UNATTENDED BAGGAGE DETECTION USING DEEP NEURAL NETWORKS

ONG YI WEI

UNIVERSITI TEKNOLOGI MALAYSIA

UNATTENDED BAGGAGE DETECTION USING DEEP NEURAL NETWORKS

ONG YI WEI

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic System)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JUNE 2018

For my beloved family and friends

ACKNOWLEDGEMENT

I would like to convey my appreciation to all the people who directly or indirectly assisted me through the completion of this project. First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Usman Ullah Sheikh for his inspiration, continuous support and guidance with ideas and advices thoughout this project. Besides that, I would also like to thank him for his patience and time allocation among his busy schedule to respond to me through email or text and even come all the way to Penang to meet with me and guide me through difficulties. His valuable motivations and suggestions have greatly contributed to the completion of this project.

Secondly, I would also like to thank all my panel of examiners for their constructive and insightful feedback during my presentation. Their comments have give me a huge help to the completion of the project and thesis.

Lastly, I would also like to thank all of my family and friends greatfully for their constant support and encouragement throughout the development of this project.

ABSTRACT

As the world becomes ever more attuned to potential security threats, the need for sophisticated surveillance system is increasing to monitor and detect any potential threats. Sophisticated surveillance system should functions as an intuitive “robotic eye” for accurate and real-time detection of threats. Unattended baggage has become a critical need for security personnel at airports, stations, malls, and in other public or crowded areas. However, an effective system for detection of objects like baggage and people with a real-time video input requires high processing power and storage to just process the video frames using the typical digital image processing technique. This will require a very high development cost and time in order to make the system work which is impractical for commercial use. Moreover, manual configuration is needed which is not flexible to be for multiple application. Therefore, the objective of this thesis is to improve the object detection accuracy and flexibility compared to existing digital image processing techniques. This proposed system uses deep neural networks approach through collection of datasets thus providing a more accurate detection and flexible application. Tensorflow framework is used as the deep neural network framework for the development of this system. This system utilizes the Single Shot multibox Detection detection algorithm to the ’MobileNet’ neural network architecture which is optimized to provide a promising performance even in embedded system. This project is developed by implementing the Tensorflow Object Detection Application Programming Interface (API). This method enables 4 main classes of detection which are suitcase, backpack, handbag and person. The datasets used for benchmarking are surveillance video sample that contain unattended baggage scenario used by most existing works like AVSS2007, PETS2006 and ABODA. The overall accuracy and flexibility of the proposed system improved up to 43% thus unattended baggage is able to be detected. The system is able to be applied in various environment due to the excellent flexibility of the system.

ABSTRAK

Komunikasi dunia yang semakin global, kita semakin berpotensi kepada ancaman keselamatan, oleh itu keperluan untuk sistem pengawasan canggih semakin meningkat untuk memantau dan mengesan sebarang ancaman. Sistem pengawasan yang lebih canggih yang dicadangkan harus berfungsi untuk mengesan bagasi yang ditinggalkan dengan tepat telah menjadi keperluan kritikal bagi anggota keselamatan di kawasan awam seperti lapangan terbang, stesen dan pusat membeli-belah. Walau bagaimanapun, sistem yang berkesan untuk mengesan objek seperti bagasi dan orang dengan input video masa nyata memerlukan kuasa pemprosesan yang tinggi dan penyimpanan untuk memproses video menggunakan teknik pemprosesan imej digital biasa. Proses ini memerlukan kos pembangunan dan masa yang sangat tinggi untuk menjadikan sistem berfungsi, oleh itu tidak praktikal untuk kegunaan komersil. Selain itu, konfigurasi manual diperlukan untuk pengesanan membuktikan bahawa sistem adalah tidak fleksibel untuk berbilang aplikasi. Oleh itu, objektif projek ini adalah untuk meningkatkan lagi ketepatan pengesanan objek dan fleksibiliti berbanding dengan teknik pemprosesan imej digital sedia ada. Sistem ini menggunakan pendekatan rangkaian neural melalui pengumpulan dataset untuk menyediakan pengesanan yang lebih tepat dan aplikasi yang fleksibel. Tensorflow digunakan sebagai rangka kerja rangkaian neural yang mendalam untuk pembangunan sistem ini. Sistem ini menggunakan algoritma pengesanan *Single Shot Multibox Detector* dengan senibina rangkaian neural 'MobileNet' yang dioptimumkan untuk memberikan prestasi yang baik walaupun dalam sistem terbenam. Projek ini dibangunkan dengan melaksanakan *Application Programming Interface* Pengesanan Objek Tensorflow. Kaedah ini akan membolehkan 4 kelas pengesanan utama iaitu beg pakaian, beg sandang, beg tangan dan orang. Dataset yang digunakan untuk menanda aras adalah sampel video pengawasan yang mengandungi senario bagasi tanpa pengawasan seperti AVSS2007, PETS2006 dan ABODA. Ketepatan dan fleksibiliti sistem meningkatkan sebanyak 43% bagasi yang tanpa pengawasan dapat dikesan. Sistem ini boleh digunakan di pelbagai persekitaran kerana fleksibiliti sistem yang amat baik.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

API	-	Application Programming Interface
CSV	-	Comma-separated Values
SSD	-	Single Shot Multibox Detector
XML	-	Extensible Markup Language
YOLO	-	You Only Look Once detection algorithm

CHAPTER 1

INTRODUCTION

1.1 Problem Background

The world is getting more connected through the years with the ease of transportation from one place to another, but it also means that it is more attuned to security threats from terrorist attacks. In the last decade, there is a sharp increase statistically in terrorist attack on crowded public places such as airport, train station, bus terminal and many more. These attacks include mass shooting, bombing and even hostage situation in crowded public places. In just 2017, these attacks have drastically increased in occurrence including those attacks that get largely reported and those which are not reported in the news. For example the bombing happened in London subway, bombing attack in Saint Petersburg subway, bombing in Brussels Airport in Belgium and many more.

There are also cases that these attempted attacks were stoped in time with the high sensitivity level of security personnel working with video surveillance. For example there was an incident that happen in North Carolina Airport which the suspect left a bag containing explosives in the airport. The security personnel were able to find the bag with explosives thus able to prevent the bombing attack. However, human error may occur where security personnel will fail to notice the unattended bag containing explosives when trying to monitor multiple video surveillance streams.

Therefore in order to help security personnel to monitor any of this suspicious unattended baggage faster and more efficiently to prevent any tragedy before it happens, an automated and accurate detection system is needed for real-time video surveillance system. The baggage detection system enables the detection of unattended baggage which may contain explosives that may cause harm to civilians in crowded public places. The detection will be able to notify and alarm security personnel to react in time to diffuse and evacuate civilians from suspected area where the unattended baggage is placed. Besides that, this system can be used to locate any lost baggage through the detection of unattended baggage in the video surveillance. This is where Deep Neural Networks implementation into unattended baggage detection system will greatly help in this context due to the accuracy and efficiency it provides.

1.2 Problem Statement

Unattended baggage detection system is highly dependent on the accuracy of the detection of baggage through the input video footage. Therefore, the system would have to dissect the video into frames then detect if any baggage exists in the frame. After the detection stage, the system will move to process the detected information through the desired algorithm and business rules defined.

Commonly used method is to use image processing method. There are a number of research papers using image processing methods for the development on this unattended baggage detection system. For image processing, techniques such as background subtraction is used to differentiate foreground and background for object extraction. After that, the extracted objects will go through classification such as stationary or dynamic objects. After the classification, post-processing of the classified objects can be done such as activity and unattended situation detection [5]. However, there are very little or close to none research on development of unattended baggage detection system using deep neural networks method. With the ability to train deep neural networks, it should be able to detect objects such as baggage in a more accurate manner compared to what is available in typical image processing techniques.

Besides that, the accuracy of image processing techniques have high dependency on algorithm and threshold settings on the processes mentioned. Therefore, the accuracy of image processing method has a low accuracy compared to deep learning. This is due to various research using image processing techniques did not clearly mentioned about the system accuracy on each detection with the variation of datasets [5]. This might be due to the nature of some image processing technique are not as flexible and only cater to some specific environment thus unable to be applied on variety of situation and environment changes.

Another drawback of typical image processing techniques is that it requires significant amount of computational effort to process the video frames. This is due to the need of going through multiple processing techniques with different algorithm to every single frame of the video input. Besides that, the complexity of algorithm may also increase the computational effort and datasets needed. This drawback is known to be an issue for image processing.

1.3 Objectives

The objective of this project is to propose the development of unattended baggage detection system with the implementation of deep neural networks approach. By utilizing deep neural networks approach, it is expected that the system will be more robust and accurate compared to typical image processing techniques. Besides that, this proposed system also aims to be more flexible than typical image processing method with variations in situations and environment in the video surveillance.

1.4 Scope of Study

The scope of this project is focused on the development of unattended baggage detection using deep neural networks. The baggage type to be detected is the type of bags that commonly used by travelers and commuters. This project focuses on 3 main types as listed below:

- backpack
- handbag
- suitcase

The input video datasets used for this work will be sample of video surveillance footage that most of the researches in this field are using. This work will use all of those datasets from different sources to prove the flexibility of this system. The surveillance video should be located in a public places where the scope mentioned such as airport, train station or mall. The footage should also be best case scenario if it is captured or recored from surveillance camera or some similar type of camera to replicate the real-time scenario as close as possible. Besides that, the surveillance video footage should also contain the scenario where people are abandoning their baggage in various distances with people walking around. The datasets name are listed as below:

- Advanced Video and Signal based Surveillance (AVSS 2007) [6]
- Performance Evaluation of Tracking and Surveillance (PETS 2006) [7]
- ABandoned Objects DAtasets (ABODA) [8]

This project will be developed in Tensorflow, which is a machine learning platform that enables the development of deep learning solutions. This platform enable students and developers to explore and develop deep learning solutions with minimum development cost as development can be done locally on personal computer or using available Google cloud server. Single Shot MultiBox Detector (SSD) will be used as the deep neural network algorithm, and a pre-trained SSD network with large collection of baggage datasets trained from MS-COCO datasets will be used. This method is able to save resources thus further improve development time, cost and efficiency without sacrificing accuracy or performance.

1.5 Organization

This project report consists of five chapters. Chapter 1 includes problem background, problem statement, objectives and scope of study. In Chapter 2, the background of deep learning and reviews of related works will be discussed in detail. While in Chapter 3, project methodology is explained which include the proposed algorithm, techniques and platform used. The results will be presented in Chapter 4, where detailed analysis on the unattended baggage detection results will be done. Discussion and comparison on the result analysis will be explained in this chapter as well. Finally in Chapter 5, conclusion and suggestions for future work will be discussed.

REFERENCES

1. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. and Berg, A. C. Ssd: Single shot multibox detector. *European conference on computer vision*. Springer. 2016. 21–37.

2. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, 2017. abs/1704.04861.

3. He, K., Zhang, X., Ren, S. and Sun, J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. 770–778.

4. Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A. A. Inception-v4, inception-resnet and the impact of residual connections on learning. *AAAI*. 2017, vol. 4. 12.

5. Mathur, G. and Bundele, M. Research on Intelligent Video Surveillance techniques for suspicious activity detection critical review. *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*. 2016. ISSN 5090-2806. 1–8. doi:10.1109/ICRAIE.2016.7939467.

6. 2007 IEEE Conference on advanced video and signal based surveillance. *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*. 2007. c1–c1. doi:10.1109/AVSS.2007.4425267.

7. Thirde, D., Li, L. and Ferryman, F. Overview of the PETS2006 challenge. *Proc. 9th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2006)*. 2006. 47–50.

8. Lin, K., Chen, S. C., Chen, C. S., Lin, D. T. and Hung, Y. P. Abandoned Object Detection via Temporal Consistency Modeling and Back-Tracing Verification for Visual Surveillance. *IEEE Transactions on Information Forensics and Security*, 2015. 10(7): 1359–1370. ISSN 1556-6013. doi:10.1109/TIFS.2015.2408263.

9. LeCun, Y., Bengio, Y. and Hinton, G. Deep learning. *Nature*, 2015. 521(7553): 436–444.

10. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural networks*, 2015. 61: 85–117.

11. Sarle, W. S. Neural Networks and Statistical Models, 1994.

12. Bengio, Y. and LeCun, Y. Scaling learning algorithms towards AI. *Large-scale kernel machines*, 2007. 34(5): 1–41.

13. Deng, L. and Yu, D. Deep learning: methods and applications. *Foundations and Trends® in Signal Processing*, 2014. 7(3–4): 197–387.

14. Zou, X. and Wen, J. Detection of object security in crowed environment. *2015 IEEE International Conference on Communication Problem-Solving (ICCP)*. 2015. 34–37. doi:10.1109/ICCPS.2015.7454083.

15. Hargude, S. and Idate, S. R. i-surveillance: Intelligent surveillance system using background subtraction technique. *2016 International Conference on Computing Communication Control and automation (ICCUBEA)*. 2016. 1–5. doi:10.1109/ICCUBEA.2016.7860046.

16. Carvajal-GonzÃ¡lez, J. P., Valencia-Aguirre, J. and Castellanos-DomÃ¯nguez, G. Silhouette classification by using manifold learning for automated threat detection. *2013 47th International Carnahan Conference on Security Technology (ICCST)*. 2013. ISSN 1071-6572. 1–5. doi:10.1109/CCST.2013.6922080.

17. Hoang, T. V. and Tabbone, S. The generalization of the R-transform for invariant pattern representation. *Pattern Recognition*, 2012. 45(6): 2145 – 2163. ISSN 0031-3203. doi:https://doi.org/10.1016/j.patcog.2011.11.007. Brain Decoding.

18. Lai, T. Y., Kuo, J. Y., Liu, C. H., Wu, Y. W., Fanjiang, Y. Y. and Ma, S. P. Intelligent Detection of Missing and Unattended Objects in Complex Scene of Surveillance Videos. *2012 International Symposium on Computer, Consumer and Control*. 2012. 662–665. doi:10.1109/IS3C.2012.172.

19. Zin, T. T., Tin, P., Hama, H. and Toriu, T. Unattended object intelligent analyzer for consumer video surveillance. *IEEE Transactions on Consumer Electronics*, 2011. 57(2): 549–557. ISSN 0098-3063. doi:10.1109/TCE.2011.5955191.

20. Wang, W. and Liu, Z. A New Approach for Real-Time Detection of Abandoned and Stolen Objects. *2010 International Conference on Electrical and Control Engineering*. 2010. 128–131. doi:10.1109/iCECE.2010.40.

21. Tian, Y., Feris, R. S., Liu, H., Hampapur, A. and Sun, M. T. Robust Detection

of Abandoned and Removed Objects in Complex Surveillance Videos. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2011. 41(5): 565–576. ISSN 1094-6977. doi:10.1109/TSMCC.2010.2065803.

22. Ren, S., He, K., Girshick, R. and Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017. 39(6): 1137–1149. ISSN 0162-8828. doi:10.1109/TPAMI.2016.2577031.

23. Long, J., Shelhamer, E. and Darrell, T. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015. 3431–3440.

24. Kim, H., Lee, Y., Yim, B., Park, E. and Kim, H. On-road object detection using deep neural network. *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*. 2016. 1–4. doi:10.1109/ICCE-Asia.2016.7804765.

25. Bhargava, M., Chen, C.-C., Ryoo, M. S. and Aggarwal, J. K. Detection of abandoned objects in crowded environments. *2007 IEEE Conference on Advanced Video and Signal Based Surveillance*. 2007. 271–276. doi:10.1109/AVSS.2007.4425322.