IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

Weapon Detection in Real-Time CCTV Videos Using Deep Learning

Akshay Ahire, Tharun Maddale, Kuthala Reddy, Ezaldeen Amair, Prof. Dnyanshwari Umap
Sandip University, Nashik, India.

Abstract: *Security and safety is a big concern for today's modern world. For a country to be economically strong, it must ensure a safe and secure environment for investors and tourists. Having said that, Closed Circuit Television (CCTV) cameras are being used for surveillance and to monitor activities i.e. robberies but these cameras still require human supervision and intervention. We need a system that can automatically detect these illegal activities. Despite state-of-the-art deep learning algorithms, fast processing hardware, and advanced CCTV cameras, weapon detection in real-time is still a serious challenge. Observing angle differences, occlusions by the carrier of the firearm and persons around it further enhances the difficulty of the challenge. This work focuses on providing a secure place using CCTV footage as a source to detect harmful weapons by applying the state of the art open-source deep learning algorithms. We have implemented binary classification assuming pistol class as the reference class and relevant confusion objects inclusion concept is introduced to reduce false positives and false negatives. No standard dataset was available for real-time scenario so we made our own dataset by making weapon photos from our own camera, manually collected images from internet, extracted data from YouTube CCTV videos, through GitHub repositories, data by university of Granada and Internet Movies Firearms Database (IMFDB) imfdb.org. Two approaches are used i.e. sliding window/classification and region proposal/object detection. Some of the algorithms used are VGG16, Inception-V3, Inception-ResnetV2, SSDMobileNetV1, Faster-RCNN Inception-ResnetV2 (FRIRv2), YOLOv3, and YOLOv4. Precision and recall count the most rather than accuracy when object detection is performed so these entire algorithms were tested in terms of them. Yolov4 stands out best amongst all other algorithms and gave a F1-score of 91% along with a mean average precision of 91.73% higher than previously achieved..*

Keywords: Gun detection, deep learning, object detection, artificial intelligence, computer vision

I. INTRODUCTION

The crime rate across the globe has increased mainly because of the frequent use of handheld weapons during violent activ- ity. For a country to progress, the law-and-order situation must be in control. Whether we want to attract investors for investment or to generate revenue with the tourism industry, all these needs is a peaceful and safe environment. The crime ratio because of guns is very critical in numerous parts of the world. It includes mainly those countries in which it is legal to keep a firearm. The world is a global village now and what we speak or write has an impact on the people. Even if the news they heard is crafted having no truth but as it gets viral in a few hours because of the media and especially social media, the damage will be done. People now have more depression and have less control over their anger, and hate speeches can get those people to lose their minds. People can be brainwashed and psychological studies show that if a person has a weapon in this situation, he may lose his senses and commit a violent activity.

High incidents were recorded in past few years with the use of harmful weapons in public areas. Starting with the past year's attacks on a couple of Mosques in New Zealand, on March 15, 2019 at 1:40 pm, the attacker attacks the Christchurch AL-Noor Mosque during a Friday prayer killing almost 44 innocent and unarmed worshippers. On the same day just after 15 minutes at 1:55 PM, another attack hap- pened killing seven more civilians [1]. Active shooter

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

inci- dents had also occurred in USA and then in Europe. The most significant cases were those at Columbine High School (USA, 37 victims), Andreas Broeivik's assault on Uotya Island (Norway, 179 victims) or the Charlie Hebdo newspaper attack killing 23. According to stats provided by the UNODC, among 0.1 Million people of a country, the crimes involving guns are very high i-e. 1.6 in Belgium, United States having 4.7 and Mexico with a number of 21.5 [2].

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. [3]. CCTVs are installed in every public place today and are mainly used for providing safety, crime investigation, and other security measures for detection. CCTV footage is the most important evidence in courts. After a crime is committed, law enforcement agencies arrive at the scene and take the recording of footage with them [4]. If we look at the surveillance system of different countries around the world, UK has about 4.5 million cameras, which are used for surveillance. Sweden has about 50000 cameras installed around 2010. The government of Poland was able to reduce drug cases by 60% and street fights by 40% by installing just 450 cameras in the city of Poznan [5]. China has the world's biggest surveillance system and 170 million cameras around the nation, and these are expected to expand three times, through an additional 400 million to be connected by 2020. It took only seven minutes for Chinese officials to find and apprehend BBC reporter John Sudworth using their strong CCTV cameras network and facial recognition technology and put the criminal behind the bar [6].

In previous years, though having surveillance cameras installed, to use them for security purposes was not an easy and dependable method. A human has to be there all the time to monitor screens. CCTV operator has to monitor 20- 25 screens for 10 hours. He has to look, observe, identify, and control the situation that can be harmful to the individuals and the property. As the number of screens increases, the concen- tration of the person decreases considerably to monitor each screen with time. It is impossible for the person monitoring the screens to keep the same level of attention all the time [7]. The solution to aforementioned problem is to install surveillance cameras with the ability to automatically detect weapons and raise alarm to alert the operators or security per- sonals. However, there is not much work done on algorithms for weapon detection in surveillance cameras, and related studies are often considering concealed weapon detection (CWD), mostly using X-rays or millimeter waves images employing traditional machine learning techniques [8]–[12]. In the past few years, deep learning in particular convo- lutional neural network (CNN) has given groundbreaking results in object categorizing and detection. It has achieved finest results thus far in classical problems of image pro- cessing such as grouping, detection and localization. Instead of selecting features manually, CNN automatically learns features from given data.

This article presents an automatic detection and classifi- cation method of weapons for real-time scenario using state of the art deep learning models. For real-time implemen- tation relating the problem question of this work ''detect- ing weapons in real-time for potential robbers/terrorist using deep learning'', detection and classification was done for pistol, revolver and other shot handheld weapons as in sin- gle class called pistol and related confusion objects such as cell phone, metal detector, wallet, selfie stick in not pistol class. A major reason behind this was our research done on weapons used in robbery cases and it further motivated us to choose pistol and revolver as our target object. We go through several CCTV captured robbery videos on YouTube and found that almost 95% of cases have pistol or revolver as the weapon used. With the implementation of this system, many robbery crimes, and other incidents like what happened last year in New Zealand's Christchurch mosque could be controlled using early alarm system by alerting the operator and concerned authorities so action can be taken immediately. Gun detection in real-time is a very challenging task.

As our desired object has a small size so, detecting it in an image is also very challenging in presence of other objects, especially those objects that can be confused with it. Deep learning models faced several below mentioned challenges for detection and classification task:

The first and main problem is the data through which CNN learn its features to be used later for classification and detection.

No standard dataset was available for weapons.

For real-time scenarios, making a novel dataset manu- ally was a very long and time-consuming process.

Labeling the desired database is not an easy task, as all data needs to be labeled manually.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

Different detection algorithms were used, so a labeled dataset for one algorithm cannot be utilized for the other one.
Every algorithm requires different labeling and pre- processing operations for the same-labeled database.
As for real-time implementation, detection systems require the exact location of the weapon so gun block- ing or occlusion is also a problem that arises frequently and it could occur because of self, inter-object, or back- ground blocking.
Different approaches are used in this work for weapon classification and detection purpose but all have deep learning and CNN architecture behind them because of their state of the art performance. Training from scratch took very much time so the Transfer learning approach was used and Ima- geNet and COCO (common objects in context) pre-trained models are used. Different datasets were made for classifi- cation and detection. For real-time purposes, we made our dataset by taking weapon photos from the camera, data was extracted manually from robbery CCTV videos, downloaded from imfdb (internet movie firearm database), data by univer- sity of Granada and other online repositories. All the work has been done to achieve results in real-time.
The main contributions of this work are: presentation of a first detailed and comprehensive work on weapon detection that can achieve detection in videos from real-time CCTV and works well even in low resolution and brightness because most of the work done earlier is on high definition training images but realtime scenario needs realtime training data as well for better results, finding of the most suitable and appropriate CNN based object detector for the application of weapon detection in real-time CCTV video streams, making of a new dataset because real-time detection also needs real- time training data so we made a new database of 8327 images and preprocessed it using different OpenCV filters i.e. Equal- ized, Grayscale and clahe that helped in detecting images in low brightness and resolution, introducing the concept of related confusion classes to reduce false positives and nega- tives, training and testing of our novel database on the latest state of the deep learning based classification and detection models among them Yolov4 performed best in terms of both speed and accuracy and our selected trained model predict images at almost every orientation, angle, and view, achieving the highest mean average precision of 91.73% along with a F1-score of 91% on Yolov4.
The rest of the paper is organized as follows: related work
is discussed in Section II. The implementation methodology based on deep learning algorithms is explained in Section III. The dataset construction, annotation, and preprocessing using different filters have been discussed in section IV, which follows the experiments and results in Section V. Finally, the conclusion and future work is discussed in Section VI.

II. RELATED WORK

The problem of detection and classification of objects in real- time started after major developments in the CCTV field, processing hardware, and deep learning models. Very little work has been done in this field before and most of the previous effort was related to concealed weapon detection (CWD).
Starting with concealed weapon detection (CWD), before its use in weapon detection, it was used for luggage control and other security purposes at airports and was based on imaging techniques like millimeter-wave and infrared imag- ing [8]. Sheen *et al.* suggested CWD method based on a three- dimensional millimeter (mm) wave imaging method, for detecting hidden weapons at airports and other safe locations in the body [13]. Z. Xue *et al.* suggested a CWD technique based on a fusion-based technique of multi-scale decomposi- tion, which combines color visual picture with infrared (IR) picture integration [14]. R. Blum *et al.* suggested a CWD method based on the inclusion of visual picture and IR or mm wave picture using a multi-resolution mosaic technique to highlight the hidden weapon of the target picture [15].
E. M. Upadhyay et. al. suggested a CWD technique using image fusion. They used IR image and visual fusion to detect hidden weapons in a situation where the image of the scene was present over and under exposed area. Their methodology was to apply a homomorphic filter captured at distinct exposure conditions to visual and IR pictures [16].
Current techniques attain high precision by using various combinations of extractors and detectors, either by using easy intensity descriptors, boundary detection, and pattern matching [9] or by using more complicated techniques such as cascade classifiers with boosting.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

CWD though had worked for some sort of cases but it had many limitations. These systems were based on metal detection; non-metallic guns cannot be detected. They were costly to use in many locations because they need to be coupled with X-ray scanners and conveyor belts and responds to all metallic objects, so were not accurate. Economic cost and health risks limited the practical implementation of such methods. Furthermore, video-based firearm detection was a preventive measure for acoustic detection of gunshot and can be combined with it for implementation [17], [18].

The idea of automated image processing for public security purposes in many fields has been well recognized and studied. CCTV was the ultimate need for this kind of work to progress. CCTV was first used back in 1946 in Germany and at that time, these cameras were installed to observe the launch of a rocket named V2 [19]. Although it had been used earlier, major improvements happened in the last two decades. With the advancement in CCTV technology, visual object recog- nition and detection for surveillance, control, and security were performed. In 1973, Charge-Coupled Device (CCD) was developed, which made the deployment of surveillance cameras possible by 1980 [20]. If we go a bit forward in time, a company named Axis Communication developed the first-ever network camera, which enabled the transformation of surveillance cameras from analog to digital [20]. This transformation of analog to digital video made it possible for everyone to apply image processing, machine learning, and computer vision techniques on videos recorded from surveillance cameras. In 2003, Royal Palm Middle School in Phoenix used facial recognition for the first time for tracking missing children.

Several object detection algorithms were proposed in the field of computer vision to make surveillance system better. Object detection algorithms were used in several sectors like anomaly detection, deterrence, human detection, and traffic monitoring [21]. R. Chellapa et.al. discussed briefly object tracking and detection in surveillance cameras [22]. The authors had explained the tracking of an object using mul- tiple surveillance cameras. Another author addressed tech- niques for detecting objects that come into contact with another object and are occluded. They also wrote regarding the segmentation of mean fluctuations. They outlined how mean segmentation of shifts can help detect objects. They used a Bayesian Kalman filter with a simplified Gaussian blend (BKF-SGM) algorithm to track the detected object [23]. J.S Marques proposed distinct techniques for evaluating the efficiency of distinct algorithms for object recognition [24]. B. Triggs et.al. described histogram oriented gradient (HOG). HOG became a novel architecture for feature extrac- tion. It was used mostly in applications involved in human detection [25]. In 2005, the sliding window technique was proposed for the recognition of number plates [26]. They had used a sliding window for the purpose of segmentation and a neural network for character recognition on the number plate. As described above, objection detection for the computer vision tasks was used for some applications with big objects to identity like a person, transport or traffic monitoring, etc. Literature review on weapon detection left me with the opinion that regardless of many object detection algorithms, the algorithms proposed for weapon detection are very few. At last, the idea of firearm detection using the images and videos was proposed and false alarms were reduced by clas- sifying neural networks with region-based descriptors and determining region of interest (ROI) using the sliding window technique and then trained the neural network classifier with image pixels [27].

With the development in CCTV's, object detection for different computer vision problems for real-time were per- formed and the idea to detect firearms were introduced first by L. Ward *et al.* in 2007 [28] and a surveillance system was also implemented by them a year later in 2008 [29]. In the aforementioned work, writers created an accurate pistol detection model for RGB pictures. However, in the same scene, their method did not detect various pistols [11], [10], [29]. The approach used comprises of first removing non- related items from the segmented picture using the K-mean clustering algorithm and then applying the SURF (Speed up Robust Features) method to detect points of interest. Darker gave the concept of SIFT based weapon detection algorithm and for ROI estimation, used the motion segmen- tation method [30]. SIFT algorithm is prone to false alarms, so for estimating ROI, authors used motion segmentation rather than using SIFT on complete image. When ROI was determined, then SIFT was applied to detect firearms in their case.

Different approaches then used for weapon detection using sliding window and region proposal algorithms. HOG (His- togram of oriented Gradient) models were used to predict the objects in the frame. HOG significant work used low-level features, discriminative learning, and pictorial structure along with SVM [25], [31], [32]. These algorithms were

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

slow for real-time scenarios with 14s per image. Although these classifiers gave good accuracies, the slowness of the sliding window method was a big problem, especially for the real- time implementation purpose.

This work focuses on the state of the art deep learning net- work rather SIFT and HOG features which use handcrafted rules for feature extraction, selection, and detection in real- time visual scenario using CCTV cameras. X. Zhang *et al.* concluded an important finding that helped my work. They concluded that the automatic feature representation gave improved results rather than manual features [33]. Not only the learned features were better in performance, they also had learned the deep representation of the data and reduced a lot of manual work, and saved time and energy.

Rohith Vajhala *et al.* proposed the technique of knife and gun detection in surveillance systems. They had used HOG as a feature extractor along with backpropagation of artificial neural networks for classification purposes. The detection was performed using different scenarios, first weapon only and then using HOG and background subtraction methods for human before the desired object and claimed to have an accuracy of 83% [34]. The aforementioned work uses the CNN along with non-linearity of ReLu, convolutional neural layer, fully connected layer, and dropout layer of CNN to reach a result for detection with multiple classes and implemented their work using the Tensor flow open-source platform. Their system achieved a test accuracy of 90.2 % for their dataset [35]. MichałGrega *et al.* proposed knives and firearm detection in CCTV images. They had applied MPEG-7 and principle component analysis along with the sliding window approach, which made their work slower for real-time scenarios, although they claimed to achieve good accuracy on their test dataset. [5].

Verma *et al.* had also used the deep learning technique to detect weapons and used the Faster RCNN model. The work was performed on imfdb, which in my opinion is not suitable to train a model for real-time case. They claimed to have an accuracy of 93.1% on that dataset but in the case of weapon detection, only achieving higher accuracy is not enough, and precision and recall must the considered [36]. Siham Tabik *et al.* work was very much related to the real-time scenario. They used Faster RCNN to detect weapons in real-time using sliding window and region pro- posal methods. Best results were obtained by using the region proposal technique. The sliding window was also very time- consuming and took 14 s/image, on the other hand, the region proposal method processed the image in 140ms with 7 fps [37]. They trained the network on Faster RCNN using only one class focusing on reducing the false positive. Recent past objection detection work with the application to firearms was proposed in 2019, where a group of researchers, Javed Iqbal *et al.* proposed orientation aware detection of the object. This system is more suitable for long and thin objects like rifles etc. The predicted bounding box in their case was aligned with the object and had the less unnecessary area to deal with. Images of very high quality were used for training and testing purposes, which may make it less suitable for real-time scenarios [38]. Jose Luis Salazar Gonz'alez *et al.* work was very much related to achieve real-time results. They did immense experimentation using different datasets and trained Faster–RCNN using Feature Pyramid Network with Resnet50 and improves the previous state of the art by
3.91 % [39].

III. METHODOLOGY

Deep learning is a branch of machine learning inspired by the functionality and structure of the human brain also called an artificial neural network. The methodology adopted in this work features the state of art deep learning, especially the convolutional neural networks due to their exceptional per- formance in this field. [40]. The aforementioned techniques are used for both the classification as well as localizing the specific object in a frame so both the object classification and detection algorithms were used and because our object is small with other object in background so after experi- mentation we found the best algorithm for our case. Sliding window/classification and region proposal/object detection algorithms were used, and these techniques will be discussed later in this section.

We had started by doing the classification using different deep learning models and achieved good precision but for the real-time scenarios, the low frame per seconds of classifica- tion models were the real issue in implementation. Oxford VGG [41], [42], Google Inceptionv3 [43] and Inception- Resnetv2 [44], [45] were trained using the aforementioned approach.

To achieve high precision, increase number of frame per seconds and improve localization, we moved to the object detection and region proposal methods. The differ- ent state of the art deep learning models for object detec- tion were

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

used and the results were compared in terms of precision, speed, and standard metric of F1 score. State of the art deep learning based SSDMobileNetv1 [46]–[48], YOLOv3 [49], FasterRCNN-InceptionResnetv2 [50]–[52], and YOLOv4 [53] were trained and tested.

Different datasets were made keeping in mind the classifi- cation and detection problem as both have a separate require- ment for performing the tasks to achieve high accuracy, mean average precision as well as frame per second for the real- time implementation. To understand object classification and detection let us first briefly understand object recognition as both the aforementioned types come under the umbrella of this and combined classification and localization make detection possible for any kind of detection problem giving class name as well as the region where our desired object is in the frame.

OBJECT RECOGNITION

As the name suggests, it is the process of predicting the real class or category of an image to which it belongs by making probability high only for that particular class. CNN's are used to efficiently perform this process. Many state of the art Classification and Detection algorithms uses CNN as a backend to perform their tasks.

Fig. 1 depicts that classification and localization come under the category of recognition and combined classification and localization is performed to do object detection. Let us have a brief overview of the object classification, localization, and detection.

IMAGE CLASSIFICATION

The classification model takes an image and slide the ker- nel/filter over the whole image to get the feature maps. From

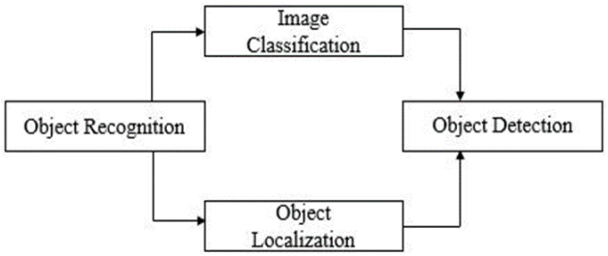

FIGURE 1. Object Recognition to detection Hierarchy.

the feature extracted, it then predicts the label based on the probability.

OBJECT LOCALIZATION

This method outputs the actual location of an object in an image by giving the associated height and width along with its coordinates.

OBJECT DETECTION

This task uses the properties of the aforementioned algo- rithms. The detection algorithm tells us the bounding box having x and y coordinates with associated width and height along with the class label. Non-max suppression is used to output the box with our desired threshold [54]. This process gives the following results altogether:

Bounding Box

Probability

In past object detection was very limited because of less data and low processing power of computers but with the passage of time the computing power of computers increased and world moved from CPU's to Graphic Processing Units (GPU). GPU's were firstly made for increasing the graphic quality of the systems and for gaming but later GPUs were used extensively for deep learning. In ImageNet, competi- tions started and contained about 1000 classes [55]. This was the evolution of machine learning and deep learning. In the beginning, the models were not very deep, means there were not many layers as they are now in an algorithm. Because of the aforementioned developments, in 2012 A.Krizhevsky presented a model called Alex Net trained on ImageNet and got the first position in that competition.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

This was the begin- ning of object detection in deep learning. It gave a way to researchers and then every year the algorithms and models keep on coming. All these algorithms contain layers that work on the principle of the convolutional neural network (CNN).

CLASSIFICATION AND DETECTION APPROACH

There are many ways to generate region proposals, but the simplest way of generating them is by using the sliding win- dow approach. The sliding window method is slow because filter slides over the entire frame and has limitations, which were tackled by the region proposal approach, so we have the following two approaches used in our work for both classification and detection models are:
Sliding window/Classification Models
Region proposal/Object Detection Models

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. This is because models are usually trained on a particular range. The outcomes are in tens of thousands (10^4) of picture spots being classified [56]. The sliding window method is com- putationally very costly because of the search with various aspect ratios and especially for each pixel of an image if the stride or step value is less.

REGION PROPOSAL/OBJECT DETECTION MODELS

This technique takes an image as the bounding boxes of input and output proposals related to all areas in a picture most probable to be the object. These regional proposals may be noisy; coinciding not containing the object flawlessly, but there is a proposal among these region proposals related to the original target object. 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 [57]. Region-based CNNs (R-CNN) was the first detection model to introduce CNNs under this approach [58]. The selective search method of this approach produces 2000 boxes having maximum likelihood.
Selective search is a widely used proposal generation method because it is very fast having a good recall value. It is dependent on the hierarchical calculation of desired areas established on the compatibility of color, texture, size, and shape [59].
Yolo series is among the state of the art object detection models. Unlike the other region proposal-based methods it divides the input image into an SxS grid and then simulta- neously predicts the probability and bounding boxes for an object with a center falling into a grid cell [49], [53].

TRAINING MECHANISM

Fig. 2 describes the general methodology used in training and optimization. It starts with defining a problem, finding the required dataset, applying pre-processing methods, and then finally training and evaluating the dataset. If the evaluation is correct then we save those weights as a classifier but if it's incorrect then comes the process of backpropagation algorithm along with the gradient descent algorithm [60]. In backpropagation, weights are optimized by subtracting the partial derivative of cost function J(O) with a multiplier of the learning rate alpha α from the old or previous weight value. Gradient descent is the main weight optimization algorithm. It is used as a base in all optimizers used for the modeling and it helps in converging the model and reaching the minima where we get the best and desired weights values.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

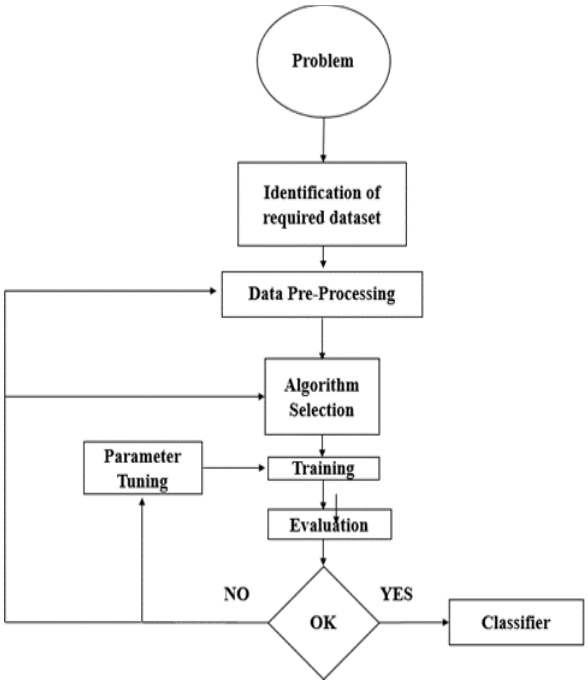

FIGURE 2. Training and Optimization Flow Diagram.

CONFUSION OBJECT INCLUSION

We have formulated the problem to reduce the number of false positives and negatives by adding relevant confu- sion object. The weapon category includes all the handheld weapons such as, pistol, revolver, shotgun and other than weapon includes the objects that can most be confused with pistol classes e.g. mobile, metal detector, selfie stick, purse, etc.

By understanding the differences between classification and detection algorithms, sliding window, and region pro- posal methods, let's now look at the algorithms used for both approaches.

CLASSIFIERS AND OBJECT DETECTORS

The classifiers used under the sliding window approach:

VGG16

InceptionV3

Inception ResnetV2

The object detectors used for real-time detection are:

SSD MobilNetV1

YoloV3

Faster RCNN-Inception ResNetV2

YoloV4

Three databases named database1, database2, and database3 were created one by one after experimentation on different algorithms with diverse images, first for classifi- cation and then for object detection. Although the results obtained from the classification algorithms were not bad but the frames per second were very slow for real-time implementation. Detail for each database will be discussed in the next section.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

DATASET CONSTRUCTION, ANNOTATION AND PRE-PROCESSING (D-CAP)

Data plays a key role in the development of any deep learning model as the model learns and extract feature from it. For a real-time model to detect weapons with minimized pro- cessing time and high precision, the importance of accurate and relevant data increases further as all other processes are dependent on it.

When we study the stats and goes through almost 50-60 videos of robbery on available online resources, we come to know that 95 percent of the videos have revolver or pistol as a weapon, so we focused on binary classification with pistol and revolver to be in a single class called pistol. Besides, to make the system more precise and to reduce the false positive and false negative values we added objects that can be confused with a weapon such as a wallet, cell phone, metal detector etc and put them in a separate class named it as not pistol.

Let's now discuss the datasets used in our case because, in a supervised learning case, the network learns the repre- sentation of the input data with given true answers, so the data must be clean, preprocessed, and properly annotated to make the network learn and predict better.

DATASET CONSTRUCTION AND SELECTION

The task of dataset construction and collection was very important and tough as well because there was no benchmark dataset available for this. Dataset for real-time detection was collected and constructed in different phases and data was collected from the internet, extracted from YouTube CCTV videos, through GitHub repositories, data by the University of Granada research group, and internet movie firearm database imfdb.org.

WEAPON DATASET CLASSES

The weapon dataset for real-time weapon detection is divided into the following two classes:

Pistol

Not-Pistol

WEAPON DATASET CATEGORIES FOR PISTOL CLASS

Dataset for this class includes weapon samples of the follow- ing categories:

Pistol

Revolver

Other shot handheld weapons

FIGURE 3. Dataset samples for pistol Class- Top left to bottom right [a-d]:
(a) CCTV image (b) Medium Resolution Image (c) Image with Dark background and Low Resolution, (d) Filtered Image.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

REASON OF CHOOSING DATA CATEGORIES OF PISTOL CLASS

The reason we choose pistol and revolver in the pistol class is because of our study and analysis after watching many robberies and shooting incident CCTV videos. We concluded that almost 95% of the weapon used in those cases were either pistol or revolver. Fig. 3 shows some sample images for real- time from the collected dataset of the pistol class.

WEAPON DATASET CATEGORIES FOR NOT-PISTOL CLASS

Datasets for this class include objects that can most likely be confused with pistol class objects. Following are some samples categories for the not pistol class:

Wallet
Metal Detector
Cell phone
Selfie stick

REASON OF CHOOSING DATA CATEGORIES OF NOT-PISTOL CLASS

We introduced this relevant confusion object concept because these are the objects that can mostly be confused with our desired weapon object, so predicting them correctly results in reducing the number of false positives and false negatives, hence increasing overall accuracy and precision.

Some previously done work though had objects other than weapons used for the background or class other than a weapon but they had samples like cars, airplanes, cats, etc and there are very fewer chances for them to be confused with our desired weapon, which is very small as compared to them. As our desired objects of pistol class are small so there are lot of chances for them to be confused with some other objects having some features like that. Fig. 4 shows some sample images from the collected dataset of the not pistol class which helps in reducing false positives and negatives.

FIGURE 4. Dataset samples for not-pistol Class-Top left to bottom right [a-d]: (a) Cell Phone (b) Metal Detector (c) Selfie Stick (d) Wallet.

DATASETS FOR REAL-TIME DETECTION

This work deals with the binary classification for a real-time scenario so two classes were made and pistol and revolver images were included in pistol class and not pistol class include confusion classes like mobile phone, metal detector, selfie stick, wallet, purse, etc. For the pistol and not pistol classes, we have made three datasets, which are explained below.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

DATASET 1

This was the initial dataset used while starting this work. In this dataset, we had 1732 images in total, with 750 images in pistol class and 950 in not pistol class. Dataset was divided by the separation criteria described in Table 1 of train and test. Images were collected from online sources and imfdb database and sliding window classification algorithms were trained and tested on it.

DATASET 2

This was the second dataset made for the real-time scenario. This dataset contains 5254 images and classification, as well as object detection algorithms, were trained on this dataset to meet the task. Images were extracted for real-time scenario with the desired object in hand from online, sources, imfdb database, and ImageNet website. Dataset was divided by the separation criteria of test and train explained in Table 1.

DATASET 3

This was the third dataset constructed for the real-time scenario and object detection algorithms were performed on it. This database was made by enhancing dataset 2 by overcom

TABLE 1. Data Distribution.

Sr.No.	Category	Total Data	Training Data	Test Data	Split Size
1.	Dataset 1	1732	1251	260	15%
2.	Dataset 2	5254	3797	784	15%
3.	Dataset 3	8327	7328	999	12%

ing the shortcomings and problems of the previous dataset. The need for this dataset arises because though we got a reasonable accuracy from classification models but the frames per second were very few. To detect images from CCTV videos, similar kinds of training data must be included so we made our own dataset to tackle this issue.

This dataset contains 8327 images divided into the pis- tol and not pistol class. In this case, a related confusion data concept was introduced to reduce false positives and false negatives in real-time detection. Dataset images were extracted from several online sources, from CCTV videos for the particular robbery scenario, made our own dataset with a weapon in hand for the diverse scenario, did data augmentation, and finally, it was separated for test and train case.

DATA DISTRIBUTION

Each of the aforementioned datasets are divided into the following categories mentioned in Table 1 with split size defining the separation percentage of the total data into test and train.

DATA PRE-PROCESSING AND ANNOTATION

Many things affect the performance of a Machine Learning (ML) model for a specified job. First, the representation and quality of the data are essential. If there are many irrelevant and redundant data existing or noisy and unreliable data, then it is harder to discover representation during the train- ing stage. Data preparation and filtering steps take significant processing time in ML issues [61]. The pre-processing process involves data cleaning, standardization, processing, extraction and choice of features, etc. The final training dataset is the result of pre-processing processes applied to the collected dataset.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

Pre-processing is necessary for better training of a model, so the first step is to make the same size or resolution of the dataset. The next step is to apply the mean normalization. The third step is making bounding boxes on these images, which is also called annotation, localization, or labeling. In data, labeling a bounding box is made on each image. The value x, y coordinates, and width, height of the labeled object was stored in xml, csv or txt format. Following are the four main steps of data preprocessing:

Image scaling

Data-augmentation

FIGURE 5. Image Augmentation and Scaling.

FIGURE 6. Image Annotation and Labelling.

(b) Equalized Filter Result

(c) Gray Scale Filter Result

Image labeling

Image Filtering using OpenCV

RGB to Grayscale

DOI: 10.48175/IJARSCT-25798

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

Equalized
Clahe
Fig. 5, 6 and 7 shows the results after applying the afore-mentioned pre-processing techniques

EXPERIMENTS, RESULTS AND ANALYSIS

We have detected weapons in real-time CCTV streams in low resolution, dark light with real-time frame per second. Most of the work done before was on detecting images and videos of high quality and because those models were trained on high-quality datasets, it is not possible to then detect an object of low resolution in real-time. The results are analyzed after training and testing models on datasets mentioned in Table 1. As described in the methodology section the results for different approaches are evaluated. Our main problem state- ment is of real-time detection because 97% of weapon used in robbery cases were pistol or revolver, so different dataset results have been evaluated here for sliding window and region proposal approach.

The performance of these models was analyzed by comparing them in terms of the standard metrics of F1-score and frame per seconds along with mean average precision (mAP) for the best performed model and these terms are calculated by using the below equation 1,2 and 3. F1 score is ratio of the precision and recall functions.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives + False Positives}}$$

(a) Actual Image

(b) Equalized Filter Result

(c) Gray Scale Filter Result

FIGURE 7. Image Filtration using OpenCV Filters- (a) Original Image
(b) Equalized Filter Result (c) Gray Scale Filter Result (d) Clahe Filter Result.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recal}}$$

DATASET-1 EXPERIMENTATION AND RESULTS

Dataset 1 contains 1732 images distributed between two classes of pistol and not-pistol with 750 and 982 images in each class respectively. Experimentation on dataset-1 has

TABLE 2. Sliding Window Results Comparison Dataset-1.

Sr.No	Algorithms	Precision	Recall	F1-score
1	VGG16	71%	66.66%	69.09%
2	Inceptionv3	74.11%	96.18%	83.71%
3	Inception-ResNetV2	79.24%	89.54%	84.07%

been performed using the sliding window/classification mod- els of VGG16, Inceptionv3 and InceptionResNetv2.

After experimentation, we have analyzed that the results obtained are not good because most of the images of this dataset have white or the same kind of background which lead to a point where the model also starts learning the background as its region of interest (ROI) and in real-time background varies so a new dataset was required to train and test the model on images with diverse cases and background. Table 2 shows the results for the aforementioned models using this dataset giving precision, recall, and F1-score.

DATASET-2 EXPERIMENTATION AND RESULTS

This dataset contains two classes of pistol and not- pistol with 3000 and 2254 images in each class respec- tively. Table 3 shows results based on it. Experimenta- tion on dataset-2 has been performed using the sliding window/classification models of VGG16, Inceptionv3, and InceptionResNetv2.

Experimentation results show that though we get a reason- able accuracy from classification models using this dataset but the frames per second were very few and which was a big problem in making a real-time weapon detector. Among these classification models, InceptionResnetV2 performed best and achieves the best results. Table 3 shows the results under the sliding window methods using dataset 2 and Fig. 8, 9, and 10 shows the accuracy, loss, and confusion matrix respectively for the best classification model under the sliding window approach.

DATASET-3 EXPERIMENTATION AND RESULTS

After experimentation on the previous two datasets and not finding satisfactory results for the real-time case a new dataset was made. Images were collected from robbery videos, our own dataset images holding a weapon in different scenar- ios, images with a dark background and low resolution, and images extracted from applying different OpenCV fil- ters are added to make real-time detection possible. A total of 8327 images are used in this case. Following object detec- tion models were trained and evaluated using this dataset:

SSD MobilNetV1

YoloV3

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

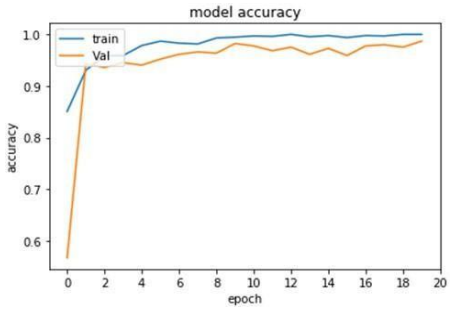

FIGURE 8. Best sliding window model accuracy graph: InceptionResNetv2.

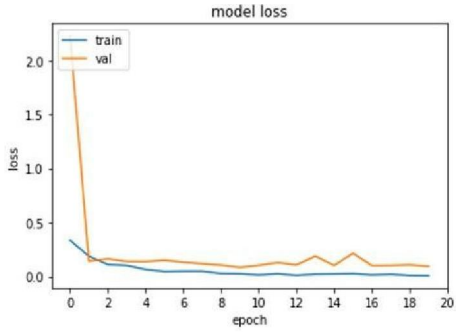

FIGURE 9. Best sliding window model Loss graph: InceptionResNetv2.

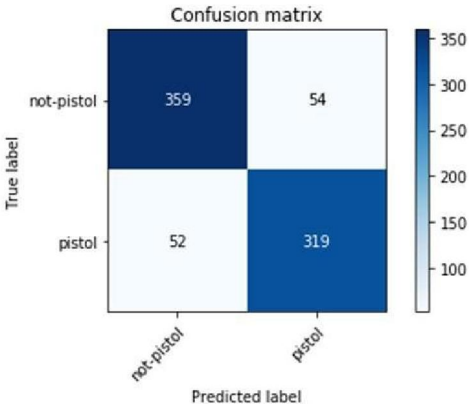

FIGURE 10. Best sliding window model Confusion Matrix: InceptionResNetv2.

Faster RCNN-Inception ResNetV2

YoloV4

Each model had its pros and cons. SSD-MobileNet is good in terms of processing frames per second. FasterRCNN-InceptionResNetv2 has good precision and recall but not processing speed. Yolo family has a series of models. It has a different approach for the detection purpose. Unlike the other region proposal based methods, it divides the input image into an SxS grid and then simultaneously predicts the probability and bounding boxes for an object with the center falling into a grid cell. We have trained the latest state of the art Yolov3 and Yolov4 on our own weapon dataset 3 for real-time detection

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

TABLE 3. Sliding Window Results Comparison Dataset-2.

Sr.No	Algorithms	Precision	Reca	F1-score
1	VGG16	80.00%	83.4	81.69%
2	Inceptionv3	84.36%	84.3	84.36%
3	Inception-ResNetV2	85.52%	85.9	85.74%

TABLE 4. Region Proposal/Object Detection Models-Dataset-3.

Sr. No	Models	IoU Threshold=50%		
		Precision	Recall	F1-score
1	SSD-MobileNet-v1	62.79%	60.23%	59%
2	Yolov3	85.86%	87.34%	86%
3	FasterRcnn-InceptionResNetV2	86.38%	89.25%	87%
4	Yolov4	93%	88	91%

and best results were obtained through YOLOv4 in terms of both processing speed and precision. Table 4 below shows the results for the aforementioned detection models for this dataset at a standard threshold score of 50%.

Yolov4 performs best among all the models of both the sliding window and region proposal approach. Performance graph for yolov4 in terms of loss and mean average precision (mAP) on a validation dataset is shown in Fig. 11. We can see that how smooth is the model loss curve and how precisely it converges to the best level giving a very good loss score of 1.062 and a mean average precision of 91.73%. The mean average precision is the mean of the average precision values for all the relevant classes. The associated values of average precision (AP) for pistol and not-pistol class for the calcula- tion of mean average precision value is given in Table 5.

The mean average precision value is calculated for the yolov4 model as it performs best in all scenario and accurately detected the desired object even when the object has a very small presense in the frame and there were lots of other objects in the background as well.

IV. ANALYSIS AND DISCUSSION

Table 2, 3, and 4 above shows the comparison between the classification and object detection models using standard metrics of precision, recall, and F1-score for evaluation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

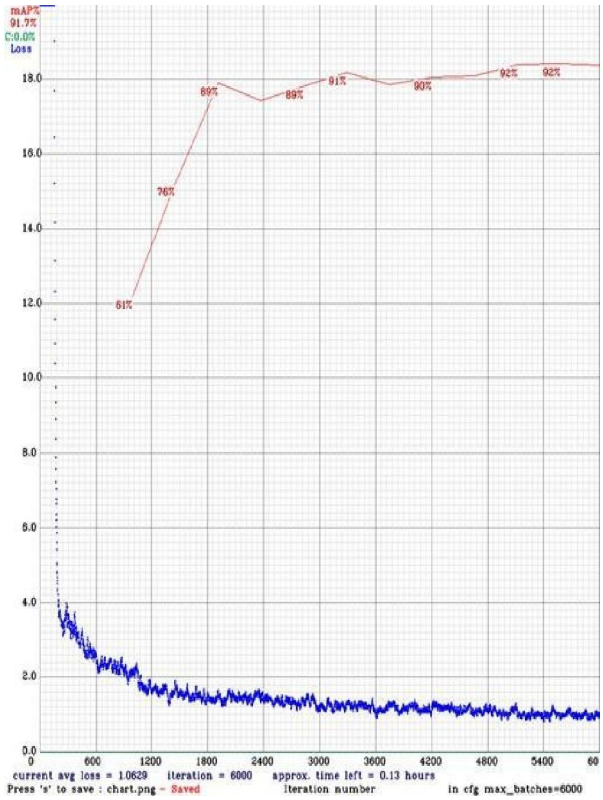

FIGURE 11. Best Object Detection Model-Yolov4: loss vs mAP

TABLE 5. Best Performed Model Yolov4: mAP Calculation

IoU_ Threshold	Average Precision Pistol (AP_P)	Average Precision Pistol (AP_{NP})	
0.35	96.91%	92.41%	94.66%
0.50	94.00%	89.47%	91.73%
0.75	60.17%	67.41%	63.79%

Some classification models showed good results but they were not suitable for a real-time scenario, were slow, not much accurate, and fast as compared to the object detection models as they performs very well and achieved high precision and recall.

The reason why some classification models have a good F1-score is the training and evaluation on initial datasets we made when starting this work, but after experimenta- tion, we come to know that these models are not suitable for real-time scenarios having the background objects.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

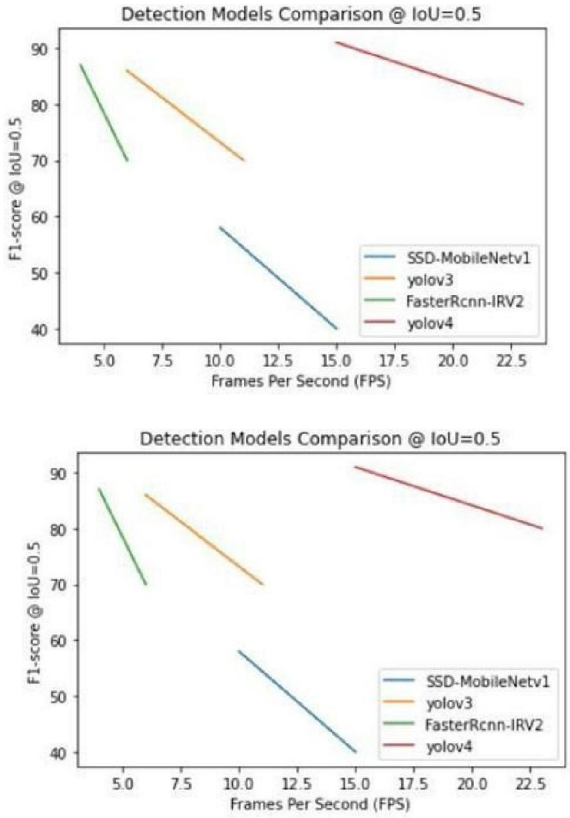

FIGURE 12. Object Detection models Performance/Comparison Graph.

Object detection models performed well for the real- time scenario and performance comparison in terms of speed and F1-score between the detection models can be seen from Fig. 12. Inference results are obtained using the NVIDIA RTX 2080ti for each model.

The standard metrics of mean average precision (mAP), recall and F1-score are calculate and all the models have been compared at a benchmark IoU threshold of 0.50 or 50%.

Yolov4 performs best amongst all models with a mean average precision and F1-score of 91.73% and 91% respectively with detection confidence of 99% in the majority of cases.

Comparison in terms of test accuracy vs F1-score for the best-performed models of both classifica- tion and detection approaches is shown in the Fig.13. Accuracy and F1-score for VGG, Inceptionv3, InceptionResNetv2, SSDMobileNet, FasterRCNN- InceptionResNetv2, Yolov3 and yolov4 are 78.20%, 85.20%, 92.20%, 79%, 96%, 94%, 99% and 81.69%,

84.36%, 85.74%, 59%, 87%, 86% and 91% respectively.

Fig. 14-19 shows the inference or detection results of our model for pistol and not pistol class on images, videos, and real-time CCTV streams.

Hyperparameters used in training the best-performed detector Yolov4 can be observed from Table 6.

It is very hard to do a comparison with studies conducted previously on this subject because each study has its own dataset, models and metrics used to evaluate performance. It should also be noticed achieve realtime detection, we also need to have a realtime dataset for traning because with high quality training images we cannot achieve results in realtime. Each study also has different testing conditions, either just on images, videos or on images with high quality but our approach from start was to achieve realtime results. In some studies, the performanece metric used is accuray, others have precisoin or mean average precision (mAP) but mostly mAP is used as standard so we have given comparison results

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

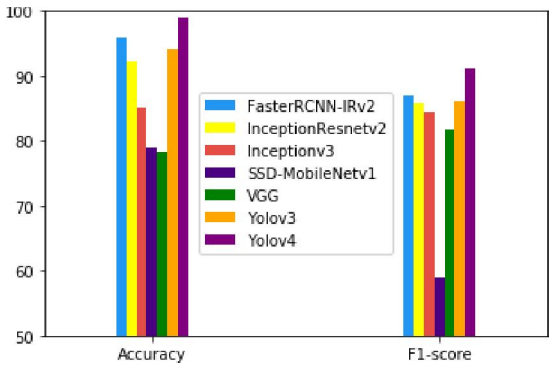

FIGURE 13. Best performed models comparison: Accuracy vs F1-score.

TABLE 6. Yolov4 Hyper Parameters.

Sr. No.	Hyper Parameter	Value
1.	Learning rate	0.001
2.	Optimizer	SGD
3.	Decay	0.0005
4.	Momentum	0.949
5.	Activation Function	Mish
6.	Batch Size	64
7.	Max Batches / Iteration	6000

TABLE 7. Comparison with some existing studies.

Study	Algorithm	Precis	mAP
Javed Iqbal et al. 2019	OOAD	N/A	85.40%
Roberto Olmos et al. 2017	Faster RCNN	84.20%	N/A
Jose Luis Salazar Gonz´alez et al. 2020	Faster RCNN using FPN	88.23%	N/A
Ours	Yolov4	93 %	91.73%

in terms of mAP and precisoin at a standard iou threshold of 50%, which ever was available.

DOI: 10.48175/IJARSCT-25798

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

DETECTION RESULTS - PISTOL CLASS WITHOUT BACKGROUND
See Figure 14.
DETECTION RESULTS - PISTOL CLASS WITH BACKGROUND
See Figure 15.

FIGURE 14. Detection Results- Only weapon in the whole frame without any background at different angles, brightness, sharpness, and quality.

FIGURE 15. Detection Results-Top left to bottom right (a-i): (a) Image with front and side view, (b) Image vertical view (c) Image with Dark background and Low Resolution fully tilted side view, (d) Low brightness image side view slightly tilted (e) Image with the back view (f) Full front view (g) Small CCTV object (h) Very small object with side view (i) Image with full side view.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

G. DETECTION RESULTS - PISTOL CLASS IN VIDEOS
See Figure 16.
H. DETECTION RESULTS - PISTOL CLASS IN REALTIME CCTV STREAMS
See Figure 17.
I. DETECTION RESULTS – NOT PISTOL CLASS
See Figure 18.
J. MISDETECTIONS
See Figure 19.

FIGURE 16. Detection Results-Top left to bottom right (a-f) - video 1 inference (a-c), video 2 inference (d-f): (a) Small object-side view tilted, (b) Small object with side view (c) Small object front view (d) side view (e) Top view double object (f) Small object with front and side view.

FIGURE 17. Detection Results-Top left to bottom right (a-i)-cctv stream1(a-c),cctv stream2(d-f), cctv stream3(g-i): (a) Small object in Low resolution (b) Tilted Object (c) Low Resolution vertical object, (d) Day light side view with slightly tilted (e) Day light side view (f) Day light side view flipped (g) Small object medium resolution (h) vertical view (i) side view.

DOI: 10.48175/IJARSCT-25798

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

V. CONCLUSION AND FUTURE WORK

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 secu- rity, 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. To achieve high precision and recall we con- structed a new training database for the real-time scenario, then trained, and evaluated it on the latest state-of-the-art deep learning models using two approaches, i.e. sliding window/classification and region proposal/object detection.

FIGURE 18 Detection Results- Top left to bottom right (a-d): (a) Cell phone (b) Metal detector(c) Wallet (d) Selfie stick.

FIGURE 19. Misdetections: False positives and Negatives.

Different algorithms were investigated to get good precision and recall.

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. It gave 91.73% mean average precision (mAP) 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.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 9, April 2025
Impact Factor: 7.67

Looking at the results, we got the highest mean average precision (mAP) F1- score as compared to the research done before for real-time scenarios.

The future work includes reducing the false positives and negatives even more as there is still a need for improvement. We might also try to increase the number of classes or objects in the future but the priority is to further improve precision and recall.

ACKNOWLEDGMENT

The authors wish to extend gratitude to Mr. Rehan Mushtaq of Ingenious Zone who provided assistance as an industrial partner in making this work possible.

REFERENCES

[1] (2019). Christchurch Mosque Shootzings. Accessed: Jul. 10, 2019. [Online]. Available: https://en.wikipedia.org/wiki/Christchurch_mosque_ shootings

[2] (2019). Global Study on Homicide. Accessed: Jul. 10, 2019. [Online]. Available: https://www.unodc.org/unodc/en/data-and-analysis/global- study-on-homicide.html

[3] W. Deisman, ‘‘CCTV: Literature review and bibliography,’’ in Research and Evaluation Branch, Community, Contract and Aboriginal Policing Services Directorate. Ottawa, ON, Canada: Royal Canadian Mounted, 2003.

[4] J. Ratcliffe, ‘‘Video surveillance of public places,’’ US Dept. Justice, Office Community Oriented Policing Services, Washington, DC, USA, Tech. Rep. 4, 2006.

[5] M. Grega, A. Matiolałski, P. Guzik, and M. Leszczuk, ‘‘Automated detec- tion of firearms and knives in a CCTV image,’’ Sensors, vol. 16, no. 1,
p. 47, Jan. 2016.

[6] TechCrunch. (2019). China’s CCTV Surveillance Network Took Just 7 Minutes to Capture BBC Reporter. Accessed: Jul. 15, 2019. [Online]. Available: https://techcrunch.com/2017/12/13/china-cctv-bbc-reporter/