ISBN 978-1-6654-3986-2

Real Time Crime Detection Using Deep Learning Algorithm

P. Sivakumar
Professor, Department of Information Technology,
Manakula Vinaygar Institute of Technology,
Pondicherry, India.
hodit@mvit.edu.in

Jayabalaguru. V
Department of Information Technology,
Manakula Vinaygar Institute of Technology,
Pondicherry, India.
jayabalaguruv@gmail.com

Ramsugumar. R.
Department of Information Technology,
Manakula Vinaygar Institute of Technology,
Pondicherry, India.
ramsugumar160300@gmail.com

Kalaisriram. S
Department of Information Technology,
Manakula Vinaygar Institute of Technology,
Pondicherry, India.
kalaisukheshram@gmail.com

2021 International Conference on System, Computation, Automation and Networking (ICSCAN) | 978-1-6654-3986-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSCAN53069.2021.9526393

***Abstract* - The Crime Rate and number of Criminals are increasing day by day, so there is great concern about the security issues, so to prevent and identify the crime before it occurs is the primary goal of the police officials. With the help of Recent Technologies, especially CCTV is normally deployed in every private and public area to control crime but it needs human supervision to monitor. It's hard for a human to monitor many screens at the same time. It leads to many errors. To overcome these problems, we proposed Real-Time Crime Detection Technique using a Deep Learning Algorithm which monitors real-time videos and alerts the nearby Cybercrime admin about the occurrence of crime with current location. In this paper, We present YOLO as our object detection algorithm. Our architecture is extremely fast and process image in real-time at 45 frames per second.**

***Keywords — Deep learning; YOLO; Object Detection; Mean Average Precision*.**

I. INTRODUCTION

YOLO is a deep learning object detection algorithm that uses a Convolution neural network, which is one of the most effective algorithms. It differs from the rest of the algorithms because it uses a single convolution neural network. Bounding boxes and their corresponding probabilities are predicted by the One Convolutional neural network. The trained model makes its detection based on its final weights trained by the dataset so that the performance of the trained model can be maximized. When the system is developed, there are many challenges faced in developing the model. To use YOLO algorithm, we need maximum utilization of the Graphical Processing Unit, so we use Google Colaboratory which comes with an in-built Tesla k80 GPU, Which is a more powerful and faster GPU. To increase the Mean Average Precision and to decrease the final average loss of the model, the database which is used to train the model must be bigger. The Trained model is able to detect many images even in a complex scene.

II. NEED OF THE SYSTEM

As we know that prevention is better than cure, preventing a crime before occurring is better than investigating what or how the crime had occurred. Just like vaccination is given to people to prevent the disease, in today's world with a way higher rate and crime happenings, it's become necessary to possess a Crime detection technique that prevents the crime happenings. These Detection Techniques can be used by police officials to detect the crime before it occurs so that the crime can be prevented.

III. RELATED WORKS

A. VIDEO SURVEILLANCE

Pratibha, Akanksha, and Gahalot presented an automated video surveillance system using statistical characteristics for crime detection. If there is any kind of abnormal situation like kidnapping, chain snatching, and fighting, the trained model alerts the situation to the nearby police station. Predicting the human behavior and moving target are the key features of the existing system. There are many advanced cameras and CCTV infrastructure installed in many areas. The main goal of the crime video surveillance system is to provide automatic detection of the crime scene by analyzing the movements for preventing future incidents[11].

B. WEAPON DETECTION

Weapon detection in real-time image and video is already existing research. We go through several image classification research including analyzing data from various weapons, then detection of violent crime scenes and detection of a gun in the

ISBN 978-1-6654-3986-2

video feed. Most of the research was not implemented in deep learning/neural network techniques so that this paper focuses on automatic detection of the gun in surveillance camera and alert if the gun is detected[4].

C. OBJECT DETECTION

Nowadays, In the field of Computer vision object detection plays a major role. The main objective of object detection is to detect and classify objects in several fields and applications like face detection and recognition. Our model success is based on the fast object detection algorithm and its efficiency. In recent times, the attention towards object detection is increasing, Several Real-time applications like security applications, automatic driving, robotics, etc are working under object detection. Computer vision and image processing are associated with object detection technology with deals with digital images and videos[12].

D. YOLO

Detection and classification of object detection is the main purpose of the You Only Look Once algorithm. There are four versions in the YOLO algorithm, each version will improve the performance of the model. When comparison with RCNN, the YOLO algorithm uses the feature of the entire image rather than a part of the captured image. YOLO algorithm performs bounding box and class prediction are done simultaneously, so it is unique from other traditional systems. YOLO treats the object detection problem as a regression problem instead of having a normal region proposal and classification. This makes YOLO perform more efficiently in real-time with some loss of accuracy.[8]

IV. EXISTING TECHNIQUES

Apporva.P, Impana H.C presented a real-time automated criminal detection system using Haar feature-based cascade classifier. This system will be very much helpful to detect and recognize the criminal face automatically in real-time.

The existing technique for criminal detection consists of 4 steps: 1. Training the Criminal Images, 2. Face Detection using Haar cascade algorithm, 3. Compare the trained images with an image captured from the real-time surveillance camera, 4. Result Comparison whether Criminal or not.

The real-time video captured from the CCTV camera which is converted into frames, then the face is detected in the frame using the haarcascade algorithm, it is preprocessed. The feature of the processed real-time image is compared with the feature of criminal images which are stored in the database. If a criminal is matched, then the alert is sent to the police personal which all the details and the time for which he was caught under the surveillance of camera [1].

V. PROPOSED TECHNIQUES

Crime is increasing considerably day by day. With the rapid increase in violent crimes, an efficient surveillance system has become necessary. In this, a model has been proposed for detecting face and object when a criminal with threatening weapons was caught under the surveillance of the CCTV camera. The server is connected with a real-time CCTV camera at each location. The criminal face and weapons are stored in the database which is connected to the server. The database is maintained by the server that contains all weapons and criminals to be found. Live feed obtained from the CCTV camera, if the Criminal is present in any feed, then the live location of that crime will be tracked and sent to the cybercrime admin. The Live CCTV image is trained with the existing trained dataset.

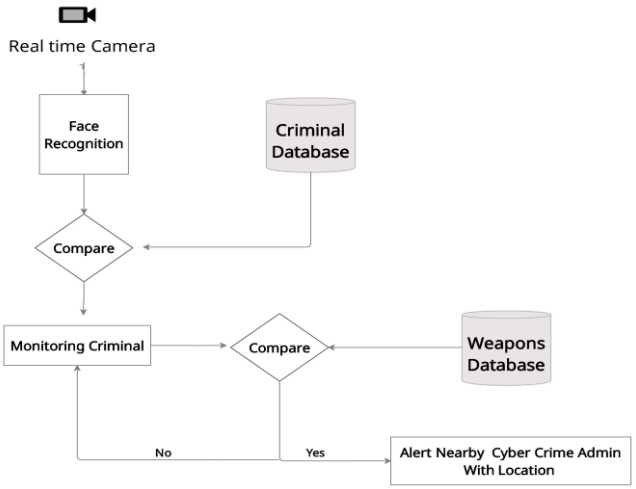

Fig.1. Architecture Diagram for proposed system

A. Dataset

Creating a Dataset of Many High-Resolution images of our desired classes and objects. The data we collected as criminals consists of 100 images for each criminal, and also we collected Images of Weapons like a knife, gun, piston, etc. Criminal Face Detection in this dataset is very challenging because of rich variations in pose, scale, face expressions, etc. To create the dataset we are using the ybat annotation tool for annotating and labeling the images. The boxes must be already be manually annotated to ensure precision and regularity. All Criminal and Weapons Dataset must be annotated with labels, object bounding boxes.

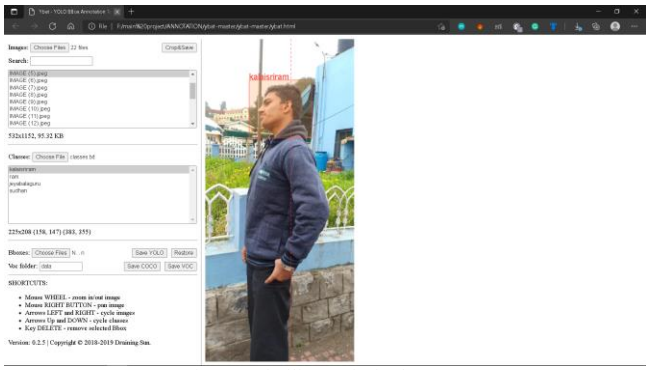

Fig. 2. Labelling Criminal Face 1

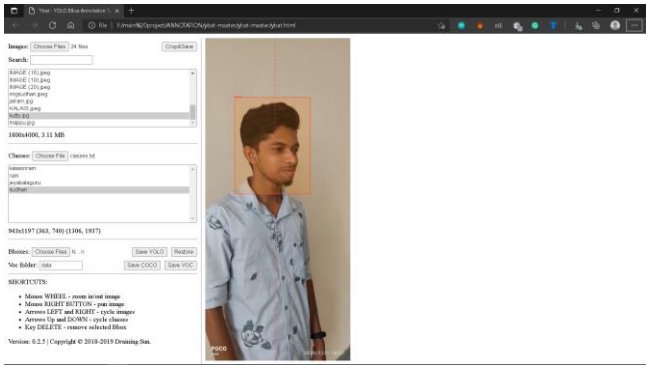

Fig. 3. Labelling Criminal Face 2

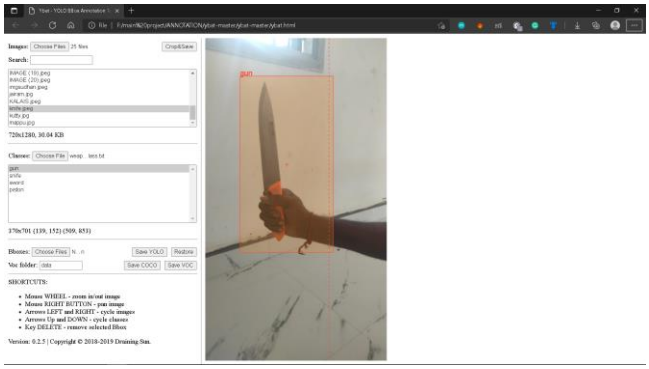

Fig. 4. Labelling Weapons

` After opening the ybat annotation tool choose all the criminal images, then the class file where are the criminal names should be entered, then click the respective criminal name in class then annotate all the images. After annotating all the images, click the SAVE YOLO button. After saving the YOLO, a text file will be generated for every criminal image which contains the coordinates of bounding boxes like XMin, XMax, Ymin, YMax. This same procedure is done for labeling the weapons also.

There are the values of the bounded box of two criminal shown in Figure 2 and 3.

1. 0.6332465277777778 0.24247685185185186 0.3151041666666667 0.3645833333333333

2. 0.5811631944444444 0.41898148148148145 0.0998263888888889 0.1527777777777778

The values of weapons of Labelled Image in Figure 4

1. 0.42916666666666664 0.38359375 0.425 0.5453125

And we store all the images with corresponding text files in the same directory for training. Once the dataset is ready, We can start our Testing on Google Colaboratory.

B. Training (Google Colaboratory)

Once done with the dataset, Google Colaboratory is used for building the YOLOv4 algorithm. The Criminal, Weapon Images and their respective text files need to be inside a zip archive. The Zip Archive has to be mounted with Google Drive. Enable NVIDIA GPU and authenticate Google drive in google Colaboratory. Also, clone and configure the Darknet on Google Colaboratory. Darknet is a Framework to train neural networks for training yolo. Object detection framework for every image in the dataset is run using Python Script. Then start the Training. Training must be done for at least 2000 iteration, the more iteration more the model efficiency will be improved and the reduction also reduced.

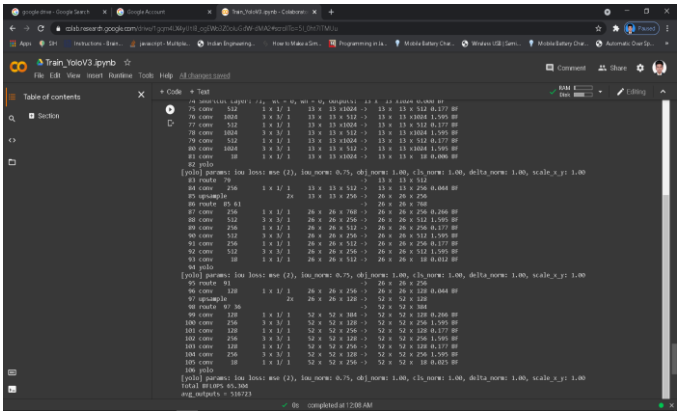

Fig. 5. Training the Dataset of Criminal and Weapons

Maximum iteration makes the system perform with better accuracy and efficiency. After all the iterations then we get weights which are trained datasets for our respective criminal's data. Once the data is trained, our system is ready to have Real-time data into it to run detections.

C. Testing

The testing process is used to validate the algorithm. The data is split in the ratio such that 80% is for training and 20% is for testing. This trained model is tested with some unseen data and the model performance is checked. Our model runs at 45

frames per second. The dataset we used for training and testing consisted of 100 images per Criminal, the dataset was divided 80:20 between training and test images for 20000 iterations.

D. Detection

After Training and Testing, the system runs on sample inputs to check whether if there is an error in identifying criminals. The time taken to detect and the confidence in identifying the criminals will be generated for every detection made by the neural network, it can identify all the criminals present in the image in just 89.6 milliseconds even though the criminal is partially covered. This model still detects the presence of criminals even though it is less confident.

VI. Performance measure

In this paper, the performance of the object and criminal detection is measured by Mean Average Precision.

Precision and Recall: Precision which is used to measure the correctness of the prediction, and we use Recall to calculate true predictions out of total predictions.

The ratio between the number of correct positive samples to the total number of positive samples, then it is called precision. The model's accuracy is calculated by the precision classifying positive samples.

$$\boldsymbol{Precision = \frac{True_{positive}}{True_{positive} + False_{positive}}}$$

TP = True Positives (Predicted as positive as was correct)
FP = False Positives (Predicted as positive but was incorrect)

The ratio between the number of correct positive samples to the total number of positive samples, which is used for detecting the model's ability to identify the positive sample. If the recall is higher than the model detects more positive samples.

$$\boldsymbol{Recall = \frac{True_{positive}}{True_{positive} + False_{negative}}}$$

TP = True Positives (Predicted as positive as was correct)
FN = False Negatives (Failed to predict an object that was there)

Average Precision: To calculate Mean Average Precision, First we have to calculate the average precision of every class.

The Average Precision is a single value that represents the average of all precision.

Mean Average Precision: YOLO manages to achieve more than twice the mean average precision (mAP) of other real-time systems. Mean Average precision is the average of the average precision taken for each class. In other words, mean average precision is the average precision taken over all classes.

$$\boldsymbol{mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k}$$
$$\boldsymbol{AP_k = the\ AP\ of\ class\ k}$$
$$\boldsymbol{n = the\ number\ of\ classes}$$

VII. Result and discussion

This paper, proposed the criminal face and weapons detection in real time. We compare the proposed model using YOLO with existing Fast R-CNN with respective Precision, Recall and Mean Average Precision.

TABLE. 1.COMPARISON OF THE RESULTS - YOLO Vs FAST RCNN

Model	Precision	Recall	mAP
Fast RCNN	92.56%	89.33%	62.4%
YOLO	98.56%	91.21%	78.3%

Table 1 shows the comparison of our proposed work YOLO and existing Fast RCNN. From the results, it shows that YOLO is very effective in running the detection of criminals and weapons. YOLO takes 44 ms to process one image, it performs the fastest. The Mean Average Precision of proposed work can achieve 78.3% on testing set while fast RCNN is just 62.4%. The Final Mean Average Precision of our proposed work is about 78.3% in our testing set, which is best compared with other object detection algorithm.

VIII. Conclusion

In this paper, we can monitor the criminals in a real-time video stream and find the weapons that are not necessary for public places and alert the nearby police cybercrime admin, which increases the security of society and decrease the crime rate with the help of technology. We proposed a Real-time Criminal detection system using the YOLO algorithm with a mean average precision of 78.3% and the final average loss of 0.6, which can successfully detect criminals even in crowded places.

References

[1] Apporva.P, Impana. H.C, Siri.S.L, Varshitha.M.R, "Automated Criminal Identification By Face Recognition Using Open Computer Vision Classifiers", International Conference on Computing Methodologies and Communication(ICCMC 2019).

[2] Piyush Chhoriya, "Automated Criminal Identification System using Face Detection and Recognition", International Research Journal of Engineering and Technology, October 2019.

[3] Pratibha, Akanksha Gahalot, Uprant, Suraina Dhiman, Lokesh Chouhan, “Crime Prediction and Analysis”, Institute of Electrical and Electronics Engineers, Stansford University, 2017.

[4] Justin Lai, Sydney Maples, “Developing a Real-Time Gun Detection Classifier”.

[5] kruti Goyal, kartikey Agarwal, Rishi Kumar, “Face Detection and Tracking”, International Conference on Electronics, Communication and Aerospace Technology(IFECA 2017).

[6] Jazeem Azeez, D. John Aravindhar ,”Hybrid Approach to Crime Prediction using Deep learning”, International Conference on Advances in Computing, Communications and Informatics (ICACCIL),2015.

[7] Nurul Iman Hassan, Nooritawati Md. Tahir, Habibah Hashim, “People Detection System using YOLOv3 Algorithm”, International Conference on Control System, Computing and Engineering (ICCSCE2020), August 2020.

[8] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi,”You Only Look Once: Unified, Real-Time Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition,2016.

[9] Weijun Chen · Hongbo Huang1,· Shuai Peng1· Changsheng Zhou· Cuiping Zhang,” YOLO-face: a real-time face detector”, Springer-Verlag GmbH Germany, part of Springer Nature 2020.

[10] A.S. Osan , V. Thutheja , M. Mulchandani,”AI Based Crime Detection System”,IOSR Journal of Engineering, May 2019.

[11] Chulyeon Kim, Liyoung Lee, Taekjin Han and Young-Min kim, “A hybrid framework combining background subtraction and deep neural networks for rapid person detection”, Journal of Big Data, 2018.

[12] Tanvir Ahmad, Yinglong Ma, Muhammad Yahya, Belal Ahmad, Shah Nazir and Amin ul Haq, “Object Detection through Modified YOLO Neural Network”,Hindawi Scientific Programming Volume 2020.