Nat. Volatiles & Essent. Oils, 2021; 8(5): 2075 - 2082

NVEO
Natural Volatiles &
Essential Oils

Development of Navigation Application for Visually Impaired

Jae Moon Lee*[1], Kitae Hwang[2], In Hwan Jung[3]

*[1, 2, 3]Professor, School of Computer Engineering, Hansung University, 02876 Korea

jmlee@hansung.ac.kr*[1], calafk@hansung.ac.kr[2], ihjung@ hansung.ac.kr[3]

Corresponding author[*]: mobile Phone: +82-010-9760-4135, jmlee@hansung.ac.kr

Abstract

Background/Objectives: This paper proposes an application that assists the visually impaired, and develops it as a prototype in a smart phone.

Methods/Statistical analysis: The proposed system is an application that guides the visually impaired to his or her destination from the departure point as a guardian might guide. To do so, the application shall store movement routes, as well as information necessary for guidance during each route. The application was developed to provide audio guidance of such stored information in accordance to the location while moving.

Findings: The application is developed to recognize various objects that appear in front and to guide -in audio- it's users of the object it encounters. It is also developed to identify traffic lights and guide it's users via voice. The application also features a notification function that informs the guardian and/or related organization of the dangers that occur to the user while moving.

Improvements/Applications: Upon testing of the application developed, the application's intended functions worked as designed. However, the precision of conventional GPS was insufficient in guiding movement to the visually impaired.

Keywords: Visually Impaired, Smart Phone, Artificial Intelligent, Image Classification, Object Detection.

1. Introduction

Worldwide, around 36 million people are blind and 217 million have a moderate to severe visual impairment. Even the everyday life that the general public take for granted may often times be difficult for the disabled. In particular, there are many obstacles that arise for the visually impaired when trying to travel. This paper designs and develops an application that can provide assistance for these visually impaired while moving. Visual impairment is a sensory and cognitive impairment, which makes it difficult to acquire information. As such, it is difficult to acquire spatial information that leads to significant inconvenience in orientation and movement. In addition, it is difficult to grasp an accurate sense through interaction with the environment and to show an appropriate response to the surroundings.

Recently, with the remarkable development of IT technology, many applications for the visually impaired have been developed. Speech-based web browser for visually impaired users was developed [1, 2]. It discussed various issues and challenges related to Web Browsing by visually impaired (blind/partially blind) Web users. The interaction of visually impaired users with assistive technologies, and in particular with screen readers, was defined as behavioral patterns, which have a sequence that occurs at specific time slot. From them, it understood user behavior and implemented the characteristics, relationship structures and functions of behavior in specific application domains. 'Eye free app' is an application for calling and messaging using high-performance smartphones for people with the visually impaired [1]. It used 'Speech to Text' and 'Text to Speech' converters for calling and messaging functions. It has a very simple user interface (UI) that provides enough multi-modes, action checks, and audio help for the visually impaired. In [2], an indoor wayfinding application was proposed. Indoor directions are difficult because GPS, the most commonly used source of location information for directions, is not installed indoors. It used a computer vision approach to indoor localization running as real-time apps on existing smartphones. This application has the feature that no physical infrastructure is required.

There are many of these applications, but they are still in their infancy. What is most needed for the visually impaired is technology that will convert visual information into auditory information. In other words, it explains vocally the information that others can obtain with their eyes to the visually impaired. Granted that this role of help is yet to be completed as perfectly as an actual human's guidance, the use of artificial intelligence technology has opened up many new possibilities. This paper is a study on the development of an application that recognizes the surrounding environment and situation of the visually impaired by using deep learning, which has made rapid technological development in recent years, and guides it through voice.

2. Related Works

2.1. *Smart phone technologies*

The evolution of the smartphone technology does not stop [3]. With the recent commercialization of 5G which supports up to 20Gpbs, various location-based services for smartphones are being developed [4]. Such 5G technology has brought life to various new methods of improving location information measured via conventional GPS. Position correction information is additional information used to increase the accuracy of satellite navigation systems (GNSS) such as GPS. Observational spatial correction (OSR) information is to be provided free of charge on the internet worldwide.

If this is used, it is expected that it will be possible to measure the exact location of the level of 20 to 30 cm even on a smartphone. The camera function is one a smartphone's most used functions. Therefore, smart phone manufacturers are pouring great efforts to improve the performance and functionalities of the camera. This is because the camera plays a significant role in the use of 5G based artificial intelligence, VR, and AR. And as such, smartphones with various zoom functions and multi-camera technology are emerging. For example, the ToF (Time-Of-Flight) -a technology that produces 3D images from camera footage by measure the depth of objects- is currently undergoing development [4, 5]. The high performance of these cameras is enough to be utilized as a tool for the visually impaired.

2.2. *Artificial intelligence technologies*

The recent advancements of artificial intelligence technology are remarkable. With such rapid development, researches that utilize the research results of artificial intelligence are being conducted in diverse fields. Deep learning is a type of machine learning, which is a representative technology of such artificial intelligence [6, 7, 8, 9, 10]. CNN and RNN are representative research results of the deep learning model. CNN is mainly used to recognize images. CNN excels in image classification. When inputting an image into CNN, it will classify whether the image is a cat, a dog, or any other possible animal. In order to compensate for the shortcomings of the existing technology, this technology derives characteristics that represent the essence of an image and puts it into a neural network instead of receiving each and every image pixel values it has. In conventional machine learning, it was upon the discretion of the user as to which characteristics should be derived for good classification accuracy. For example, when it comes to classifying apples and bananas, features such as color, length, and shape would have be derived. CNN automates this process of deriving features. Among the various models of CNN, image classification and object detection models are most widely used.

For these technologies, Tensorflow[7, 9], YOLO[7, 9], ImageAI[9], and Detectron2: A PyTorch-based modular object detection library are provided. The most commonly used one is Tensorflow provided by Google.

2.3. *Speech recognition and synthesis technologies*

Mention the statistical details, the number of samples used, statistical tools/software used/ SD/level of significance, repeatability etc. Speech recognition and synthesis technology is a technology that recognizes a speech signal and converts it into text data and vice versa- converting text data into speech data. Speech recognition is the particularly difficult part.

However, machine learning -one of the latest artificial intelligence technologies- has enabled high-quality speech recognition through various learning. Voice recognition libraries are already provided in various Cloud services by companies such as Naver and Google [11]. Speech synthesis is already well-known.

3. Design of the application

The proposed application will not and cannot rid completely of the disabilities of the visually impaired. Nonetheless, the final goal of the application is to provide a walking experience as if the user were accompanied by a guardian/assistant. It is not easy to achieve this goal with the current limitations of technology, etc. However, the application proposed in this paper is designed to have 'route setting and guidance', 'traffic light recognition and guidance', 'front object recognition and guidance', and 'danger status notification' functions in order to approach the final possible goal. The architecture of the proposed application is shown in figure 1.

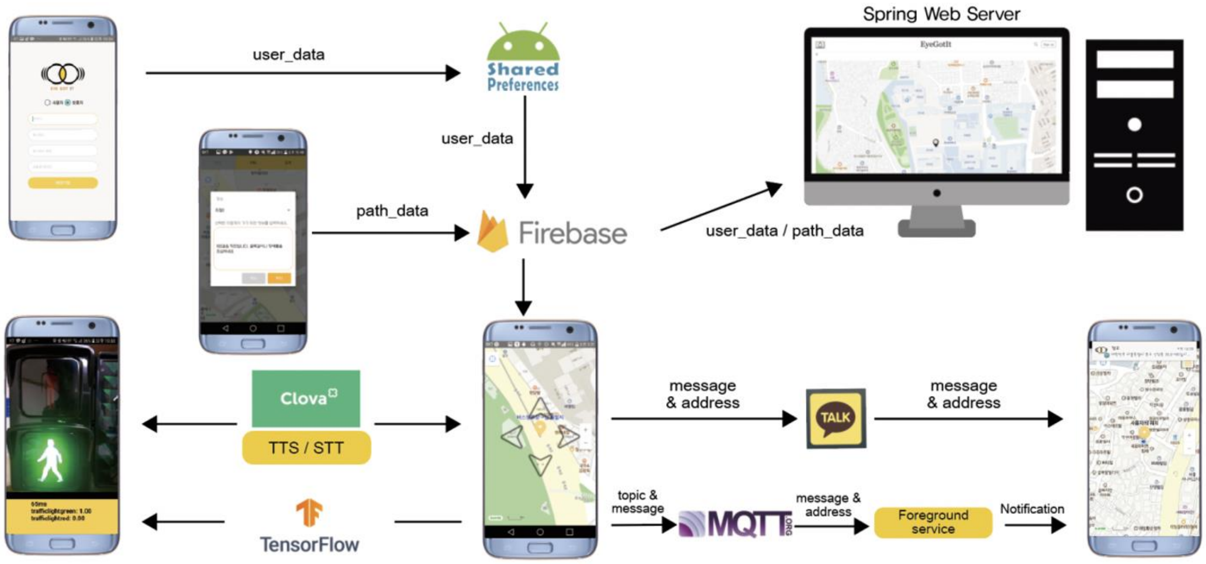

Figure 1. Architecture of the proposed application

- Route setting and guidance

Once the starting point and destination are settled, the route of travel is saved on the smartphone before departure. When a visually impaired user moves, the smart phone recognizes the GPS and uses the recognized GPS coordinates and the coordinates of the movement route to guide the movement in whichever desired direction by voice. In addition, if there is a special matter to be discussed at a specific point, the matter to be guided at that location is also stored. Then, when the user arrives at the location, the content is guided by voice. In order to perform this function, the GPS recognition technology of a smart phone, the technology to convert text to speech, and an optimal route search technology to move two arbitrary locations are required.

- Traffic light status recognition and guidance

This is a function for the visually impaired to recognize signals in places where traffic lights are installed. When the visually impaired user is looking around a traffic light, he or she uses the camera function of a smartphone to capture the state of the traffic light. The application recognizes the state of a traffic light by applying artificial intelligence to the captured image and guides it by voice. For this function to work properly, the technology to capture surrounding images using a camera built into a smart phone and image classification technology of artificial intelligence that recognizes the state of a traffic light from the captured images are required.

- Forward object recognition and guidance

One of the most important things to alert when guiding the visually impaired is the presence of obstacles ahead. In general, this is why the visually impaired use a cane for the disabled. The application proposed in this study extracts information from captured images to identify and alert the user of any potential obstacles heading his or her way. It uses object detection technology, which has achieved many accomplishments in the field of artificial intelligence throughout recent years. The visually impaired user should always look forward of the camera while moving. At this time, the camera periodically captures the front image. Through artificial intelligence object detection, the captured image is used to identify objects in front and the text is outputted. The textual results are voiced to the visually impaired. The text is then translated to voice, to be guided to the user.

- Critical situation notification

There are many instances where the visually impaired are in danger when moving in the streets. This is because it is almost impossible for the visually impaired to comprehensively judge the surrounding situation and predict what will happen in the very near future. This means that the visually impaired is most likely unable to prevent dangerous situations. The following function of the proposed app will help the visually impaired to escape such dangerous situations quickly even after falling into one. The function will automatically contact a guardian or related organization via message or call when the user detects danger and alerts the smartphone via voice. For this, voice recognition technology and push service technology are required.

4. Prototype implementation

The proposed application was created on Android. Maps Api of NaverColudPlatform was used for routing, Naver Clova Speech Recognition and Text To Speech of Android were used for voice guidance. Google's Tensorflow was used to recognize the state of traffic lights and detect objects in front, and MQTT Broker (HiveMQ) [11], Foreground Service, and Kakao API were used to notify the danger of the user.

- Implementation of route setting and guidance function

This function saves the optimal travel route in an array of (latitude, longitude) when you enter in the point of departure and destination. This function saves -as a text file- the optimal travel route in an array of (latitude, longitude) when you enter the point of departure and destination. Maps API of NaverColud Platform was used for the searching of movement routes and to obtain (latitude, longitude) information. NaverCloud Platform is a collection of cloud services provided by Naver. One of these services is Map API, which provides information regarding moving route (latitude, longitude) in real time.

- Implementation of traffic light recognition and guidance functions

Out of the various artificial intelligence technologies, image classification has made great advancements recently. Image classification technology was used to recognize the state of traffic lights. The technology was trained using three images: a red traffic light, a green traffic light, and a black traffic light with no signal. As there are diverse types of traffic lights on the streets that may appear differently according to the direction in which they are viewed, the data model was created by collecting and learning from various images of traffic lights on the Internet that appear in different shapes and forms. A total of 200 images were used for this learning. The learning algorithm was programmed in Python using a library provided by Google. The model turned out to be approximately 20 megabytes, and this is due to the fact that the model is optimized to a size suitable for a mobile environment. Tensflow-lite provided by Google [9] was used for traffic light state recognition of the learned model. This is a deep learning module optimized for mobile environments. The success rate of traffic light status recognition of the developed system was over 90%. The figure 2 shows the implementation result and traffic light recognition rate. The system was created as a prototype and simply voiced such recognition results. As errors in the recognition of traffic lights may lead to large accidents, various post-processing work on the results of traffic light recognition is required in order to develop a system that may be used in a real life environment. This will be discussed in the next chapter.

Figure 2. An Example of Traffic Light Classification

- Implementation of forward object recognition function

In the research results of deep learning, an area that has achieved as huge of a result as image classification is the detection function of objects in images. Object detection in images is detected in real time using tensflow, YOLO, and PyToch libraries [9]. In other words, it is possible to analyze and identify the types and sizes of images included in the camera in real time from images that are inputted by the camera. This paper uses deep learning-based object detection technology provided by Tensflow for this function. For this purpose, learning about objects was not conducted. It is because the objects appearing on the road are very diverse, and there is a learned model provided by Google for the learning of these general objects. The figure 3 is a prototype system, displaying results of recognition of a person in front. Results include what the object is, and the probability of the recognition being correct. In the developed system, these results are notified to the visually impaired via voice.

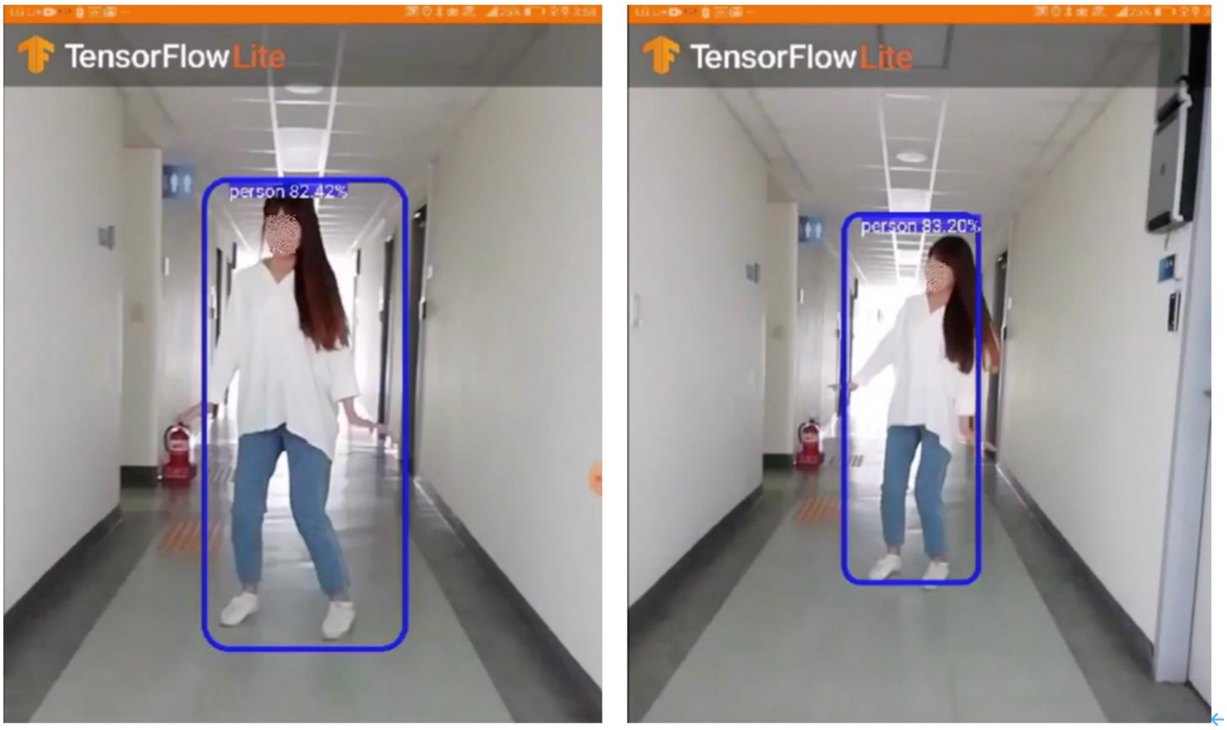

Figure 3. An Example of Front Object Detection

- Implementation of risk situation notification function

The visually impaired can be in danger from time to time while traveling to a destination. In this case, most of them have difficulty coping with these risks alone. The application, includes a function to alert and notify the guardian or related organization when a visually impaired user recognizes danger. To complete this function, the guardian or related organization can always observe the movement of the visually impaired user.

When the user begins to move, the application periodically stores the location of the user in a real-time database. In order to track the location of the user, the guardian or related organization may check the movement of the user stored in a real-time database. In particular, Google's firebase, a push service, was used to transmit movement in real time. In addition, in order to notify the guardians or related organizations of the danger, the situation in which the user is in was automatically transmitted via Kakao Talk messenger, and such information will be displayed as a notification function on the guardian's smartphone using MQTT technology [11].

5. Discussion and conclusion

The application proposed in this paper was developed as a prototype on an Android-based smart phone, and was used in a test run by participant who were acting as visually impaired users. The following contents have been improved and require further research in the future. GPS accuracy was a problem in the route setting and guidance functions. The GPS for pedestrians should be much more precise than the GPS measured in vehicle movement, but the precision of conventional GPS was insufficient in guiding the correct movement to the visually impaired. Also, there are unforeseen obstacles in pedestrian walks such as constructions. In the traffic light recognition and guidance function, the recognition of the traffic light was very accurate. However, as predicted, it was not easy for a visually impaired user to photograph a traffic light using a smartphone. In order to secure this, supplementary guidance such as voice guidance to help the photographing of traffic lights was added at the location of the traffic lights on the route, but there were still many limitations. In the forward object recognition function, the front object was recognized and the voice guidance for the object worked very well. However, there were many cases where recognition and guidance were for objects that did not interfere with movement, and voice guidance for the distance to the front object was also required.

In the application proposed in this paper, accurate distance measurement is not required, but it is necessary to calculate and inform an approximate distance. For example, if a person is recognized in front, the size of the person is approximately 1.5m to 2.0m, so it will be possible to calculate the approximate distance by recognizing the image size in the camera. The danger situation notification function worked relatively well as designed.

6. Acknowledgment

This research was financially supported by Hansung University.

7. References

1. Sagale U, Bhutkar G, Karad M, Jathar N. An eye-free Android application for visually impaired users. In Ergonomics in Caring for People. 2018; 291-297.
2. Fusco G, Coughlan JM. Indoor localization for visually impaired travelers using computer vision on a smartphone. In Proceedings of the 17th International Web for All Conference. 2020 April; 1-11.
3. Cheol-Yeong Jeong. Design and Implementation of Mobile System for Intelligence Inspection. Journal of Next-generation Convergence Technology Association. 2020;4(3):242-249.DOI:10.33097/JNCTA.2020.04.03.242
4. Al-Turjman F. 5G-enabled devices and smart-spaces in social-IoT: an overview. Future Generation Computer Systems. 2019;92: 732-744.
5. Tian Y, Ma Y, Quan S, Xu Y. Occlusion and collision aware smartphone AR using time-of-flight camera. In International Symposium on Visual Computing. 2019 October; 141-153.
6. Pouyanfar S, Sadiq S, Yan Y, Tian H, Tao Y, Reyes MP, et al. A survey on deep learning: Algorithms, techniques, and applications. ACM Computing Surveys (CSUR). 2018; 51(5): 1-36.
7. Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F., Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical image analysis. 2017; 42: 60-88.
8. Jeong-Won Ryu, Jae-Heong Lee. An Efficient Vehicle License Plate Recognition System Based on Embedded Systems. Journal of Next-generation Convergence Technology Association. 2021;5(1):22-27. DOI:10.33097/JNCTA.2021.05.01.22
9. Li Y, Zhao Z, Luo Y, Qiu Z. Real-Time Pattern-Recognition of GPR Images with YOLO v3 Implemented by Tensorflow. Sensors. 2020; 20(22): 6476.
10. Young-Jung Yu, Sang-Ho Moon, Sang-Jin Sim, Seong-Ho Park. Recognition of License Plate Number for Web Camera Input using Deep Learning Technique. Journal of Next-generation Convergence Technology

Association. 2020;4(6):565-572.DOI:10.33097/JNCTA.2020.04.06.565

11. Lee, J. M. Development of translation telephone system by using MQTT protocol. International Journal of Computational Vision and Robotics. 2019; 9(6): 599-608.