International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

WEAPON DETECTION IN VIDEO SURVEILLANCE USING CONVOLUTIONAL NEURAL NETWORK

Prema Arokia Mary G[1], Amrithaa D[2], Santhosh MG[3], Sowndarya K[4]

[1]*Assistant Professor, Department of Information Technology, Kumaraguru College of Technology [autonomous], Coimbatore, India*
[2-4]*Department of Information Technology, Kumaraguru College of Technology [autonomous], Coimbatore, India*

Abstract – *This project explores the concept of convolutional neural networks (CNNs) and Google Object Detection API and uses these two new developments in technology to retrain a pre- trained model to perform weapon detection in real time surveillance. The aim of this project is to investigate the effect of training convolutional neural networks with one extra class "non- weapon" based on two original classes "gun" and "knife" from previous work. This research attempts to find out whether the model produces different accuracy with non-weapon class. In the result of this research, faster R-CNN model with Inception feature extractor produced a significant result of 19.50% mean average precision on detecting gun, knife and non-weapon classes. Evaluation and challenges of this research will be addressed in this paper as well as the potential recommendations and solutions to solve the problem*.

Key Words: *Convolution Neural Networks, Inception-v2, Google Object Detection API, Deep Learning, TensorFlow*

1 INTRODUCTION

1.1 Background

Safety is always an issue of concern to a person, a family or even a country, it is essential to secure our personal belongings, property or even live from threats such as robbery or homicide. It is also the duty of a country to maintain their citizens' safety. According to Fox and Pettersson from CNN(208), approximately half (48%) of the estimated 650 million civilian-owned guns globally are owned by the Americans. Furthermore, British people are 51 times less likely to be killed by gunfire than the American based on the OECD 2010 data from WHO (as cited in Fox & Pettersson, 2018). In Malaysia, the country has the highest crime index in Southern-East Asia from Numbeo statistics ("South-Eastern Asia: Crime Index by Country 2018," 2018). As the issues have been highlighted in many countries, it is necessary to review our surveillance system again with the knowledge of new technology. We should prepare and prevent the dangers and not repair and repent the wounds. Current surveillance systems generally are operating passively which indicates that related authority usually is aware of crime events after it happened. For example, police only review CCTV footage after a robbery event occurred. Another Example, appointing security guards to monitor surveillance systems all day long does not give a result of low crime rate and potentially humans could make mistakes and it is not financially manageable to all the people or family.

1.2 Objectives

The aim of this paper is to design and implement an autonomous weapon detection system using Faster R-CNN and improve the accuracy with one extra class called "non-weapon", then evaluate the performance and precision of Faster R-CNN and alert in the form of email. . TensorFlow and Google object detection API will be used as a platform and library to run Inception-v2 model for object detection.

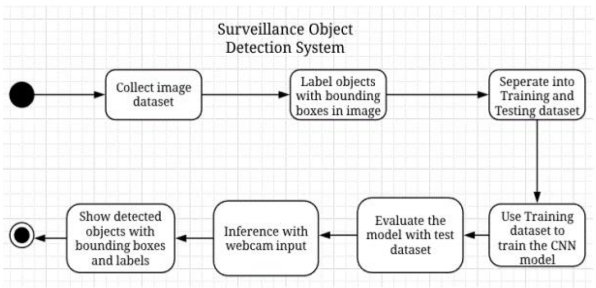
1.3. Constraints

The main constraint of this project is that the project requires high performance computing power to train the model before the model is effective to detect the specified objects. Running the project is acceptable on a personal computer, but it is time consuming without latest GPU support. Thus, computing power is the main constraint in this project.

2. PROJECT DEFINITIONS AND INFORMATION

2.1 Overview

In general, the steps of completing this project is first collecting images and pre- processing images with labels. Second, separating the dataset into training and testing, and feeding the training dataset to the model for the training process. After the model is trained, the test image dataset will be fed to the model for evaluation. Then it can be used to perform inferences on surveillance input video. Since there is no class in this project, the activity diagram is drawn to illustrate the system architecture as shown.



System design flowchart

2.2 Resource Requirements

2.2.1 Hardware:

A. 64-bit, x86 desktops or laptops with dedicated Nvidia graphics cards. The lowest compatible desktop graphic card is GeForce GT640(GDDR5). The lowest compatible laptop graphic card is the GeForce GT730M.
B. A toy gun and a knife.

2.2.2 Software:

i. TensorFlow with GPU support
a. An open-source software library used for machine learning
ii. labelImg.exe for labelling images with bounding boxes
a. A software used for annotating images and labelling objects with bounding boxes in image.
iii. Python 3.5.x or 3.6.x
Iv. Algorithms: R-CNN, Deep Learning.
v. Pycharm Professional Edition
a.An open-source tool developed for interactive programming and supported with many programming languages.
v. Libraries used: Matplot, Pandas, Keras, Scikit Learn, Numpy, Six, Tensorflow Object Detection.

2.2.3 Resources:

i. Images dataset
a. 1000 images for gun
b. 1000 images for knife

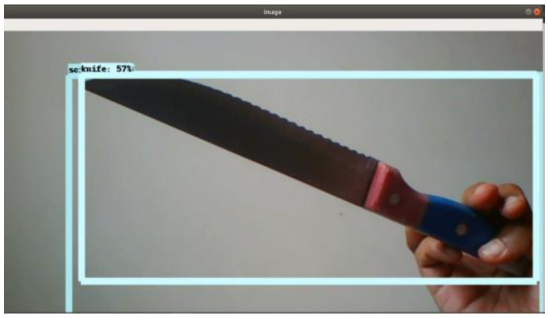
3. DESIGN AND METHODOLOGY

3.1 Image Collection:

First step is collecting sufficient numbers of images for the research. In the image collection step of system design, images collection is required to get sufficient number of images for training and testing purposes.

3.2 Image Labelling:

Labelling objects with bounding boxes and labels on the images for model training. Then converting xml (bounding boxes records) to csv is followed by generating TFrecord from csv.



Labelling demonstration

3.3 Image Separation:

Image dataset is divided into 75/25 ratio in random selection. 25% of images are used for testing in the evaluation step, 75% of the images are used to feed into the model for training purposes and this could be done by using a python script.

3.4 Model Training:

Transfer learning is a concept of reusing and transferring the knowledge and dataset from a previously trained model. Starting to train the model with the training dataset. Once the dataset is well prepared, the dataset is then fed into the model to start training.

3.5 Model Testing:

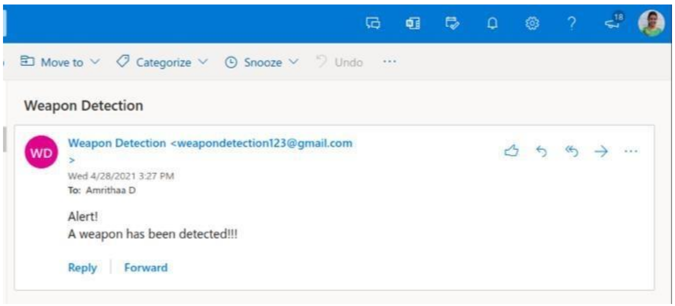
Once the model finished training. In this stage, a test image dataset is used to evaluate the model to output the average precision and map. Then the script outputs the result from the model in the command prompt. Testing process can run on the existing trained model.

3.6 Model Inferencing:

In this stage, the trained model will be tested on a new input dataset taken from a webcam to check whether the model does what it is designed for. Then the following step is to write a script to run the inference model on a webcam input and displaying detected objects with bounding boxes and labels

3.7 Mail Notification

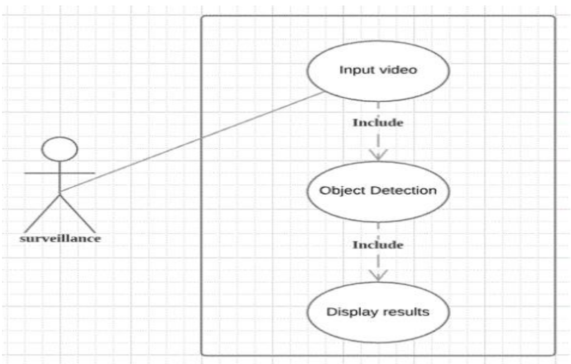
Then the following step is to write a script to run the mail notification based on the model on a webcam input. Once the weapon is detected, a mail containing the alert message is sent using the SMTP server object.



Alert Mail

3.8 Internal Design:

Firstly, the system loads the trained model and activates the webcam, and then takes input data from the webcam which is passed to the trained model. Then the models will make a prediction on how likely the image is gun, knife or non-weapon after it processed the input data from the webcam. Finally, the system displays output of the objects detected with predicted bounding boxes, labels and percentage of accuracy on the webcam video and an alert mail is sent.



Use case of the system

4. Result

The following figures are showing the inferencing results from weapon detection taken from the webcam. Figure 4-1 shows the first case of true positive (TP) for knife and non-weapon. The Inception model correctly detected the knife as a knife and the phone as a non-weapon with 99% and 56% accuracy respectively. The next figure shows the false negative (FN) case, the model itself detects nothing in the image but in fact it should actually detect the gun as a gun, thus, it is a FN case in Figure 4-2.

Figure 4-1 True Positive case of the model

Figure 4-2 False Negative case of the model

5. Recommendation of Future Work

With the experience and inspiration of this project, it is recommended to collect image dataset from various sources to give a comprehensive dataset for training CNNs. And for each class of object, it is recommended to obtain a consistent object dataset, this means in the case of this research, only assigning monitors to the monitor object class instead of using different kinds of non-weapon objects to one single class. Possible future works can be expanding the number of images to the dataset to avoid overfitting issues and separating the non-weapon class objects into their own categories is a possible way to increase the accuracy.

6. Conclusion

Weapon detection using neural networks in surveillance is one of the future solutions to our security. For the goal of autonomous weapon detection, accuracy is important. This project has shown that with non-weapon class objects, it does not improve the accuracy of the weapon detection. However, with more and categorized images in each class, a higher accuracy could be achieved. Lastly, from the exploration of this project, a great knowledge has been learned on the tool of machine learning and neural networks.

7. References

[1] Feng-Li Lian, Yi-Chun Lin, Chien-Ting Kuo, and Jong-Hann Jean" vol. 9, no. February 2018 Voting-Based Motion Estimation for Real-Time Video Transmission in Networked Mobile Camera Systems" IEEE transactions on industrial informatics

[2] Weiming Hu, Ruiguang Hu, Nianhua Xie, Haibin Ling, and Stephen Maybank, January 2019
Image Classification Using Multiscale Information Fusion Based on Saliency Driven Nonlinear Diffusion Filtering

[3]Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2020).
Rich feature hierarchies for accurate object detection and semantic segmentation.

[4]He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition.

[5]Huang, J., Rathod, V., Sun, C., ZŘu, M., Korattikara, A., Fathi, A., ... Guadarrama, S. (2017).
Speed/accuracy trade-offs for modern convolutional object detectors.

[6]Ismail, M. A. B. (2018). Detecting weapons in surveillance videos using neural networks.

[7]Karn, U. (2016, 2019). An Intuitive Explanation of Convolutional Neural Networks.