International Journal of Research in Engineering, Science and Management
Volume-1, Issue-8, August 2018
www.ijresm.com
ISSN (Online): 2581-5782

Navigation System for Blind People Using Artificial Intelligence

Pankaj Dusane[1], Vrushabh Deokar[2], Janhavi Doke[3]

[1, 2, 3]*Student, Department of Computer Science Engineering, MGM'sCET, Navi Mumbai, India*

***Abstract*: In this world of advance technologies, about 20% of the population is suffering from different kinds of disabilities such as visual, physical, or hearing problems. It becomes very difficult and hectic for the blind people to reach their destination without help for the surrounding environment and people. It is not possible for everyone to takeout time from their busy schedule to help these people. To overcome this problem, we have proposed a navigation system which would use artificial intelligence to assist the blind people.**

The proposed system uses a sensor box interfaced with an android application for intelligent navigation. The sensor box contains different kinds of sensors for identifying and detecting the obstacles. When an obstacle is detected, it sends a signal to the android application through Bluetooth. The mobile application receives this signal and tells the person about this obstacle through voice commands. The person then gets alert about the obstacle and reaches the destination safely.

Keywords: Artificial Intelligence, Technology, Bluetooth.

I. Introduction

According to the World Health Organization, around 40 million people in the world are blind, while another 250 million have some form of visual impairment and age-related disorders, like glaucoma and diabetes, mean these numbers are on the rise in the ageing populations of the UK, Europe and other countries, so developing new solutions that allow those individuals to interact with sighted people, and the sighted world, in a way that lessens any of problems that can arise from being blind is becoming increasingly important. However visually blessed people cannot takeout time from there busy schedule every time to help blind people. Due to this impairment, many road accidents happen every day. So it is important to solve this problem and improve the quality of life for these visually impaired people.

So, in this project, we use a sensor box that will detect the obstacles in the environment. The android application will act as an interface between the blind people and the sensor box. The app receives signals through a Bluetooth module. The person will hear about the obstacle through Text-to-Speech technology.

II. Existing System

In current world, visually impaired people have to either use the walking sticks to detect the obstacles and tackle them or ask for help from other people which is not always possible for them.

III. Proposed System

This project mainly contains three parts: Raspberry pi 3, camera and artificial intelligence. When the button is pressed on the device by the person, the camera starts taking pictures of the surroundings and analyze them and detect what exactly the picture is all about. Once the object is identified, the device will assist the person with voice commands about the scenario.

A. Raspberry pi 3

Fig. 1. Raspberry pi 3

The Raspberry pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote the teaching of basic computer science in schools and in developing countries.

Raspberry pi can be used for variety of purposes but in this project, we are using raspberry pi 3 for signal processing and image processing.

The Raspberry Pi 3 uses a Broadcom BCM2837 SoC with a 1.2 GHz 64-bit quad-core ARM Cortex-A53 processor, with 512 KB shared L2 cache.

Raspberry Pi 3 Model B has 1 GB of RAM. The Raspberry Pi 3 (wireless) is equipped with 2.4 GHz WiFi 802.11n (150 Mbit/s) and Bluetooth 4.1 (24 Mbit/s) based on Broadcom BCM43438 FullMAC chip with no official support for Monitor mode but implemented through unofficial firmware patching and the Pi 3 also has a 10/100 Ethernet port.

The Raspberry Pi is connected to the system via an Rj45 connector cable. Raspberry pi consists of variety of slots for performing variety of functions. Raspberry pi Foundation provided a Raspbian operating system for Raspberry Pi. Python and Scratch are the main programming language used, and also support many other languages.

B. Artificial Intelligence

Artificial intelligence is a field of computer science in which intelligent machines work like humans and respond

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-8, August 2018
www.ijresm.com
ISSN (Online): 2581-5782

quickly. Machines are able to react like humans only if they have abundant information related to the daily life. Machine learning is another one of the core parts of AI. Learning without any kind of supervision requires an ability to identify patterns in streams of inputs, whereas learning with adequate supervision involves classification and numerical regressions.

A typical AI understands its environment and takes actions that increases its chance of successfully achieving its goals. An AI's intended goal function can be simple or complex. Goals can be explicitly defined, or can be obtained. If the AI is programed for specific tasks, goals can be implicitly induced by rewarding some types of behavior and punishing others. Alternatively, a better system can induce goals by using a "fitness function" to mutate and preferentially replicate high-scoring AI systems.

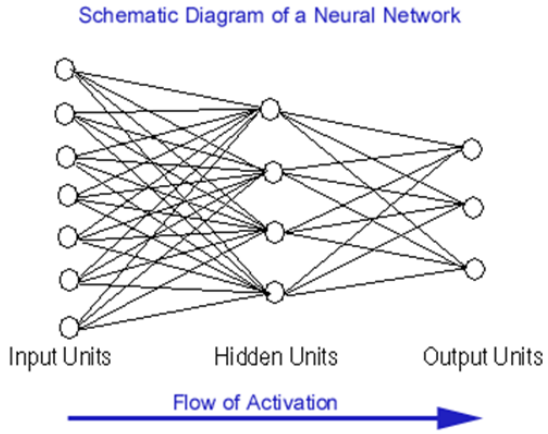

Fig. 2. Neural network used in artificial intelligence

C. *Raspberry Camera Module*

Fig. 3. Raspberry pi camera module

In April 2016, the original Camera Module was replaced by Camera Module of Raspberry pi. It consists of Sony IMX219 8-megapixel sensor. Camera module can take video and still photographs. Libraries bundled in the camera can be used to create effects. It supports 1080p30, 720p60, and VGA90 video modes, as well as still capture. It attaches via a 15cm ribbon cable to the CSI port on the Raspberry Pi.

The Raspberry Pi camera module can be used to take HD images and videos, as well as stills photographs of scenery. It's easy to use for beginners, but has plenty to offer advanced users with variety of features.

IV. The Model Used for Image Classification: Convolutional Neural NETWORK (CNN)

A convolutional neural network (CNN) is a class of deep, forward artificial neural networks that have successfully applied to analyze the visual image.

Convolutional neural network uses a multilayer perceptron's to obtain minimal preprocessing. They are also known as space invariant artificial neural networks (SIANN), due to their straight architecture and translation invariance characteristics. The deep CNN can achieve reasonable performance on hard visual recognition tasks, matching or exceeding human performance in some domains. This network that we build is a very small network that can run on a CPU and on GPU.

Convolutional neural network is composed of convolutional modules of the stack that performs extraction. Each module has a convolutional layer followed by a pooling layer (PL). The last convolutional module is followed by single layers that perform extraction and classification. The final dense layer in convolutional neural network contains a single node for each target in the model, with a SoftMax activation function to generate a binary value for each node (the sum of all these SoftMax values is equal to 1). We can translate the SoftMax values for a given image as relative measurements of how likely it is that the image fall into each target class.

Convolutional layers, which apply a specified number of filters to the image. For each sub region, some mathematical operations are performed by layer to produce a single value in the output feature map. Convolutional layers then typically apply an activation function to the output to introduce nonlinearities into the model.

The rectifier is an activation function defined as the positive part of its argument:

$f(x)= x+ = \max(0, x)$, where x is the input to a neuron.

A smooth approximation to the rectifier is the analytic function $f(x) = \log(1+ \exp x)$, which is called the soft plus function. The derivative of soft plus is $f'(x) = \exp(x) / (1+\exp x) = 1/(1+\exp(-x))$ i.e. the logistic function (LF).

Convolutional Layer: Applies 64 5x5 filters (extracting 5x5pixel sub regions), with activation function.

V. Model Architecture

This model is a multi-layer architecture consisting of convolutions and nonlinearities. These layers are followed by fully connected layers leading into a soft max classifier 0. This model achieves a peak performance of about 76% accuracy within a few hours of a GPU. It consists of 1,068 learnable parameters and requires about 19.6M multiply-add operations to compute inference on a multiple image. By using these Model the images are processed as follows:

1. They are cropped to 64 x 64 pixels, centrally for evaluation for training.
2. They are approximately blured to make the model insensitive to dynamic range.

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-8, August 2018
www.ijresm.com
ISSN (Online): 2581-5782

This is a good way to verify that inputs are built correctly.

Reading images from secondary memory and distorting them can use a much amount of processing time. To prevent these operations from slowing down processing, we can run them inside 16 separate threads which continuously fill a Tensor Flow queue. Here is a diagram of this model:

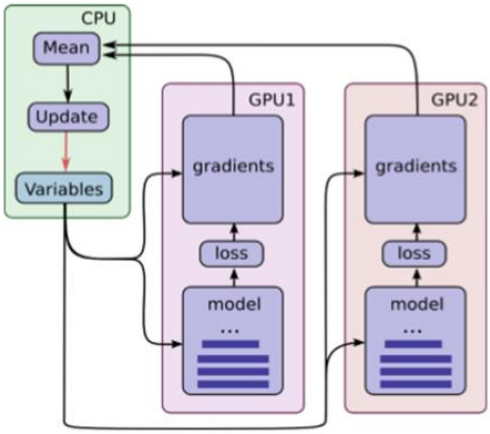

Fig. 4. Model architecture

Pooling layers, which down sample the image data extracted from the convolutional layers to reduce the dimensionality of the feature map in order to decrease processing time. We used max pooling algorithm, which extracts sub-regions of the feature map, keeps their maximum value, and discards all other values 0.

It is called average pooling if you take the average in place of taking maximum, but it's not very popular. If your input is of size w (1)*h(1)*d(1) and the size of the filter is f*f with stride S. Then the output sizes w2*h2*d2 will be

W2= (w (1)-f)/S +1
h2= (h (1) -f)/s+1
d2= d (1)

Most pooling is done with the filter of size 2*2 with a stride of 2. As you can calculate using the above formula, it essentially reduces the size of input by half.

In fully connected layers, which perform classification on the features extracted by the convolutional layers and down sampled by the pooling layers. In this layer, every node in the layer is connected to every node in the preceding layer.

Dense Layer 1: 1,024 neurons, with dropout regularization rate of 0.6 (probability of 0.6 that any given element will be dropped during training)

(Logits Layer) 2: 10 neurons, one for each digit target class (0–9).

Logits Layer, is the final layer of our neural network is the logits layer, which will return the raw values for our predictions. This is a regression model where the dependent variable (DV) is categorical. Where the output can take only binary value, which represent outcomes such as pass/fail or win/loss. Cases, where the dependent variable has more than two outcome categories, may be analyzed in multinomial regression, or, if the multiple categories are ordered, in ordinal logistic regression 0.

VI. Conclusion

The goal of this research is to provide the better image processing using artificial intelligence. By using convolution neural network image classifier, we can predict the correct answer with more than 89% accuracy rate. By doing so we achieved the state-of result on the dataset. We also use the exercise this model with real time image and can be able to obtain the correct label.

We integrated tensor flow in android studio with our trained model. And it is deployed in raspberry pi with the help of android operating system.

References

[1] Dean, Jeff; Monga, Rajat; et al. (November 2015). "Tensor Flow: Large-scale machine learning on heterogeneous systems" (PDF). TensorFlow.org. Google Research. Retrieved 12 November 2015.
[2] Hinton, G.; Deng, L.; Yu, D.; Dahl, G.; Mohamed, A.; Jaitly, N.; Senior, A.; B. (2015). "Deep Neural Networks for Acoustic Modeling in Speech Recognition --- The shared views of four research groups". IEEE Signal Processing Magazine.29 (6):8297.
[3] Krizhevsky, Alex. "ImageNet Classification with Deep Convolutional Neural Networks". Retrieved 17 November 2015.
[4] Raspberry Pi User Guide; Eben Upton and Gareth Halfacree; 312 pages; 2015;
[5] Xilinx. "HDL Synthesis for FPGAs Design Guide". Section 3.13: "Encoding State Machines". Appendix A: "Accelerate FPGA Macros with One-Hot Approach".
[6] Zoph, Barret; Le, Quoc V. (2016-11-04). "Neural Architecture Search with Reinforcement Learning". arXiv:1612.01578
[7] Hope, Tom; Resheff, Yehezkel S.; Lieder, Itay (201708-09). Learning TensorFlow: A Guide to Building Deep Learning Systems. "O'Reilly Media, Inc.". pp. 64.
[8] Hilbe, Joseph M. (2010), Logistic Regression Models, CRC Press, p. 3, ISBN 9781420064779.
[9] Sutton, R. S., and Barto A. G. Reinforcement Learning: An Introduction. The MIT Press, Cambridge, MA, 1999
[10] Getting Started with Raspberry Pi; Matt Richardson and Shawn Wallace; 180 pages; 2013; ISBN 9781449254214.
[11] Zeiler, Matthew D.; Fergus, Rob (2013-01-15). "Stochastic Pooling for Regularization of Deep Convolutional Neural Networks". arXiv:1301.3565
[12] Lawrence, Steve; C. Lee Giles; Ah Chung Tsoi; Andrew D. Back (1997). "Face Recognition: A Convolutional Neural Network Approach". Neural Networks, IEEE Transactions on. 8 (1): 98– 113.
[13] R. J. Williams and D. Zipser. Gradient-based learning algorithms and their computational complexity. In Back-propagation: Theory, Architectures, and Applications. Hillsdale, NJ: Erlbaum, 1994.