International Journal of Advanced Culture Technology Vol.9 No.4 320-328 (2021)
DOI https://doi.org/10.17703/IJACT.2021.9.4.320

IJACT 21-12-39

Development of IoT System Based on Context Awareness to Assist the Visually Impaired

Mi-Hwa Song

Prof., School of Information and Communication Sciences, Semyung University
mhsong@semyung.ac.kr

Abstract

As the number of visually impaired people steadily increases, interest in independent walking is also increasing. However, there are various inconveniences in the independent walking of the visually impaired at present, reducing the quality of life of the visually impaired. The white cane, which is an existing walking aid for the visually impaired, has difficulty in recognizing upper obstacles and obstacles outside the effective distance. In addition, it is inconvenient to cross the street because the sound signal to help the visually impaired cross the crosswalk is lacking or damaged. These factors make it difficult for the visually impaired to walk independently. Therefore, we propose the design of an embedded system that provides traffic light recognition through object recognition technology, voice guidance using TTS, and upper obstacle recognition through ultrasonic sensors so that blind people can realize safe and high-quality independent walking.

Keywords: *IoT System, Object Detection, Walking Aids, Yolo-v3, Visually Impaired*

1. INTRODUCTION

According to the statistics of the Ministry of Health and Welfare in 2020, the number of registered persons with disabilities in Korea was 2.68 million, the most in the order of retardation (45.8%) > hearing (15%) > sight (9.6%) > brain lesions (9.5%), of which the visually impaired 253,000, accounting for a high rate of about 9.6%. As the number of visually impaired people increases, interest in independent walking is also increasing. However, the current independent walking of the visually impaired mostly depends on convenience facilities and walking aids for the visually impaired, but most of them are in a poor state and there are many problems such as cases where they are not managed or damaged [1, 2].

For this reason, facilities for the disabled must be regularly inspected and repaired. Standard products complying with laws and conditions should be maintained according to logic, but such management is often left unattended. Typically, the Braille block is a facility created for the visually impaired to walk independently. However, there are many cases where the original function is lost due to illegal parking, advertisement flyers, bollards, and other objects and facilities. Facilities blocking the braille block include bus stops, fire hydrants, and civic spaces. In addition, if the protruding protrusion is less than 2 mm, the braille block needs to be replaced, but most of the braille blocks are left unattended despite almost all points being worn out. As of 2019, more than half of the audio traffic lights have been installed to help the visually impaired. However, even if

Manuscript received: October 31, 2021 / revised: November 10, 2021 / accepted: December 7, 2021
Corresponding Author: mhsong@semyung.ac.kr
Tel: +82-43-649-1645, Fax: +82-43-649-1747
Professor, School of Information and Communication Sciences, Semyung Univ., Korea

audio signal lights are not installed at all intersections, it is difficult to recognize the location on the sound signal device due to various advertisements such as flyers. In addition, the braille displayed on the traffic light does not match the spacing between the dots, making it difficult to recognize, and there are cases where the braille is damaged and unreadable [3].

With the recent development of Internet technology, the importance of IoT has increased, and the combination of IoT technologies has created a white cane that can help the visually impaired [4-6]. However, it is difficult to immediately respond to a dangerous situation because it is impossible to judge the upper obstacle and to recognize the situation when the effective distance is exceeded. There are many problems with independent walking of the visually impaired as follows. Therefore, in this paper, we design an embedded system that provides traffic light recognition through object recognition technology, voice guidance using TTS, and upper obstacle recognition through ultrasonic sensor to provide safe and high-quality independent walking for the visually impaired.

2. RELATED WORKS

2.1 Deep Learning-Based Object Recognition

Deep learning-based object recognition used in the IoT environment is typically CNN-based R-CNN, Fast R-CNN, and YOLO [7]. R-CNN creates a candidate region as shown in Figure. 1 and finds the location of an object by learning the CNN in the region [7]. The object recognition of R-CNN consists of three steps. The first step generates candidate regions from the input image using a selective search algorithm, and the second step converts each generated candidate region to the same size and then extracts features through CNN. In the third step, the extracted features are used to classify objects existing in the domain using Support Vector Machines (SVM). After performing the three steps, the position of the object's bounding box is accurately corrected through regression learning.

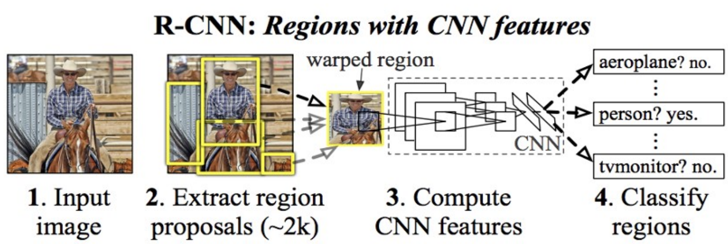

Figure 1. R-CNN

Darknet's YOLO is fast with a simple operation and has the advantage of superior accuracy and performance compared to existing real-time object recognition systems [8]. However, YOLO has the disadvantage of low accuracy for small objects, so the performance of models such as YOLO v2 and YOLO v3 has been gradually improved [8, 9]. YOLO v3 advances the last layer twice over FCN using the function combined with the previous layer to detect small objects, increasing the accuracy of small objects. In addition, YOLOv3-tiny is a faster architecture for YOLO v3, which reduces the amount of computation by convolution and pooling processes instead of FCN, increasing speed at a time through high samples and improving the performance of small objects, but with YOLO v3 and have the same drawbacks. Figure 2 shows YOLO's real-time object recognition system [8].

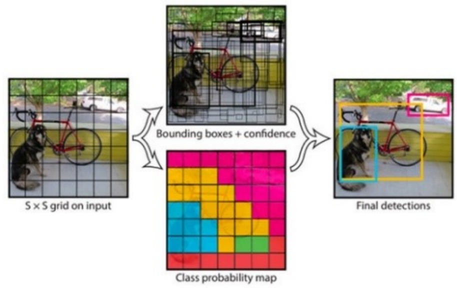

Figure 2. YOLO method

2.2 Blind Walking Assistance System

Figure 3 is a product called Sense Five designed by German industrial designer 'Tim Schutze', and it is a smart wand that delivers information about the surrounding environment to the user [10]. Sense Five detects obstacles at a distance of 5 meters through an ultrasonic sensor, and when an obstacle is close, the tactile feel of the handle changes to convey the surrounding situation to the user. It can recognize fast-moving objects, such as cars, and delivers information about the surrounding environment to the user in various patterns and intensities.

Figure 4 We-Walk is a project that started as a fundraising project of YGA (Young Guru Academy), an international non-profit organization, and is a smart cane for the visually impaired developed by the blind engineer Kursat Ceylan [11]. It is equipped with an ultrasonic sensor on the handle and vibrates to notify you if there is an obstacle above chest height. In addition, it is equipped with a function of navigation, linking with voice support and Google Maps. There is even a microphone, so you can use Amazon's Alexa artificial intelligence (AI) assistant function mounted on the wand.

Figure 3. Sense Five

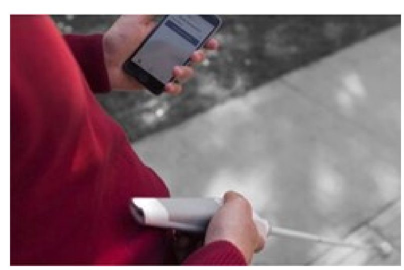

Figure 4. We-Walk

3. PROPOSED SYSTEM

In this chapter, we propose an IoT device and system to reduce the inconvenience caused by visually impaired people moving. The proposed system recognizes the surrounding situation with an ultrasonic sensor and a camera sensor, and provides a function to deliver the current gait information situation through the bone conduction earphone.

3.1 System Diagram

The proposed IoT system is configured using an ultrasonic sensor, Bluetooth module, camera module, and bone conduction earphone. Jetson Nano Board and Yolov3 Tiny model are used as real-time support hardware and implemented as a bag-type model so that blind people can use it with a white cane [12, 13]. The configuration diagram of the proposed system is shown in Figure 5.

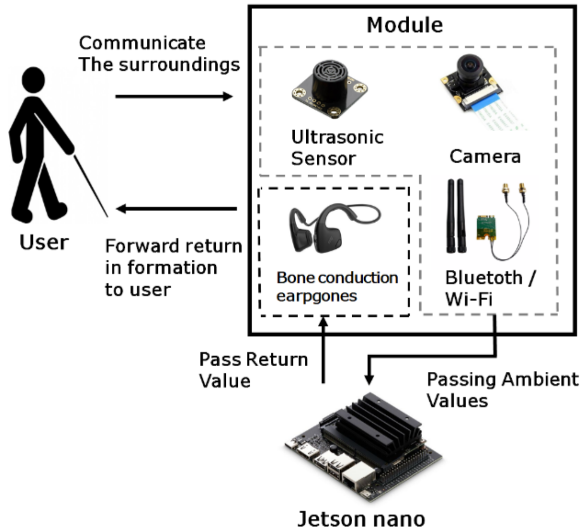

Figure 5. Structure of proposed system

After the user puts on the bag, the information of the external environment is transmitted to the Jetson Nano through the ultrasonic sensor and camera module. Jentson nano outputs the received information through the traffic light algorithm and the ultrasonic sensor algorithm to the user as a voice through TTS through bone conduction earphones. The visually impaired can check the status and status of traffic lights and the presence or absence of obstacles through the received TTS. In the case of the ultrasonic sensor, the upper obstacle of 200 cm to 250 cm and the upper obstacle within 100 cm are measured, and the user is guided by voice through the bone conduction earphone in real time. In the case of the camera module, when the signal displayed at the traffic light is recognized, the result output through the traffic light algorithm is delivered to the user through the bone conduction earphone. The Bluetooth and Wi-Fi module connects the Jetson Nano and bone conduction earphones to deliver the TTS needed to the user, and connects the Internet to Wi-Fi as needed.

3.2 Yolo v3 Tiny Based Object Recognition Modeling

3.2.1 Training Data

The training data used is mainly composed of two types. It was composed by taking photos of crosswalks and traffic lights around S University, and part of it was composed of data collected from the Internet. There are a total of 4 classes to be used for object detection, consisting of traffic lights, crosswalks, red lights, and green lights. As for the photos used in the data set, learning was carried out with 5,000 photos taken directly and photos found on the Internet. In order to improve the accuracy of the data set, various photos were collected from day and night, close and distant photos, and photos taken from various angles for each class and used for learning. Among the weights obtained after completing training, we used a dataset that performed 10,000 trainings because about 2000 trainings were appropriate at a time.

3.2.2 Labeling Framework

Object detection refers to detecting a desired object using a camera or other sensor. Object detection recognizes that input data is embedded and that an object exists as part of that data, and recognizes that an object exists in an image or photo. When all of these processes are executed, object detection is completed. In this paper, Yolo_Mark is used because new data is created and used instead of already learned data. Yolo_Mark can designate the coordinates of the box by drawing the bounding box directly on the image files. There are a total of 4 objects we need to recognize: traffic lights (class name: signal lamp), red lights (class name: red), green lights (class name: green), and crosswalks (class name: crosswalk), each using a class name Thus, the labeling operation was carried out.

3.2.3 Darknet-based Learning

To train the data set created through YOLO_MARK on the Yolo model, this paper uses the Darknet framework. Darknet is an open source neural network framework that enables very fast and accurate real-time object detection. Also, because Darknet is written in C and CUDA, it shows fast execution speed.

A. Darknet Build Configuration

In this paper, Alexeyab's Darknet is used. To build Darknet, you need to set up Makefile first. Since CUDA, CUDNN, and OPENCV are installed in Jetson Nano, change CUDA=1, CUDNN=1, OPENCV=1 in ~/Darknet/Makefile and change to GPU=1 because Jetson Nano has built-in graphics. After that, execute the Make command to proceed with Darknet build.

B. Yolov3-tiny Configuration

After the Darknet build work is completed, the object recognition model to be used by Jetson Nano, and because this paper uses the yolov3-tiny model, yolov3-tiny, the neural network configuration information file of the yolov3-tiny model that exists in the following path: ~/Darknet/cfg. cfg must be modified to match the data set created through YOLO_MARK. The number of classes (objects) that the Yolov3-tiny model should recognize is 4 in total: red light (class name: red), green light (class name: green), traffic light (class name: signal_lamp), and crosswalk (class name: crosswalk) In the neural network configuration information file, the value of [classes] of [Convolutional Layer] is set to 4, which is the total number of classes, and the value of [filter] is 3*(4+5) = 27 through the filter formula 3* (classes+5) to complete the neural network information configuration.

C. Training Process

After completing the setting of the neural network information configuration file, training proceeds to the weight file called yolov3-tiny.weights. Move the created data set file to the following path in ~/Darknet/data/ and enter the train command in the ~/Darknet path to proceed with learning.

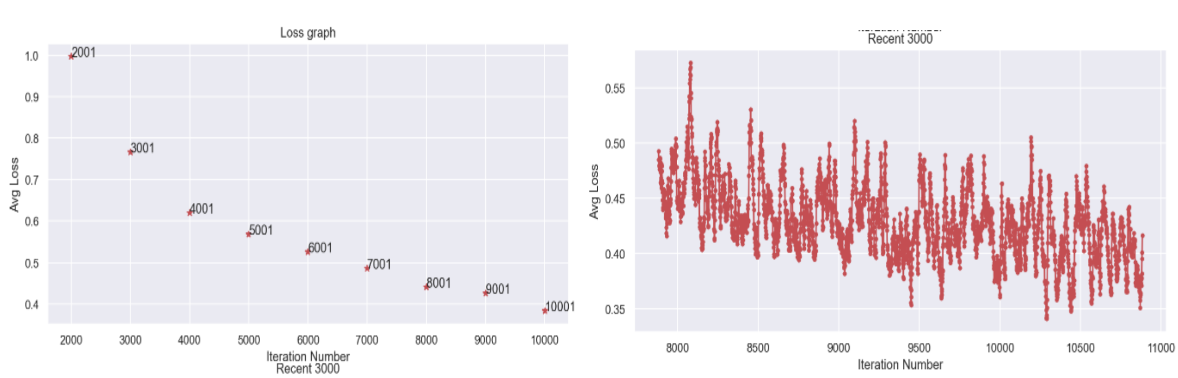

Figure 6. Training processing visualization

As the learning progresses, the change in loss value and accuracy according to the number of learning is shown in Figure. 6. In order to increase the recognition rate of the specified object, 2000~3000 learning times per class are required, so (2000 x 4 + 3000 x 4) / 2 average of 10000 learning is carried out. When training is in progress, various states such as loss value and accuracy are displayed in the terminal in text form for each number of training times. Figure 6 shows the average loss value according to the number of learning times.

3.3 IoT System

3.3.1 Traffic Light Algorithm (Object Recognition)

The object recognition model obtained through learning recognizes a crosswalk with a red light and a green light traffic light as shown in Figure 7. In this paper, the object recognition result is processed through an algorithm, and information to determine whether to walk at a safe traffic light is provided to the visually impaired. The traffic light algorithm is shown in Figure 8. If the camera recognizes a red light at first, it outputs a TTS not to cross because it is a situation in which pedestrians should not cross. Therefore, it prints a TTS telling you not to cross it. If a green signal is recognized after recognizing a red signal, the TTS is output stating that it can be crossed because the change from red light to green light has been confirmed. We implement the following traffic light algorithm.

Figure 7. Object recognition

First, declare red and green variables to determine whether red and green lights are recognized. And to prevent unnecessary repetition of notifications, declare red_count and green_count variables, and increment the count value by one whenever red and green lights are recognized. If the red_count and green_count values reach a certain number of times, the current signal state is output to the user through TTS.

Declare a list named signal_list to store the red and green light signals. And whenever a signal is recognized, a variable name of red or green is put. If the red light is recognized and the variable name red continues to be entered in the list, then the signal changes to green and the variable green enters, deletes all the red variable names in the front based on the index containing the green variable and deletes unnecessary storage space. . If you configure this algorithm, it will output TTS that you can skip when it changes from red to green.

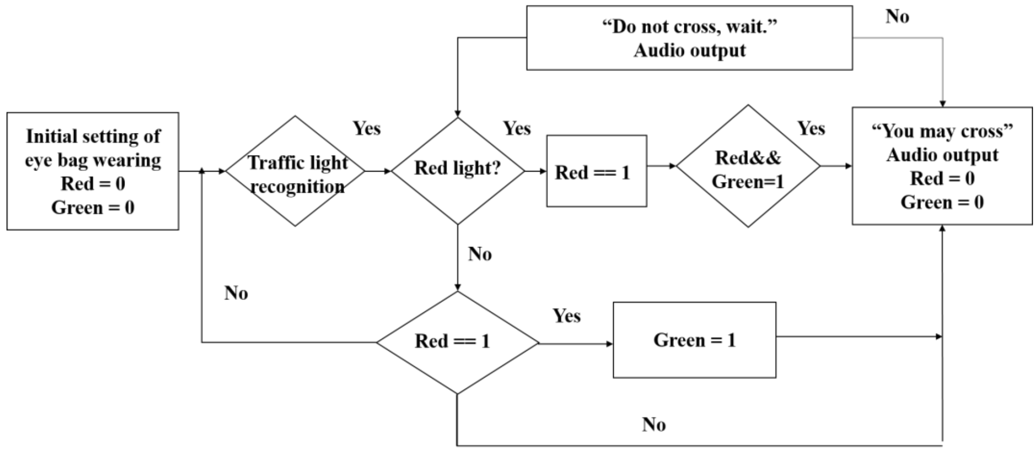

Figure 8. Traffic light algorithm

3.3.2 Ultrasonic Sensor-Based Obstacle Detection

The visually impaired measure the distance of the upper obstacle using an ultrasonic sensor to minimize obstacle collision. In addition, through the processing process through the ultrasonic sensor algorithm, information to determine the presence of obstacles is informed to the visually impaired through TTS. The ultrasonic sensor can measure a distance of 20 cm to 750 cm and is divided into two algorithms according to the effective distance, as shown in Figure 9.

The first algorithm measures the upper obstacle by setting a distance of 200cm to 250cm from the ultrasonic sensor that recognizes up to 750cm, and outputs an alarm sound to the user at regular intervals as TTS. The second algorithm measures the upper obstacle within 100 cm. Since there is an obstacle closer than the first algorithm, the risk factor is high, and an alarm sound that rings at a narrower predetermined time to the user is output as TTS.

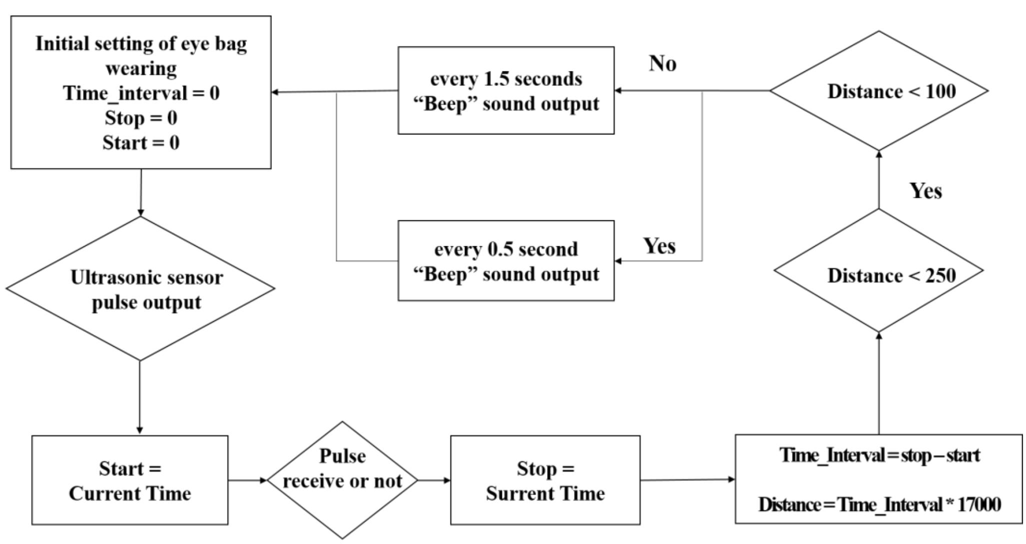

Figure 9. Ultrasonic sensor algorithm

3.4 IoT System Hardware External Design

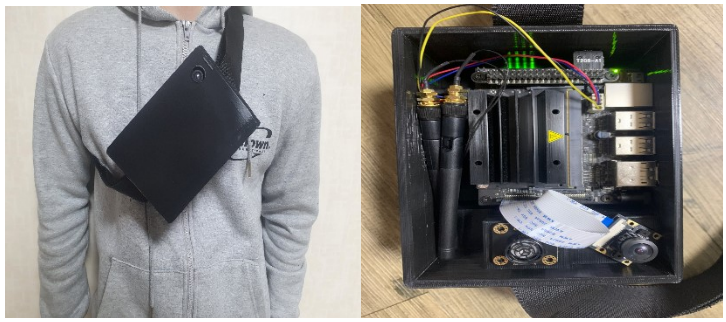

Figure 10. Hardware appearance

Figure 10 shows Jetson Nano, ultrasonic sensor, camera module, and Wi-Fi module implemented in the form of a bag. Considering the angle of the camera, the camera installation angle was adjusted so that the camera could face forward without tilting the angle when worn. It was designed and implemented to protrude to a certain extent, and the UPS module, which is an auxiliary battery module, was designed to be portable by mounting it under the Jetson Nano module.

4. CONCLUSION

In this paper, we designed and implemented an IoT device equipped with a camera module and ultrasonic sensor so that the visually impaired can safely walk in a crosswalk by focusing on one of the most inconvenient factors when walking independently of the visually impaired. The IoT device proposed in this paper solves the recognition of obstacles above and outside the effective distance, which is a disadvantage of the white cane, which is one of the existing walking aids for the visually impaired, through the ultrasonic sensor. The inconvenience of walking in a crosswalk caused by damage or incompleteness was solved by outputting information to the user through TTS and bone conduction earphones for object recognition and traffic light algorithm processing results through a camera. For the IoT device proposed in this paper, NVIDIA's Jetson Nano board and the YOLO v3 model were used as the object recognition model for real-time implementation. This device is expected to provide a better quality of life by resolving the difficulties in social activities caused by the limitations of independent walking for the visually impaired. It is expected to solve more problems by understanding this surroundings, recognizing more kinds of objects, and implementing corresponding algorithms.

REFERENCES

[1] S. Park, J. Park, H. Ryu, H. Lee, and K. Lee, “Development of an auxiliary IoT device and pedestrian guidance service app for visually impaired people,” *Journal of the Korean Digital Contents Association*, 21(2), pp. 269-275, 2020.

[2] S. Park, T. Jeon, S. Kim, S. Lee, and J. Kim, “Deep learning-based symbol recognition for blind people,” *The Journal of Korea Institute of Information, Electronics, and Communication*, 9(3), 249, 2016.

[3] S. Lee, M. Kang, “Visual impairment auxiliary object detection and voice guidance system using object recognition technology,” *IEIE Transactions on Smart Processing & Computing*, 55 (11), pp. 65-71, 2018.

[4] J. Kim, S. Jung, and Y. Yoo, “Implementation of walking deep learning cane for blind people,” *The Journal of the Institute of Internet, Broadcasting and Communication*, 24(1), pp.343-345, 2020.

[5] S. Oh, G. Jeong, H. Kim, and Y. Kim, “Development of a visually impaired road crossing assisted embedded system using machine learning,” *Journal of the HCI Society of Korea*, 14(2), 41-47. 2019.

[6] S. Lee, K. Lee, S. Lee, J. Ko, and W. Yoo, “Technology Trends and Analysis of Deep Learning Based Object Classification and Detection," *Electronics and Telecommunications Trends*, 33(4), pp.33–42, 2018. DOI: 10.22648/ETRI.2018.J.330403

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 580-587, 2014.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. of the IEEE conference on computer vision and pattern recognition*. pp. 779-788, 2016.

[9] J. Redmon, A. Farhadi, "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[10] Sense Five, https://www.behance.net/gallery/100682827/Sense-Five (accessed September 7, 2021)

[11] 'WeWork', a smart cane for the visually impaired, http://www.bizion.com/bbs/board.php?bo_table=gear&wr_id=1441&sca=Outdoor%2CLeports (accessed September 7, 2021)

[12] NVIDIA Jetson Nano, https://developer.nvidia.com/embedded/jetson-nano-developer-kit (accessed September 8, 2021)

[13] NVIDIA Jetpack SDK, https://developer.nvidia.com/embedded/jetpack (accessed September 8, 2021)