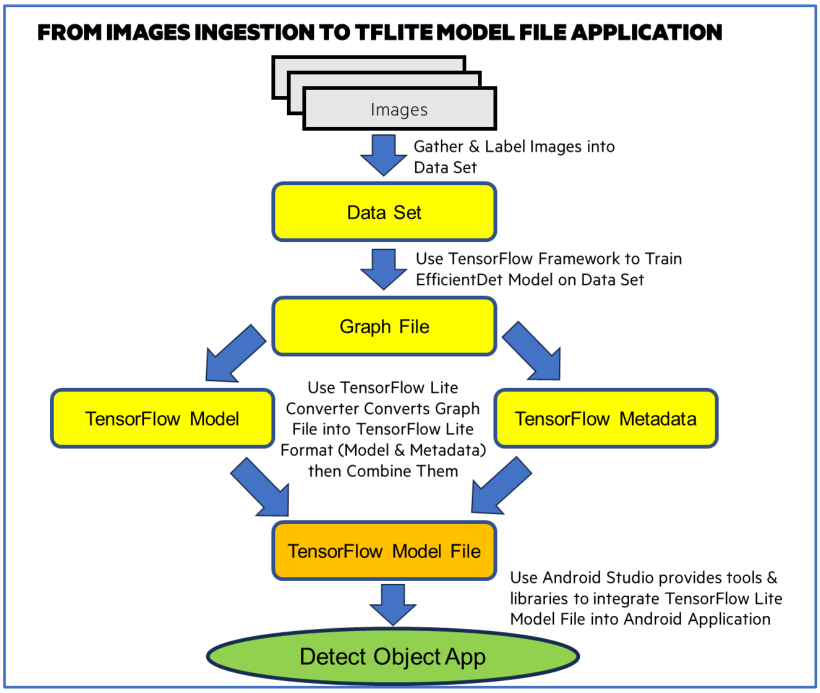

Figure 3. Process Flow for Object Recognition Model

Once the above TensorFlow Model File with file extension *.tflite is generated, Android Studio will provide tools and libraries to integrate the model file into Android applications. The tools will load the model file and perform object detection inferences and display the results on the mobile device screen as well as converting the result into speech.

3.3 Text Recognition Model Preparation

For text recognition, TextRecognizer from ML Kit under Google is adopted to perform Optical Character Recognition (OCR) on sign images or documents to extract the required text. Extensible Markup Language (XML) is used to store and transfer data in text-based format readable by both machine and human as well as to support information exchange between various computer systems to produce data structures used for application processing.

3.4 Mobile Application Design

In order to help AMD patients to recognize objects or texts in their surrounding environment, an Android-based mobile application will be developed to detect both the object and text and then convert them into speech. This mobile application will be developed using Kotlin programming language which makes coding concise, cross-platform and interoperate fully with Java. In addition, Kotlin is Google preferred programming language for Android applications development [30].

The following Figure 4 shows the wireframing of the mobile application.

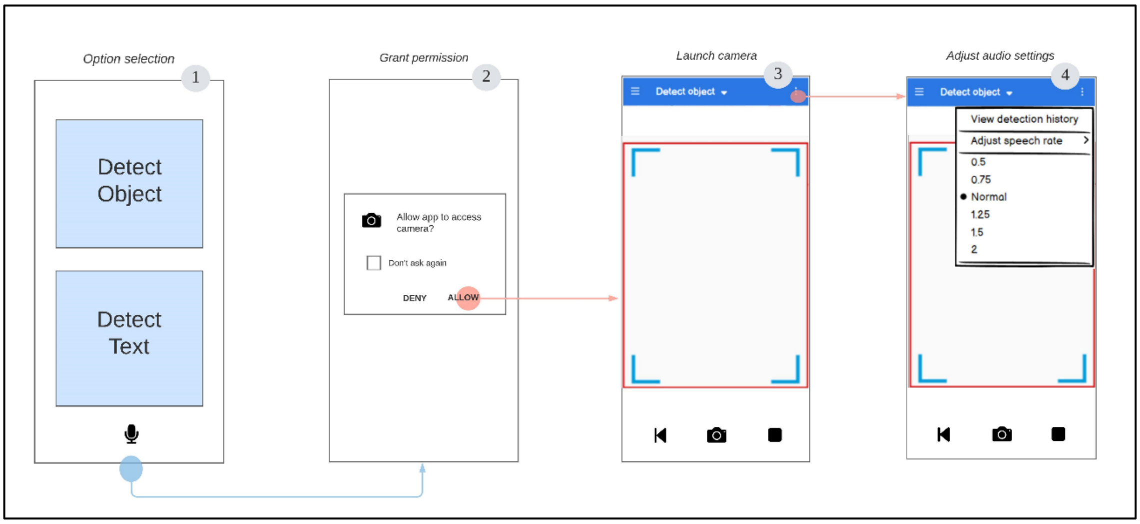

Figure 4. Wireframe of the Mobile Application

Upon the application is launched from the user mobile phone without Internet connection, a menu of two options is displayed i.e. detect object or detect text. Once an option is selected, the user can select “allow” to activate the camera function of the mobile phone before an object or text can be detected. Upon successful detection and recognition, a voice message will be generated to inform the AMD patient. For object recognition, confidence score also will be read out.

Figures below show the user interface design of the mobile application to detect object and text respectively.

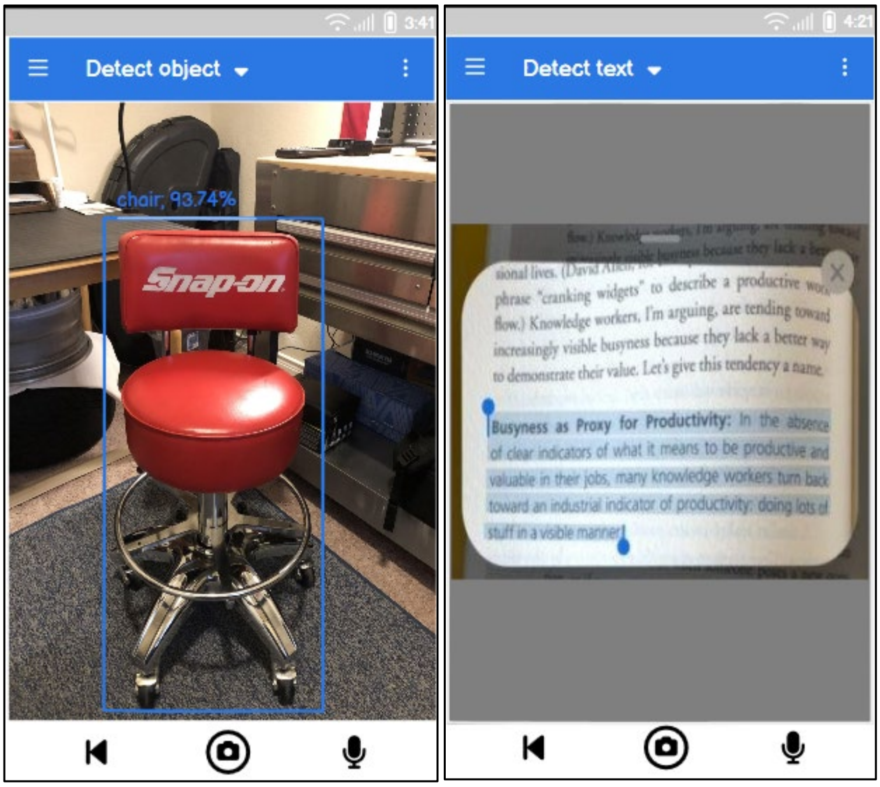

Figure 5 & 6. User Interface Design for Object and Text Recognition Screens

For the object detected as shown in Figure 5, a bounding box appears around the object on screen and labels it with the recognized name in addition to the computed confidence score. Similarly for text detection as shown in Figure 6, detected text will be picked out and highlighted on screen which will be read out by the application.

After the mobile application is developed, pilot tested and deployed into AMD patients' mobile phones successfully, three AMD patients are conveniently selected to use and evaluate the mobile application for both object and text recognition. Out of the three AMD patients, only two have Android phones. A separate Android phone was lent to the third patient for testing. All three patients initially took a while to adapt using the mobile phone to point and detect any object or text as they can't see clearly through their central vision. But after several attempts they managed to do so without assistance.

4.0 Research Operationalization

4.1 Object Recognition Model Selection

As previously explained, the chosen object recognition models being considered for implementation all have varying file sizes to determine if file size is a factor influencing model accuracy. Larger model file might generally have been trained on a larger dataset and may have captured more complex patterns in the data, this does not fully guarantee it will always outperform a smaller model [32]. However, since EfficientDet Lite models are swifter and efficient variants of the original EfficientDet family, it can be less accurate since it is operating on smaller input sizes as shared by [32].

The accuracy of a TFLite model might also depend on other factors including model architecture, weight precision, quality and quantity of training data, task complexity and training methods [32]. In certain situations, a larger TFLite model might indeed provide better accuracy due to its capability to capture more intricate details within the data. However, there is no direct relationship between file size and accuracy [32]. Based on the above explanations, all four models are selected and ingested one at a time into the mobile application to better evaluate their overall performance.

4.2 Text Recognition Model Selection

After comparing between "Play Services ML Kit Text Recognition" and "Firebase ML Kit Text Recognition", both of their libraries have its advantages and disadvantages but to meet this research requirement that is to develop a light application for AMD patients to detect texts using their mobile phones, the library should be smaller in size and can work offline without Internet or cloud connections. Hence, the "Play Services ML Kit Text Recognition" library is selected as pre-trained model for text detection and recognition.

4.3 Mobile Application Development

This mobile application is divided into three modules as depicted in the following Figure 7

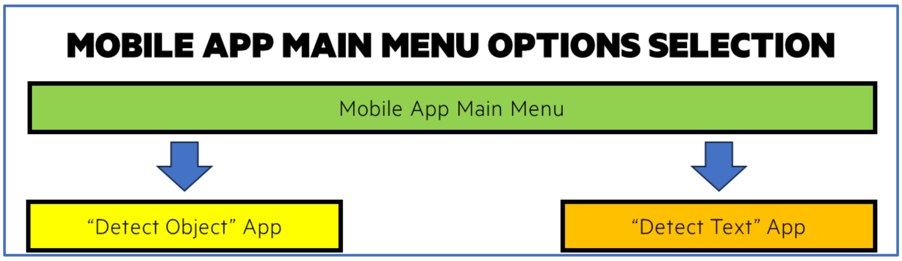

Figure 7. Process Flow of Main Menu Application

When an AMD patient activates the mobile application, the main menu module will be launched first to receive selection from the patient. Depending on the patient's selection, the main menu will connect to either object detection application or text detection application.

Following Figure 8 describes the process flow of the object detection application module.

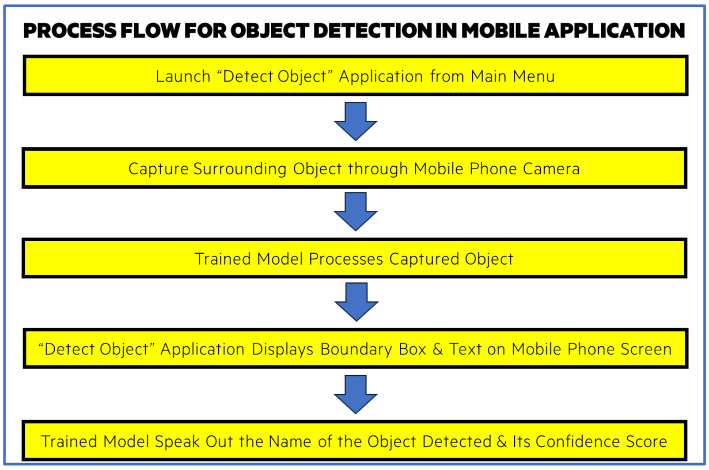

Figure 8. Process Flow of Object Detection Application

The following Figure 9 also describes the process flow of the text detection application module.

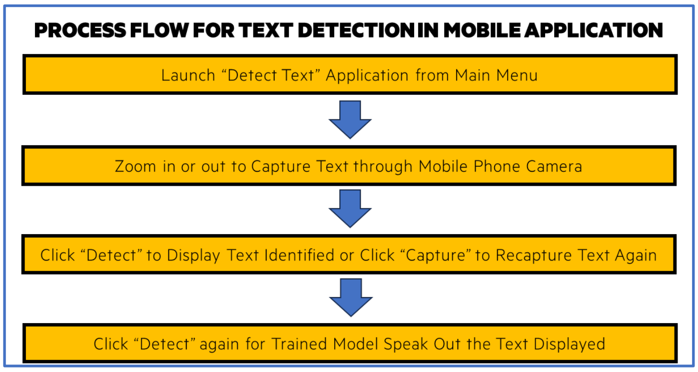

Figure 9. Process Flow of Text Detection Application

Upon the mobile application is successfully developed and synchronized into the user's mobile phone, a pilot test will be performed to test the application rigorously with some enhancements to the source codes. After the pilot test was completed successfully, three AMD patients were selected to test and evaluate the mobile application performance.

5.0 Results and Discussion

5.1 Object Recognition Model Performance

Following Table 2 depicts the Object Recognition Performance for all the 4 models tested.

Table 2. Object Recognition Model Performance

Model	Size (MB)	Latency (ms)	Average Precision (mAP)
EfficientDet-Lite2	7.4	69	33.97%
EfficientDet-Lite3	11.7	116	37.70%
EfficientDet-Lite3x	13.6	208	40.98%
EfficientDet-Lite4	20.4	260	41.96%

Above shown EfficientDet-Lite4 model exceeds other models in terms of mean Average Precision (mAP) despite having the largest model file size, highest latency and mAP value. After using the developed mobile application to test individual model, it shows file size and higher latency of a model doesn't significantly hinder its effectiveness in detecting and recognizing objects.

5.2 Text Recognition Model Performance

For the text recognition model performance based on Google's Play Services ML Kit Text Recognition library, mean Average Precision (mAP) of 82.73% which is considered acceptable as the library typically achieves mAP value in the range of 0.8 to 0.9 for text recognition tasks. TextRecognition library that process an image of text and return the recognized text yields latency of 221 milliseconds which is within the norm of less than few hundred milliseconds depending on the mobile phone capability and image size of the text. Above result findings indicate we can continue to adopt Play Services ML Kit model as the TextRecognition library in the "Detect Text" Application for this mobile application.

5.3 Mobile Application Testing

The following Table 3 and 4 showing unit testing for application module pertaining to object recognition and its audio feedback.

Table 3. Unit Testing of Object Recognition Module

Test Case	Test Steps	Test Data	Expected Result	Status
1) *Scan user's surrounding*	a) Launch application	User's current surrounding	App can consistently capture and display user's surrounding clearly.	Pass
2) *Detect* and *recognise* an *object* which yields a confidence score *higher than set threshold*	a) Capture a scene with one object	A scene of a chair, returning confidence score higher than 70%	App will produce a bounding box around the detected chair in addition to displaying the object's name and confidence score.	Pass
3) *Detect* and *recognise* an *object* which yields a confidence score *lower than set threshold*	a) Capture a scene with one object	A scene of a doll, returning a confidence score lower than 70%	Item is not detected by the system; app will not display any object name, bounding box and confidence score.	Pass
4) *No object* to be *detected* and *recognised*	a) Capture a scene without any item	A scene of a blank wall.	App will not display any object name, bounding box and confidence score.	Pass

Table 4. Unit Testing of Audio Feedback for Object Recognition Module

Test Case	Test Steps	Test Data	Expected Result	Status

1) *Detect* and *recognise* one *object* in a frame which yields a confidence score *higher than set threshold*	a) Capture user's current surrounding b) Process the captured frame	Detection result returns a chair where the confidence score is higher than 70%	"Chair with ≥70% accuracy" is spoken as audio feedback	Pass
2) *Detect* and *recognise* an *object* which yields a confidence score *lower than set threshold*	a) Capture user's current surrounding b) Process the captured frame	Detection result returns a doll with a confidence score lower than 70%	No audio feedback is provided by the system.	Pass
3) *No object* to be *detected* and *recognised*	a) Capture user's current surrounding b) Process the captured frame	Detection result returns null	App will not display any object name, bounding box and confidence score.	Pass

Following Table 5 and 6 presenting unit testing for application module pertaining to text recognition and its audio feedback.

Table 5. Unit Testing of Text Recognition Module

Test Case	Test Steps	Test Data	Expected Result	Status
Clearly *detect* and *recognise* text phrases *without spelling error*	a) Capture text phrases found within user's current surrounding b) Process the captured frame	Detection result returns the phrase "Puzzle for the Secret Seven" from a book cover	App is able to clearly display the phrase from captured photo without any spelling error.	Pass

Table 6. Unit Testing of Audio Feedback for Text Recognition Module

Test Case	Test Steps	Test Data	Expected Result	Status

Clearly *read aloud* detected text phrases *without reading error*	a) Capture text phrases found within user's current surrounding b) Process the captured frame	Detection result returns the phrase "Puzzle for the Secret Seven" from a book cover.	Audio feedback is provided where the system reads out "Puzzle for the Secret Seven" in English.	Pass

Result findings from the above tables confirmed the developed mobile application can perform the functions as per intended in terms of both object and text recognition. Moreover, positive feedback also received from the three AMD patients about no issue faced in accessing the application from their Android mobile phone as well as pointing their mobile phone at the right direction before an object or text being detected and scanned. In addition, the lightweight mobile application is operational without Internet connection making it useable at a time anywhere if there is enough battery to power the mobile phone. Hence, the unit testing is completed successfully.

5.4 Mobile Application Usability Test

A usability test is conducted involving three AMD patients. They are asked to answer a set of questionnaires based on their interaction with the mobile application, assigning scores ranging from 1 (strongly disagree) to 5 (strongly agree) when answering each question. Following Table 7 describes the recorded answers from the AMD patients.

Table 7. Results from Usability Test

Questions	User 1	User 2	User 3
1) The application assists you in *better identifying objects or texts* in your surroundings	4	3	4
2) The application *accurately identifies objects or texts* found in your surroundings	4	2	3
3) The application can *quickly identify objects or texts* found in your surroundings	3	4	4

4) The audio feedback provided is *clear* and *comprehensible.*	5	5	4
5) This application is *over complicated* than necessary.	1	2	2
6) This application is *easy to use.*	5	4	4
7) I can *quickly learn* to use and navigate this application.	4	4	4
8) I *do not need to remember* too many steps to use this app.	4	3	3
9) I feel *more confident* in navigating my surroundings when using this application.	4	4	3
10) What about the application did you *like the most?*	Easy and straightforward to use. Simple user interface.	It is good that the app can be used without Internet.	No need use internet connection. Also good voice feedback.
11) What about the application did you *like the least?*	It is helpful in identifying objects but audio feedback when reading texts are not always very clear.	Robotic voice feedback provided. Detection can be confusing sometimes since it also reads out items from very far away.	App sometimes struggle to detect objects when camera is shaky.
12) How would you *describe this application to others*?	Good way of assisting people in their daily life. Thank you to developers for this app.	This app is not just a tool but act as stability for visually impaired people (Age-related Macular Degradation). It will make a big difference in peoples' lives.	App is a good start, has potential. But requires more enhancements for better convenience for people with poor vision.

13) Final comments	I found this app very helpful in my daily life. It can help me identify objects I usually cannot see clearly like food labels and clothing tags.	Impressed with the accuracy of this mobile app. Would be better if can provide audio feedback in other languages (for object recognition) as not everyone is fluent in English language.	More improvements can be made to the mobile app. Would like a version that uses less time for detection and is not easily affected by shakiness of camera.

The result findings of the above usability test generally indicate the mobile phone application developed based on the selected object and text recognition models are well received by the three AMD patients quantitatively and qualitatively despite administered in English. AMD patients understood the English used from the mobile application be it screen output or audio generated. This empirical study provides evidence that this simple mobile application is useful and serves as an augmented vision to AMD patients in Malaysia. Moreover, for some of the Malaysian AMD patients that are at retirement age or above the age of 60 might find this application is cost effective to manage and not burdening them while incurring rehabilitation cost for their AMD medical treatments. In the contrary, for AMD patients aged 60s in other developed and western countries whereby they are still under employment or before retirement age, they still have the financial means to seek better medical treatment which may not be the scenario for AMD patients in Malaysia. Hence, seeking a low-cost mobile application solution e.g. comparatively lower purchasing cost of Android mobile phone and capability of the mobile application to function without Internet connection are critical. In summary, the mobile application is user friendly to the AMD patients whereby this application can help reduce their financial burden while confronting with the AMD disease.

6.0 Conclusion

Patients suffering from AMD have impacted their lives due to partial vision loss. This applied research through the development of a mobile application that can detect, recognize and convert object or text into speech can help Malaysian AMD patients to perceive their surrounding environment more accurately. By utilizing this application from their mobile phone, AMD patients can continue to carrying out their daily activities in addition to navigating more freely and safely. Through amalgamation of deep learning, machine learning and mobile application development, this applied research contributed a small step into relieving AMD patients to have a better grasp of their surroundings.

In answering the two research questions postulated earlier and based on the comment input from the three AMD patients, this mobile application with object and text recognition capabilities generally is user friendly, useful and did enlighten them in perceiving the surrounding environment better.

One of the research limitations include too few the mobile phone users involved in the application testing whereby AMD patients are not easily available. Given ample time in the future, adequate random sample of AMD patients will be sourced to go through the usability testing again. The second limitation is the low mean average precision (mAP) of 42% for object recognition. There is planning to literally ingest more images into data sets to increase mAP and using TensorFlow algorithms to train EfficientDet Model on those data sets to generate better Graph files.

One of the future enhancements after completion of this research includes expanding object and text recognition as well as speech generation to cover other common languages aside English in Malaysia i.e. Bahasa Malaysia, Chinese, Tamil, Iban, Kadazan et al. Moreover, the mobile application also will be developed in iOS version so that other ADM patients using iPhone can also benefits from this application. Finally, in future we can include transfer learning for object recognition because current images within existing models might be different from those taken in Malaysia directly due to cultural and geographical differences.

In conclusion, this applied research and mobile application developed with object and text recognition capabilities have made a modest contribution in helping patients contracting the AMD disease in Malaysia.

7.0 Acknowledgement

The author received no funding from any party for the research and publication of this article.

8.0 References

[1] Crincoli, E., Sacconi, R., Querques, L. *et al.* (2024) Artificial intelligence in age-related macular degeneration: state of the art and recent updates. *BMC Ophthalmol* 24, 121 (2024). https://doi.org/10.1186/s12886-024-03381-1

[2] Lombardo, M., Serrao, S., & Lombardo, G. (2022) Challenges in Age-Related Macular Degeneration: From Risk Factors to Novel Diagnostics and Prevention Strategies. *Front. Med.* 9:887104. doi: 10.3389/fmed.2022.887104

[3] Moshfeghi, A. A., Lanitis, T., Kropat, G., Kuznik, A., Gibson, A., Feng, H. et al. (2020). Social cost of blindness due to AMD and diabetic retinopathy in the United States in 2020. *Ophthalmic Surg Lasers Imaging Retina*. 51:S6-14. doi: 10.3928/23258160-20200401-01

[4] Jeffery R. C. H., Mukhtar, S. A., Lopez, D., Preen, D. B., McAllister, I. L., Mackey, D. A., et al. (2021). Incident of newly registered blindness from age-related macular degeneration in Australia over a 21-year period: 1996-2016. *Asia Pac J Ophthalmol.* 10:442-9.

[5] Macnamara, A., Chen, C. S., Davies, A., Sloan, C., & Loetscher, T. (2021). Low vision devices for age-related macular degeneration: a systematic review. *Disability and Rehabilitation: Assistive Technology*, *18*(7), 998–1010. https://doi.org/10.1080/17483107.2021.1966523

[6] Wang, J., Wang, S., & Zhang, Y. (2023). Artificial intelligence for visually impaired. Elsevier, Displays 77 (2023) 102391. https://doi.org/10.1016/j.displa.2023.102391

[7] Honavar, S. G. (2023). Eye of the AI storm: Exploring the impact of AI tools in ophthalmology. *Indian Journal of Ophthalmology* 71(6):p 2328-2340, June 2023. | *DOI:* 10.4103/IJO.IJO_1478_23

[8] Bali, J., & Bali, O. (2023). Eyes of the Future: A Comprehensive Mapping of the Evolving Landscape of Artificial Intelligence in Ophthalmology. *Delhi Journal of Ophthalmology* 33(3):p 237-245, Jul–Sep 2023. | *DOI:* 10.4103/DLJO.DLJO_135_23

[9] Leonard, V. K., Jacko, J. A., & Pizzimenti, J. J. (2006). An investigation of handheld device use by older adults with age-related macular degeneration. *Behaviour & Information Technology*, *25*(4), 313–332. https://doi.org/10.1080/01449290600636512

[10] Hakobyan, L., Lumsden, J., O'Sullivan, D., & Bartlett, H. (2012). Understanding the IT-related attitudes and needs of persons with age-related macular degeneration: a case study. *IN: Proceedings of BCS HCI 2012 - People and Computers XXVI.* GBR: UNSPECIFIED.

[11] Reeves, B. C., Wickens, R., O'Connor, S. R., Gidman, E. A., Ward, E., Treanor, C., ... & Hogg, R. E. (2024). Descriptive study of the challenges when implementing an app for patients with neovascular age-related macular degeneration to monitor their vision at home. *BMJ open*, *14*(3), e077196.

[12] Lavric, A., Beguni, C., Zadobrischi, E., Cailean A., Avatamanitei, S. (2024). A Comprehensive Survey on Emerging Assistive Technologies for Visually Impaired Persons: Lighting the Path with Visible Light Communications and Artificial Intelligence Innovations. *Sensors* 2024, 24, 4834.

[13] StatCounter Global Stats. (2024, August 4). *Mobile Operating System Market Share Malaysia*. https://gs.statcounter.com/os-market-share/mobile/malaysia

[14] Akkara, J. D., Kuriakose, A. (2019). Role of Artificial Intelligence and Machine Learning in Ophthalmology. *Kerala Journal Ophthalmology*, Volume 31, Issue 2, May-August 2019, 150-160. DOI: 10.4103/kjo.kjo_54_19

[15] Hwang, D. K., Hsu, C. C., Chang, K. J., Chao, D., Sun, C. H., Jheng, Y. C., ... & Chiou, S. H. (2019). Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics*, *9*(1), 232.

[16] Romond, K., Alam, M., Kravets, S., Sisternes, L. D., Leng, T., Lim, J. I., ... & Hallak, J. A. (2021). Imaging and artificial intelligence for progression of age-related macular degeneration. *Experimental Biology and Medicine*, *246*(20), 2159-2169.

[17] Lin, Y. (2023, January 26). *10 Mobile Usage Statistics Every Marketer Should Know in 2023*. Oberlo. https://www.oberlo.com/blog/mobile-usage-statistics#:~:text=There%20are%206.8%20billion%20smartphone,approach%20an%20in%2Dstore%20employee.

[18] Ud Din Arshad, Q. K., Kashif, A. U., & Quershi, I. M. (2019, January 8-12). *A review on the evolution of cellular technologies* [Paper presentation]. Proceedings of 2019 16th International Bhurban Conference on Applied Sciences & Technology (IBCAST), Islamabad, Pakistan.

[19] Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science,* 35**,** 982-1003.

[20] TensorFlow. (2022, September 3). *TensorFlow Lite for Android*. https://www.tensorflow.org/lite/android

[21] Hisham, Z. A. N., Faudzi, M. A., Ghapar, A. A., & Rahim, F. A. (2020, August 24-26). *A systematic literature review of the mobile application for object recognition for visually impaired people* [Paper presentation]. 2020 8th International Conference on Information Technology and Multimedia (ICIMU), Selangor, Malaysia.

[22] Nasir, H. M., Brahin, N. M. A., Aminuddin, M. M. M., Mispan, M. S., & Zulkifli, M. F. (2021). Android based application for visually impaired using deep learning approach. *IAES International Journal of Artificial Intelligence (IJ-AI), 10*(4), 879-888. https://doi.org/10.11591/ijai.v10.i4.pp879-888

[23] Salunkhe, A., Raut, M., Santra, S., & Bhagwar, S. (2021). Android-based object recognition application for visually impaired. *ITM Web of Conferences, 40*, Article 03001. https://doi.org/10.1051/itmconf/20214003001

[24] Martinez-Alpiste, I., Golcarenarenji, G., Wang, Q., & Alcaraz-Calero, J. M. (2022). Smartphone-based real-time object recognition architecture for portable and constrained systems. *Journal of Real-Time Image Processing, 19*, 103-113. https://doi.org/10.1007/s11554-021-01164-1

[25] Jia, J., Fu, M., Liu, X., & Zheng, B. (2022). Underwater object detection based on improved EfficientDet. *Remote Sensing, 14*(18), Article 4487. https://doi.org/10.3390/rs14184487

[26] Tan, M., Pang, R., & Le, Q. V. (2020, June 13-19). *EfficientDet: Scalable and efficient object detection* [Paper presentation]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA.

[27] Gonçalves, M. P. (2022). MLKit Text Recognition Evaluation.

[28] Charishma, K., Umar, M., Chettri, R., Singh, R. K., Hussain, T., & Ravindranath, R. (2021). Review of Android Based Portable Sign and Text Recognition System. *International Journal of AH Research Education and Scientific Methods (IJARESM),* ISSN: 2455-6211 Volume 9, Issue 5, May -2021, Impact Factor: 7.429

[29] Boesch, G. (2022, July 9). *LabelImg for Image Annotation.* Viso. https://viso.ai/computer-vision/labelimg-for-image-annotation/#:~:text=LabelImg%20is%20a%20lightweight%20and,success%20of%20computer%20vision%20applications.

[30] Góis Mateus, B., & Martinez, M. (2019). An empirical study on quality of Android applications written in Kotlin language. *Empirical Software Engineering*, *24*(6), 3356-3393.

[31] Saabith, A. S., Fareez, M. M. M., & Vinothraj, T. (2019). Python current trend applications-an overview. *International Journal of Advance Engineering and Research Development*, *6*(10).

[32] Macenauer P. (2022, February 4). *EfficientDet-IMX*. GitHub. https://github.com/nxp-imx/efficientdet-imx