COMPARATIVE STUDY OF OBJECT RECOGNITION ALGORITHMS TO FIND THE OPTIMAL ALGORITHM FOR WEAPONS DETECTION

Aastha Sakshi, Aritra Ghosh, Bhakti Barve, Vasudha Agrawal, Dr. Girija Chiddarwar

Abstract: Based on the current situation of gun violence around the world, there is a need for automated surveillance for the detection of handguns the objective of this proposed technique is to visually detect the handgun in images captured in video surveillances. The proposed method is to conduct a comparative study between the available object recognition algorithms and choosing an optimal algorithm for a video surveillances system. A dataset of images of guns and knives is used. The data will be trained using the most optimal algorithm. There are available surveys that show that YOLO-V3 (You Only Look Once version 3) can be used as an alternative of Faster RCNN (Region Based Convolutional Neural Networks), but there are no available papers or surveys assessing the function of YOLO-V4 (You Only Look Once version 4) so that will be tested by this paper.

Introduction

The number of crimes involving guns violence is increasing in many parts of the world, especially in that place where the possession of guns is legal. From statistics, it can be assumed that violence rate concerning guns are increasing every year becoming a challenge for law enforcement agencies to overwhelm this issue timely. The solution of this problem is constant monitoring and hasty detection using camera-based surveillance systems which can help to prevent this kind of violence and assist policemen or security agents to act timely. Research in computer vision focusing on object detection is growing rapidly. Object detection is a technique associated with computer vision and image processing that performs the task of detecting instances of certain objects such as a human, vehicle, banner, building from a digital image or a video.[1] Object detection combined with other advanced technology integrations allows us to perform face detection or pedestrian detection, popularly known as person from a video. Object detection is used in a variety of applications like detecting a broken bone from X-ray images, detect brand logo from image or video, and video surveillance for weapon detection.

Originally, the traditional image processing system included pixel by pixel object matching. While moderately effective, this is an inefficient and time-consuming process. Eventually with time and improvement in available technology, machine learning and deep learning technique such as R-CNN (Region-based conventional neural network) were developed. To overcome the limitation of R-CNN, Fast R-CNN, and Faster R-CNN has invented. Faster R-CNN was the best algorithm out of all of the above by evaluating its performance on the COCO (Common Objects in Context) dataset which is a benchmark dataset for object detection in terms of accuracy and training time. YOLO (You Only Look Once) object detection algorithm was developed using the darknet framework and the latest version i.e., the V4 of YOLO which is advertised as one that outperforms other versions of YOLO in terms of accuracy and inference time on the benchmark dataset.[2]

We have emphasized the different object detection algorithms Such as RCNN, YOLOv3, YOLOv4 and their implementation details and comparative analysis of our case-study on the custom dataset. an automatic gun detection system survey is presented using deep learning mainly YOLOv4 which is compared with the results of gun detection systems using a deep Convolutional Neural Networks (CNN) classifier and the number of false positive and false negative with Faster RCNN algorithm. Detecting gun in a scene is very challenging issue because of various subtleties linked with it. The challenges in gun detection are particularly caused by occlusion. Gun to scene and gun to object are the two types of gun occlusion. Real time processing is another main problem in gun detection system that arises during detection.

Literature Survey

Each of these papers are selected because of a unique purpose that these play for our study

The work presented by Roberto Olmos et al [3] shows a novel automatic pistol detection system in videos appropriate for both, surveillance and control purposes. They reformulated this detection problem into the problem of minimizing false positives and solved it by building the key training data-set guided by the results of a VGG-16 (Very Deep Convolutional Networks for Large-Scale Image Recognition) based classifier, then assessing the best classification model under two approaches, the sliding window approach and region proposal approach. The most promising results in this have been obtained with Faster R-CNN based model, trained on new database, providing zero false positives, 100% recall, a high number of true negatives and good precision 84,21%. Among 30 scenes, it successfully activates the alarm after five successive true positives within an interval of time smaller than 0.2 seconds, in 27 scenes. The detector by Roberto Olmos et al can be used in several applications, e.g., i) real time detection of guns in places monitored by cameras and ii) control whether the videos uploaded to social media contain scenes with guns. As present and future work, it is evaluated to reduce the number of false positives, of Faster R-CNN based detector, by preprocessing the videos, i.e., increasing their contrast and luminosity, and also by enriching the training set with pistols in motion.

Arif Warsi et al [4], have evaluated the performance of the YOLOv3 based detector on four different videos. The objective was to minimize the false positive using YOLOv3 algorithm. YOLOv3 based model has been trained with a dataset containing ImageNet dataset. It is clear from the results that YOLOv3 has a good detection performance even in low quality videos as than faster RCNN. The advantage of YOLOv3 over Faster RCNN is its speed. The processing speed of YOLOv3 is 45 frames per seconds while Faster RCNN has 8 frame per seconds. Two out of four videos showed better accuracy than Faster RCNN algorithm.

Michael Grega et al [5], have focused on the two specific tasks of automated detection and recognition of dangerous situations. They proposed, implemented and tested algorithms for the detection of a dangerous tool held in a hand. A knife or a firearm (the most frequently-used weapons in assaults) held in a person's hand is an example of a sign of danger. The specificity and sensitivity of the knife detection algorithm are 94.93% and 81.18%, respectively. These results are significantly better than others published recently. Their solution to the knife detection problem deals with poor quality and low-resolution images. This is important because many CCTV systems only provide such quality of footage. It should be noted that the algorithm is processed in real time.

Harsh Jain et al [6] demonstrate that, in terms of speed, SSD (Single Shot Multi Box Detector) algorithm gives better speed with 0.736 s/frame. Whereas Faster RCNN gives speed 1.606s/frame, which is poor compared to SSD. With respect to accuracy, Faster RCNN gives better accuracy of 84.6%. Whereas SSD gives an accuracy of 73.8%, which is poor compared to faster RCNN.

It is shown that a Tensor flow-based implementation of the Over feat network as an integrated network for detecting and classifying weapons in images. The best performance was achieved by Justin Lai et al [7] on Overfeat-3 with 93% training accuracy and 89% test accuracy with adjusted hyperparameters.

Lei Pang et al [8] demonstrate that the YOLO algorithm is an excellent real-time detection method of great development potential. This project focused on the real-time metal contraband detection from human body for PMMW (Passive Millimetre Wave) images with a small sample dataset and YOLOv3 algorithms. The Yolov3-13 and Yolov3-53 target detection models with different convolutional layers were trained, and their advantages and disadvantages were analysed, and the experiment results were compared with that of SSD algorithms. The experiments shown that the YOLOv3-based contraband detection method for the PMMW images and can meet the real-time detection requirements during large passenger flows.

Alexey Bochkoviskiv et al [9] demonstrates the high level of accuracy and speed demonstrated by the YOLOv4 algorithm. It also focuses on the computer environment required to run the yolov4 algorithm.

The following is a short table of all the papers and the strengths and limitations, summarize.

Sr No	Title of paper	Authors	Year of publication	Publication	Strengths	Limitations	Future Scope
1	Automated Detection of Firearms and Knives in a CCTV Image [5]	Michał Grega, Andrzej Matiolanski, Piotr Guzik and Mikołaj Leszczuk	2016	IEEE	In this study, they focused on the two specific tasks of automated detection and recognition of dangerous situations. Their solution to the knife detection problem deals with poor quality and low-resolution images. This is important because many CCTV systems only provide such quality of footage.	Has not specified what algorithm it is using just the defining state of the video used for recognition	Shows how the video can be used in such a way that the software will recognize low quality video.
2	Automatic Handgun Detection Alarm in Videos Using Deep Learning [3]	Roberto Olmos, Siham Tabik, Francisco Herrera	2017	IEEE	Considering the accuracy of its chosen algorithms which is calculated on the basis of percentage, while training the algorithm, they have considered the possibility of false positives.	Does not consider adding a comparative study to demonstrate the working of rcnn with other available algorithms in 2017. While faster RCNN is considered, other more functions like YOLO are ignored.	Can be used as a structure for training future algorithms.
3	Developing a Real-Time Gun Detection Classifier [7]	Justin Lai and Sydney Maples	2017	IEEE	Tensor flow-based implementation of the Over feat network as an integrated network for detecting and classifying weapons in images. The best	The study is conducted in an optimal situation with multiple resources that are not widely available.	By this study it is observed that the upper limit achievable that can be obtained and can act as a guide in the comparison to whatever optimal

					performance was achieved on Overfeat-3 with 93% training accuracy and 89% test accuracy with adjusted hyperparameters.		algorithm that is decided on.
4	Gun detection system using YOLO v3 [4]	Arif Warsi, Munaisyah Abdullah, Mohd Nizam Husen, Muhammad Yahya, Sheroz Khan, Nasreen Jawaid	2019	IEEE	They evaluated the performance of the YOLOv3 based detector on four different videos. The objective was to minimize the false positive using YOLOv3 algorithm. It is also demonstrated how yolov4 does in comparison to RCNN	The number of Algorithms is very limiting.	Adding more recent algorithms will help with deciding the most optimal algorithm for the software.
5	Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications [6]	Harsh Jain, Aditya Vikram, Mohana, Ankit Kashyap, Ayush Jain	2020	IEEE	This Study demonstrates how SSD and RCNN perform compared to each other in speed by frame per second and how they perform.	Does not consider add an comparative study to demonstrate the working of rcnn with other available algorithms in 2020.	Consider the working of YOLOv3 and Yolov4 in a similar comparative study.
6	Real-time Concealed Object Detection from Passive Millimeter Wave Images Based on the	Lei Pang, Hui Liu, Yang Chen and Jungang Miao	2020	IEEE	Focuses on the real-time metal contraband detection from human body for PMMW images with a small sample dataset and YOLOv3 algorithms. It is a good demonstration of YOLOv3 algorithm and it compares its	The comparison of YOLOv3 with SSD wasn't the smartest idea since SSD is one of the slowest algorithms.	YOLOv3 is a very versatile algorithm that can be used for the software hence will be added to our comparative study.

	YOLO v3 Algorithm [8]				functioning with SSD algorithm.		
7	YOLO v4: Optimal Speed and Accuracy of Object Detection [9]	Alexey Bochkovskiy,Chien-Yao Wang, Hong-Yuan Mark Liao	2020	IEEE	Demonstrates the high level of accuracy and speed demonstrated by the YOLOv4 algorithm. It also focuses on the environment required to run the yolov4 algorithm.	There is no comparative study even though an approximate reference in difference by the SSD is given.	YOLOv4 should be run with other algorithms to see the scope and speed of the algorithm.

Proposed Technique

The proposed work consists of three modules:

1. Object Detection: It takes video as the input. The video is converted into frames in the frame conversion block. Each frame from the frame conversion block is sent to the image processing module, where the edge distortions and high-quality frames are produced. These frames are then processed using the Convolution Neural Networks (CNN). After the detection of the object is alone taken and sent to the second module to classify if the object is knife or a gun.

2. Behavioural Analysis: It takes the detected frame as input. The input is given to classification sub stage where it uses support CNN. Classification sub-stage gets the supporting reference frames from training dataset. The Alexnet dataset consists of various annotated objects. Based on this it classifies the given frame as normal or abnormal events. The detected activity is sent to the third module.

3. Alert Modules: The detected activity is sent to the alert system module for classification. In the activity classification sub stage, the activity is classified as abnormal activity based on the training data set. Whenever an abnormal activity is found, it is sent to the alert system.

Flow of the proposed model:

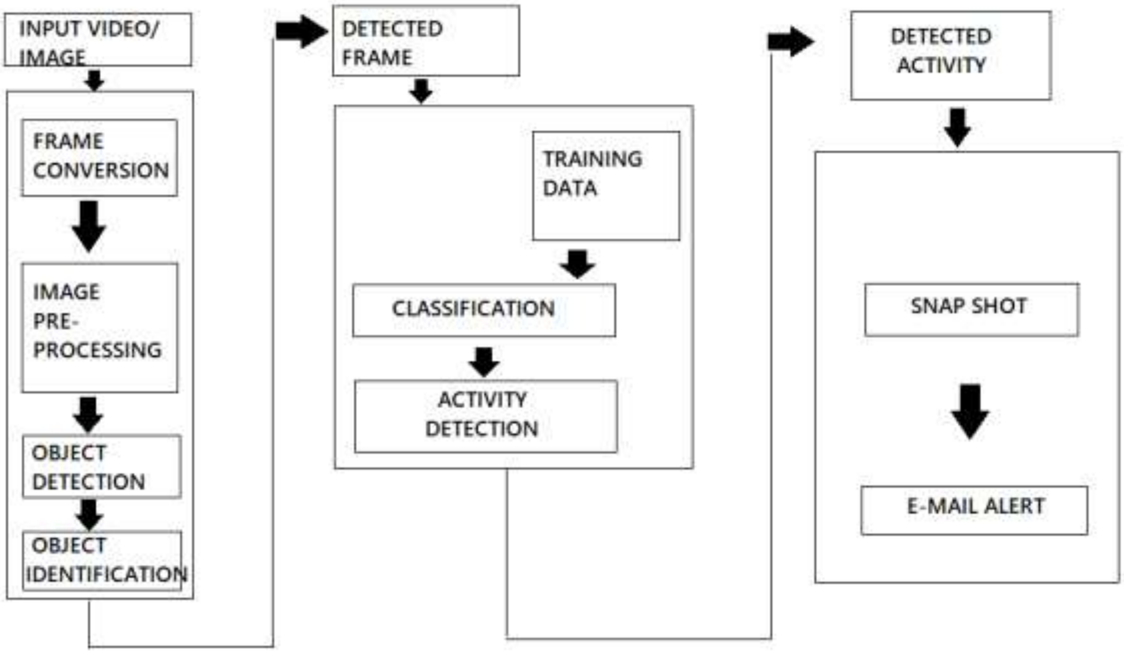

Fig [1] proposed model

Experiment results

Dataset

The chosen database contains 2 aspects, first is a set of images and their adjoining coordinate numbers for the location of the gun in the aforementioned image. The image is in jpg form
Such as

Fig [2] example photo of the dataset

And the location of the image is given in text form such as 0 0.387833 0.356771 0.167300 0.140625 which are all four values of the location of the gun on the image. The first two numbers are the x axis and the y axis location of the first corner and the next two numbers are the axis's if the ending corner of the box. Similarly for all the data, there is an image for learning and the location of the weapon given in the form of coordinates.

Method Analysis

The algorithms of SSD, R-CNN, YOLO-v3, YOLO-v4 are run to check the speed and accuracy of each of the algorithms and we record the output that is observed according to the working of each algorithm.

Output

The output photograph when a weapon is detected is shown labelled like this.

Fig [3] the weapon is detected and labelled as knife and the snapshot is shown

The output after running the algorithm is the detection of weapon

Observation

Table [1]: Comparison of different algorithm in terms of Accuracy and Speed

	Approximate Accuracy (in %)	Speed (at millisecond to recognize Image) image
SSD	43	76.8
R-CNN	81	73.2
YOLO-v3	78	66.4
YOLO-v4	98	52.8

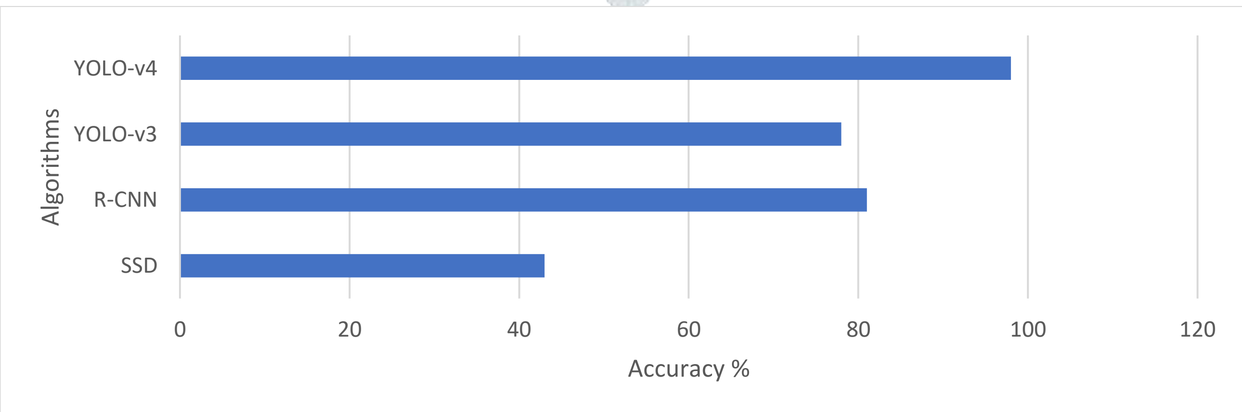

Fig [4] approximate Accuracy (in %)

As we can see by graph [1] that YOLO-v4 displayed the highest accuracy followed by R-CNN algorithm.

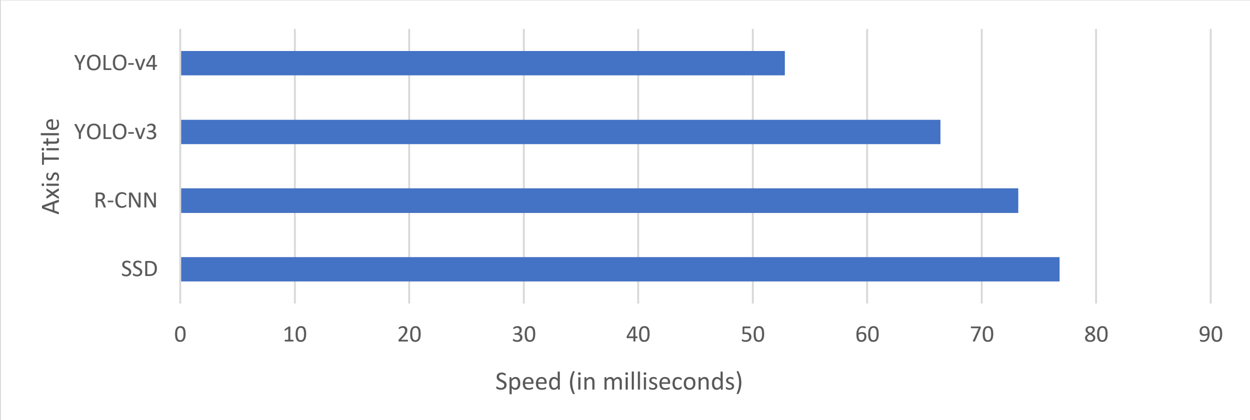

Fig [5] speed (at millisecond to recognize image) of image

In the graph [2] we can see the order at which the speed of the algorithms recorded is Slowest in SSD then R-CNN, YOLO-v3, YOLO-v4 in that order.

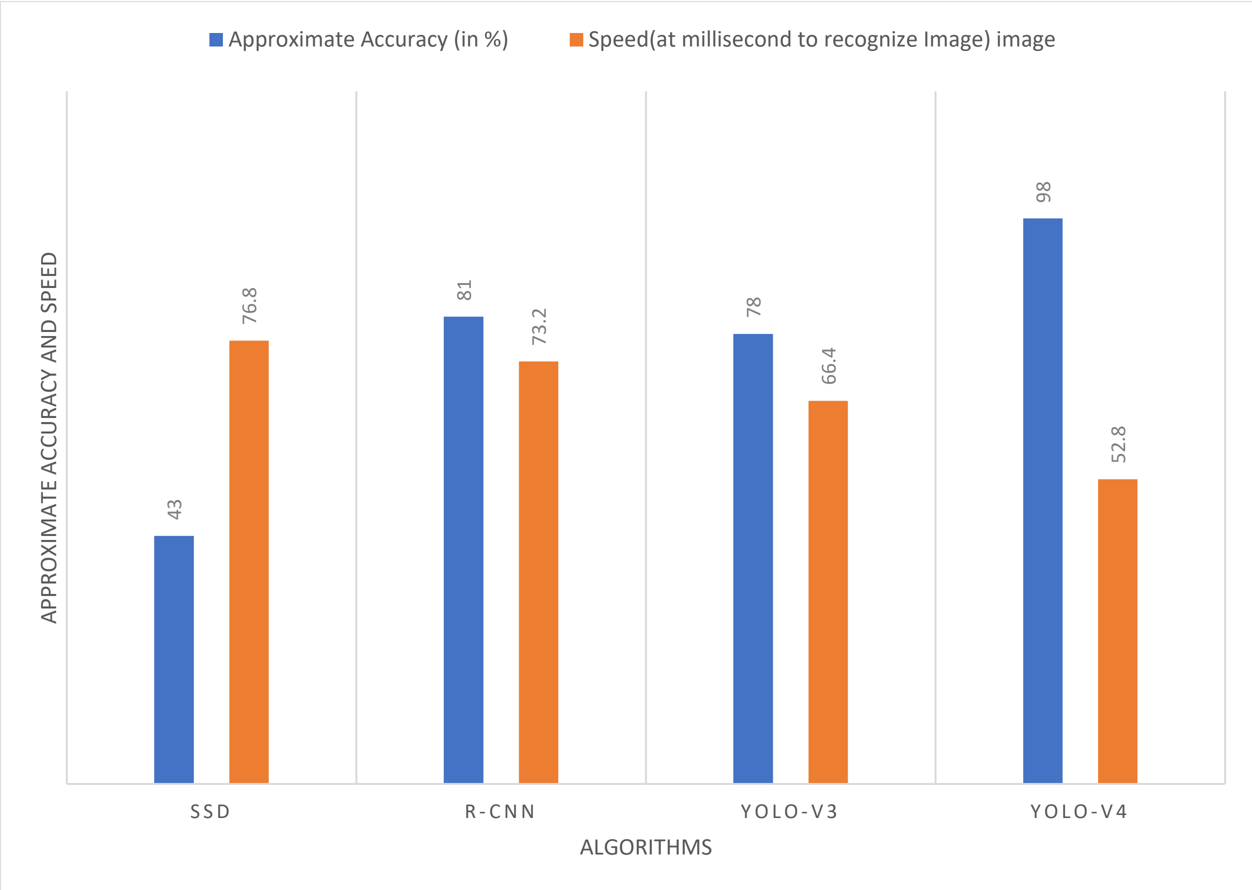

Fig [6] comparison of accuracy and speed

Conclusion

We have presented a comparative evaluation of object detection algorithms for determining weapons from CCTV images. The SSD provides the worst performance of the available algorithms, It requires least amount of computational power. The R-CNN and the YOLO-v3 algorithms showed similar results even though R-CNN uses selective search after dividing the image into boxes and YOLO-v3 uses the you look only once method. YOLO-v4 algorithm by far trumps the other algorithms in the experiment with 98% accuracy and 52.4 millisecond speed. This effectively declares the YOLO-v4 as the algorithm with the higher speed and highest accuracy.

References

1. https://intellicaai.medium.com/acomparativestudyofcustomobjectdetectionalgorithms9e7ddf6e765e
2. https://intellica.ai/blog/blog.html?a-comparative-study-of-custom-object-detection-algorithms-9e7ddf6e765e
3. Olmos, Roberto, Siham Tabik, and Francisco Herrera. "Automatic handgun detection alarm in videos using deep learning." *Neurocomputing* 275 (2018): 66-72.
4. Warsi, Arif, et al. "Gun detection system using YOLOv3." *2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*. IEEE, 2019.
5. Grega, Michał, Andrzej Matiolański, Piotr Guzik, and Mikołaj Leszczuk. "Automated detection of firearms and knives in a CCTV image." *Sensors* 16, no. 1 (2016): 47.
6. Jain, H., Vikram, A., Kashyap, A., & Jain, A. (2020, July). Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 193-198). IEEE.
7. Lai, Justin, and Sydney Maples. "Developing a Real-Time Gun Detection Classifier." Course: CS231n, Stanford University (2017).
8. Pang, Lei, et al. "Real-time Concealed Object Detection from Passive Millimeter Wave Images Based on the YOLOv3 Algorithm." Sensors 20.6 (2020): 1678.
9. Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934 (2020).