ICIC Express Letters
Part B: Applications
Volume **11**, Number **7**, July **2020**

ICIC International ©2020 ISSN 2185-2766
pp. 691–696

A PROPOSAL OF WALKING SUPPORT SYSTEM FOR VISUALLY IMPAIRED PEOPLE USING STEREO CAMERA

Hikaru Takefuji, Ran Shima, Pornthep Sarakon and Hideaki Kawano

Faculty of Engineering
Kyusyu Institute of Technology
1-1, Sensui-cho, Tobata-ku, Kitakyusyu 804-8550, Japan
kawano@ecs.kyutech.ac.jp

Received January 2020; accepted April 2020

Abstract. *The number of visually impaired tends to increase in the future. They have many accidents when walking outdoors and the demand for walking support system development is high. They walk on a Braille block by relying on a white stick or guide dog. However, there is a problem that they collide without noticing the obstacle on Braille block. In this paper, we propose the construction of a system that detects and notifies obstacles by combining existing devices and models to prevent collision accident. The proposed device is built on Jetson TX2 and uses a stereo camera ZED Mini that provides depth image. The object detection part is real-time object detection system that is the learned model. Voice from a wearable speaker through a smartphone notifies the location of obstacle. Demonstration test confirms the usefulness of this device.*
Keywords: Visually impaired, Walking support, Object detection, Depth image, Jetson, ZED

1. **Introduction.** Advances in medical care have made people live longer. Therefore, the population is aging and the number of visually impaired is expected to increase from the current 36 million to 115 million by 2050 [3]. It is important to develop equipment for daily assistance for the visually impaired. Walking outdoors is very dangerous for them. There are accidents that cause injuries due to collisions with obstacles such as people, bicycles and cars. A blind person walks on a Braille block relying on a white stick or guide dog. However, it is difficult for the two assistances to notice crossing people and parked bicycles and cars. If the situation around they can be grasped, they can prevent a collision. Thus, we developed a device that finds and notifies existing obstacles.

The obstacle detection in this paper is performed by object detection using images. Research on object detection is actively conducted and its development is remarkable in recent years. You Only Look Once (YOLO) [1] was announced in 2015 using a convolutional neural network proven in image recognition. It is an approach that does not detect the part that seems to be an object, divides the image into grids, performs label classification of the detected part for each area, and performs regression of the position coordinates surrounding the accurate object. More accurate YOLOv2 was announced in 2017 [2]. In the next year the more accurate YOLOv3 was announced. However, the speed is inferior to YOLOv2. In this paper, learning data is small and with low accuracy, but use Tiny-YOLOv2 with emphasis on speed.

In this research, we use object information detected by YOLO and distance information obtained by a stereo camera to grasp the position of the object. By using a stereo camera, you can obtain new distances and directions of objects that cannot be obtained by YOLO alone. And the location information is notified to the user and investigate the usefulness of this device. This device is built on Jetson and uses a stereo camera to acquire information

DOI: 10.24507/icicelb.11.07.691

and detect objects. The detected information is uploaded to Google's cloud. And the mobile phone acquires the information and notifies by voice through the speaker.

2. Related Works.

Several systems have been proposed to help blind and visually impaired people live better. Systems assisting object reaching [5-7] and navigation [8,9] have also been proposed. These systems typically consist of a perception module and a guidance module. Some of them propose to mount camera close to head [5], wrist [6] or both [7]. The notification method suggests using tactile [5], audio [6], or both [7]. However, there is no such thing as distance acquisition by a camera and notification of the position of an object by voice as in this study. It is guided by vibration in a direction free of obstacles using a camera capable of acquiring distance [8,9], but unlike this system, the surrounding environment is unknown.

In [10], authors develop a deep learning-based wearable vision-system with vibrotactile-feedback to guide blind and visually impaired people to reach objects. This study puts a fisheye lens on the arm and uses YOLOv2 for object detection. The direction of movement is notified by vibration of the detected object coordinates. However, you have to keep moving to get information of distance because it is a monocular lens.

The image is acquired by the RGB camera, transmitted to the server through the client, and the object is detected by YOLO. Create 3D audio from the information detected by Unity and notify the position of the object from the earphone [11]. However, it is not accurate because the distance is calculated from the RGB image.

3. The Proposed System.

3.1. Equipment configuration.

The proposed device is built on Jetson TX2 which is an AI computing platform that accelerates the parallel processing of NVIDIA's mobile embedded systems with GPU. This device has high computing power despite its low power consumption. The Jetson TX2 developer kit on the market is a board with various interfaces, but it is large and inconvenient as a carrying device. Therefore, we replaced Jetson with the Jetson TX1/2 Carrier Board manufactured by Macnica Implementation. Figure 1 shows the state of miniaturization.

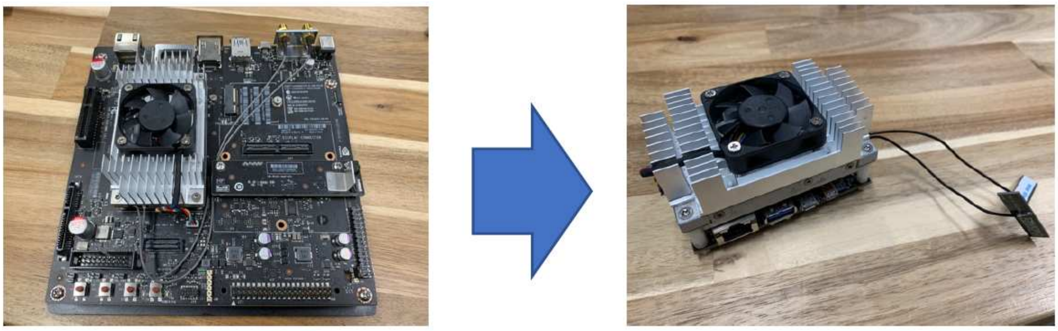

FIGURE 1. The Jetson TX2 development kit board (170 mm × 170 mm) was changed to Macnica Jetson TX1/2 Carrier Board (87 mm × 50 mm).

The stereo camera is selected ZED Mini of STEREO LABS. ZED Mini can easily perform spatial grasping and mapping. The device can instantly acquire a depth image that includes distance information for each pixel. Information can be acquired up to 20 meters, and at angle can be acquired horizontally up to 60° and vertically up to 90°. As shown in Figure 2, a clip fixed to clothes or backpack was attached to the camera. In addition, information of obstacles is notified by voice from the wearable neck speaker via the mobile phone. The wearable neck speaker is shown on the top right of Figure 2. This is used around the neck. In the proposed system, the information of obstacles is put on

Figure 2. ZED Mini, SHARP wearable neck speaker, image of wearing a device

the network, and the information can be obtained by mobile phones. The information is notified by voice through a speaker connected to the mobile phone. Since the visually impaired understand the surrounding situation from the sound, we selected this device to be able to notify the surrounding sound without blocking.

As shown on the bottom of Figure 2, Jetson and the battery were put into the backpack, and the stereo camera was clipped to the shoulder belt of the backpack.

3.2. **Overview of the proposed system.** This system was created on Robot Operating System (ROS). ROS was adopted because it performs distributed processing. Since ROS does not work unless it is connected to the net, it connects from the cell phone to the net. The flow of the proposed system is shown in Figure 3. The distance and direction are obtained by adding depth image information from the stereo camera to the detected object information.

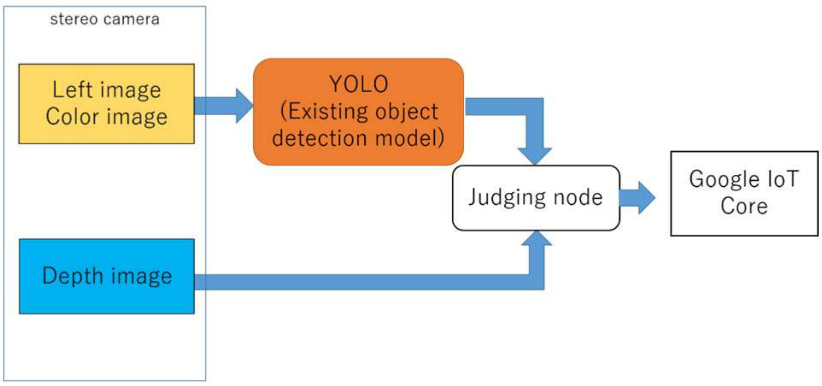

Figure 3. The proposed system flow on ROS

First, the stereo camera ZED Mini acquires left and right color images. At the same time, a depth image is created in the ZED. The color image acquired by the left lens is put into YOLO to perform object detection.

Second, the Judging node obtains the information and depth image of the object detected by YOLO. The detection image of YOLO is shown in Figure 4. The detection

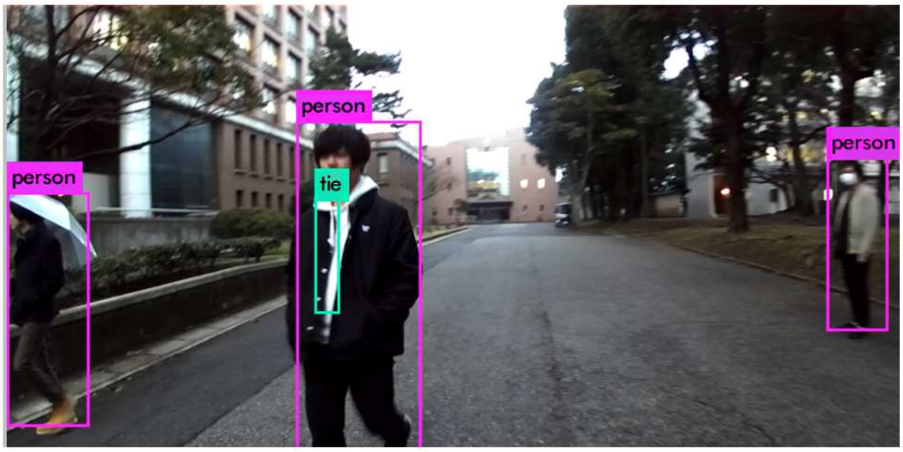

FIGURE 4. Detection image by YOLO

result is given a class such as person or car surrounded by a rectangular shape called Bounding box. The information obtained from YOLO is the class of the object and probability of the class, and the coordinates of the four corners of Bounding box. When all the information of the detected object is acquired, the number of detections may be too large. Therefore, class is limited to people, bicycles and cars, and only those with a probability of 70% or more are notified. The center coordinate of the object satisfying this condition is calculated from the coordinates of the four corners of Bounding box. The distance information at the obtained center coordinates is acquired from the depth image. On the other hand, the direction in which the object exists is determined from the position of the central coordinates. It is classified into three directions, 11 or 12 or 1 o'clock. Information of class, direction and distance is uploaded onto Google's cloud via MQTT.

Finally, the information on the cloud is obtained from the mobile phone and notified by voice through the speaker. The class, direction and distance of the detected object are notified.

4. Demonstration Experiment.

4.1. Accuracy check of object detection before demonstration experiment. We actually walked outside and measured the accuracy of YOLO under this system environment. The measurement results are as shown in Table 1.

TABLE 1. YOLO accuracy on the proposed system

	person	car	bicycle
detection	209	168	156
false detection	3	22	44
total flame	211	190	200
accuracy	99.1%	88.4%	78.0%

Under the present system environment, object detection was performed at a speed of 14 to 17 frames per second. We took about 15 seconds of video and measured the accuracy.

4.2. Experimental conditions. In this research, demonstration experiments were conducted in Oita city and Beppu city, Oita prefecture. Oita Prefecture is one of 47 Japanese prefectures located on the east side of Kyushu. Two types of demonstration experiments were conducted: "staff demonstration experiment" and "accompanying demonstration experiment".

Staff demonstration experiment for 2 days.

Prior to the confirmation of the visually impaired, the blindness association staff should check whether the equipment configuration is defective or whether it can withstand the use of persons with disabilities. In fact, two staff members participated.

Accompanying demonstration experiment for 6 days.

Select one of the multiple walking courses set in the demonstration area, and perform verification while walking as a group of visually impaired and two verification personnel. One of the verifiers mainly checks the verification items and the other checks the safety. When it rained, we conducted a survey of umbrellas and guide dogs. At the destination, we interviewed the feeling of use. A total of 69 visually impaired people were able to participate.

After each demonstration experiment, various questionnaires were conducted, such as the availability of the device and whether increase or decrease of going out.

4.3. **Experimental result.** According to the questionnaire results, 70% of users can go out with confidence, and 71% of users answered that the time to go out will increase. For example, they answered "I want to go to places I have never been to", "I want to go to various places", "More opportunities for walks", "Felt more comfortable", and "Don't feel good even when I feel bad it's easy to go out". For the same visually impaired person we measured changes of walking time with and without the proposed support system at the demonstration experiment. As a result, with the system one way average 26 min 06 s, without the system one way average 30 min 39 s, it was speed improvement of 15%.

It turned out that attaching this device can increase the confidence and motivation of the visually impaired to go out. The walking speed could also be increased. Therefore, the usefulness of the proposed device has been confirmed.

5. **Conclusions.** In this paper, we proposed a walking support system for the visually impaired by object detection. The proposed system also added distance information to the information detected by YOLO, calculated directions and notified them. Demonstration experiments have confirmed the usefulness of the proposed device. However, the information in YOLO is greatly affected by frames. It is considered to be improved by using filtering or object detection with less change in information. It is also necessary to consider the development of a new object detection method using a stereo camera that does not use YOLO. The Jetson is a GPU-equipped device that consumes high power and has a short battery life. So there is room to change to low power devices.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You only look once: Unified, real-time object detection, *arXiv preprint arXiv:1506.02640*, 2015.

[2] J. Redmonand and A. Farhadi, YOLO9000: Better, faster, stronger, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[3] R. R. A. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das and J. B. Jonas, Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis, *Lancet Glob Health*, vol.5, no.9, pp.e888-e897, 2017.

[4] J. Redmon and A. Farhadi, YOLOv3: An incremental improvement, *CoRR*, abs/1804.02767, 2018.

[5] Z. Yu, S. J. Horvath, A. Delazio, J. Wang, R. Almasi, R. Klatzky, J. Galeotti and G. D. Stetten, *PalmSight: An Assistive Technology Helping the Blind to Locate and Grasp Objects*, Tech. Rep. CMU-RI-TR-16-59, Carnegie Mellon University, Pittsburgh, PA, 2016.

[6] K. Thakoor, N. Mante, C. Zhang, C. Siagian, J. Weiland, L. Itti and G. Medioni, A system for assisting the visually impaired in localization and grasp of desired objects, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Germany, Springer Verlag, 2015.

[7] P. A. Zientara, S. Lee, G. H. Smith, R. Brenner, L. Itti, M. B. Rosson, J. M. Carroll, K. M. Irick and V. Narayanan, Third eye: A shopping assistant for the visually impaired, *Computer*, vol.50, no.2, pp.16-24, 2017.

[8] H.-C. Wang, R. K. Katzschmann, S. Teng, B. Araki, L. Giarré and D. Rus, Enabling independent navigation for visually impaired people through a wearable vision-based feedback system, *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp.6533-6540, 2017.

[9] S. Mattoccia and P. Macr, 3D glasses as mobility aid for visually impaired people, in *Visual Interaction Including Biometrics Information for a Socially Assistive Robotic Platform*, 2014.

[10] M. Shih, Y. Chen, C. Tung, C. Sun, C. Cheng, L. Chan, S. Varadarajan and M. Sun, DLWV2: A deep learning-based wearable vision-system with vibrotactile-feedback for visually impaired people to reach objects, *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018.

[11] R. Jiang, Q. Lin and S. Qu, *Let Blind People See: Real Time Visual Recognition with Results Converted to 3D Audio*, Technical Report, Stanford University, http://cs231n.stanford.edu/reports/2016/pdfs/218_Report.pdf, 2016.

[12] E. Aoki, S. Otsuka, T. Ikenaga, H. Kawano and M. Yatsuzuka, Study on autonomous outing support service for the visually impaired, *Conference on Complex, Intelligent, and Software Intensive Systems (CISIS-2019)*, pp.713-722, 2019.