International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 01 | Jan 2023 www.irjet.net p-ISSN: 2395-0072

Object and Currency Detection for the Visually Impaired

Ms.Aditi Aiyar[1], Dr. Rachana Dhannawat[2]

Department of Computer Science and Technology
Usha Mittal Institute of Technology, SNDT Womens University, Juhu, Mumbai

---

Abstract - *The partial or complete loss of vision can create many difficulties in movement and other daily activities. People with such problems need to recognize objects around them to support navigation and avoid obstacles. The solution to this is to build a system that recognizes objects and provides voice response for accessibility. The simplest way is to use object localization and classification. Deep learning based approach employs CNN to perform end to end unsupervised object detection. The proposed system aims to serve as an object detector (thereby detecting obstacles) and currency detector. The system adds tags to objects in line of the camera and provide voice feedback for the same. The detection model is built using Tensorflow and Keras on pretrained network Mobilenet*

Key Words: **CNN, Mobilenet, Tensorflow, Transfer Learning, Detection**

1. INTRODUCTION

T HE biggest challenge for a blind person, especially the one with the complete loss of vision, is to navigate around places. Visually impaired are those who have a reduced ability or are completely unable to see. Object detection is a field related to computer vision and image processing that finds real world objects in real time or images. Computer vision uses a combination of machine learning algorithms and image processing algorithms for mimicking the human brain. Object Detection for visually impaired is the need of the hour as they need to recognize day to day objects around them, save themselves from obstacles and navigate better. Visually impaired people can then lead a better and more independent life. Braille (pattern based translation), canes and tactile are some accessibility equipment currently available. This system uses convolutional neural networks for object detection. The main aim of object detection is to locate the object in the image and classify it using the labelling classes. Here it works better than it's simple image processing counterpart as it describes the scene under the camera instead of returning a processed image. This system will use the mobile camera to capture the scene in real time while simultaneously detecting objects in the frame and providing voice output. This system will be used for obstacles detection as well as currency detection

1.1 Existing System

There are different approaches used for object detection and currency detection by different researchers.

Mishra, Phadnis,Bendale[1] use Google's Tensorflow object detection API, a framework for deep learning and used it to train on their custom dataset. SSD MobileNet, a predefined model offered by TensorFlow is used as the base and fine-tuned to improve the accuracy and the range of objects that can be detected. The downside of this tracking project was the privacy invasion. Semary et al.[2]in their case study demonstrate simple image processing techniques for currency detection. Contrarily Zhang and Yan[3] use deep learning techniques for currency detection. Four different models were tested using the Single Shot Detector (SSD) algorithm and the model with best accuracy was picked for deployment.

In Object Detection with Deep Learning: A Review [4], deep learning based object detection frameworks which handle different sub-problems are reviewed, with different degrees of modifications on R-CNN. The review starts on generic object detection pipelines which provide base architectures for other related tasks. Then, three other common tasks, namely salient object detection, face detection and pedestrian detection, are also briefly reviewed. Through [1], we are introduced to a class of efficient models called MobileNet for mobile applications. MobileNets [5] are based on a streamlined architecture that uses depthwise separable convolutions to build light weight deep neural networks. This paper presents extensive experiments to show the strong performance of MobileNet compared to other popular models on ImageNet classification.

1.2 Proposed System

All existing accessibility applications for the visually impaired contain object detection module and currency recognition modules as separate applicants. This means multiple apps have to be downloaded for different purposes. This project combines two such modules into one project application for easier use. This project will use two neural networks for detection - one for objects and obstacles and the other for currency detection. These two neural networks will be used in one android application for detection and voice output will be provided for alerts and recognitions. The system hardware is composed of a smart phone with a rear end camera for real time capturing of

frames. The model is trained and tested on Google Colaboratory with GPU using Python.

2. Methodology

2.1 MobileNet

This system uses two neural networks, both using MobileNet. It uses depthwise separable convolutions which basically means it performs a single convolution on each colour channel rather than combining all three and flattening it.[5] This has the effect of filtering the input channels. It is also very low maintenance thus performing quite well with high speed. Different models like ResNet, AlexNet, DarkNet as pre-trained model can also be used but will conclude that MobileNet, due to light weight is very quick to run on smartphones, hence apt for the project.Figure1 describes the Network Training plan used for building the system.

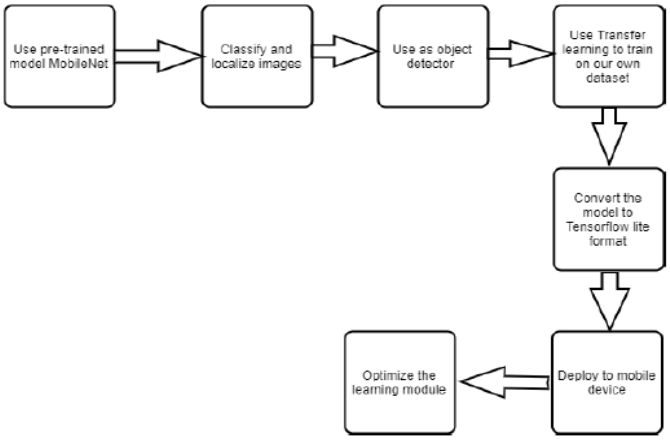

Figure1: Network Training Diagram

2.2 Object Detection

The first one is using Google's Tensorflow object detection API. The detection module is built over Pre trained Mobilenet using the SSD [8] algorithm over the COCO dataset. Therefore it is trained over 90 classes. This network will be used for object and obstacles detection. The model is picked from the Tensorflow Model Zoo which consists of several models with different algorithms. Single Shot means that the tasks of object localization and classification are done in a single forward pass of the network. MultiBox is the name of a technique for bounding box regression developed by Szegedy et al.[8].The network is an object detector that also classifies those detected objects.

2.3 Currency Detection

The second neural network is the same Mobile net trained over Imagenet which has 1000 classes. This network will be retrained using currency dataset for currency detection. The process of transfer learning will be used for retraining. Keras, which is a neural network library capable of running on tensorflow is used. Transfer learning is used to manipulate the MobileNet architecture by training it on a collected currency dataset of different denominations. This is done by freezing the base layer and adding new layers. The training data is loaded into the ImageDataGenerator function which sends the data for training in batches. The model is then compiled over 10 epochs using GPU using Adam [6] optimizer algorithm and then the trained model can be used for predictions. Epoch is a hyperparameter that is defined for training. One epoch means that the training dataset is passed forward and backward through the neural network.

2.4 User Interface

The medium to run the neural networks will be a mobile application. When the app is opened and camera is turned on, the network will locate objects in the line of the camera and create bounding boxes around it followed by the class that is identified. The other network will detect currency denominations using only classification. Both networks will use Google's Text to Speech to convert into voice output.

2.5 Results

The proposed system detects the object through mobile camera and provides voice output for objects, warns against obstacles and detects right currency. The project solves the basic problems faced by the visually impaired i.e.to recognize objects around and to support navigation and avoid bumping into obstacles.

The system creates bounding box with appropriate class labels and confidence score. Confidence score here means the probability that a box contains the object represented in percentage. Figure 2 and Figure 3 depict indoor objects and obstacles like bottle, chair, remote etc.

Figure2:Detection of day to day objects

Figure3:Detection of obstacles

Figure 4 and Figure 5 show outdoor object detection like vehicles and potted plants.

Figure4:Detection of obstacles

Figure4:Detection of obstacles

3. CONCLUSION

This application will help visually impaired people both indoors and outdoors to save themselves from harm. It will also eliminate the need to carry any extra equipment (stick, cane, etc.) as well as minimize the cost of purchase. The only object required will be a basic smartphone. This project can be extended to self-driving cars, object tracking, pedestrian detection, anomaly detection, people counting and face recognition.

REFERENCES

[1] R. Phadnis, J. Mishra and S. Bendale, "Objects Talk - Object Detection and Pattern Tracking Using TensorFlow," 2018 Second International Conference on Inventive Communication and Computational Technologies

(ICICCT), Coimbatore, 2018, pp. 1216-1219.

[2] N.. A. Semary, S. M. Fadl, M. S. Essa and A. F. Gad, "Currency recognition system for visually impaired: Egyptian banknote as a study case," 2015 5th International Conference on Information Communication Technology and Accessibility (ICTA), Marrakech, 2015, pp. 1-6.

[3] Q. Zhang and W. Q. Yan, "Currency Detection and Recognition Based on Deep Learning," 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Auckland, New Zealand, 2018, pp. 1-6.

4] Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu, Fellow,IEEE.Object Detection with Deep Learning: A Review

[5] A. Howard,M. Zhu,B. Chen,D. Kalenichenko,W.Wang T.Weyand,M. Andreetto, H. Adam.MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

[6] Y. Sun et al., "Convolutional Neural Network Based Models for Improving Super-Resolution Imaging," in IEEE Access, vol. 7, pp. 43042-43051,2019.

[7] Mane,S.,Prof.Mangale,M . Moving object detection and tracking Using Convolutional Neural Networks

[8] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Berg AC (2016) "Ssd: Single shot multibox detector. European conference on computer vision" , Springer: 21-37.

[9] A. Nishajith, J. Nivedha, S. S. Nair and J. Mohammed Shaffi, "Smart Cap - Wearable Visual Guidance System for Blind," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA),Coimbatore, 2018, pp. 275-278.

[10] https://towardsdatascience.com/transfer-learning-using-mobilenet-andkeras-c75daf7ff299