Utilizing Artificial Intelligence Techniques for Assisting Visually Impaired People: A Personal AI-based Assistive Application

Samah Alhazmi, Mohammed Kutbi, Soha Alhelaly, Ulfat Dawood, Reem Felemban, and Entisar Alaslani
College of Computing and Informatics
Saudi Electronic University
Riyadh, Saudi Arabia 11673

Abstract—Nowadays, the Artificial Intelligence (AI) field has made a significant change in the real life. Numerous applications use the AI techniques for the purpose of assisting people in different life aspects. Furthermore, with the increased number of people with visual difficulties around the world, there is a need for such AI assistive applications which provide them an independent life. Limited affordable and appropriate solutions developed so far. In this paper, we present a personal AI-based assistive application called (Vivid) that supports visually impaired people being more independent. Vivid has many features such as identifying objects, objects' colors, recognizing text, and faces detection. It relies on using the mobile camera to sense the environment, and the machine learning techniques to understand the environment. By translating a meaningful information in audible sound for those users, Vivid does not require to have any visual ability. Moreover, the whole interaction with the user is only based on voice commands. The input from the user is captured as finger gestures on tablet or cell phone touch screen. In addition to Vivid, we also shade the lights on a supplementary application that notify/alarm visually impaired people of any nearby objects using sensors. These personal assistive applications were developed then tested on the real world and showed promising results.

Keywords—Artificial intelligence; machine learning; assistive technology; visually impaired

I. INTRODUCTION

Globally, there are around 285 million people who are considered visually impaired. Thirty-nine million of them are blind and the rest have considerably low visions abilities [1]. In the US, there were four million visual impairments cases in 2010, and are projected to be seven million cases and thirteen million cases in 2030 and 2050, respectively (NIH-NEI). Those people can benefit greatly and improve their life-independently by using AI-based assistive solutions. Even though many smartphone applications developed, limited applications focused on visually impaired people. It is very difficult for a blind or visually impaired user to use a smartphone effortlessly. However, there are several features that can allow those special users to utilize such technology seamlessly as regular users.

In this paper, we propose an AI-based assistive application (Vivid) for those targeted users which is affordable, accessible, and easy to use. Many of the required assistive features were combined such as: (1) colors identifier, (2) objects labels, (3) text reader, (4) facial expression, and (5) distance notifier. The first four features were implemented together in single camera-based application which allow the user to use (Vivid) without any assistance. The use of this application is based merely on finger gestures as an input, and voice feedback as an output. The user interface was made incredibly simple to provide seamless user experience for targeted users. Whereas the last feature "distance-notifier" was developed as a supplementary assistive application with Vivid. The reasons of separating these features in two different applications are: (1) reducing complexity; (2) the first application "Vivid" can be used totally by the blind person without the need for an assistant help, while the second application "distance-notifier", the user might need assistance from someone else to avoid wireless connection errors; (3) the user might only need to use the features provided by "Vivid" which are camera-based, thus, they only download "Vivid", which will provide more flexibility. Table I shows a brief description of the five features.

II. RELATED WORK

The development of modern technologies helped to make all these technologies accessible by all categories of people. Modern technologies are not limited to be used by normal visually people only, however, they can be used by blind people as well. A few years ago, smartphones have been widely used by community – by normal people – and became most popular which touch our daily lives. Whereas, for visually impaired and blind people, technologies are still limited for them; however, the new technologies and smart solutions provided by smartphones encourage blind people to be more independent and self-reliance completely. Authors in [3], proposed a system that is based on Morse-Code, which is a code in which letters were represented by combinations of long and short light or sound signals.

Researchers in [4] raised a question in their study, how do blind users use smartphones? Usually, it depends on a screen reader that exists in its operating system, such as, Google operating system which known as Android, or Apple's system which known as IOS. But more than that, it depends on the presence of some other services, such as screen magnification and the development of night mode feature to suit some other visual disabilities. Moreover, there are other settings for disabled people within "General" menu under the name of "Accessibility". This feature allows disable people to choose what is suitable to them. In addition, smartphones have a feature which called "screen reader" to help in reading what is shown in the screen.

TABLE I. A Brief Description the Five Features of the Vivid Application

Feature	Description
Colors Identifier	To identify colors of any object by capturing an image of that object. Basically, the implemented algorithm detects the most dominant RGB colors, and identifies colors values; then, converts these values to colors names that are known by human. At the end, voice feedback is generated saying the identified color's name. This feature is not only assisted people with blindness, but also helps people who suffer from color-blindness.
Objects Labels	A machine learning tool, which is 'MLT kit' API [2] was used. The ML Kit mobile SDK brings all the Google's expertise on machine learning techniques to mobile developers in an easy-to-use package, our application can allow the user to identify any object name/label by taking an image of that object. For example, there is a "ball" object which the blind person cannot identify, by taking an image of that object, the application can identify it and tell the user it is a ball.
Text Reader	Again, by using the 'ML kit' API [2],this feature allows visually impaired people to be able to detect any written text and hear it; not only the text that is in braille. The user can take an image of the text by bringing the camera closer to where the text is written and taking an image of that text. The application then generates a voice version of that text.
Facial Expression	The 'ML kit' API [1] is used, this feature allows visually impaired people to be able to detect if there is a person in front of them with the use of face detector and then facial expression detector. User can take an image of the person in front of him/her by bringing the camera closer to where of his/her moving direction and take an image. The application then generates a voice version of that text.
Distance Notifier	To use this feature, the user will need a hardware sensor. The sensor is attached to the user as a belt and will notify/alarm him/her of any object that is getting closer or might hit him/her. The sensor used was Ultrasonic Sensor HC-S04, which is a hardware for identifying distances. The way how this hardware works, was by sending sound waves from the transmitter, which then, bounce off an object, and return to the receiver; the user can determine how far away something is, by the time it takes for the sound waves to get back to the sensor. To connect this sensor to the smartphone, we developed a separate application called "Distance-Notifier" that can handle wireless Bluetooth connection to this sensor and send alarms/notifications to the user.

Another study [5] discussed how smartphones can be used by blind people. In IOS, for example, it provides feature of screen reader known as "Voice Over", and it supports most of the languages in the world. When this option is activated, the use of this device is entirely different; the device turns into a speaker device. Any touch on the screen, will tell the user what is that touched point. For example, if the user touched an application icon i.e. (Facebook), the screen reader will say "Facebook twice to open" and will say this sentence in other languages as well. Next, the application is opened only after pressing it twice – this is similar to double-click the mouse when using a computer to move from one item to another. The screen is swiped from the left to right to go to the next item or application, or also to go to the previous item or application. With each swipe, a screen reader will utter the name of the item or application and the mechanism of its activation. The device is easy to use and depends completely upon gestures with fingers on the screen; it seems at the beginning a difficult job for blind people to use it, but it is actual requires some training to use it [5]. While IOS devices provide the service of Voice Over, Android devices which are widely used by people around the world provide the service of "Talkback" which is an accessibility service. It allows visually impaired people to interact with smartphones and use them regularly as everyone else does. It is based on spoken words, and other audible feedback that will give the user a full experience of what they are doing, and what the output produced by the device [6].

Authors in [7] said that we still convinced that modern technologies allow disable people to live in all life aspects with a realist and an effective manner. This enhanced the continuous development of modern technologies for those people. As discussed in [8], there were many text reader applications, such as, the scanner through the smartphone camera, and the conversion of written texts to audio. Moreover, authors in [9] implemented an Optical Character Recognition (OCR) program, which provides the opportunity to scan books and letters. The program works once any text has scanned, then, reads the text loudly. The OCR consists of a camera that captures the text, which then converted to speech through the program. The drawback of this program is that, this technique required a hardware in order to work sufficiently; which made it difficult to be available to everyone. However, by involving this technique into smartphones, it becomes more useful and usable [9].

Furthermore, an application developed by researchers in [10] which was based on two hardware devices. One for text input and the other for speech output. Basically, it works using a sensor component - as an eye - that captures any printed text. Then, it extracts the recognized text area, and produced a speech output - by the audio output device - for blind users. The main drawback of this application was the cost. Another study [11] showed the automation of text-to-audio. Basically, a pen-like device is used to convert any non-Braille text to audio. Any piece of text that a person would like to read, is converted to an audio signal; after that, by using the Bluetooth technology, these audio signals are transmitted to Bluetooth earphones. The authors in [11] believed that pen technology is lighter in the sense of it can be easily portable. This can change blind people life by allowing them to read whatever they desire.

In the International Conference on Computer and Information Technology [12], an Android system assistant for visually impaired people was developed which called Eye Mate which can help users to know where the obstacle is through vibrations. This android application provides navigation to a blind person and track his/her movement as well. This was based on a voice command; the application will generate a voice command according to the obstacle object position. The movement of a person are measured by using GPS which tracks the user position latitude and longitude. Furthermore, in [13] a sensor-based assistive device for visually impaired people was proposed. This device has a sensor to identify the distance between the person and a harmful object. Other studies [14][15] showed some devices that have been used by visually impaired people, however, these devices are big and not comfortable to use.

In [16], a proposed product called "Self-Energized Smart Vision Stick" developed for blind people, as shown in Fig. 1; the stick uses Arduino Ultrasonic Sensor. It is basically a sticklike a tool that provides safety and privacy to those people. This stick is attached to distance sensors to identify the distance between the stick and any nearby objects, then,
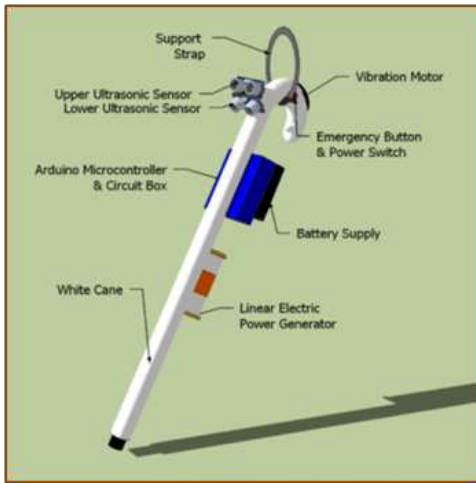
notify the user accordingly.

Fig. 1. General 3D Model of the Self-Energized Smart Vision Stick [16].

Authors in [17] declared that colors are a concern to blind people because it is a very important to them to know what the colors of their clothes are. There are many color-identifier applications proposed, such as "Color Reader" sensor. An American foundation for blind people invented a "color teller" an easy-to-use device, which helps visually impaired people to identify the color of any object. Moreover, the researchers in [18] highlighted the importance of colors, and implemented a device called "Coloresia". This device could convert any color into music or words. In addition, the American council for blind people [19] made a comparison between the color detectors applications in the market, and then, provided better solutions for these applications if needed. A list of other applications discussed that help visually impaired people in several ways, such as, a reader to identify colors and objects which work with QR technique [20].

Google Firebase is the newest technology of image labelling process. Image labelling provides information of what the images contain. With the use of ML kit API, it will enable the application to recognize objects within an image. Here, are some examples of objects that can be detected: people, activities, places, things, and so on. When it recognizes an object, it indicates a score of confidence level to show how confidently the machine learning could detect the object.

In [21], a video-based application was developed, where number of frames can be generated from the video, then, converting the images from RGB to Gray scale. This was done by applying ML algorithm, which set the key points of an object, then, match the object with the database object. If (object=database), then, convert the text to speech. Table II shows the advantages of "Vivid" as opposed to existing applications.

III. MATERIALS AND METHODS

As stated in the previous section, our goal is to develop a comprehensive solution to help people with visual difficulties with low cost. Thus, we proposed a mobile application to assist visually impaired and/or color-blind people to be more independent by providing five features: Colors Identifier, Objects Labels, Text Reader, Facial expression, and Distance Notifier. We proposed two mobile applications which were linked together. The main application is "Vivid" and the second application is a "Distance-Notifier". The first application "Vivid" can be used totally by the blind person without the need for an assistant help, while the second application "distance-notifier", the user might need assistance from someone else to avoid wireless connection errors. The "Distance-Notifier" is connected to a hardware sensor that alarms/notifies the users of any nearby objects. We separated these features into two independent applications to reduce the complexity.

A. Vivid Application

Vivid is a camera-based application that captures objects/things and translates them into audible sound. Thus, considering our targeted users who will not be able to see the screen or to use the regular interface that consists of buttons and other controllers, Vivid was built with a simple interface. The features of colors identifier, objects labels, and anything reader are combined in single camera-based application that is called "Vivid". Thus, a blind/ visually impaired user can use "Vivid" application totally by himself/herself without any assistance because the use of this application is based merely on finger gestures for an input, and voice feedback for an output. The user interface was made incredibly simple to provide seamless user experience for blind/ visually impaired people.

The color identifier works over two steps: (1) find the dominant color D; (2), find what color is that. It estimate the dominant color by averaging the color of all pixels per channel $C_1, 2, 3 (1 = red, 2 = green, 3 = blue)$:

$$C_i = \frac{\sum_{k=0}^{n} p_{ik}}{n} \qquad (1)$$

where i is the channel number, k is the pixel index in the image and n is the total number of pixel in the image. The average RGB color is then refer to as the dominant color

$$D = (C_1, \ C_2, C_3) \qquad (2)$$

After deciding the dominant color of the picture, RGB color is converted to HSV format. First, we get the maximum

$$C_{\max} = \max (C_1, C_2, C_3) \qquad (3)$$

and minimum

$$C_{\min} = \min (C_1, C_2, C_3) \qquad (4)$$

of the values from the three channels. Then,

$$a = C_{\max} - \ C_{\min} \qquad (5)$$

Second, we calculate the value of Hue (H) and Saturation(S). H is calculated as follow:

$$H = \begin{cases} 0, & a = 0 \\ 60 \ x \ (\frac{C_2 - C_3}{a} \bmod 6, & C_{\max} = C_1 \\ 60 \ x \ (\frac{C_3 - C_1}{a} + 2, & C_{\max} = C_2 \\ 60 \ x \ (\frac{C_1 - C_2}{a} + 4, & C_{\max} = C_3 \end{cases} \qquad (6)$$

TABLE II. Advantages of "vivid" as Opposed to Existing Applications

Existing Applications	Vivid
An Optical Character Recognition (OCR) program, which provides the opportunity to scan books and letters [9].	The mentioned technique requires a hardware in order to work sufficiently to convert text to audible sound, which made it difficult to be available to everyone. Vivid provides a solution to this drawback by involving this technique into smartphones. Since smartphones are accessible to almost everyone, getting access to this technology is only a few clicks away.
Portable camera based assistive text reading from handheld objects for blind person [10].	Vivid takes the advantage of already existing audio input/out means that are built in the smartphone. No additional hardware required. All the three tasks (scanning text, processing text, producing audio) are done instantaneously in one single device.
Automated electronic pen aiding visually impaired in reading, visualizing, and understanding textual contents [11].	The mentioned pen technology is light and portable. However, Vivid provides better advantages in a sense that despite of being available on a light, portable mobile phone, it is also accessible easily within a few clicks.
Android assistant EyeMate for blind and blind tracker [12].	The proposed technology uses GPS to track coordinates and notify of any obstacles. Vivid enhances this technology by using an Arduino sensor. The sensor measures the distance instead of using GPS system to provide better accuracy and more precise distance measurements.
Android assistant EyeMate [12].	In this application, the impaired person needs to deal with two different applications, (1) to capture contextual (distance of an obstacle, position of the sensors, environment around the user), and then, (2) to communicate with the other application to deliver this information to the user. Our Vivid's algorithm solved this issue by eliminating the need to communicate to any outsider. Thus, the algorithm itself handle the captured information and translate it to audible sound.
Optical devices for distance [14].	The technologies used by developing a device which made it too large and uncomfortable to use as compared to Vidid.
Self-energized smart vision stick for visually impaired people [15].	Using a sick that sense the nearby environment and obstacle is a smart idea. However, the stick is heavy, and the user needs to carry it around. Vivid provides a smart belt instead of a stick so that the user does not need to carry, they only need to wear it and the application will take care of notifying the user of any nearby obstacle.
Assistive technology products by the American Foundation for the blinds [16].	The foundation invented a "color teller" which is an easy-to-use device. It helps visually impaired people to identify the color of any object in front of them. However, Vivid is allowing the users to use the color detection services free of charge while the proposed device costs around 205 dollars.
Bilingual wearable assistive technology for visually impaired people [18].	The wearable hardware requires additional costs compared to Vivid which is merely a downloadable mobile application.
Color to sound converter for blind people [22]	The study suggests using a sensor called "Color Reader". Vivid on the other hand, does not require any additional sensors other than the camera of the mobile. The "color detection" algorithm integrated within Vivid uses the information captured by the camera and acts as a sensor that can identify colors.

Lastly, S is calculated using the following formula:

$$S = \begin{cases} 0 \,, & C_{\max} \neq 0 \\ \frac{a}{C_{\max}}, & C_{\max} \neq 0 \end{cases} \qquad (7)$$

Then, identifier takes the dominant color and classifies it to one of 12 predefined colors. Each of those colors has a range of numbers if the dominant color fill in the range, then it is classified with it. Table III shows the ranges of HSV values for color identification. Then, find what color is that by:

$$Color = f(H, S) \qquad (8)$$

To make sure we are using state-of-the-art techniques, we used Android ML Kit called Firebase [2]. For object recognizer, it offers objects classification of 10000+ classes and 400 classes when working with the online version. The model's performance are measured in term of accuracy is the number of times the model correctly classifies an object. The accuracy for the model is 60% as reported by authors of ml kit library [2].

The model's architecture is MobileNet [23] to use minimal resources on the phone or tablet. For text reader, it recognizes the test in an image and split it into blocks which later read word-by-word out loud to the user. Lastly, facial expression recognizer, it has two features; first, it detects faces in the image then, second, it classifies the expression into smiling or not. The system output "There are no person" if there were no people detected in the captured image. If there is a person and its face is detected, it will classify the expression of the person. Then, it output the expression either smiling person or not smiling person.

Vivid application has a simple workflow which is described in Fig. 2. After launching the application, user manual will appear on the screen only if it is the first time the app has been launched. Otherwise, camera feed will be presented on the screen. Then, user press on the screen to capture specific picture of interest. Then, users have five core actions: swiping up to get object label recognition, swiping down for re-capturing picture of interest, swiping left for color detection, swiping right for text recognition and, finally, long-press for face expression.

Vivid application consists of two activities: the main activity and the image processing activity. The main activity as shown in Fig. 3(a) is the first interface the user will expect, which is responsible to capture the image. As shown in the figure below, the whole screen is merely the camera display, it would not be reasonable to display any sort of text output, graphic or buttons for user interaction with the app because the target users are visually impaired ones. Thus, user will interact by figure gestures. On this activity user can capture the desired object by bringing it close to the camera lens and then tabbing anywhere on the screen. Tabbing gesture will tell the camera to capture the scene at that moment and then send that captured content to the Image Processing activity for processing. However, the image processing activity as shown in Fig. 3(b) is responsible for processing the content captured in the previous activity. The interface again doesn't have any form of typical user interaction controls. Interaction is based on audible outputs and finger gestures. Once user hears the word "swipe" that indicate that the image is processed, and the application is ready to produce the output. User then can get the output using figure gestures of swiping in different directions to get different outputs as defined in the flowchart figure above.
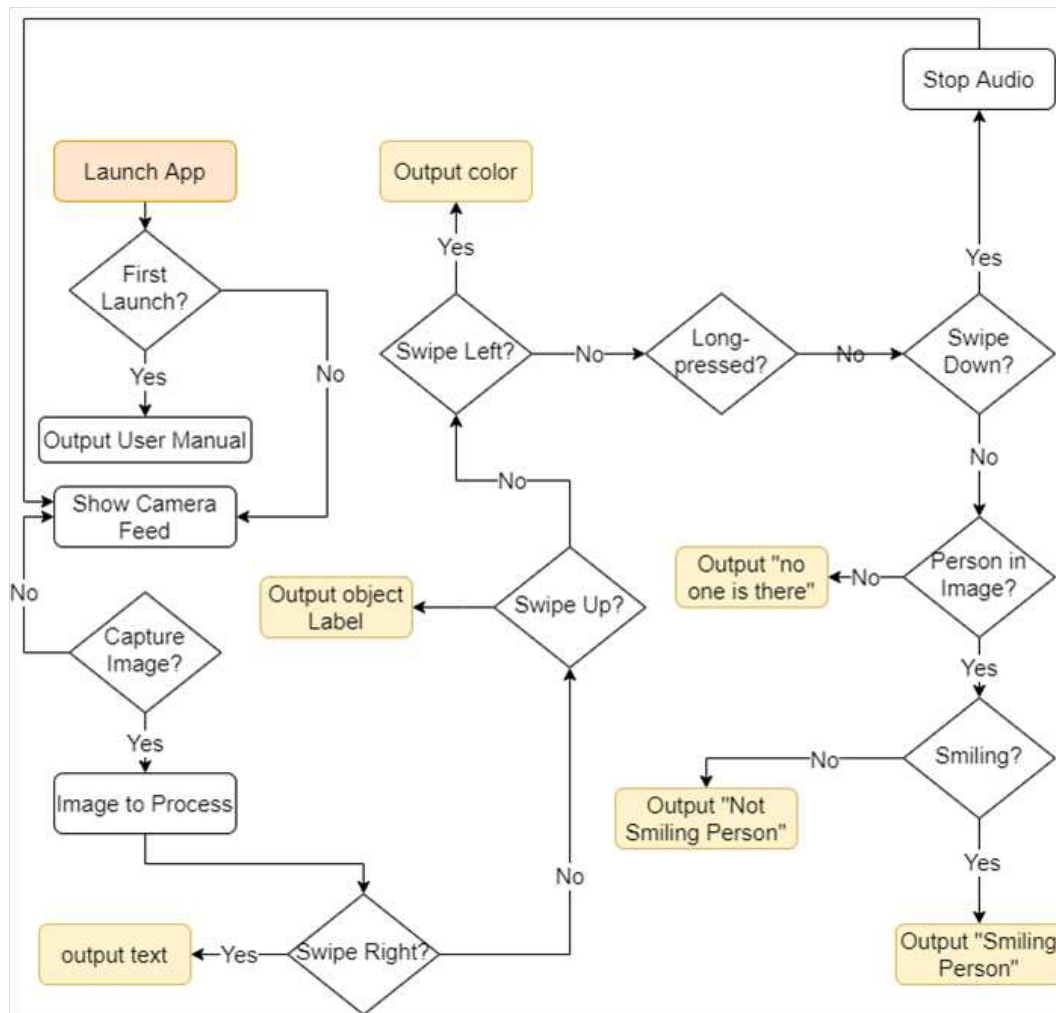
Fig. 2. Flowchart of Vivid Application. All Features and their Workflow of the Vivid Application is Explained in the Figure.

Fig. 3. Overview of the Interfaces of Vivid Application. (a) Main Activity. Camera Feed Showing before Any Interaction is Done by the User. (b) Image Processing Activity. After Taking a Picture.
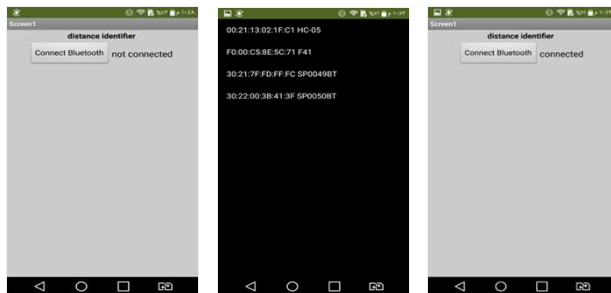
B. Distance-Notifier Application

Distance Notifier application uses a hardware sensor. The sensor is attached to the user as a belt and will notify/alarm him/her of any object that is getting closer or might hit him/her. The used sensor is Arduino Ultrasonic Sensor HC-S04, which is a hardware for identifying distances. The sensors would send sound waves from the transmitter, which then, bounce off of an object, and return to the receiver; you can determine how far away something is, by the time it takes for the sound waves to get back to the sensor. Therefore, to connect this sensor to the smart phone, we developed a separate application called "Distance-Notifier" that can handle wireless Bluetooth connection to this sensor and send alarms/notifications to the user. It is preferred that, to use an assistant person to setup and connect this application to the hardware, to avoid any connection errors. Afterwards, the blind user can receive notifications and alerts from this application.

Distance-Notifier application is a standalone which requires additional hardware. The additional hardware is Ultrasonic Sensor HC-SR04 and an Arduino [24]. The phone is connected to the Arduino using Bluetooth. The app receives the signals from the Arduino and output voice notification for the user. The voice notifications alert user about obstacles on her way. Fig. 4(a, b and c) shows in the interface of the distance notifier application. For the first interface we had a button that allows the user to select Bluetooth device, then, it will show to the user the list of Bluetooth devices, after the list, the user can select the desirable device. Finally, after selecting the device, the used connected and it will be ready to be used.

TABLE III. PREDEFINED COLORS FOR COLORS IDENTIFIER (COLORS ARE CAPTURED IN RGB FORMAT THEN CONVERTED TO HSV FORMAT)

Color	Condition Using Hue (H) and Saturation (S) values f(H,S)
White	$H + S > 59$
Black	$H + S < 35$
Red	$H >= 345 or H < 45$
Orange	$45 <= H < 75$
Spring green	$75 <= H < 105$
Green	$105 <= H < 135$
Turquoise	$135 <= H < 165$
Sky Blue	$165 <= H < 195$
Blue	$195 <= H < 225$
Purple	$225 <= H < 285$
Pink	$285 <= H < 315$
Crimson	$315 <= H < 345$

(a) Start Screen. (b) Selecting Device. (c) After Connecting with Device.

Fig. 4. The Interface of the Distance Notifier Application.

Distance notifier application workflow is shown in Fig. 5. After launching the app, users choose the distance identifier device from the list. Upon succession of connection, the app start notifies the user about obstacles in her/his way.
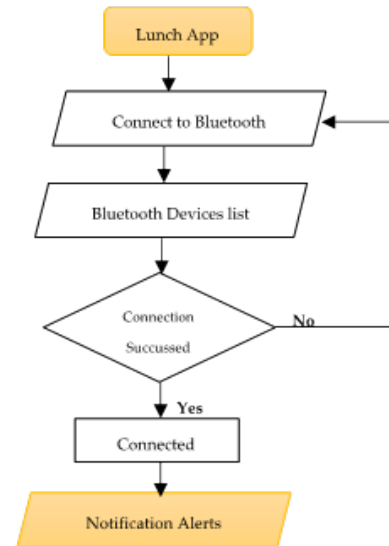

Fig. 5. Flowchart of Distance-Notifier Application.

IV. RESULTS

In this section, we present the evaluation of our experiment which conducted to evaluate the proposed applications' features in the real-world. The design of experiment and reporting of performance results are inspired by [25].

A. Test Cases

To test the proposed applications, we used it in multiple scienaros with various lighting conditions and poses. Our applications have five features: (1) color identifier, (2) object labels, (3) text reader, (4) facial expression, and (5) distance notifier. Fifteen test cases were desgined for each feature to be tested. Fig. 6 shows samples for the test cases with images. Test cases were designed to have veriaty in fabric, color, shape, camera pose and lighting conditions. For each feature 15 test cases was designed.

B. Experimental Results

An experiment is conducted in the real world using the developed two applications. Generally, the application worked with high accuracy as reported in Table IV ranging from as low as 33% and as high as 100%. Reasons vary for this some of which are due very long text, low light, blurred from moving camera or zooming too close or too far in an image. The color identifier is affected by lighting condition a lot as it shows in the accuracy. Future investigation on improving the accuracy of the color identifier is needed.

V. DISCUSSION

There are many ways to improve the accuracy of each feature. Research about each of them is extensive with rapid improvements. However, all this is out of scope of this work.
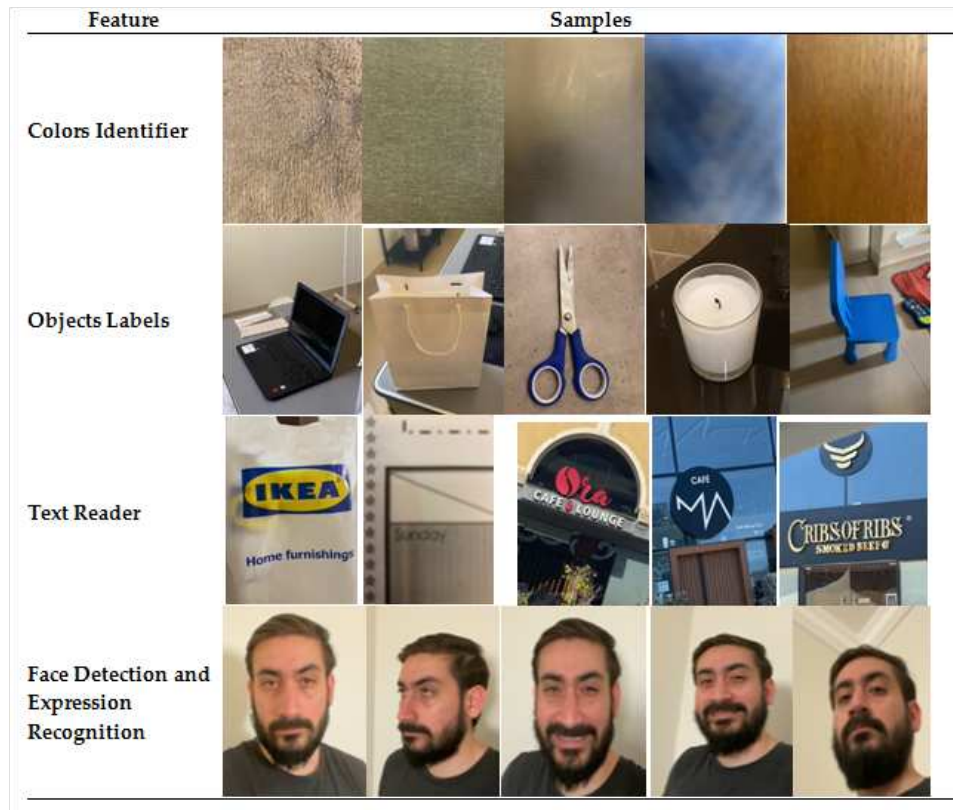
Fig. 6. Test Cases Samples for the Feature Tested.

TABLE IV. Experimental Results for All Test Cases for each Feature (Indicates that the Results are Represented based on the Percentage of Letters Recognize Correctly as Percentage of All Letters in the Image.).

Feature	# Test Cases	Obtained Results	Causes for Errors
Colors Identifier	15	33%	Low light - Blurred image - The image is too close or too far
Objects Labels	15	80%	Low light - Blurred image - The image is too close or too far
Text Reader	15	90%	Very long text - Low light - Blurred image - The image is too close or too far
Face detection	15	100%	
Facial Expression	15	93%	Head pose - Low light - Blurred image - The image is too close or too far
Distance Notifier	N/A	100%	

The focus of this work is to integrate state-of-the-art features that we believe are most beneficial for the application users at low cost. It is also important to note that this application has an interface for interaction without requiring the user to have visual abilities. This expand our application target audience from only partially blind people to complete blind people. One more group of people that could benefits who do not suffer from blindness or short of sight but not color blindness can benefits from this application as well. They can use color identifier when in doubt about the color they see. To this end, our experiment proof the validity of the application and the robustness to various condition in real-world environment.

VI. Conclusions and Future Work

An affordable solution for people with visual impairments was proposed and implemented. Vivid application provides those users the ability to recognize objects, detect people in the scene and their facial expression, assist in identifying colors and help in reading texts. Additionally, an extra feature for obstacle avoidance was implemented using a secondary standalone application with attachable hardware which have relatively low cost. These applications were tested in the real world and provided very good results. The results of the experiments indicate that such an application is a viable option for assisting people in need at an affordable price.

Future work on this research includes improving the use of machine learning to identify colors instead of predefining the range and enhancing the text reader feature to include long texts and more languages for non-English speakers. More work will be done to improve the sensor by shrinking the device size to enhance its portability. Lastly, adding the navigation features to the distance-notifier application which will help the users not only to avoid obstacle, but also to navigate well based on the shorter route. Lastly, more experiments are needed to be conducted with subjects that are from the target user base which can highlight additional challenges and areas of improvements from user experience or robustness of the application.

REFERENCES

[1] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 9, pp. e888–e897, 2017.

[2] J. Echessa, "A look at android ml kit - the machine learning sdk," https://auth0.com/blog/a-look-at-android-ml-kit-the-machine-learning-sdk/, 2018, [Online; accessed 24-July-2018].

[3] A. S. SINGH D, "Advanced human-smartphone interface for the blind using morse code," *International Journal of Advances in Electronics and Computer Science*, 2017.

[4] R. Kuber, A. Hastings, and M. Tretter, "Determining the accessibility of mobile screen readers for blind users," *UMBC Faculty Collection*, 2020.

[5] GEORGIA, "How to use voiceover for iphone," https://www.imore.com/voiceover-tutorial, 2010, [Online; accessed 17-November-2018].

[6] J. Hildenbrand, "What is google talkback?" https://www.androidcentral.com/what-google-talk-back, 2014, [Online; accessed 11-September-2018].

[7] T. Mayisela, "The potential use of mobile technology: Enhancing accessibility and communication in a blended learning course." *South African Journal of Education*, vol. 33, no. 1, pp. 1–18, 2013.

[8] D. Ahuja, J. Amesar, A. Gurav, S. Sachdev, and V. Zope, "Text extraction and translation from image using ml in android," *International Journal of Innovative Research in Science, Engineering and Technology*, 2018.

[9] S. View, "How do people who are blind or visually impaired read printed text?" https://sandysview1.wordpress.com/2016/09/15/how-do-people-who-are-blind-or-visually-impaired-read-printed-text/, 2016, [Online; accessed 8-September-2018].

[10] A. V. Mhaske and M. S. Sadavarte, "Portable camera based assistive text reading from hand held objects for blind person," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 5, 2016.

[11] A. K. Joshi, T. P. Madhan, and S. R. Mohan, "Automated electronic pen aiding visually impaired in reading, visualizing and understanding textual contents," in *2011 IEEE INTERNATIONAL CONFERENCE ON ELECTRO/INFORMATION TECHNOLOGY*. IEEE, 2011, pp. 1–6.

[12] M. S. R. Tanveer, M. Hashem, and M. K. Hossain, "Android assistant eyemate for blind and blind tracker," in *2015 18th international conference on computer and information technology (ICCIT)*. IEEE, 2015, pp. 266–271.

[13] W. Elmannai and K. Elleithy, "Sensor-based assistive devices for visually-impaired people: Current status, challenges, and future directions," *Sensors*, vol. 17, no. 3, p. 565, 2017.

[14] C. Willings, "Optical devices for distance," https://www.teachingvisuallyimpaired.com/optical-devices-for-distance.html, 2015, [Online; accessed 12-November-2018].

[15] S. Z. Mohammad Mohammadi, Saif AlAmeri, "Self-energized smart vision stick for visually impaired people," *IEEE Conference on Electromagnetic Field Computation*, 2013.

[16] A. K, "American foundation for the blind," https://www.afb.org/prodProfile.asp?ProdID=746, 2014, [Online; accessed 22-October-2018].

[17] A. Gonzalez, R. Benavente, O. Penacchio, J. Vazquez-Corral, M. Vanrell, and C. A. Parraga, "Study of color importance for blinds, and pda device," *International Journal of Innovative Research in Science*, 2013.

[18] H. Rashid, A. R. Al-Mamun, M. S. R. Robin, M. Ahasan, and S. T. Reza, "Bilingual wearable assistive technology for visually impaired persons," in *2016 International Conference on Medical Engineering, Health Informatics and Technology (MediTec)*. IEEE, 2016, pp. 1–6.

[19] C. Willings, "Identification apps," https://www.teachingvisuallyimpaired.com/identification-apps.html, [Online; accessed 10-October-2018].

[20] "Image labeling features," https://firebase.google.com/docs/ml-kit/label-images, 2018, [Online; accessed 18-October-2018].

[21] K. B. Tharkude, A. K. Wayase, P. S. More, and S. S. Kothey, "Smart android application for blind people based on object detection," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 4, 2016.

[22] P. J. Neha Patil1, "Color to sound converter for blind people," *International Research Journal of Engineering and Technology*, 2017.

[23] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[24] Isaac100, "Getting started with the hc-sr04 ultrasonic sensor." https://www.hackster.io/Isaac100/getting-started-with-the-hc-sr04-ultrasonic-sensor-036380, 2017, [Online; accessed 20-September-2018].

[25] G. Senarathne, D. Punchihewa, D. I. Liyanage, G. Wimalaratne, and H. De Silva, "Blindaid-android-based mobile application guide for visually challenged people," in *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2021, pp. 0039–0045.