DESIGN AND DEVELOP OBJECT DETECTION SYSTEM FOR BLIND PEOPLE BASED ON CNN IMAGE RECOGNITION

LAI JUN ANN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DESIGN AND DEVELOP OBJECT DETECTION SYSTEM FOR BLIND PEOPLE BASED ON CNN IMAGE RECOGNITION

LAI JUN ANN

This report is submitted in partial fulfilment of the requirements for the degree of Bachelor of Electronic Engineering with Honours

Faculty of Electronic and Computer Engineering
Universiti Teknikal Malaysia Melaka

2018

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJUTERAAN ELEKTRONIK DAN KEJURUTERAAN KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek : Design and Develop Object Detection System for Blind People Based on CNN Image Recognition

Sesi Pengajian : 2017/2018

Saya LAI JUN ANN mengaku membenarkan laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (✓):

☐	**SULIT***	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
☐	**TERHAD***	(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan.)
☑	**TIDAK TERHAD**	

Disahkan oleh:

(TANDATANGAN PENULIS)

(COP DAN TANDATANGAN PENYELIA)

Alamat Tetap: 4, Dataran Tasek Timur 15, Tmn Mujur, 31400 Ipoh,Perak

Tarikh : 25 Mei 2018

Tarikh : 25 Mei 2018

*CATATAN: Jika laporan ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali tempoh laporan ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I declare that this report entitled "Design and Develop Object Detection System for Blind People Based on CNN Image Recognition" is the result of my own work except for quotes as cited in the references.

Signature : 

Author : LAI JUN ANN

Date : 25/5/2018

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Bachelor of Electronic Engineering with Honours.

Signature	:	
Supervisor Name	:	Dr. Wong Yan Chiew
Date	:	25/05/2018

DEDICATION

Special dedicated to my beloved parents for their caring, understanding and encouragement

ABSTRACT

Object detection is a popular topic in visual recognition and plays a significant role in many fields. On the other hand, visually impaired people or blind people are usually unaware of danger that they are facing in their daily life. They faced many difficulties in their activity even in their familiar environments. This project proposes a smart object detection system based on Convolutional Neural Network (CNN) to provide a smart as well as safer living to visually impaired people. The region proposals from edge maps for each image is produced by using edge box algorithm. Then the proposals is passed through a fine-tuned CaffeNet model. The object is detected by the webcam and the feature of the image is extracted, if the object is matching with the trained model in the database which is the cloud storage, then output audio will generate by the system to let the visually impaired people to identify the object. The result is evaluated by using mean average precision (mAP) as well as frame-per-second. As a result, SSD reduced the complexity and archives higher accuracy and faster speed in object detection compared to Fast R-CNN.

ABSTRAK

Pengesanan objek merupakan topik yang hangat dan memainkan peranan yang amat penting dalam pelbagai applikasi. Selain itu, golongan yang cacat penglihatan atau golongan yang buta biasanya tidak menyedari bahaya di sekeliling mereka dalam activity seharian mereka. Mereka menghadapi pelbagai jenis masalah walaupun meraka berada di suasana biasa. Projek ini mencadangkan sistem pengesanan objek yang pintar berdasarkan Convolutional Neural Network (CNN) untuk memberi kehidupan yang lebih pintar and selamat kepada golongan cacat penglihatan. Algoritma edge box menjanakan rantau cadangan daripada rantau peta untuk setiap gambar dan melewati model CaffeNet. Kamera digunakan untuk mengesankan objek dan ekstrak ciri-ciri gambar, jika objek tersebut memadankan dengan gambar dalam pangkalan data dan pangkalan data itu ialah penyimpanan awam, sistem ini akan menjanakan suara untuk golongan cacat penglihatan supaya mereka dapat mengenalpasti objek tersebut. Purata ketepatan dan bingkai sesaat telah digunakan untuk menilaikan keputusan project ini. Kesimpulannya. SSD dapat mencapai ketepatan yang lebih tinggi and lebih laju berbanding dengan Fast R-CNN.

ACKNOWLEDGEMENTS

As a student of Universiti Teknikal Malaysia Melaka (UTeM), first and foremost, I would like to thank UTeM for giving me this opportunity to involve in PSM which it is an eye-opening experience for me.

I also would like to extend my gratitude to my supervisor, Dr. Wong Yan Chiew for her concerned, kindness, advised and helped during the time I am doing my project. Without her guidance from the beginning until the completion of this report, the progress of this project may have stuck and cannot be done and documented.

Finally, greatest credit goes to my parents and my friends for their encouragement that make me strong when I encounter with any problems. I also would like to thank all my course mates who have been helping me directly from the beginning until the final stage in this project. Their opinion and suggestion are very useful.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

The purpose of this project is to implement convolutional neural network (CNN) for image recognition application. CNN has the advantage over conventional approaches since CNN able to extract features individually, decrease dimensional of the data as well as category in one network structure. As a result, CNN uses relatively little computational load compared to other conventional algorithms. Hence, many people choose CNN as their choice for overcoming the image classification problems. The purpose of object detection is to recognize the objects from the scene and predict the corresponding bounding boxes. The plan is to train the image data by using CNN algorithm and the trained caffe model will be stored in the Microsoft Azure cloud storage database. The system will be on cloud platform so that the trained caffe model can be retrieve and perform the image recognition process anywhere anytime.

In this work, the complexity of RCNN approach in object detection is reduced. As shown is Figure 1.1, edge box is chosen to generate region proposals, rather than that of selective search used in rCNN. Generally, edge boxes produces and outcomes the proposal based on the edge map of the image. Edge boxes can achieved much faster speed over selective search in RCNN. The average runtime is 0.3 seconds for edge box while for selective search is 10 seconds. Therefore, the edge boxes able to decrease the computational load with excellent performance. Each pixel carries a magnitude and orientation data of the edge in edge detector. An edge map is produced with a structured edge detector. Although the mean average precision (mAP) for both method are almost the similar, edge boxes took shorter time to process compared to selectively search. In addition, all the class specific SVM are removed and the output of SoftMax is used. In this project, an intelligent object detection system for blind people based on Convolutional Neural Network is developed to achieve safer and better quality of life. Video scene in real time will be take on object detection systems and activity of interest will be determine by computer vision-based techniques which the primary task is to produce a high-level understanding of the imaged scene and generate application-specific data to be used in an autonomous and intelligent system. Figure 1.1 shows the model used in this project.

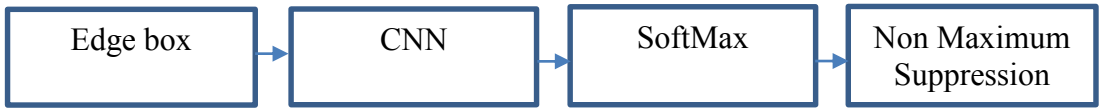

Figure 1.1: The overview of the model

The aim of this project is to present a smart living to the blind people by assisting them in their daily life through object detection system. For example, the system will detect the object through the camera in real time and process the video by comparing

with the trained model in the cloud database. Blind people who need more attentions will be the target of the proposed object detection system.

1.1 Project Background

Recently, visual recognition becomes a popular topic and grows drastically. There are many features that used in visual recognition in the pass such as scale-invariant feature transform (SIFT) and histogram of oriented gradients (HOG). These features are collaborated with Support Vector Machine (SVM) which implementing into multiple classes and object detection in a single image. Hence, many people are replaced by Convolutional neural network (CNN) in object detection which having significant accuracy by utilizing a deep CNN due to the poor performance of the previous models.

CNN has become an interesting topic in visual recognition since 2012 because CNN able to classify the images with high accuracy and it becomes the best method in solving image classification problems in visual recognition as well as object detection. The connections of CNN between the neuron networks have numeric weights that are tuned during the training process, so that a properly trained network will respond correctly when presented with an image or pattern to recognize. The network consists of multiple layers of feature-detecting “neurons”. Each layer has many neurons that respond to different combinations of inputs from the previous layers. The layers are built up so that the first layer detects a set of primitive patterns in the input, the second layer detects patterns of patterns, and the third layer detects patterns of those patterns, and so on.

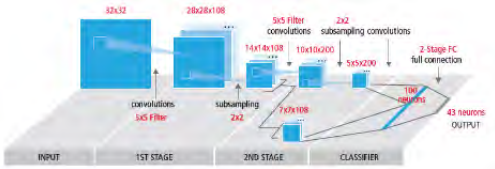

Figure 1.2: Block diagram

Feature extractors are hand designed in conventional algorithm. On the other hand, the weights of the convolutional layer in CNN is used for feature extraction and the fully connected layer is used for classification during the training process. As shown in Figure 1.2, the improved network structures of CNN results in saving in term of memory requirements and computation complexity requirements as well as giving better performance for applications where the input has local correlation. The advantage of using CNN for image recognition are ruggedness to shifts and distortion in the image, fewer memory requirements and faster as well as better classification training.

1.2 Problem Statement

In the past decade, object detection has become an interesting topic in visual recognition. The people requires fast detection and recognition speed, high accuracy as well as low power consumption for greener environment. Since the object detection of this work is applying on the blind people which is portable, hence the power consumption is one of the issue in designing the algorithm. In the pass, we tend to design features from raw image to improve the performance of the detection. For conventional recognition approaches, the raw input will undergo three different type of process in order to achieve the class score, the process includes preprocessed input, compressed input and feature vector process. As we can see, the conventional recognition approach will increase the computational load and design complexity

which leading to low power efficiency. Conventional recognition approach also tends to take longer time to train compared to other and increase the complexity load. In this work, the complexity of the algorithm will be reduced to achieve the high power efficiency, high accuracy as well as high speed.

1.3 Objective

The aim of this project is to design and develop object detection system for blind people based on CNN image recognition. The objective of this project are listed down as follows:

- To investigate the key design parameters for image detection system
- To design convolutional neural network (CNN) of image recognition system for object detection system
- To develop a real-time CNN for image recognition system on cloud platform
- To analyze the performance and accuracy of CNN

1.4 Scope

This project mainly focuses on training the image by using convolutional neural network (CNN) and import to the cloud database for image recognition. The training process will undergo by using Microsoft Azure Machine Learning Studio. The image will store in the cloud storage that created. The application of this project mainly focuses on using Raspberry Pi3 model B board as microcontroller and webcam to capture the scene to perform the object detection process.

1.5 Importance of work

The purpose of this project is to provide a smart living to the blind people by assisting them in their daily life through object detection system. In general, the system will detect the object through the camera in real time and process the video which compare with the trained model in the cloud database. This system assists the blind people on identifying the object in their daily life.

1.6 Chapter organization

This thesis comprises of five chapters:

Chapter 1 introduces the project, starting from background, objectives, problem statement, scopes of project, and chapter organization.

Chapter 2 explains the theory that is related to this project, how Convolutional Neural Network works, and different method used on object detection for blind people from variety sources of the articles and papers.

Chapter 3 illustrates the process flow of this project. This also includes the tools of Python 3.5, Visual Studio Code, Visual Studio, and Microsoft Azure Cloud Storage software that are used in this project.

Chapter 4 discusses the result obtained with the image trained with edge box method as well as selective search method. The trained image is uploaded to the cloud storage and the result is evaluated by using mean average precision (mAP) as well as