International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 6, Nov-Dec-2024, ISSN (Online): 2395-566X

Weapon Detection Using Yolo

Assistant Professor Ms. Monika, Nikhil Tiwari
Department of Computer Science & Engineering,
HMR Institute of Technology & Management, GGSIPU, Delhi

Abstract- In light of the increasing gun violence incidents worldwide, there is a pressing need for automated visual surveillance systems capable of detecting handguns. This paper presents a method for real-time handgun detection in video streams using the YOLO algorithm, comparing its performance in terms of false positives and false negatives against the Faster CNN algorithm. To enhance detection accuracy, we compiled a custom dataset featuring handguns from various angles and merged it with the Roboflow dataset. The YOLO model was trained on this combined dataset and validated using four different videos. The results indicate that YOLO effectively detects handguns across diverse scenes, demonstrating superior speed and comparable accuracy to Faster CNN, making it suitable for real-time applications.

Index Terms- Machine Leaning, Deep Learning, YOLO, Tensorflow, Efficient v1

I. INTRODUCTION

The rise in gun-related crimes poses significant challenges for law enforcement agencies globally, especially in regions where gun ownership is prevalent. Reports indicate a troubling trend in gun violence, with the United States experiencing hundreds of thousands of incidents in recent years. Similarly, Malaysia ranks among the top ten countries with the highest gun violence rates in Southeast Asia.

To combat this issue, automated surveillance systems capable of timely detection and monitoring are essential. Recent advancements in machine learning and deep learning have revolutionized object detection and image segmentation, with the YOLO (You Only Look Once) algorithm outperforming traditional methods in identifying objects within images.

This paper introduces an automated handgun detection system employing the YOLO-v1 algorithm, contrasting its results with those obtained from Faster R-CNN. Detecting firearms in various environments presents challenges, particularly due to occlusion and the need for real-time processing capabilities.

II. LITERATURE SURVEY

Mostly, gun detection research specially emphases on hidden weapon detection and knife detection. Hidden weapon detection is based some techniques of imaging like millimetre wave imaging, infrared imaging used in airport for luggage (containing gun and knife) control applications.

The research work in [5] proposed a visual gun detection system based on Harris interest point and SIFT. They used color based segmentation to select dissimilar object from an image by deploying K-mean cluster algorithm. In [6], researchers used 3D millimetre wave imaging technique to detect the weapon concealed in the body and other hidden location.

In another work [6], researchers deploy a real time gun classifier that detect and classify guns. They also used imageNet dataset for training their model and acquired a mAP of 89% using overfeat-3 algorithm. The research work [7], is based on terahertz human dataset used in deep learning to detect the concealed weapon. The achieved a competitive accuracy compared to other concealed weapon detection system.

[8], used numerous networks like YOLOv2, Sliding window-based CNN, region-based fully convolutional networks (R-FCN), faster region-based CNNS (F-RCNNS) with transfer learning for image classification and detection problems. They showed empirically that fine-tuned CNN features give greater performance than conventional algorithms. The image dataset used in training the model is based on X-rays images.

In [4], the researchers used faster RCNN with VGG-16 based classifier for detecting guns in in YouTube videos and achieved a maximum mean average precision (mAP) of 84.21%. They used ImageNet dataset for training a modelling a handguns detection system.

III. METHODOLOGY

1. Dataset

We have taken the dataset containing guns with different position and orientation from ImageNet dataset. YOLO algorithm has been trained and validated to evaluate our gun detection system for better results. Figure 1 shows some of the examples from the dataset being used for training our model.

2. YOLO Algorithm

The YOLO (You Only Look Once) algorithm is a real-time object detection system that uses a single neural network to predict bounding boxes and class probabilities directly from the entire image, making it both fast and efficient for applications requiring real-time analysis. YOLO divides the image into a grid, where each cell predicts bounding boxes, confidence scores, and class probabilities for any object whose center falls within that cell.

By streamlining detection into a single pass, YOLO achieves high-speed detection, ideal for applications like autonomous driving, surveillance, and live video analysis. Despite challenges with small objects and precise localization, YOLO's architecture continues to evolve, with versions like YOLOv3, YOLOv4, and YOLOv5 introducing feature pyramids, anchor boxes, and advanced optimizations that enhance accuracy and ease of use.

With benefits like end-to-end training, high accuracy, and non-maximum suppression (to reduce redundant boxes), YOLO remains a leading choice for real-time object detection in many fields.

Fig 1: images used for training YOLO model

Only a single boundary box prior is associated with each ground truth object. If bounding box prior is not allocated, no classification and localization lost incurs, only confidence loss on objects incurs. We compute the loss by using the following equations.

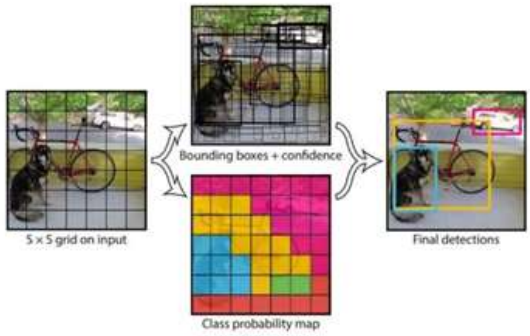

Fig 2: YOLO model working

The models detection as a regression problem. It divides the image into an S × S grid and for each grid cell predicts B bounding boxes, confidence for those boxes, and C class probabilities. These predictions are encoded as an S*S* (B*5 + C) tensor.

3. Algorithm Used

Data Collection and Annotation: Gather a dataset containing images of various weapons and non-weapon objects. Annotate the images using bounding boxes around the weapons, labelling each box with the corresponding class (e.g., gun, knife).

Data Pre-processing: Resize images to the input size required by YOLO (typically 416x416 pixels). Normalize pixel values to a range of 0 to 1 by dividing by 255. Augment the dataset with techniques like rotation, scaling, and flipping to improve model robustness.

Training the Model: Split the dataset into training and validation sets. Use the annotated training set to train the YOLO model, utilizing a suitable loss function. Monitor training progress through metrics like mean Average Precision (mAP) and adjust hyper parameters (learning rate, batch size) as needed.

Model Evaluation: Evaluate the trained model on the validation set to assess its performance. Calculate precision, recall, and mAP to quantify detection accuracy.

Inference and Detection: Load the trained YOLO model and perform inference on new images or video streams.

Post-Processing and Visualization: Annotate detected weapons on the original images with bounding boxes and labels.

Development and Improvement: Integrate the model into a suitable application or system for real-time weapon detection. Periodically retrain the model with new data to improve accuracy and adapt to changing conditions.

International Journal of Scientific Research & Engineering Trends
Volume 10, Issue 6, Nov-Dec-2024, ISSN (Online): 2395-566X

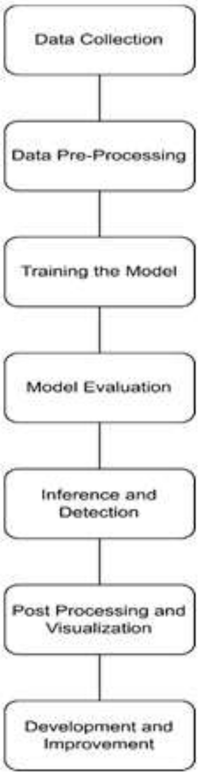

Fig 2: Algorithm of YOLO

Result Analysis

Fig 3 shows an epoch accuracy graph from Tensor Board, displaying the training and validation accuracy of a model over several epochs. The curve (in dark blue) starts lower and increases rapidly over the initial epochs, plateauing around an accuracy of 0.9844. The validation accuracy curve (in light blue) closely follows the training accuracy curve and stabilizes at a similar accuracy level, around 0.9841. Both the training and validation accuracy curves converge around 0.984, suggesting the model is well-trained and generalizes well to validation data without signs of overfitting.

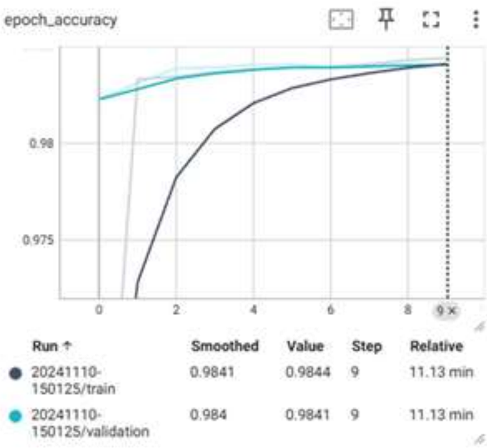

Fig 3: epoch_accuracy

The image shows a graph with a title "epoch_loss" and two lines, one for "train" and one for "validation". The "train" line is on top of the "validation" line at the beginning of the graph and then both lines decrease over time, with "validation" line decreasing at a slower rate than "train" line.

The y-axis of the graph is labeled with numbers from 0.03 to 0.07, representing the loss value. The x-axis is labeled with numbers from 0 to 9, representing the epoch number.

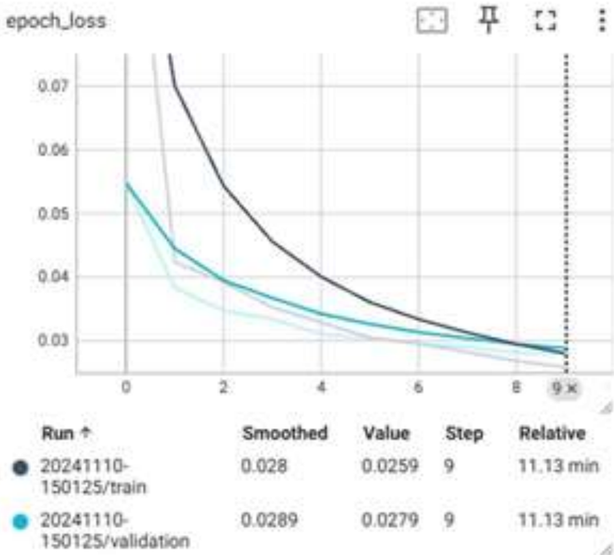

Fig 4: epoch_loss

Fig 5: Detected weapon by model

IV. CONCLUSION

In conclusion, the development of a weapon detection system utilizing deep learning techniques demonstrates significant potential for enhancing security and safety in various environments. Through the implementation of advanced algorithms and data augmentation methods, the model effectively identifies and classifies firearms, such as pistols, rifles, and shotguns, with high accuracy. This research not only contributes to the growing field of computer vision but also addresses critical societal issues related to weapon-related violence. The integration of this technology into surveillance systems can aid law enforcement agencies in real-time threat assessment and response. Future work may focus on improving model robustness and expanding the dataset to include diverse weapon types and scenarios, ultimately leading to more comprehensive solutions for public safety. This project serves as a stepping stone toward innovative applications in security technology, underscoring the importance of continued research in this vital area.

Future Work

The current research on handgun detection using YOLO opens numerous promising avenues for future exploration and technological advancement. One critical area of future research involves expanding the dataset to encompass a more diverse range of scenarios, including complex urban environments, low-light conditions, and varied occlusion levels.

Researchers could develop more comprehensive datasets by incorporating images from multiple geographical contexts, different camera angles, and varying weapon orientations to enhance the model's generalizability. Additionally, exploring multi-class weapon detection capabilities could significantly expand the system's practical applications, enabling simultaneous identification of different types of firearms and potential threats.

The integration of advanced deep learning architectures, such as transformer-based models or hybrid neural network approaches, could potentially improve detection accuracy and reduce false positive rates. Computational improvements could focus on developing more lightweight versions of the algorithm that maintain high accuracy while reducing computational requirements, making the system more adaptable to resource-constrained environments like mobile devices or embedded surveillance systems.

Another promising direction involves incorporating real- time tracking mechanisms that can continuously monitor and predict potential threat movements, transforming the detection system from a static identification tool to a dynamic threat assessment platform. Machine learning techniques like few-shot learning and transfer learning could be employed to improve the model's performance with limited training data, particularly in specialized or sensitive environments.

The research could also explore privacy-preserving techniques that maintain detection capabilities with anonymised individual identities, addressing potential ethical concerns in surveillance applications. Furthermore, integrating advanced sensor fusion techniques, combining visual data with thermal, infrared, or millimetre-wave imaging, could create a more robust multi-modal detection system. Interdisciplinary collaborations between computer vision experts, security professionals, and law enforcement agencies could provide practical insights for refining the detection algorithm.
Ultimately, the goal is to develop an intelligent, adaptable, and ethically responsible weapon detection system that can significantly enhance public safety while minimizing false alarms and preserving individual privacy..

REFERENCES

1. Gun Violence Archive, “Past Summary Ledgers: Gun Violence Archive,” [Online].
2. Muhammad Amin, B., K. Mohammad Rahim, and M. S. GeshinaAyu."A trend analysis of violent crimes in Malaysia." Health 5, no. 2 (2014): 41-56.
3. Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788. 2016.
4. Olmos, Roberto, SihamTabik, and Francisco Herrera. "Automatic handgun detection alarm in videos using deep learning." Neurocomputing 275 (2018): 66-72.
5. Tiwari, Rohit Kumar, and Gyanendra K. Verma. "A computer vision-based framework for visual gun detectionusing harris interest point detector." Procedia Computer Science 54 (2015): 703-712.
6. Arceda, V. Machaca, K. FernándezFabián, and Juan Carlos Gutíerrez. "Real time violence detection in video." (2016): 6-7.
7. Akcay, Samet, Mikolaj E. Kundegorski, Chris G. Willcocks, and Toby P. Breckon. "Using Deep Convolutional Neural Network Architectures for Object Classification and Detection within X-Ray Baggage Security Imagery." IEEE Transactions on Information Forensics and Security 13, no. 9 (2018): 2203-2215.