GUN DETECTION IN REAL TIME CCTV USING DEEP LEARNING ALGORITHM

Mr. Anil P Jawalkar
Asst.Professor
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

S.Deepika
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

S. Rajeshwari
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

P. Srivani
Student
Department of Information Technology
Malla Reddy Engineering College for Women
(UGC-Autonomous)
Maisammaguda, Hyd-500100, Telangana, India.

Abstract— Current surveillance and control systems still require human supervision and intervention. This work presents a novel automatic handgun detection system in videos appropriate for both, surveillance and control purposes. We reformulate this detection problem into the problem of minimizing false positives and solve it by building the key training data-set guided by the results of a Deep Convolutional Neural Networks (CNN) classifier, then assessing the best classification model under two approaches, the sliding window approach and region proposal approach. The most promising results are obtained by Faster R-CNN based model trained on our new database. This can be helpful to make informed decisions be it regarding identification of intent, promotion of offers or security related threats. Recognizing emotions from images or video is a trivial task for human eye, but proves to be very challenging for machines and requires many image processing techniques for feature extraction. The best detector show a high potential even in low quality YouTube videos and provides satisfactory results as automatic alarm system.

INTRODUCTION

The crime rates caused by guns are very concerning in many places in the world, especially in countries where the possession of guns is legal or was legal for a period of time. The last statistics reported by the United Nations Office on Drugs and Crime (UNODC) reveals that the number of crimes involving guns per 100,000 habitants are very high in many countries, e.g., 21.5 in Mexico, 4.7 in United States and 1.6 in Belgium. In addition, several psychological studies demonstrated that the simple fact of having access to a gun increases drastically the probability of committing a violent behavior.

In the last five years, deep learning in general and Convolutional Neural Networks (CNNs) in particular have achieved superior results to all the classical machine learning methods in image classification, detection and segmentation in several applications. Instead of manually selecting features, deep learning CNNs automatically discover increasingly higher-level features from data. We aim at developing a good gun detector in videos using CNNs.

We have seen that security is the main reason for identifying any person. It can be based on finger-print matching, voice recognition, passwords, retina detection etc. Identifying the intent of the person can also be important to avert threats. This can be helpful in vulnerable areas like airports, concerts and major public gatherings which have seen many breaches in recent years.

One way to reducing this kind of violence is prevention via early detection so that the security agents or policemen can act. In particular, one innovative solution to this problem is to equip surveillance or control cameras with an accurate automatic handgun detection alert system. Related studies address the detection of guns but only on X-ray or millimetric wave images and only using traditional machine learning methods.

The role of deep learning in improving task performance in security controls systems is considered indisputable. Deep learning is a sub-field of machine learning. It uses many layers of non-linear processing units for deep learning and feature extraction and conversion. The deep learning structure is based on the learning of more than one feature level of data. Deep learning is based on learning from the representation of the main data. The representation of an image can be considered to comprise a vector of density values per pixel or features such as clusters of edges and custom shapes, with some representing the data better. The basic architecture of the deep learning concept is the convolutional neural network (CNN), which consists of convolution, pooling, activation function, dropout, fully connected, and classification layers. In the last few years, deep learning has become a mainstay in the field of object detection and classification and image segmentation. To date, CNNs have achieved the best results for classical image processing problems, such as image segmentation, classification, and detection.

Today, most criminal activities are carried out using handheld weapons. Many studies have revealed that handheld weapons are the most important criminal elements used for various crimes, such as theft, illegal hunting, and terrorism. The solution to reduce such criminal activities is installing a surveillance system or control cameras so that security units can take appropriate measures at early stages. Weapon detection is challenging due to the various subtleties associated with it. The most important problems in weapon detection are self-occlusion and the similarities between objects and background structures. Self-occlusion occurs when a part of the gun is blocked on one side. Similarity between objects occurs when different objects such as hands and clothes look like weapons. Background problems refer to those related to the background against which the gun is located. to predict students' academic performance and based onhomework submission data, the academic performance of students enrolled in a blended learning course was predicted. According to their predicted academic performance, early feedbacks and interventions could be individually applied to at-risk students.

In a weapon classification study developed using a deep CNN, two new approaches were presented. In the proposed approach, the weights of the pre-trained VGG-16 model were taken. Using this model, the effects of changing the number of neurons in the fully connected layer on classification were investigated. In a study aiming to detect firearms in surveillance videos, the focus was placed only on only areas where human beings were found and a weapon detection system was implemented using the separate components of the weapons.

A study on multilevel security management presented a system for the management of multimedia security in Internet of Things systems. This system automatically analyzed multimedia events and calculated security levels. In another real-time object detection study, the authors detected handheld weapons (pistols and rifles). In that study, a TensorFlow-based application of Over feat, a CNN-based image classifier and feature extractor, was used to detect and classify weapons in the image.

EXISTING SYSTEM

A study on multilevel security management presented a system for the management of multimedia security in Internet of Things systems. This system automatically analyzed multimedia events and calculated security levels. In another real-time object detection study, the authors detected handheld weapons (pistols and rifles). In that study, a TensorFlow-based application of Over feat, a CNN-based image classifier and feature extractor, was used to detect and classify weapons in the image. In another study on the automatic detection of firearms and knives, algorithms were proposed to warn human operators when guns or knives were detected in the closed-circuit television system. In the same study, in order to apply the system in real life, the number of false alarms was reduced and a system that was capable of real time warning when a dangerous situation was detected was developed.

Clustering algorithm and color-based segmentation were also previously used to eliminate irrelevant objects from an image for the purpose of automatic visual weapon detection. The Harris detector of interest points and the fast retina key point descriptor were used to detect the relevant object (weapon) in segmented images. By applying this system to the collected weapon sample images, the partial jamming, scaling, rotation and presence of multiple weapons were detected.

The use of deep learning techniques to solve computer vision problems has achieved great popularity in the last decade in comparison with traditional machine learning techniques. This popularity is due to both its excellent results and the lack of necessity for the manual selection of features to solve the problem. These networks are based on adjusting or learning the parameters (weights) during their training using the gradient descent algorithm, which aims to minimize the network's response error or loss function. In this optimization process, the error is backpropagated through the network to adjust its parameters across all its layers. This process is also known as error backpropagation through the network. The use of convolution operations allows considering the process of adjusting the network weights as that of obtaining filters that focus on the characteristics that solve the problem, even when dealing with heterogeneous datasets.

The problem of the automatic detection of firearms and bladed weapons hidden inside luggage has been tackled for some years using images obtained with X-ray scanners. To this end, the classical cascade-based learning techniques of Haar feature detectors and AdaBoost classifiers have been applied. Indeed, those methods can only work with expensive X-ray scanners and cooperative individuals. A very interesting complementary context is the detection of visible weapons in images captured by CCTV systems, since these systems are already common in video surveillance of public spaces and allow detecting weapons held by noncooperative individuals, regardless of the construction material of such weapons.

PROPOSED SYSTEM

CCTV cameras play an important role to overcome this problem and are considered to be one of the most important requirements for the security aspect.

Video surveillance is an inexpensive method that allows covering large areas without interfering with the flux of people. However, it faces major limitations such as those arising from image capture speed, image resolution, scene light quality, and occlusions. In addition, the task of monitoring images captured by CCTV systems requires a high level of attention over long periods of time, which leads to unnoticed events because of human operator fatigue.

To propose a system that meets the two characteristics previously noted, this work presents a study of three firearm (handgun) detectors in images based on the application of convolutional neural networks (CNNs). While "classical" methods require the manual selection of discriminant features, CNNs are able to automatically extract complex patterns from data.

The Faster R-CNN was proposed to achieve the required processing speed for real-time applications. In the Faster R-CNN, the non-learning-based selective search algorithm is substituted by a region proposal network (RPN), which "learns" how to determine regions in which the objects are located. To propose the regions where each object is located, the RPN network slides an n×n spatial window of the input convolutional feature map obtained by the convolutional layers of a backbone network.

A witness call might be a great way to encourage a faster dispatch and prevent crime, but some people might still be reluctant to call the officers or going through a lot of panic thus not providing the right demographics. Providing the police with accurate information during a gunfire situation is key to a rapid response.

When integrated with video surveillance, GDT allows researchers to review video feeds of street block attributes, run onsite observations and conduct interviews.

GDT provide information that crime analysts use to spot trends, locate concentrations of gun violence and produce maps to help authorities make tactical decisions about where to deploy patrols and other resources.

Images are downloaded in bulk using Fatkun Batch Image Downloader (chrome extension) which can download multiple Google Images at once. Then the downloaded images are labelled. 80% of total images used for training and 20% images for testing. The created ammunition dataset was then trained using Single Shot Detector (SSD) model.

SSD algorithm reached new milestones in terms of precision and performance detection. SSD speeds up the process by eliminating the need of region proposal net

SYSTEM ENVIRONMENT

Python:

It is an interpreted, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python interpreters are available for many operating systems. Python is managed by the non-profit Python Software Foundation. Python features adynamic types system and automatic memory management. It supports multiple programming paradigms, including object oriented, functional and procedural and has a large and comprehensive standard library. Python is easy to learn yet powerful and versatile scripting language, which makes it attractive for Application Development.

Python IDLE:

IDLE stands for Integrated Development and Learning Environment. The story behind the name IDLE is similar to Python. Guido Van Rossum named Python after the British comedy group Monty Python while the name IDLE was chosen to pay tribute to Eric Idle, who was one of the Monty Python's founding members. IDLE comes bundled with the default implementation of the Python language since the 01.5.2b1 release. It is packaged as an optional part of the Python packaging with many Linux, Windows, and Mac distributions.

DJANGO:

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent thewheel. It's free and open source. Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasize reusability and "pluggability" of components, rapid development, and the principle of don't repeat yourself. Python is used throughout, even for settings files and data models.

IMPLEMENTATION

1. Sliding window/Classification Models

In the method to the sliding window, a box or window is moved over a picture to select an area and use the object recognition model to identify each frame patch covered by the window. It is an exhaustive search over the whole picture for objects. Not only do we need to search in the picture for all feasible places, we also need to search on distinct scales.

2. Region proposal/Object Detection Models

As this method takes a picture as the bounding boxes of input and output related to all patches in a picture most probable to be a category, so it proposes a region with the maximum score as the location of an object. Instead of considering all possible regions of the input frame as possibilities, this method uses detection proposal techniques to select regions. Region-based CNNs (R-CNN) was the first detection model to introduce CNNs under this approach. approach produces 2000 boxes having

CONCLUSION

This work presented a novel automatic pistol detection system in videos appropriate for both control purposes. We reformulate this detection problem into the problem of minimizing false positives and solve it by Building the key training of Data Sets.

The most promising results have been obtained with Faster R-CNN based model, trained on our new database, providing zero false positives, 100% recall, a high number of true negatives and good precision 84,21%. The best detector has shown a high potential even in low quality YouTube videos and provides very satisfactory results as automatic alarm system.

Through a series of experiments, we concluded that object detection algorithms with ROI (Region of Interest) perform better than algorithms without ROI. We have tested many models but among all of them, the state-of-the-art Yolov4, trained on our new database, gave very few false positive and negative values, hence achieved the most successful results.

For both monitoring and control purposes, this work has presented a novel automatic weapon detection system in real-time. This work will indeed help in improving the security, law and order situation for the betterment and safety of humanity, especially for the countries who had suffered a lot with these kind of violent activities.

This will bring a positive impact on the economy by attracting investors and tourists, as security and safety are their primary needs. We have focused on detecting the weapon in live CCTV streams and at the same time reduced the false negatives and positives.

It gave 91.73% mean average precision and a F1-score of 91% with almost 99% confidence score on all types of images and videos. We can say that it satisfactorily qualifies as an automatic real-time weapon detector.

REFERENCES

[1] Xiao, Z.; Lu, X.; Yan, J.; Wu, L.; Ren, L. Automatic detection of concealed pistols using passive millimeter wave imaging. In Proceedings of the 2015 IEEE International Conference on Imaging Systems and Techniques (IST), Macau, China, 16–18 September 2015.

[2] Tiwari, R.K.; Verma, G.K. A Computer Vision based Framework for Visual Gun Detection Using Harris Interest Point Detector. Procedia Comp. Sci. 2015.

[3] Grega, M.; Matiola´nski, A.; Guzik, P.; Leszczuk, M. Automated Detection of Firearms and Knives in a CCTV Image.

[4] Verma, G.K.; Dhillon, A. A Handheld Gun Detection using Faster R-CNN Deep Learning. In Proceedings of the 7th International Conference on Computer and Communication Technology-ICCCT-2017 Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014.

[5] Elmir, Y.; Laouar, S.A.; Hamdaoui, L. Deep Learning for Automatic Detection of Handguns in Video Sequences. In Proceedings of the 3rd edition of the National Study Day on Research on Computer Sciences (JERI 2019), Saida, Algeria, 27 April 2019.

[6] Lim, J.; Jobayer, M.I.A.; Baskaran, V.M.; Lim, J.M.; Wong, K.; See, J. Gun Detection in Surveillance Videos using Deep Neural Networks. In Proceedings of the 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Lanzhou, China, 18–21 November 2019.

[7] Wei Liu et al., "SSD: Single Shot MultiBox Detector", European Conference on Conputer Vision, Volume 169, pp 20-31 Sep. 2017.

[8] D. Erhan et al., "Scalable Object Detection Using Deep Neural Networks," IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2014.

[9] Ruben J Franklin et.al., "Anomaly Detection in Videos for Video Surveillance Applications Using Neural Networks," International Conference on Inventive Systems and Control,2020.

[10] H R Rohit et.al., "A Review of Artificial Intelligence Methods for Data Science and Data Analytics: Applications and Research Challenges,"2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2018.

[11] Abhiraj Biswas et. al., "Classification of Objects in Video Records using Neural Network Framework," International conference on Smart Systems and Inventive Technology,2018.

[12] Pallavi Raj et. al.,"Simulation and Performance Analysis of Feature Extraction and Matching Algorithms for Image Processing Applications" IEEE International Conference on Intelligent Sustainable Systems,2019.

[13] Mohana et.al., "Simulation of Object Detection Algorithms for Video Survillance Applications", International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud),2018.

[14] Yojan Chitkara et. al.,"Background Modelling techniques for foreground detection and Tracking using Gaussian Mixture model" International Conference on Computing Methodologies and Communication,2019.

[15] Rubner et.al, "A metric for distributions with applications to image databases", International Conference on Computer Vision,2016.

[16] N. Jain et.al., "Performance Analysis of Object Detection and Tracking Algorithms for Traffic Surveillance Applications using Neural Networks," 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2019.