Intelligent Offline Multi Object Recognition Walking Stick for The Blind

Ibrahim Mohammed Abdullahi[1], Olayemi Mikail Olaniyi[1], Jacob Omokhafe Irefu[1], Sangwon Oh[2], Ibrahim Aliyu[2]

[1] Department of Computer Engineering, Federal University of Technology, Minna, Nigeria

{amibrahim, mikail.olaniyi}@futminna.edu.ng, omotsemoje@gmail.com

[2] Department of ICT Convergence System Engineering, Chonnam National University, Gwangju, 61186, South Korea

osw0788@gmail.com, 187282@jnu.ac.kr

Abstract. Vision is one of the most important characteristics of a human that aid their day to day activities. Loss of vision however affects the ability of humans to freely navigate their environment and recognized objects along their path. Existing object recognition systems for the blind are mostly cloud based and its performance depends on reliable internet access. This makes them unsuitable in places with unreliable internet. Therefore, in this paper, a multi-object recognition intelligent walking stick for the blind that is completely independent of the internet was developed. The system is divided into three units, detection, recognition and communication units. The detection unit make use of an ultrasonic sensor and a buzzer, for informing the user of an impending obstacle. The recognition system makes use of a camera for capturing images with Convolutionary Neural Network architecture and Mobile Network Single Shot Multi-Box Detector (MobileNet SSD) for detecting objects in images. The communication unit transmits the recognised objects through voice to the user in English Language. The entire system was deployed in a Raspberry Pi microcontroller due to its processing power. The result obtained from testing of the device on the field showed that the recognition unit achieved an average sensitivity, specificity, precision and accuracy of 87.26%, 67.45%, 89.07%, 82.50% respectively. This shows that the system is reliable and can be used in recognizing objects for the blind.

Keywords: Blindness, Deep Learning, Object Recognition, Single Shot Multi-Box Detector, Open Computer Vision

---

JCC ⓒ Journal of Contents Computing

Vol. 3, No. 2, pp. 351-364, Dec. 2021

Received 29 August 2021
Revised 06 November 2021
Accepted 14 November 2021

***Corresponding Author;** Ibrahim Aliyu

Tel: +82-062-530-0407
E-mail: 187282@jnu.ac.kr

1 Introduction

Vision is the most significant part of human physiology as 83% of information human beings receives from the environment is through sight [1]. The 2019 statistics by the World Health Organization (WHO) on visual impairment and blindness estimated at least 2.2 billion cases globally [2]. The number of visual impaired persons in low and middle-income countries is four times higher than high income countries [3]. Visual impairment are mostly caused by the following: cataracts, uncorrected refractive errors, corneal opacity, glaucoma, age-related macular degeneration, diabetic retinopathy and trachoma [2, 4]. Some of the challenges faced by the visually impaired are their inability to move independently in their environment. Traditional aids like, walking sticks, human and well-trained dogs are commonly used to help the visually impaired people to ease their movement around their environment [5]. However, traditional walking stick only detects obstacles on contact making it inconvenient to use, while Guide dogs requires the individual to be trained by orientation and mobility specialists to be able to comprehend the dog's movement.

Electronic Travel Aids (ETAs) devices were introduced to serve as a mobility aid for the Visual Impaired. ETAs devices consist of sensors that alert the blind about obstacles in their way through vibration and sound. Such devices increase their perception and provide safety, as they provide the necessary aid to assist the visually impaired to facilitate their movement in unfamiliar environment. Some of such devices includes a GPS base Electronic white cane for blind, that helps locate the blind when they get lost [6], Infra-Red (IR) Radio Frequency Identification Walking Stick (RFIWS) device used to guide the visual impaired but can only detect obstacles [7] and an ultra-cane that uses ultrasonic sensors for guiding the visually impaired and detecting obstacles [8]. Some advanced ETA's that are developed to recognize objects and guide visually impaired persons are cloud based and therefore requires a reliable internet access for optimal operation [1]. This therefore means that such systems cannot be effective in regions with unreliable and unavailable internet access.

To overcome these challenges, this paper proposes an offline multi-object recognition system that operates without the use of internet using deep learning and Raspberry pi microcontroller. The microcontroller was used to deploy the system because of its processing power is sufficient to handle deep learning and computer vision processes. The remainder of this paper is structured as follows: Section 2 presents the related electronic walking aids and related systems in literature, section 3 presents the method and materials used in the systems development. Section 4 presents the simulation and deployment test results as well as the discussion of the results. Finally, section 5 presents the conclusion and future research directions.

2 Related Works

Technological advancement has brought about radical changes into our immediate environments in many fronts especially for daily activities in particular and people with various challenges. Various technologies have been developed to ease navigation within known and unknown paths for visually impaired persons. These include obstacle detection, path and obstacle recognition systems which can be grouped into offline based and online based systems. The online systems are mostly cloud based where heavy computations are deployed in the cloud or an online server. On the other hand, the offline systems rely on microcontrollers for systems deployment. Due to the high computational requirement of computer vision and deep learning applications, the type of microcontroller and algorithm must be carefully selected. [9] designed a guidance system for the visually impaired using different sensors. The designed system works perfectly on a clear path and known environment. This system neither be deployed to an unknown environment nor can it be used without the path being indicated. It lacks Global Position System (GPS) and advanced sensors such as accelerometers, PIR motion detector and digital compass that will communicate exact location of the user. [10] proposed a novel walking stick with two distinct mechanisms: gripping and lifting in order to be able to grip and lift objects from the ground conveniently to the user's height without losing contact with the ground. The importance of the proposed mechanism is to reduce the frequency of users having to stoop down to pick their belongings at as when due hence the proposed system would be useful to the society not only blind users but also for conventional walking stick users. An intelligent ultrasonic-based walking stick designed for blind users to enhance and ease individual navigation was presented in [11]. The system uses combination of ultrasonic and water sensors for obstacles detection and water ahead. It has wireless RF remote control that distinct sounds when pressed which aids to locate the walking stick whenever it gets misplaced. Accordingly, different sound of buzzer was used to differentiate type of obstacles when approached. A smartphone-based online and offline guiding system for visually impaired people was proposed in [12]. The system which can be operated in online and offline mode uses deep learning for recognition of obstacles and a smartphone for capturing images. The systems recognition is 60% which can be improved. Also, the systems dependent on smartphone makes it an expensive solution.

[13] developed an internet of things mixed with Radio frequency identification-based strolling stick system for visually challenged users to aid their navigation without stumbling on obstacles on their path. It used ultrasonic detector to transmit supersonic waves into the air and detects mirrored waves from an object on it path. The speed of wave transmission is measured by multiplication of frequency and wavelength which translates to higher resolution as a potential to induce higher measure and accuracy. The system can be used in any terrain, detects obstacles as well as environmental changes easily, however, it lacks vital features like narrator

and GPS capabilities. [14], designed an intelligent voice navigation walking stick based on GPS and GSM using ultrasonic obstacle avoidance, night warning and emergency assistance. A smart stick for the blind was proposed in the work of Romadhon and Husein [15]. The proposed smart stick was designed to detect obstacles using ultrasonic sensor, detect the presence of puddle or flooded area using water sensors and it can measure the pulse of the user while reporting the location of the user using GPS. The system, however, lacks obstacle recognition capability. An IoT based smart walking stick for blind or aged people was proposed in [16]. This system is capable of detecting obstacles and moisture but cannot recognize the nature of the obstacle. Similarly, an IoT based voice controlled blind stick for blind people guidance was proposed in [17]. Obstacles within four meters can be detected using infrared sensors and a voice message is initiated to warn the blind of impending obstacles. However, it lacks obstacle recognition capabilities. In the work of Singh and Singh, [18], an intelligent walking stick for elderly and blind people was proposed. The stick which is capable of detecting obstacles, ground water, pit and location of the user. It can also measure the users' temperature, pulse rate and alert others on the health status of the user. It however lacks recognition capabilities.

From the reviewed works, most systems lack obstacle recognition capabilities and systems with real-time obstacle detection and recognition capabilities require internet which is not reliable and most case, unavailable in most developing countries where visual impairment is mostly prevalent, hence, the need for the proposed system.

3. Materials and Methods

The developed system is composed of two different modules with each module consisting of both hardware component and its associated software component. The Object Recognition module uses the pi camera to capture images in user's path and sends it to the Raspberry Pi3 model B+ for processing and recognition. The Feedback module uses speech to inform the user through any audio outlet, about objects known by that system that was detected in the image captured. Raspberry Pi is the Micro Controlling Unit (MCU) used in the system, it possesses the desired processing capacity required for object recognition. It interfaces with pi camera, Control Buttons and Ear piece. The Object Recognition module software makes use of a Mobile Network Single Shot Multi-Box detector which is a light weight deep neural network build from depth wise separable convolutions and outputs class probabilities directly from the proposal generator over several aspect ratios and scales per feature map location. Java and python programming language together with open CV framework were used in developing the software of the system. Figure 1 shows the architecture of the system. Figure 2 shows the block diagram of the system. It explains the methodology of the system adopted and the various components of the system. The system development follows a precise, step by step model to accomplish the developed system.

3.1 Data Acquisition and Preprocessing

The Image data were acquired using an 8 Mega pixel Raspberry pi camera with 1920*1080 resolution. The image Data acquired includes: Human Beings, Cars, Bus, and Motor Bikes. Several images were acquired which was used to train and test the system. 100 images for each image object were taken, of which 70 of them were used for training, while the remaining 30 were used for testing. The images were taken in the Joint Photographic Experts Group (.jpeg) format, with a size of 1920*1080 in RGB colour format. The images that were acquired were normalized using histogram equalization to adjust their contrast and resized to 300 x 300 to reduce pixel data for training and test phase.

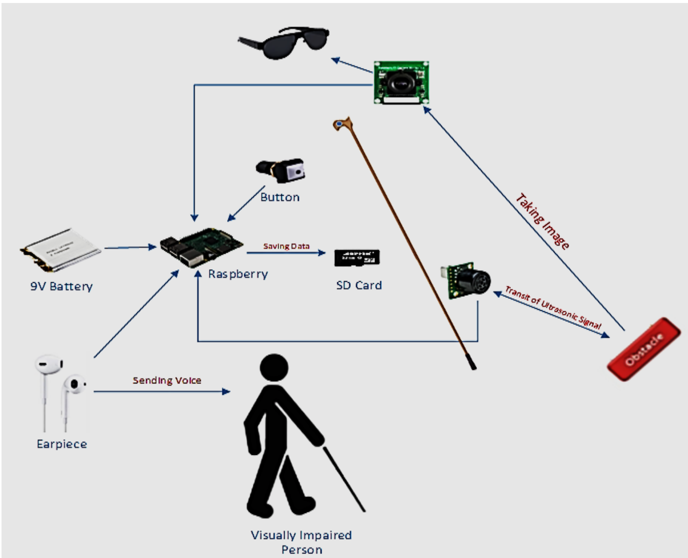

Fig. 1 Multi-object detection system architecture

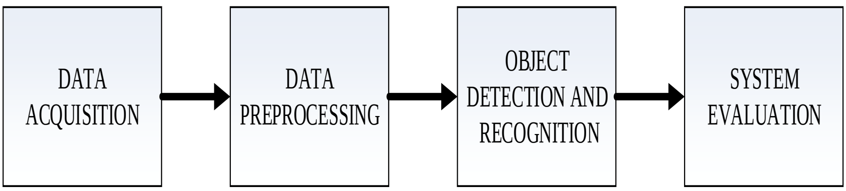

Fig. 2 System Block Diagram

3.2 Model Training and Testing

The images that were pre-processed were used to train a Convolutionary Neural Network (CNN). The Neural Network adopted was a Mobile Network [19], Single Shot Multi-Box Detector (MobileNet SSD) [20] for object detection. This model was used to develop classifiers for detecting and recognizing objects from images. The cascaded network layers used is shown in Figure 3

$$INPUT > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > CONV > RELU > FULLY\ CONNECTED\ LAYERS > FULLY\ CONNECTED\ LAYERS > OUTPUT$$

Fig. 3 Object Recognition Convolutional Naural Network Architecture

A configuration file was created that describe the layers in the neural network architecture and its associated parameters, known as the prototxt file. The MobileNet SSD that was obtained in the training stage was used to recognize objects from the testing dataset. Figure 4 shows the algorithm used in detecting objects from images.

3.3 Performance Evaluation

The performance of the trained model was evaluated over the following evaluation metrics:

Sensitivity: The number of true positives relative to the sum of the true positives and the false negatives. It is a measure of the percentage of true targets that is detected.

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

Specificity: The true false detections relative to the sum of the true false detections and the false positive. It provides a measure of the likelihood of a negative response given the total number of actual negative detections.

$$SPC = \frac{TN}{FP + TN} \times 100\% \tag{4}$$

Precision: The number of true positives relative to the sum of the true positives and the false positives. It is the fraction of detected items that were correct.

$$PPV = \frac{TP}{TP + FP} \times 100\% \tag{5}$$

Accuracy: The sum of the true positives and the true negatives relative to the total number of objects detected during evaluation.

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{6}$$

Where, TP (True positive) is correctly detected positive objects from an image. TN (True negative) is correctly detected negative objects from an image. FP (False positive) is incorrectly detected negative objects from an image and FN (False negative) is incorrectly detected positive objects from an image.

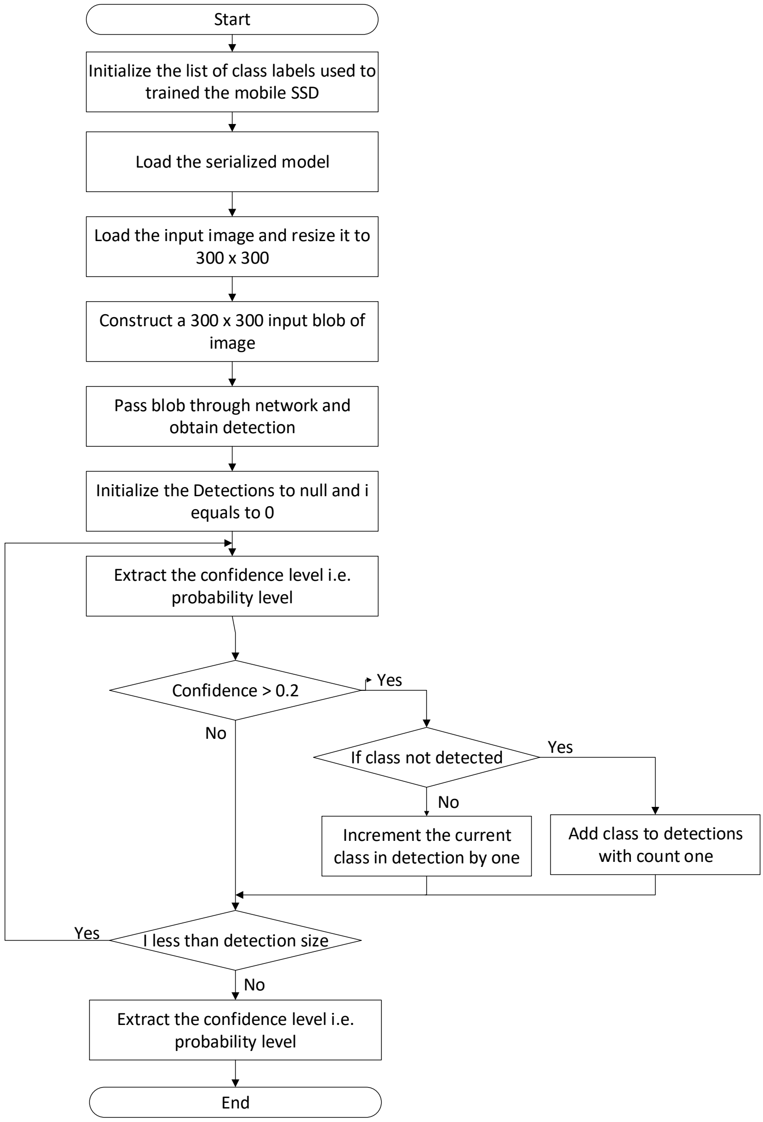

Fig. 4 Flow Chart for Image Recognition Module

3.4 System Deployment

Several setups and configurations were made on the Raspberry pi micro controller before deploying the discussed software components onto it. Some of such setups and configurations are:

i. Installation of Raspbian Operating System, a Linux distribution, on the Raspberry.
ii. Installation of Virtual Network Computing (VNC) software, a remote computing software, used for sharing the raspberry operating system on a Computer.
iii. Installation of PI4J API library on the raspberry, for seamless communication between Java and the Raspberry GPIO pins, serial interfaces, and other peripherals of the Raspberry.
iv. Disable the Raspberry Serial console in other to enable its Universal Asynchronous Receiver-Transmitter (UART) interface use by the GPS receiver.
v. Installation of Open Computer Vision (Open CV) Image Processing library and framework containing efficient algorithm used for Deep Neural Network computing like Convolutionary Neural Network computing.

4 Results and Discussion

The performance of the system was evaluated both at the training and testing phase of the development process. At training phase, the accuracy yielded per iterations made was observed and this is shown in Figure 5, also known as the training curve. Training curve shows the maximum accuracy that could be attained by the neural network. The results in Figure 5 shows that the accuracy of the detector increases as the number of iterations increases while the training and test loss decreases, in order words every iteration yields a gain in accuracy. This continues till the four thousandth iterations, where any increase in the number of iterations doesn't result to an increase in accuracy, thereby yielding no decrease in training lose. Table 1 shows the result obtained from evaluating the performance of the detector in recognizing the following objects: Person, Car, Bus and Motor Bike in terms of sensitivity, specificity, precision and accuracy.

Fig. 5 Convolution neural network training curve

Table 1: Performance evaluation result for object recognition

Objects	Sensitivity	Specificity	Precision	Accuracy
Person	83.33%	71.43%	88.24%	80.00%
Car	88.73%	75.86%	90.00%	85.00%
Bus	87.50%	60.00%	89.74%	82.00%
Motor Bike	89.47%	62.50%	88.31%	83.00%
Mean	87.26%	67.45%	89.07%	82.505

The results show that the average fraction of detected obstacles that are correctly detected by the developed system is 89.1%. This indicate an efficient response by the system. Other results show that the system has an average accuracy, specificity and sensitivity of 82.5%, 67.45% and 87.26% respectively. This result indicates that the system's ability to recognize untrained objects is lower than trained objects. Figure 6 shows the prototype of the developed system which is implemented on a local walking stick.

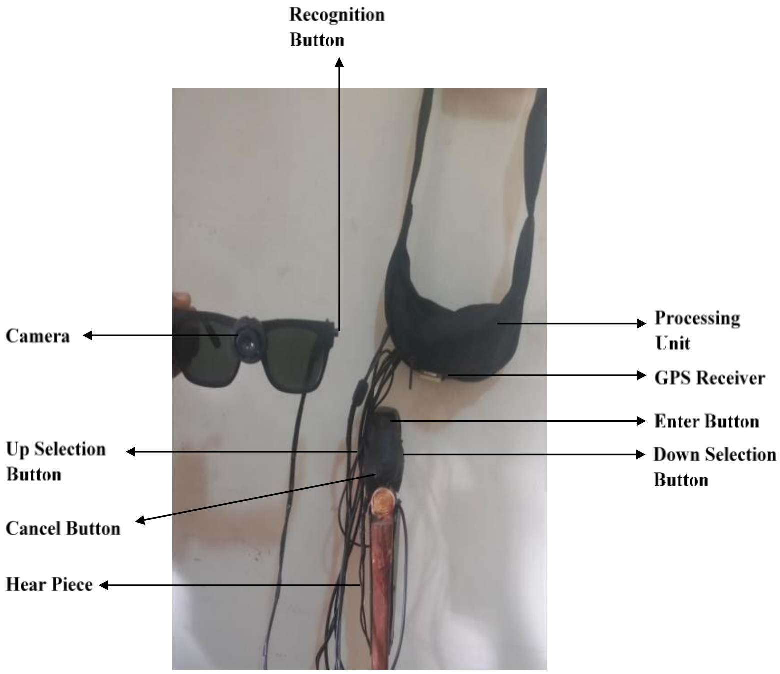

Fig 6. Prototype of the Developed System

5 Conclusions and Recommendation

In this paper, an intelligent offline multi-object recognition walking aid for visually impaired persons was developed. The system uses a single shot multi-box object detector and Convolutional Neural Network (CNN) deep learning model for object detection and recognition respectively. The trained CNN model was deployed in a Raspberry pi microcontroller. The performance of the system was evaluated using sensitivity, specificity, precision, and accuracy metrics. The test results indicated that the system achieved an average sensitivity of 87.26%, specificity of 67.45%, a precision of 89.07% and an accuracy of 82.50% over all the trained objects. The system also responded fast with an average of 20s between image capture, recognition and communication. The varied performance of the system can be attributed to the changing background of the target objects. The performances of the system can be improved by improving the object detection process, possible the use of deep learning models for object detection and recognition. For example, deploying Multi-Task Cascaded Convolutionary Neural Network (MTCNN) deep learning object detection algorithm.

References

[1] Tripathi M et al (2014) Electronics Guidance For The Navigation of Visually Impaired Person. International Journal for Research in Applied Science and Engineering Technology (IJRASET) 2(6):335-342
[2] WHO (2019) Blindness and vision impairment. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment. Accessed 14 Oct 2021
[3] Bourne RR et al (2017) Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. The Lancet Global Health 5(9):e888-e897
[4] Fricke TR et al (2018) Global prevalence of presbyopia and vision impairment from uncorrected presbyopia: systematic review, meta-analysis, and modelling. Ophthalmology 125(10):1492-1499
[5] Villanueva J, Farcy R (2011) Optical device indicating a safe free path to blind people. IEEE transactions on instrumentation and measurement 61(1):170-177
[6] Fernandes H et al (2013) RFID mesh network as an infrastructure for location based services for the blind. In: Kurosu M (ed) HCII/HCI 2013, vol 8008. Springer, Heidelberg, pp 39-45
[7]] Saaid MF, Ismail I, Noor MZH (2009) Radio frequency identification walking stick (RFIWS): a device for the blind. In: 5th International Colloquium on Signal Processing and Its Applications, Kuala Lumpur, Malaysia, 6-8 March 2009
[8] Hoyle B, Waters D (2008) Mobility AT: the Batcane (UltraCane). In: Hersh MA, Johnson MA (eds) Assistive technology for visually impaired and blind people. Springer, London, pp 209-229
[9] Chaurasia S, Kavitha KVN (2014) An electronic walking stick for blinds. In: International Conference on Information Communication and Embedded Systems (ICICES2014), Chennai, India 27-28 Feb 2014
[10] Yeshwant SS, Mukati A, Unal A (2016) Novel Walking Stick with Gripping and Lifting Mechanism. International Journal of Advanced Technology in Engineering and Science 4(10):518-531
[11] Olanrewaju RF, Radzi MLAM, Rehab M (2017) iWalk: intelligent walking stick for visually impaired subjects. In: the 4th IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), Putrajaya, Malaysia 28-30 November 2017
[12] Lin BS, Lee CC, Chiang PY (2017) Simple smartphone-based guiding system for visually impaired people. Sensors 17(6):1371.
[13] Kollathodi MA, Drolia R, Kumar A (2018) Smart Walking Stick for the Visually Challenged. Indonesian Journal of Electrical Engineering and Computer Science 12(3):1282-1288
[14] Wang Y, Long L (2018) The Design of Intelligent Speech Walking Stick Based on GPS. In: MATEC Web of Conferences, Aachen, Germany, 27-29 September 2018
[15] Romadhon AS, Husein AK (2020) Smart stick for the blind using Arduino. Journal of Physics: Conference Series 1569(3):1-7

[16] Saisubramanyam K, Rani JA (2020) IOT Based Smart Walking Stick for Blind or Old Aged People with GPS Tracking Location. Journal of Engineering Science 11(1):501-505
[17] Adil M et al (2020) An IoT based Voice Controlled Blind Stick to Guide Blind People. International Journal of Engineering Inventions 9(1):9-14
[18] Singh S, Singh B (2020) Intelligent Walking Stick for Elderly and Blind People. International Journal of Engineering Research and Technology (IJERT) 9(3):19-22
[19] Liu W et al SSD: single shot multibox detector. In: ECCV 2016, Amsterdam, Netherlands, 11-14 October 2016
[20] Howard AG, Zhu M, Chen B et al (2017) MobileNets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861