iWalkSafe - Wearable Navigation Assistance for the Visually Impaired Based on Miniaturized Edge AI

1st Wei En Tsai
National Tsing Hua University

2nd Kuo Cheng Chin
DeepMentor Technology

3rd Borhan Lee
National Tsing Hua University

4th Wei Chung Chen
National Tsing Hua University

5th Ren Song Tsay
National Tsing Hua University

***Abstract*—We propose a computer vision-based, wearable navigation assistance based on a miniaturized AI model, utilizing the benefit brought by a specialized edge AI box. This study contributes to the literature in two ways. First, we implement a navigation system with high mobility, helping the visually impaired eliminate most of the common dangers in their daily lives. Second, we demonstrate how our previous work — a patented CNN miniaturization method — can be put into a critical application such as offline navigation assistance. The miniaturization technology enables the system with high mobility and low power consumption, with a negligible precision trade-off.**

***Index Terms*—wearable, edge AI, computer vision, AI miniaturization**

I. Introduction

Vision impairment severely impacts the quality of life in both physical and mental aspects. Studies have shown that vision impairment can contribute to social isolation, difficulty walking, a higher risk of falls, and a greater likelihood of early entry into nursing homes [1]. Furthermore, visually impairment has been a serious global issue; according to VLEG/GBD, over 338 million people in the world are visually impaired, out of which 43 million are blind and 295 million have low vision [2]. Dependency on others, reduced social interactions, and difficulty walking caused by vision impairment may lead to higher rates of depression and anxiety [1]. Thus we propose a computer vision-based system to aid visually impaired people with mobility.

Most of the assistive devices on the market use cloud computing to reduce the computation load and power consumption. However, it makes the product highly rely on the network connection. They may fail to operate in areas that do not have good coverage of internet (e.g., on the subway, in an elevator), which may lead to severe accidents. Therefore, there is a trend of using edge devices as replacements. However, new challenges emerge when using a edge device. A standard AI model consumes a vast amount of energy and thus leads to short battery life, whereas some quantization approaches for Edge AI miniaturization sacrifice an unneglectable amount of precision.

With our dedicated AI box and a patented miniaturization approach, computer vision algorithms can run offline in real-time with a low-power supply and neglectable precision loss (1% loss compared to the standard model). Our system would inform the user with audio output and haptic motor regarding "danger avoidance" or "environment perception" events. As a wearable navigation assistance, our system features offline availability, low power consumption, and comprehensive event detection.

II. Related Work

Various computer vision-based aid systems are proposed in the literature. Some of these convert depth information captured by the sensor into human auditable sound as a clue. For example, the vOIce system [3] converts captured image into sound to give information regarding environments, whereas it is slow and physically uncomfortable. Most importantly, the learning cost of using such device is unrealistically high.

Some literature takes advantage of the advanced wearable technologies, attaching multiple sensors and the computing unit on daily-used devices. For example, NavBelt [4] is an intelligent waist belt with eight sensors with a computing unit located in a backpack. GuideCane [5] is a wheel-based cane with ultrasonic sensors, and the computing unit is also embedded in the cane. However, the images captured by these devices are inherently unstable due to the high degree of movement and thus may lead to blurry images. Therefore, these wearables are experimental and not suitable for practical applications.

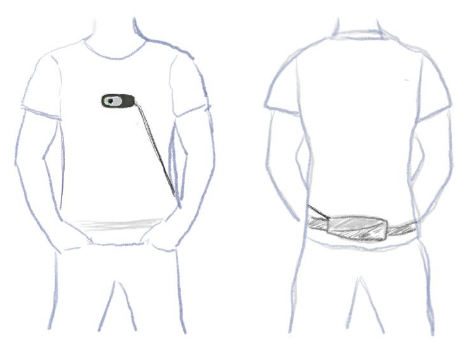

Fig. 1. Our proposed wearable design

III. System Design

In this paper, we propose a novel and practical design for a visual assistance system using a wearable RGB camera attached to a cloth (See fig [1]) and an edge AI device as the computing unit that can overcome the limitations of existing approaches. We utilize the patented miniaturization technology in our previous work [6], [7] to deploy miniaturized AI models on the edge device, to perform advanced computer vision tasks, including object-detection, people tracking, and facial recognition.

Our goal is to enable the visually impaired to walk independently, addressing most common problems, including obstacle avoidance, obstacle/object detection, and communication. As a wearable navigation assistance, our system features offline availability, low power consumption, and comprehensive event detection.

Our hardware consists of a general computing platform and 4 Coral TPU (Tensor Processing Unit) for deep learning acceleration. Therefore, it can achieve 16 TOPS (tera operations per second) for deep learning tasks. The computing platform connects to the TPU through the PCIe interface and connects to the RGB camera through the USB 2.0 interface. The GPIO is used to actuate the haptic motor and control the external LED for the user interface. (See fig [2])

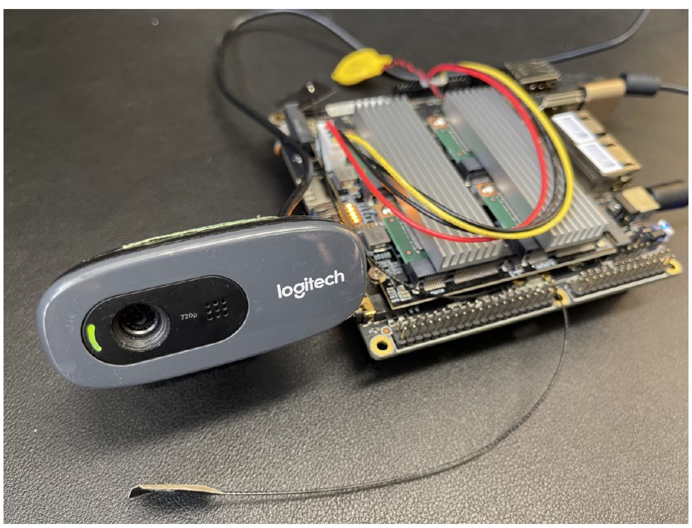

Fig. 2. Hardware: RGB camera and AI box

From a software perspective, our tasks mainly involve object detection, pose estimation, and face recognition.

The software architecture of our design can be separated into three fields of tasks. Through our patented miniaturization technology, we deploy AI models on this entry-level hardware, with low data movement (90% data are reduced) and low power consumption (about 20W in total), whereas 99% of model accuracy is retained. Given four TPUs on the AI box, we deployed three deep learning algorithms on our system, and we left 1 TPU for future use. These three algorithms are object-detection-and-tracking, pose-estimation, and face recognition. Our primary functions are powered by these algorithms but not merely bonded to these algorithms; we have the main program that summarizes the outputs of these algorithms and gives the final logical output as the reaction to the environment to utilize these data more comprehensively. For example, while the pose-estimation algorithm may give information about a person's pose, we would combine it with face recognition to inform the user of the person's identity.

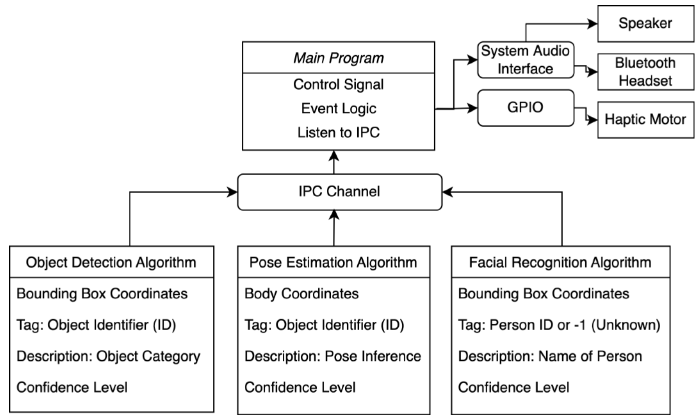

Fig. 3. System context diagram

The data pipeline is illustrated in fig 3, and these three algorithms connect to three virtual cameras duplicated from the actual camera. Moreover, the results of the three algorithms are sent to the IPC channel created by the operating system API. The main program then listens to these IPC channels in real-time and gives a logic output, which triggers the following actions such as informing or warning notification. Notification is sent to the user by haptic motor and speaker, as the computer-human interface.

IV. Result

We use common scenarios on campus from a visually impaired student's perspective to demonstrate and evaluate our work.

The purposes of our event-triggered notifications can be mainly categorized into two goals, danger avoidance, and environment perception.

Danger avoidance is the most critical goal of a safety assistance device; therefore, we provide runner-detection functions and crowd-detection (Fig. 4). When one or more runners are present in the field of vision, the system would actively alert the user through the haptic motor and system audio output The user can optionally choose to output through a speaker or a Bluetooth headset.

Environment perception also contributes to the welfare of the visually impaired because it gives people a sense of where they are and therefore brings them confidence and feel of control over their lives. Users can push the device's external button, and the system would passively give a verbal description of the current scene. For example, "There is a dog in front of you." (Fig. 5), "The man in front of you raises his right hand." (Fig. 6), and "Your friend Wayne is at your right-hand side." (Fig. 7)

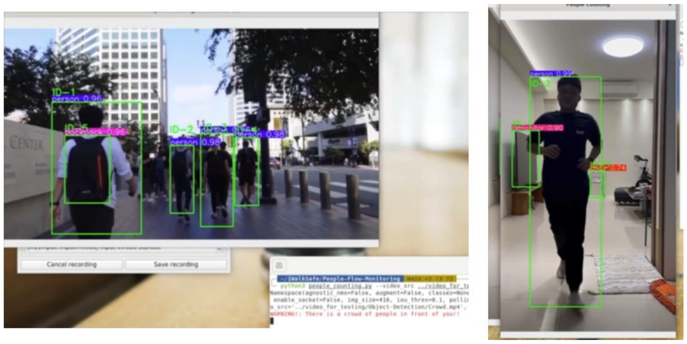

Fig. 4. Result of crowd-detection and runner-detection

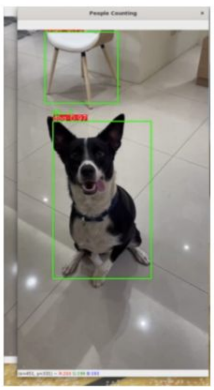

Fig. 5. Object Detection

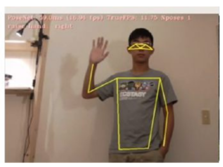

Fig. 6. Pose Estimation

Fig. 7. Face Recognition

V. Conclusion

In this project, we proposed a novel, comprehensive wearable navigation system for the visually impaired. The system leverages previous work of our lab, a patented miniaturization technology, to build an efficient, low-powered, high-precision edge AI device. The system provides proof of how edge AI devices can be deployed in personal assistive equipment and shows that the miniaturized models perform evenly well as the original models.

References

[1] Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years: evaluating the prevalence of avoidable blindness in relation to "VISION 2020: the Right to Sight". Lancet Global Health 2020. doi.org/10.1016/S2214-109X(20)30489-7

[2] Bourne R, Steinmetz J, Flaxman S, et al., Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Glob Health. 2020. Accessed via the IAPB Vision Atlas (https://www.iapb.org/learn/vision-atlas)

[3] P. B. Meijer. An experimental system for auditory image representations. In IEEE Trans. Biomed. Eng., vol. 39, no. 2, Art. no. 2, 1992.

[4] S. Shoval, J. Borenstein, and Y. Koren. The Navbelt-A computerized travel aid for the blind based on mobile robotics technology. In IEEE Trans. Biomed. Eng., vol. 45, [32] no. 11, Art. no. 11, 1998.

[5] I. Ulrich and J. Borenstein. The GuideCane-applying mobile robot technologies to assist the visually impaired. In IEEE Trans. Syst. Man Cybern.-Part Syst. Hum., vol. 31, no. 2, Art. no. 2, 2001.

[6] Tsung-Yin Lu and Ren-Song Tsay. A Normalization-based Logarithmic Quantization Approach for Convolution Neural Network Miniaturization. [Master's Thesis, National Tsing Hua University], 2019. https://hdl.handle.net/11296/243mk5

[7] Hsu-Hsun Chin and Ren-Song Tsay. A High-Performance Adaptive Quantization Approach for Edge CNN Applications. [Unpublished master's thesis, National Tsing Hua University], 2021. https://hdl.handle.net/11296/ur5mzv