Complex Support System for Visually Impaired Individuals

Yavuz Selim TASPINAR [a],* iD, Murat SELEK [b] iD

[a] *Selcuk University, Doganhisar Vocational School, Konya, Turkiye*
[b] *Konya Technical University, Vocational School of Technical Sciences, Konya, Turkiye*

ARTICLE INFO

ABSTRACT

It is very difficult for visually impaired individuals to avoid obstacles, to notice or recognize obstacles in distance, to notice and follow the special paths made for them. They continue their lives by touching these situations or finding solutions with the help of a walking stick in their hands. Due to these safety problems, it is difficult for visually impaired individuals to move freely and these situations affect individuals negatively in terms of social and health. In order to find solutions to these problems, a support system has been proposed for visually impaired individuals. The vision support system includes an embedded system with a camera with an audio warning system so that the visually impaired individual can identify the objects in front of him, and a circuit with an ultrasonic sensor so that he can detect the obstacles in front of him early and take precautions. The object recognition system is realized with convolutional neural networks. The Faster R-CNN model was used and in addition to this, a model that we created, which can recognize 25 kinds of products, was used. With the help of the dataset we created and the network trained with this dataset, the visually impaired individual will be able to identify some market products. In addition to these, auxiliary elements were added to the walking sticks they used. This system consists of a camera system that enables the visually impaired individual to notice the lines made for the visually impaired in the environment, and a tracking circuit placed at the tip of the cane so that they can easily follow these lines and move more easily. Each system has been designed separately so that the warnings can be delivered to the visually impaired person quickly without delay. In this way, the error rate caused by the processing load has been tried to be reduced. The system we have created is designed to be wearable, easy to use and low-cost to be accessible to everyone.

1. Introduction

With the increase in the number of people in the world, the number of people per square meter in public living areas, especially in cities, is also increasing. The world population is approximately 7.3 billion and the number of visually impaired individuals in this population is 253 million in total. About 36 million of this number are completely blind [1]. In Turkey, this number is around 220 million. It is very difficult for these individuals to perceive obstacles in their living spaces, to recognize the objects in front of them, and to take precautions by noticing the auxiliary tools made for them. Visually impaired individuals use a cane to detect obstacles and recognize their surroundings. They may need other people to learn new environments they are not familiar with. In order to get rid of these problems, the use of sensors has become widespread. However, having information about the type of disability will enable the visually impaired individual to progress more safely. With the development of deep learning, progress has been made in the field of computer vision and studies to facilitate the lives of visually impaired individuals have accelerated [2]. It has been developed in systems with sensors capable of real-time object recognition and audible warning [3]. Applications that can run on low-cost smartphones have been developed so that they can find their way indoors and outdoors [4]. Some studies have been carried out so that they can move freely and recognize and identify the obstacles in front of them. Systems capable of GPS, obstacle detection and object detection have been designed. In these studies, ultrasonic sensors were generally used for obstacle detection [5]. In another study, a GPS assisted system was designed to enable visually impaired individuals to move by using ultrasonic sensors and vibration motors. Distance sensors and liquid sensors are used to make the walking sticks they use more functional. When these systems come into contact with liquid, they warn the user with a buzzer [6]. There are studies in the literature that have multiple

wearable obstacle sensors and can warn the user in obstacle detection and liquid detection. There are studies that can detect the localization of objects with sensors that can detect depth using a Kinect or another RGB-D camera [7]. Computer vision is used in object detection and identification. However, the fact that these mobile systems run object recognition algorithms also brings some problems. Due to their processor and memory capacities, they cannot show an advanced level of success. It is foreseen that these problems can be overcome in time. In order for visually impaired people to recognize objects, systems that segment the objects in the scene by using depth sensor cameras and the depth of the images and transmit this to the user have been developed [8]. Although real-time image and video processing can be done by computers, in embedded systems these processes are often not possible or are performed slowly. Therefore, more powerful embedded systems have emerged. Embedded systems with GPU processors can identify objects faster than other embedded systems by performing real-time image and video processing. Existing deep learning models are tested on embedded systems and the fastest way to detect objects is carried out [9]. In these developments, it makes a great contribution to the work done for visually impaired people. In order to solve the problems by examining the studies in the literature, a deep learning-based camera vision support system, an obstacle detector with an ultrasonic sensor, a camera system that detects the roads made for the visually impaired and a sensor system that allows these roads to be followed easily have been proposed. The system includes some differences and improvements in addition to the studies in the literature. The embedded system used in the vision support system was created with Nvidia Jetson TX2, which can easily run deep learning algorithms. The circuit at the tip of the cane is a color sensor circuit that allows to follow the roads made for the visually impaired. There is a camera system in the middle of the cane and this system is a system that allows the user to notice the lines on the road. Multiple ultrasonic sensors are also designed to ensure that the user does not crash into surrounding obstacles. All these systems are designed to be operated easily by the user and are designed independently of each other in order to minimize the error rate. The materials and methods used in the second part of our study, the experimental results in the third part, and the results in the fourth part are given.

2. Material and Methods

In this section, the modules that make up the system are given in order. Firstly, the object identification system (ODS) sub-module, which enables the user to identify and vocalize objects, secondly, the obstacle detection sub-module with multi-sensors that allows the user to recognize the obstacles in front of them, and thirdly, the

yellow line detection sub-module on the cane, which allows to detect the lines made for the visually impaired on the road. module, and finally, the yellow line tracking module at the tip of the cane, which enables the user to follow the yellow line. The modules that make up the system are shown in Figure 1. The use of the system created is shown in Figure 2.
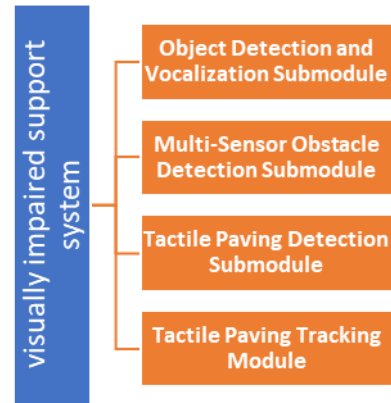

Figure 1. Visually Impaired Support System Modules
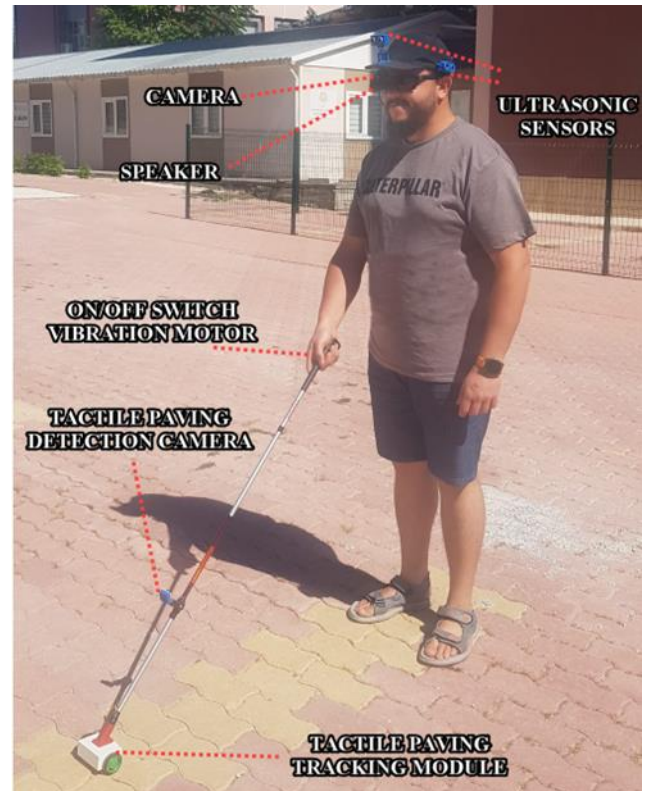

Figure 2. Visually Impaired Support System

2.1. Object Detection and Vocalization Submodule

This module consists of a hat, a camera placed on this hat, and the Nvidia Jetson TX2 embedded system to which the camera is connected. Real-time object identification is made with the images coming from the camera. In order to make this detection, the SSD MobilNet V1 [11] model, which was trained with the COCO dataset [10], was used. There are several reasons why we use this model. The first of these is that this model has a shorter image resolution

time than other trained models. On the other hand, classification success is lower than other models. However, since our object recognition study will be performed on an embedded system, this model has been used to avoid delays in object recognition and vocalization. In addition to this model, our model that we trained with our own dataset was used. In our dataset, there are 9000 images in 25 categories containing market products in Turkey. With this dataset, it is aimed that visually impaired people can recognize the products in the markets. Python programming language and Tensorflow Object Detection API [12] are used to run these models. This tool is used more frequently in object detection operations to be made with bounding box than other applications. The models we use in our study perform the object recognition process by working one after the other. Figure 3 shows the flow chart of object recognition and vocalization processes.
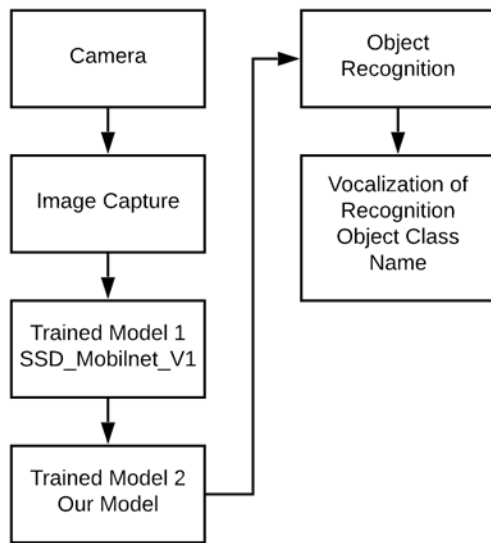

Figure 3. Object Recognition and Voiceover Process Flow Chart

MobilNet model structure was used in the training of our own model. The reason for using both models in the system can be shown as being able to recognize more objects, the ready model can be easily replaced with models with higher classification success when desired, and the model we have created can be continuously improved. Object detection is made by processing the frames coming to the SSD MobilNet V1 model. There may be objects that cannot be defined in this model. The same frame that entered the first model is processed by the second model, that is, by our model, and the objects are defined. Object class names defined in both models are voiced. Vocalization of class names was made with the Python TTS library, and because the names of the classes were in English, they were translated into Turkish and voiced. By selecting the desired language, voice over in

that language can be done easily. A high resolution camera was chosen so that the images from the camera could be clear, but the high resolution brought with it processing load, causing the system to run slowly in some cases. The module contains Logitech C930e camera, Nvidia Jetson TX2 and headphones. The Logitech C930e camera is capable of capturing images in full HD quality and has a 90-degree viewing angle. Nvidia Jetson TX2 has a 4-core ARM57 CPU, 256 Nvidia Cuda-core GPU, 8GB of 128-bit LPDDR4 memory. Thanks to the GPU, image processing can be done quickly. By using these tools, object recognition is performed and voiced by performing operations on the frames coming from the camera in real time [13]. By setting the number of frames per second to 1, the user is prevented from being constantly disturbed by noise. If desired, the number of frames per second can be reduced. However, this can create a security vulnerability for the user.

2.2. Multi-Sensor Obstacle Detection Submodule

Ultrasonic sensors are sensors consisting of two modules that can detect an object, wall or other obstacle in front of it with sound waves. They can calculate distances using high-frequency sound waves that the human ear cannot hear. By vibrating one of the modules with an electrical signal, the module emits a sound wave. The second module creates an electrical signal at its output by vibrating with the sound waves reflected from the obstacle. The distance of the obstacle is calculated by calculating the time it takes for the sound wave to hit an obstacle and return. Because ultrasonic sensors can measure distance, they are also used in different areas such as measuring liquid levels in the tank. Ultrasonic sensors can measure distances between 2 cm and 400 cm, but the sound waves produced must hit an obstacle and bounce back. The sound waves produced may not return for some reason. The softness of the surface on which the sound wave will reflect can make it difficult for ultrasonic sensors to detect the obstacle. The speed of the sound wave emitted by the sensor is 343 m/s. Accordingly, the time required to measure a distance of 1 meter is 6 milliseconds. That's quite enough time for our work. The angle at which the ultrasonic sensor used in our study measures is 30 degrees. For this reason, many sensors are used in our obstacle detection module. It has been tried to minimize the possibility of the user hitting an obstacle without creating any blind spots. The sensor placed in front of the flat is placed on a servo motor, and the servo motor is rotated 30° at 2 second intervals, allowing the sensor to scan a wider area. Figure 4 shows the placement and detection areas of the ultrasonic sensors used in the obstacle detection module.
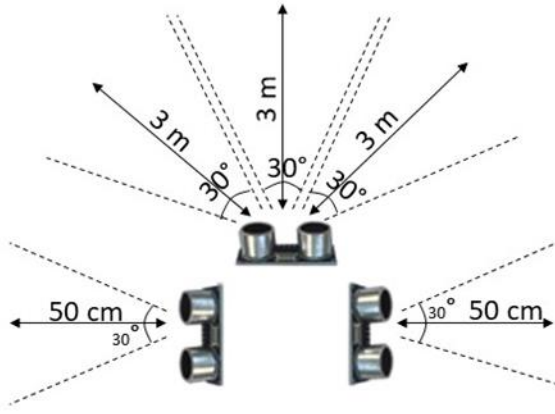
Figure 4. Placement and Detection Areas of Ultrasonic Sensors

There are tree sensors placed on the hat. The front sensor is designed to prevent the visually impaired person from hitting the obstacle in front of him. The object recognition module warns the user by recognizing 80 objects, but this warning system is needed for all obstacles other than these 80 objects. As seen in Figure 2, the viewing angle of each sensor is limited to 30°. For this reason, ultrasonic sensors have been placed on the right and left of the hat against the dangers from the right and left. The sensing distance of the previous ultrasonic sensor is 3 meters, and the sensing distance of the adjacent sensors is 1 meter. These distances can be easily changed if desired. Ultrasonic sensors on the sides are placed in order to prevent the user from hitting his head left and right, and to enable him to notice the objects passing by. The user is warned by 3 vibration motors, right, left and front. Right vibration motor for warnings coming from the right, left vibration motor for warnings coming from the left, front vibration motor for warnings coming from the front works according to the distance of the obstacle. As the obstacle approaches, the vibration motors vibrate more frequently, informing the user that the obstacle is approaching. It is designed in such a way that the warnings coming from the ultrasonic sensor in the front can be changed with the help of a button when desired, with the voice warning system. Right, left and front sensors can be disabled separately.

### 2.3. Tactile Paving Detection Submodule

Tactile paving are placed on the sidewalks, shopping malls and walking areas of various places so that visually impaired individuals can walk more quickly and safely. Visually impaired people can easily move forward by touching the reliefs on these paving with their walking sticks. However, in order to follow these paving, they must first be aware of their existence. Various studies have been carried out in the relevant literature on the determination of these paving. These paving are generally produced in yellow and yellow tones. Colors and designs may vary from country to country. However, it is used as yellow in Turkey. The reason why it is made in these colors is to attract the attention of people who are not visually impaired so that they do not put different objects on these roads. In this study, Raspberry Pi 3 B+ embedded system, Pi Camera compatible with this system, power supply, headphones and buzzer were used for the detection of paving. The images coming from the camera are processed by the Python OpenCV image processing library and the presence of paving is detected. As a method, background subtraction technique and color extraction were used. The background subtraction technique is generally used to detect or track objects by removing them from the background. It is a faster method for object detection than other methods. Along with this method, the color sorting method was also used. With the color extraction method, only objects in the specified color range are displayed by thresholding on the image. When paving in yellow and yellow tones are detected, an audible warning is given to the user. There are two options for the audible warning. The first option is buzzer and buzzer sounds when paving is detected. The second option is an audible warning with a headset, and it warns the user in the presence of paving by playing the desired mp3 sound. A mini speaker can be used instead of a headphone. The user can start and stop this system at any time with the help of a key. It is inevitable that the paving will get dirty over time and the color scale has been kept wide in order to reduce the detection error rate. The color tones used are shown in Figure 5.

Figure 5. Color Scale Used in Tactile Paving Detection

The camera is placed in the middle of the cane and is positioned so that it can easily see the paving on the floor. The area that the camera scans continuously is shown in Figure 6.
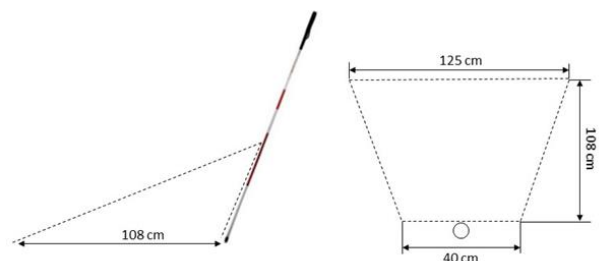
**Figure 6.** Camera Scan Area

When the tactile paving are detected, the user can follow the paving by running the tactile paving tracking module.

2.4. Tactile Paving Tracking Module

The yellow tactile paving on the pavements are made so that visually impaired citizens can find their way by feeling the small notches on the paving that they touch with their walking sticks on the pavement. In this way,

they can continue on the road without falling into the pits, without getting on the road, without hitting something on the curves and they can be protected from other harmful effects. Tactile paving applications are made in order to facilitate the life of visually impaired individuals and to enable them to travel on the roads without being harmed. They try to walk by moving their canes on the paving and feeling the notches on it. However, due to this walking cane process, their progress on these paving is slower. In addition, in the parts of the pavement where there are no tactile paving, the floor may not be smooth and the visually impaired person may mistake them for tactile paving and go in different directions. Studies have been carried out to follow these paving with image processing techniques. Tracking systems have also been developed in order to provide easy access to frequently used places such as toilets and sinks with RFID. In this study, a tactile paving tracking module with a color sensor is recommended so that visually impaired individuals can progress on tactile paving much faster. While designing this module, cost and accessibility came to the fore as in other modules. Arduino Pro Mini, color sensor, power supply and vibration motor were used in the design of the circuit. The module is designed to be easily disassembled and attached to the cane tip. There are strong neodymium magnets on the module and the tip of the cane. The user module is enough to attach it to the tip of the cane. The module is stored in a protective case made in a 3D printer so that it is not affected by external factors. Two wheels are placed on this case. By means of these wheels, the module can move quickly on tactile paving. When the module goes out of the tactile paving, the vibration motor placed on the handle of the cane works and warns the user. In the parts where there is no tactile flooring or when the user does not want to use this module, he can disable the module with the button on the handle of the cane. Since the sensor in the module does not need any external light source, it can also be used in unlit environments. The circuit diagram of the module is shown in Figure 7. The module attached to the cane tip is shown in Figure 8.
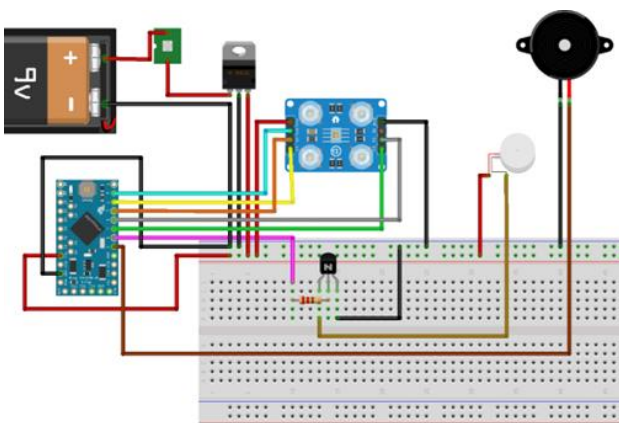
Figure 7. Tactile Paving Tracking Module Circuit Diagram

Figure 8. Tactile Paving Tracking Module

In case the color sensor used in this module cannot detect color due to external factors, the part of the module that touches the ground is covered with a soft material from the sides to touch the ground.

3. Experimental Results

All modules of our support system, which consists of many modules, are designed to not slow down each other's speeds and to give the user the fastest notification. It has been designed in such a way that visually impaired people can meet their daily needs by using equipment that will require minimum power, and they can easily change their power units when necessary. A support system was created with the aim of providing a support system that all visually impaired people can easily purchase, taking into account the minimum cost principle. Each module has been tested within itself. Experiment results are given in this section.

3.1. Object Detection and Vocalization Submodule Experimental Results

This module is the one that will help the user in the support system the most and is open to development. Thanks to the Nvidia Jetson TX2 embedded system GPU processor in this module, it provides a fast operation on the frames taken from the videos. However, it does not comply with the minimum cost policy considered in the design of the support module. However, the Nvidia Jetson Nano system, which is cheaper than this device, can be used instead of this embedded system. This embedded system is used in our system so that the tests can be done easily and quickly. Two models were used in our visual system and these models were tested for success. Our first model is the SSD Mobilnet V1 model, which includes weights previously trained with the COCO dataset. With this model, 80 objects can be defined. The object recognition success of the SSD Mobilnet V1 model is 72.4%. This model has been used because it can run faster on the embedded system. It is important for the detection of

objects as well as the classification rate. The object identification rate was obtained by taking the arithmetic average of the identification rates of the objects. In the same way, tests were carried out with our model created with our own dataset containing 24 classes. The results of the tests performed with SSD Mobilnet V1 and our own model are shown in Table 1. There is no object that both models can define in common. 85% of the generated dataset was used to train the model and 15% for testing. When using models, the process of placing them in the processor, that is, the allocate process, is performed. Since the two models work one after the other, there is a time to allocate the models on the processor. Each model is allocated once on the first run of the program on the processor. These times are shown in Table 1.

Table 1. Success Rates and Allocate Times of Models

Models	Accuracy (%)	Allocation Time (second)
SSD MobilNet V1	72.4	3.19
Our Model	90.2	3.48

A large number of objects can be defined by taking advantage of the pre-trained model, and more objects can be defined with the model trained using the dataset we have created. The product can be defined and voiced from 25 different product packages.

3.2. Multi-Sensor Obstacle Detection Submodule Experimental Results

The sensor placed on the front of the hat in the obstacle detection module rotates at an angle of 30 degrees every 2 seconds with the help of a servo motor. Ultrasonic sensors placed on the sides of the hat scan a scanning area of 30 degrees. In this way, it is ensured that the user can take precautions by noticing all the obstacles in front of him, to his right and to his left. Since the movable sensor placed at the front rotates every 2 seconds, it has been tested against the possibility of missing the obstacles. In addition, the error rate was determined by measuring the differences between the distance measured by the ultrasonic sensor and the actual distance. The distance measured by the ultrasonic sensor and the actual distance are shown in Table 2.

Table 2. Distance Measured by Ultrasonic Sensor and Actual Distance

Distance (cm)	200	150	100	75	50	25	10
Distance Measured by Ultrasonic Sensor (cm)	196	146	99	74	49	24.5	9.75
Error rate (%)	2	2.66	1	1.33	2	2	2.5

When the data obtained in Table 2 are examined, it is seen that the error rate is the least in the measurements made at 75 and 100 cm. The distances measured by the ultrasonic sensor may vary depending on weather conditions, electronic noise in the environment where the sensor is located, and the type of obstacle. The sensors placed on the right and left sides of the hat, which enable the user to notice the objects on the right and left, are adjusted to measure 50 cm. According to the data in Table 2, it can warn the user by measuring with an error of 1 cm at a distance of 50 cm. This value is acceptable for our system. The error rate for the ultrasonic sensor placed on the front of the hat and checking the distance of 200 cm every 2 seconds is 2%. All microcontrollers, sensors and servo motor on the flat are fed from the same power source in order not to weigh the user down. The power consumed when the servo motor is connected and the ultrasonic sensor is connected without any auxiliary elements on the control circuit are shown in Table 3.

By measuring the power consumed by the Arduino and its components, it was observed how much power the obstacle detection system would consume. In this way, how long the user will need the battery can be calculated. As a result of the tests, one battery can operate the obstacle detection system for 8 hours.

3.3. Tactile Paving Detection Submodule Experimental Results

Successful results have been achieved with minimum cost and minimum hardware understanding in this module, which we designed in order to recognize the tactile paving made for the visually impaired. Raspberry Pi 3 B+, the cheapest embedded system, easily ran the OpenCV image processing library. With the camera connected to the Raspberry Pi 3 B+, there was no need for extra cooling in our system, which constantly takes images from the ground. This is an advantage in terms of power

Table 3. Power Consumption of Arduino and Components

	Volt		Amper		Watt	
	Min.	Max.	Min.	Max.	Min.	Max.
Arduino without component	2.800	2.860	0.011	0.014	0.040	0.069
Arduino with Servo Motor	2.800	2.860	0.020	0.065	0.054	0.199
Arduino with Ultrasonic Sensor	2.800	2.860	0.014	0.017	0.069	0.084

consumption.

Raspberry Pi 3 B+ can run smoothly in the module where we use a rechargeable power supply. When operating under maximum power, the power supply module can operate for 9 hours. Real field tests were carried out in order to test the reliability of the module's process of detecting tactile paving and informing the user. Tactile paving are made in shades of yellow in Turkey, and their colors can change over time by being affected by natural conditions. Our preference for a wide color scale that can detect paving with changing colors reduces the error rate of the system.

3.4. Tactile Paving Tracking Submodule Experimental Results

This module has been proposed in order to follow the tactile paving in the fastest way. There are image processing and tracking systems in the literature, but we have designed a simpler, very low cost and accessible module for visually impaired individuals to move on tactile flooring. This module is designed to fit between bubbles on tactile paving. In this way, the user will be able to feel these bubbles and proceed safely. In order for the module to detect tactile paving, it must be at a certain distance from the paving. As a result of the measurements and tests, the distance of the sensor on the module should be a maximum of 25 mm from the paving. Considering this value, the position of the sensor on the module has been adjusted. Although the sensor located at the bottom of the module has its own illumination, it is affected by the external light and gives false warning. In order to prevent this, the part where the sensor is located is closed so that it will not receive light from the outside, as seen in Figure 7. In field tests, tactile paving could be followed successfully with the help of the module at the tip of the cane.

4. Conclusions

In order for visually impaired individuals to live their social lives without any problems, environmental barriers should be minimized. With the support system we have created based on this problem, the work of visually impaired individuals will be easier. In this way, they will be able to devote more time to social life. With the camera placed on the cane, the tactile paving on the road made for the visually impaired can be detected. The presence of tactile paving is reported audibly to the visually impaired individual. With the tactile paving tracking module placed on the tip of the cane, it can easily follow the yellow lines on the road without moving the cane. When it goes out of the yellow line, the cane vibrates and gives a warning. This module can be disabled with the button on the walking stick. With the help of a hat equipped with ultrasonic sensors and a camera, the visually impaired individual can move forward without hitting an obstacle. Whichever direction the obstacle is approaching, the part of the hat where the obstacle is located vibrates and warns the visually impaired individual. With the help of the camera on the front of the hat, 80 different objects are detected and voiced. In addition, a dataset consisting of 9000 images containing market products in Turkey was created. With the model created using this dataset, 25 different market products can be identified and voiced.

The proposed system has been tested in real life and its operability has been tested. Object recognition models are flexible. For this reason, their training can be carried out so that they can recognize more objects. Embedded systems used in the created system can be easily changed by users in case of failure due to time. The proposed system can be transformed into a product so that all visually impaired people can access it.

References

[1] Islam, M.M., M.S. Sadi, K.Z. Zamli, and M.M. Ahmed, *Developing walking assistants for visually impaired people: A review.* IEEE Sensors Journal, 2019. **19**(8): p. 2814-2828.

[2] Kuriakose, B., R. Shrestha, and F.E. Sandnes, *Tools and technologies for blind and visually impaired navigation support: a review.* IETE Technical Review, 2022. **39**(1): p. 3-18.

[3] Simões, W.C., G.S. Machado, A. Sales, M.M. de Lucena, N. Jazdi, and V.F. de Lucena, *A review of technologies and techniques for indoor navigation systems for the visually impaired.* Sensors, 2020. **20**(14): p. 3935.

[4] Choi, J., S. Jung, D.G. Park, J. Choo, and N. Elmqvist. Visualizing for the non-visual: Enabling the visually impaired to use visualization. in Computer Graphics Forum. 2019. Wiley Online Library.

[5] Real, S. and A. Araujo, Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. Sensors, 2019. **19**(15): p. 3404.

[6] Zhang, J., K. Yang, A. Constantinescu, K. Peng, K. Müller, and R. Stiefelhagen. Trans4Trans: Efficient transformer for transparent object segmentation to help visually impaired people navigate in the real world. in Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.

[7] Tapu, R., B. Mocanu, and T. Zaharia, *Wearable assistive devices for visually impaired: A state of the art survey.* Pattern Recognition Letters, 2020. **137**: p. 37-52.

[8] Manjari, K., M. Verma, and G. Singal, *A survey on assistive technology for visually impaired.* Internet of Things, 2020. **11**: p. 100188.

[9] Aruna, M.A., M.B. Mol, M. Delcy, and P.D.M. ME, *Rduino Powered Obstacles Avoidance For Visually Impaired Person.* International Journal of Engineering and Information Systems (IJEAIS), 2018. **3**(2).

[10] Lin, T.-Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C.L. Zitnick. *Microsoft coco: Common objects in context.* in *European conference on computer vision.* 2014. Springer.

[11] Li, Y., H. Huang, Q. Xie, L. Yao, and Q. Chen, *Research on a surface defect detection algorithm based on MobileNet-SSD.* Applied Sciences, 2018. **8**(9): p. 1678.

[12] Sai, B.K. and T. Sasikala. Object detection and count of objects in image using tensor flow object detection API. in 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT). 2019. IEEE.

[13] Taspinar, Y.S. and M. Selek, *Object recognition with hybrid deep learning methods and testing on embedded systems.* International Journal of Intelligent Systems and Applications in Engineering, 2020. **8**(2): p. 71-77.