International Journal of Multidisciplinary Engineering in Current Research

ISSN NO: 2456 - 4265

Volume 6, Issue 3, March 2021, http://ijmec.com/

SCALABLE OBJECT DETECTION USING ARTIFICIAL INTELLIGENCE AND DEEP LEARNING

Mr.P.Premchand[1], Mrs.M.Alekhya[2], Mrs.B.Vaila[3], Ms.G.Bhargavi[4]

Assistant Professor[1,2,3,4]

Department of CSE

NRI INSTITUTE OF TECHNOLOGY, Visadala X Road, Medikonduru(M), Guntur (Dist), Andhra Pradesh 522438.

Abstract:-Due of the increased crime rate during busy events or in suspiciously quiet regions, security is always a major worry in every sector of life. Computer vision has a wide range of applications in the identification and monitoring of abnormalities. Video surveillance systems that can detect and analyse the scene and anomalous occurrences play an important role in intelligence monitoring because of the rising need for safety, security, and personal property protection. The SS D and Faster RCNN methods, which are based on convolution neural networks (CNNs), are used in this study to create automated gun (or weapon) identification. Two datasets are used in the proposed implementation. One dataset featured pre-labeled photographs, while the other had images that had to be manually labelled by the researcher. Algorithms may be used in real-world scenarios depending on the tradeoff between speed and accuracy, but results are summarized

Keywords— The use of artificial intelligence in surveillance systems, the detection of armed intruders, and the development of a faster RCNN (AI).

INTRODUCTION

A weapon or anamoly detection is the discovery of anomalous occurrences or things, which are not deemed to be a regular event or item in a pattern or objects contained in a dataset, and are therefore distinct from the existing patterns. Patterns that are out of the ordinary are known as anomalies. Since anomalies are linked to specific phenomena, it's difficult to generalise about them. Object detection employs feature extraction and learning techniques or models to identify different types of things [6]. [6] Gun detection and categorization are the primary goals of the proposed implementation. Because a false warning might trigger unfavourable reactions, accuracy is also an issue [11]. [12]. To have the best of both worlds, it's necessary to choose the appropriate strategy. Figure 1 illustrates how deep learning may be used to identify weapons. The input video is broken down into its individual frames. Anmethod for frame differencing is used, and a bounding box is first generated [7, 8, 14].

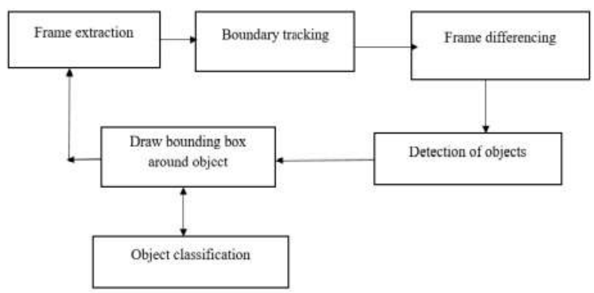
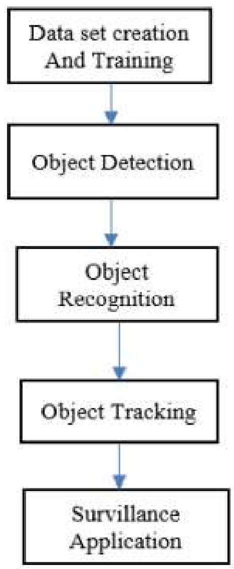


Fig.1.Methodology

International Journal of Multidisciplinary Engineering in Current Research
ISSN NO: 2456 - 4265
Volume 6, Issue 3, March 2021, http://ijmec.com/



T racking and detection in Fig. 2 Figure 2 depicts the process of detecting and tracking an item. An object detection algorithm is given data from a newly constructed dataset that has been trained and refined. The detection technique for guns is selected according on the application (SSD or fast RCNN). The method uses a variety of machine learning models, such as the Region Convolutional Neural Network (RCNN) and Single Shot Detection (SSD), to tackle the detection issue [2][9][15].

II. IMPLEMENTATION A.

Resources or components used for implementation

3.4 of the OpenCV computer vision library is now available as an open source project.

High-level programming language for image processing applications such as Python 3.5.

In this dataset, you'll find a collection of common items, each with a unique name.

• Tensorflow 1.1 and Anaconda 1.1

GeForce is a brand of graphics processing units created by Nvidia. Case I: Video Specifications in the Dataset Specifications

• Intel i5 7th Generation System Configuration (4 Cores)

An NVIDIA GeForce 820M graphics card powers this system.

There are 29.97 frames per second for the input.

the output frame rate is 0.20 frames per second.

•.mov is a video formatCOCO and self-created picture datasets are included in the video.

In this case, the number of courses taught was five.

• Intel i5 7th Generation processor (4 Cores)

• Clock Speed: 2.5 GHz • GPU: NVIDIA GeForce 820M •

Dimensions of the input image: 200-300 KB; Training Time: 0.6 seconds (SSD)

• 1.7 milliseconds (RCNN)

Dataset of COCO and self-created images, in.JPG image format

Five Cs: Assumptions and Constraints for the implementation

• The weapon is in full or partial view of the camera and has been completely or partly exposed.

• The ammo may be seen against a light backdrop.

In order to speed up the identification of ammo, a GPU with high-end computational capacity was utilised.

• The system is not fully automated. There will be a person in charge of verifying each and every gun detection warning that is sent.

D. FASTER R-CN

International Journal of Multidisciplinary Engineering in Current Research

ISSN NO: 2456 - 4265

Volume 6, Issue 3, March 2021, http://ijmec.com/

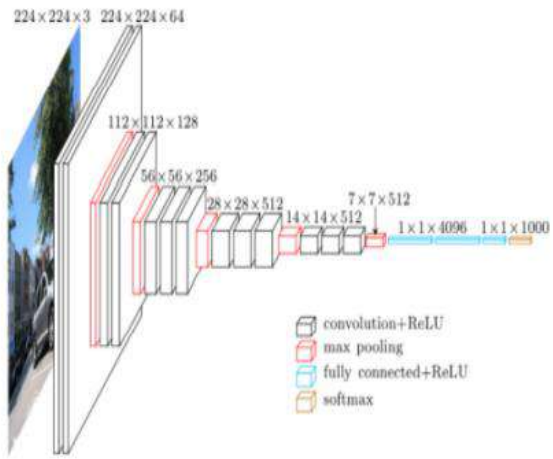


Fig 3. Layers in CNN Architecture [5]

As indicated in figure 3, the CNN and RCNN architectures are shown in their respective layers. In order to create region suggestions and to identify objects, it has two networks: RPN and network. Selective search is used to create the region suggestions. The RPN network assigns a ranking to anchors and region boxes.

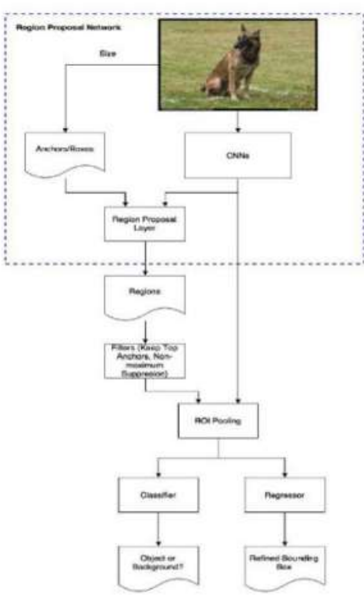


Fig 4. Faster R-CNN [5]

Creating and Educating a Dataset Fatkun Batch Image Downloader (chrome extension) may download several Google Images at once and download them all at once. After that, the photos are downloaded and labelled. Training pictures make up 80% of the total number of photos, while testing images make up 20%. Single Shot Detector (SSD) was used to train the ammo dataset, and 2669 iterations/steps were done on the model to confirm that the loss was less than 0.05. Test and training photographs are shown in the folder shown in Figure 5. The picture in Figure 6 has labels added.

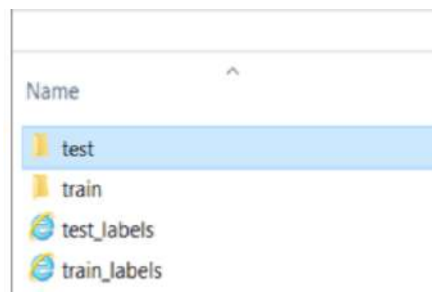


fig.5. Folder with test and train images

Fig.6. Image along with its label

International Journal of Multidisciplinary Engineering in Current Research
ISSN NO: 2456 - 4265
Volume 6, Issue 3, March 2021, http://ijmec.com/

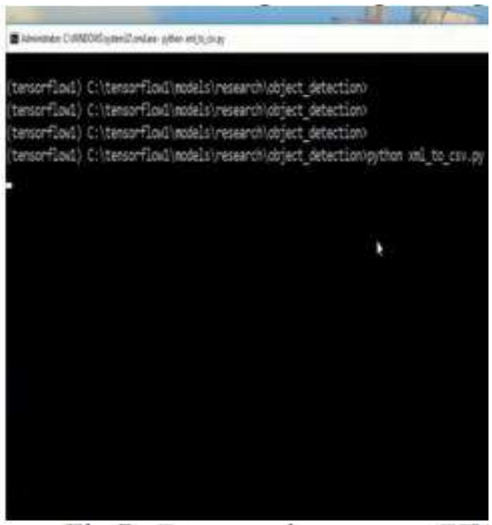
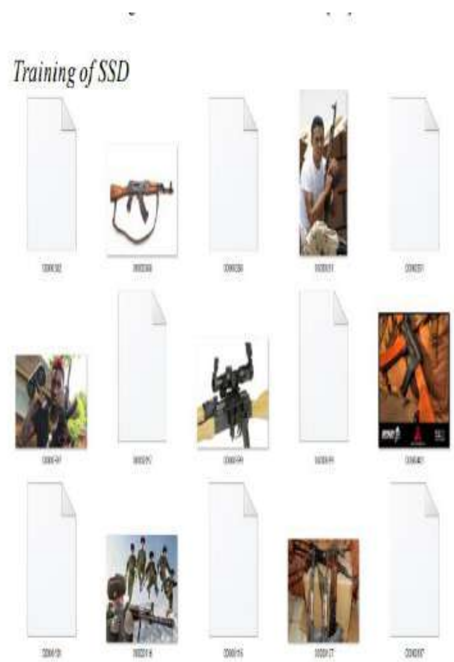


Fig.7. Command to create CSV files for the image labels XML data is converted into CSV file by executing this command in Anaconda Prompt: python xml_to_csv.py as shown in figure 7.

	A	B	C	D	E	F	G	H
1	filename	width	height	class	xmin	ymin	xmax	ymax
2	00000022.	600	450	ak47	142	197	567	300
3	00000028.	600	439	ak47	251	11	388	392
4	00000030.	600	900	ak47	90	221	467	374
5	00000034.	500	389	ak47	56	42	444	322
6	00000038.	600	450	ak47	19	9	597	402
7	00000039.	600	600	ak47	160	240	380	382
8	00000039.	600	600	ak47	245	288	400	434
9	00000039.	600	600	ak47	6	160	367	381
10	00000052.	600	438	ak47	325	11	388	101
11	00000052.	600	438	ak47	383	1	435	191
12	00000055.	482	200	ak47	263	147	318	180
13	00000079.	480	480	ak47	2	332	480	436
14	00000079.	480	480	ak47	1	198	478	310
15	00000098.	240	240	ak47	5	94	235	147
16	00000099.	600	427	ak47	259	73	417	206
17	00000112.	600	800	ak47	175	358	494	503
18	00000112.	600	800	ak47	1	200	293	323
19	00000112.	600	800	ak47	379	293	545	587
20	00000121.	600	376	ak47	1	46	599	259
21	00000122.	300	257	ak47	119	54	200	103
22	00000127.	380	570	ak47	195	218	372	570
23	00000130.	480	480	ak47	21	246	176	295
24	00000130.	480	480	ak47	11	12	194	70
25	00000130.	480	480	ak47	13	87	158	165
26	00000144.	600	450	ak47	21	19	597	359
27	00000147.	360	170	ak47	8	59	344	163
28	00000151.	600	337	ak47	1	43	305	301
29	00000163.	600	963	ak47	238	423	419	869
30	00000169.	480	480	ak47	4	132	478	330

test_labels

Fig.8. CSV file of testing dataset



Photos that have been manually labelled are shown in Fig.13. Figure 13 depicts xml-formatted images, as demonstrated (xml). Each picture was assessed using 838 and 241 photographs, respectively, as part of training (22 percent testing and 78 percent training). Anaconda Prompt's python xml to csv.py command turns XML data into a CSV file. The truth box and the size of the actual world are shown in this way.

$$predicted_{center} * variance0 = \frac{real_{center} - prior_{center}}{prior_{hw}}$$

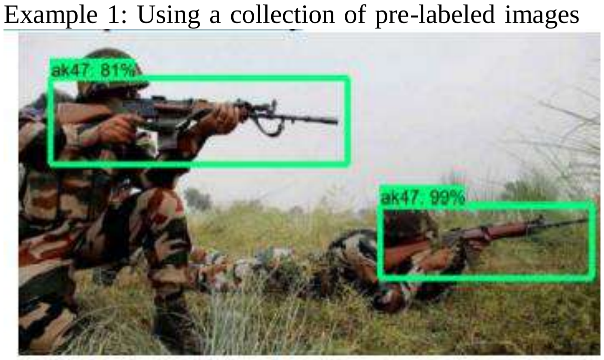
$$\exp\left(predicted_{\square_{hw}} * variance1\right) = \frac{real\_hw}{prior_{hw}}$$

In order to verify that the item is detected, the label map and tf record file must be changed.There is a file called a "label map" that contains information on every potential category of thing that may exist. Labels for the AK47 have been added to the map.Object identification and detection

International Journal of Multidisciplinary Engineering in Current Research
ISSN NO: 2456 - 4265
Volume 6, Issue 3, March 2021, http://ijmec.com/

We update the label map and TFF records to make sure the item is detected. A file called a label map will be used tocontain information on all potential types of objects. The AK47 has been added to the label map. A sensitivity of 72 percent and 67 percent, respectively, were found for the COLT M1911 and the Smith & Wesson Model handgun in Figures 19 and 20. B. Detection of weapons with a faster R-CNN Using pre-labeled photos as an example

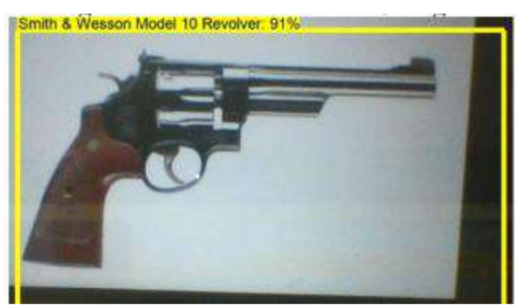
Both the COLT M1911 and the Smith & Wesson Model pistol were detected with a sensitivity of 72% and 67%, respectively, in Figures 19 and 20, respectively. B. Faster R-CNN for weapon detection Example 1: Using a collection of pre-labeled images



A faster R-CNN is used to detect the AK47 (Fig. 22). Figure 22 demonstrates the identification of an AK47 in the hands of the military with a 99 percent and 81 percent accuracy rate, respectively.

Recognizing the AKM47 assault rifle in the video stream (Fig. 23) As you can see in Figure 23, we were able to accurately recognise an AK47 rifle from the video stream. An picture dataset produced by the user is used in Case 2.

Fig.24. Detection of Colt M1911 gun



ISSN NO: 2456 - 4265

Volume 6, Issue 3, March 2021, http://ijmec.com/

Fig.25. Detection of Smith & Wesson Model 10 gun

Figures 24 and 25 show how the Faster R-CNN approach correctly identified a Smith & Wesson Model 10 and a Colt M1911. C. Analyzing the Data Table I shows the results of a quicker R-CNN algorithm's performance study.

IV. Conclusions

Pre-labeled and self-created photo datasets are used to simulate weapon (gun) detection. When employing these algorithms in real time, the trade-off between speed and accuracy must be considered. The SSD algorithm has a frame rate of 0.736 s/s. Faster RCNN delivers a frame rate of 1.606s/frame when compared to SSD. 84.6 percent more accurate than Slower RCNN is Faster RCNN. Comparing SSD's performance against that of the more accurate RCNN, it comes up short with an accuracy rate of just 73.8%. Due to SSD's faster speed, real-time detection was made feasible, however the Faster RCNN was more accurate. Large datasets may be trained using GPUs and high-end DSP and FPGA systems.

REFERENCES

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

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

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

[4] 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.

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

[6] 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.

[7] 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.

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

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

[10] 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.

[11] A. Glowaczet.al., "Visual Detection of Knives in Security Applications using Active Appearance Model",Multimedia Tools Applications, 2015.

International Journal of Multidisciplinary Engineering in Current Research
ISSN NO: 2456 - 4265
Volume 6, Issue 3, March 2021, http://ijmec.com/

[12] S. Pankantiet.al.,“Robust abandoned object detection using regionlevelanalysis,”International Conference on Image Processing,2011.

[13] Ayush Jain et.al.,“Survey on Edge Computing - Key Technology in Retail Industry” International Conference on Intelligent Computing and Control Systems,2019.

[14] Mohana et.al., Performance Evaluation of Background Modeling Methods for Object Detection and Tracking,” International Conference on Inventive Systems and Control,2020.

[15] J. Wang et.al., “Detecting static objects in busy scenes”, Technical Report TR99-1730, Department of Computer Science, Cornell University, 2014.

[16] V. P. Korakoppa et.al., “An area efficient FPGA implementation of moving object detection and face detection using adaptive threshold method,” International Conference on Recent Trends in Electronics, Information & Communication Technology,2017.

[17] S. K. Mankaniet.al., “Real-time implementation of object detection and tracking on DSP for video surveillance applications,”International Conference on Recent Trends in Electronics, Information & Communication Technology,2016.