PAPER NAME

Final research.docx

WORD COUNT

2147 Words

CHARACTER COUNT

11196 Characters

PAGE COUNT

6 Pages

FILE SIZE

1.2MB

SUBMISSION DATE

Jan 30, 2023 11:56 PM GMT+5:30

REPORT DATE

Jan 30, 2023 11:57 PM GMT+5:30

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 1% Internet database
- 2% Publications database
- Crossref database
- Crossref Posted Content database
- 4% Submitted Works database

A comparative analysis of weapons detection using various deep learning techniques.

***Abstract*—** In crime scene analysis, object detection can be used to identify and track objects and people, which can help investigators to recreate the events and understand the sequence of actions that took place during a crime. The images are manually annotated, which is a process where a human expert goes through each image and marks the location and class of objects within the image. This process is important for training object detection algorithms as it provides the necessary ground truth data for the algorithm to learn from.
In the case of crime scene analysis, high accuracy is crucial, as it can help ensure that no evidence is missed, but speed is also important, as time is often a critical factor in investigations. The proposed system in our study attempts to balance this trade-off by using algorithms like YOLOv5, SSD, and RCNN, which are known for their real-time performance while maintaining a high accuracy level.

I. INTRODUCTION

Assault is a serious crime that can have severe consequences for the perpetrator.It is important to take steps to prevent an assault from occurring, particularly in public places where it can be more difficult to escape or get help. It is certainly important to notify the police as soon as possible when a crime

Weapon detection is an important issue for public safety and security. By using deep learning algorithms, it is possible to develop models that can accurately identify weapons in real time. These algorithms can analyze large amounts of data and identify patterns or features that may not be immediately apparent to humans. It involves identifying the presence of weapons, such as guns or knives, in a given area or location. This can be done using an image or video analysis, where the algorithm is trained on a dataset of images or videos that include weapons and other objects.

The proposed system aims to assist the police in detecting and identifying weapons in a variety of settings, including outdoor scenes. By using machine learning algorithms to analyze visual information, the system can potentially help the police to more efficiently and accurately identify weapons. This could be a valuable tool for law enforcement, as it could help them to more quickly and effectively respond to situations involving weapons, and potentially prevent crimes from occurring.

II. RELATED WORKS

The work by Laurie et al. [1] covered a real-time object recognition system that employed deep learning to identify objects in an image. This article mainly focused on the use of convolutional neural networks for object detection, which is the core technology used in YOLO architecture. The model was able to recognize items in photos using this technique with maximum accuracy and little latency. The use of depth-separable convolutional layers in this paper increased object detection accuracy while lowering computational model expenses. The feature extraction layer, the bounding box prediction layer, and the classification layer are just a few of the elements that make up the YOLO architecture as well as the technique that it employs.
The work by Redmon et al. [2] discussed the architecture of YOLO, a real-time object detection system capable of recognizing objects in an image with high accuracy and low latency. This study focused on various components of the YOLO architecture such as a feature extraction layer, a bounding box prediction layer, and a classification layer. The article also discussed the use of convolutional neural networks for object detection. The aim of the project was to investigate the use of Transfer Learning using Convolutional Neural Networks (CNN) to improve the accuracy of an object detection model and its application to new image datasets.
The paper by Fan et al. [3] provided an enhanced pedestrian detection system based on the SSD model of object detection. As an extra layer to the SSD concept, this system featured the Squeeze-and-Excitation (SE) model. The SE model allowed the system to learn from its flaws and enhance the precision of small-scale pedestrian identification. The enhanced model's accuracy was tested using the INRIA dataset. This study served as a resource for comprehending the SSD concept and future applications.
Several works have shown that CNNs can also be used as feature extractors for various vision tasks. Razavian et al. [4] proposed an approach to extract CNN features from an off-the-shelf network to perform object recognition. Similarly, Donahue et al. [5] used the activation of convolutional layers to improve object recognition accuracy. In another work, Oquab et al. [6] used pre-trained CNNs for image retrieval tasks.
In order to anticipate the hourly wind speed, Ling Chen and Xu Lai (2011) [7] compared the experimental results generated using an Artificial Neural Network (ANN) and

Autoregressive Integrated Moving Average (ARIMA). Comparatively speaking, the ANN model outperforms the ARIMA model.

Shiju Sathyadevan, Devan M. S., et al. (2014) [8] identified places with a high possibility of experiencing crime and depicted crime-prone zones. The authors used the Naive Bayes classifiers algorithm, a supervised learning and statistical approach for classification that has a 90% accuracy rate, to classify the data.

III. METHODOLOGY

A) Dataset

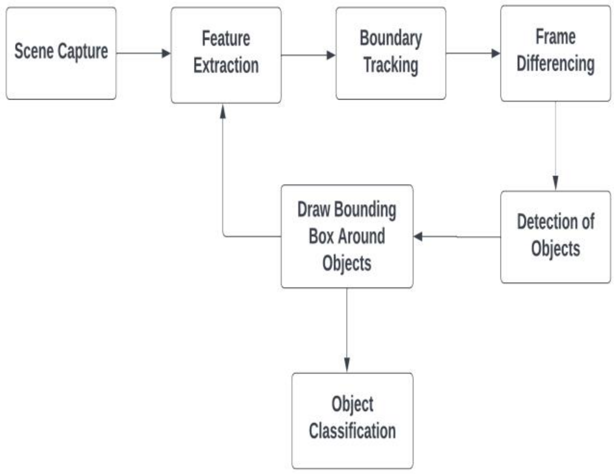
The dataset is used to train the model. Raw images often need to be pre-processed before they can be used for further analysis or model training. The YOLO v5, SSD, and R-CNN models were trained on a publicly available weapon image dataset provided by the University of Granada research group, with a specific split of 2208 training images, 660 testing images, and 174 validation images. The aim of the training was to detect pistols, rifles, and knives, and the number of training epochs was set to 100.

B) Separation of the dataset into train and test data:

We have three classes of gun, knife, and person, and based on position; weapons are marked in the image. It will generate a text file for each image which consists of coordinates used for markings. The label file will store the classes for which the images will be marked. You can then upload the zip file to a cloud storage service such as Google Drive and then you can split it utilizing the chosen split ratio, into two sets: a training set and a test set.

C) Workflow Diagram

Fig.1 explains the workflow of the system.



1. Scene Capture: This is the initial stage where we must capture a view of a real environment that is most probably a crime scene

2. Feature Extraction: Extraction of features is done from the initial dataset these new features have this little redundant information in them and therefore features are often less intuitive to understand.

3. Boundary Tracking: In object detection, boundary tracking refers to the process of identifying and locating objects in video sequences or images.

4. Frame Differencing: Frame differencing is a method for finding objects to detect changes in a video stream by subtracting the current frame from a previous frame.

5. Detection of Object: In this process, we must detect the instance of the object in the image.

6. Draw a Bounding Box Around the Object: Bounding Boxes act as the reference points for weapon detection and they are normally in rectangular shape. We must first draw rectangles over images and identify X and Y coordinates for each image.

7. Object Classification: We have three different classes in our dataset i.e., knife, person, and guns so will determine which object from these three classes is present in the given scene or image it refers to our model to find out which class is present

D) Pre-processing and Augmentations

Data augmentation can be a useful technique to enhance the performance of a deep learning model by increasing the size and diversity of the dataset, which can lead to better learning and generalization.

From the information provided, it seems that the project explored several data augmentation techniques to improve the performance of the model. These include:

These techniques are commonly used data augmentation methods in deep learning to increase the diversity of the training dataset and prevent overfitting. By applying these techniques to the training data, the model is exposed to more variations of the same object, which helps to make it more robust and better at generalizing to new, unseen data. Auto-Orient helps to orient the image correctly, Resize adjusts the size of the image, Grayscale converts the image to grayscale, Auto-Adjust Contrast improves the contrast of the image, and Outputs per training example refers to the number of bounding box predictions made per image. Rotation and Shear add rotational and shearing transformations to the image, respectively, which help to increase the diversity of the data even further.

IV.IMPLEMENTATION

A) Algorithms

1. In YOLOv5, these were the two unified blocks that turned into a single monolithic block.
   1. feature extraction
   2. object localization

   YOLOv5 has three most important components model Backbone, Neck and Head.

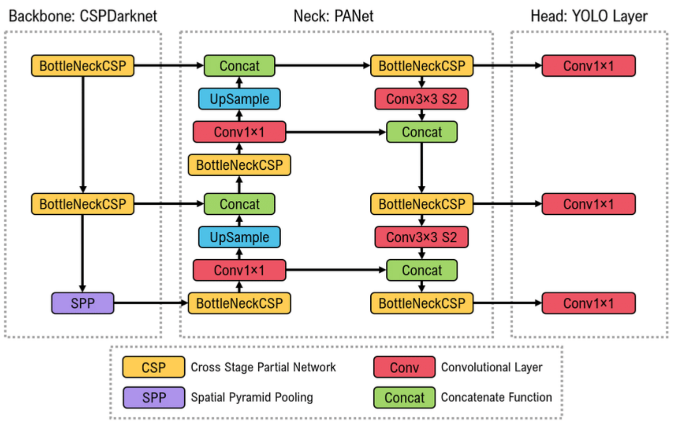


Fig.2. YOLOv5 Architecture[9]

YOLOv5 employs Cross Stage Partial Networks for the purpose of obtaining instructive information from the input image (CSP). The model neck creates feature pyramids (FP). Anchor boxes are used to apply feature class probabilities.

2. The object detection technique known as Region-based Convolutional Neural Network (RCNN) is based on the visual data of images.

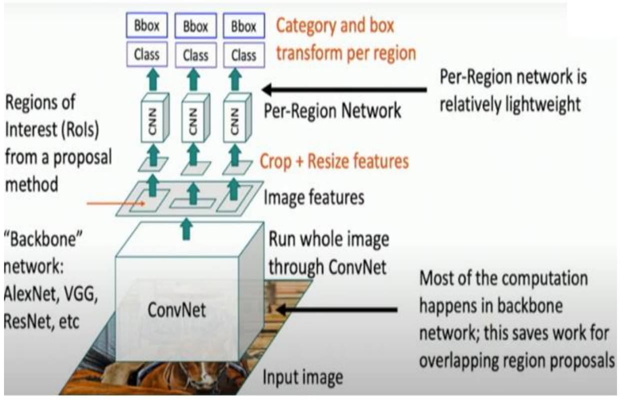


Fig.3 R-CNN Architecture[18]

Figure 3, shown above,describes the RCNN architecture.After determining the region where objects might be present ,we extract the feature of input image provided using fearture extractor.

The following steps are often used by R-CNN to classify objects.
1. Take note of the n areas (Region Proposals) where the original image's items will be located.
2. Use CNN to extract choices from regions that are 227x227 (AlexNet) or 224x224 (VGG16) in size.
3. The 4096-dimensional Ultimate Output Layer (UOL) forecasts class adjustments using SVMs and prediction boxes.

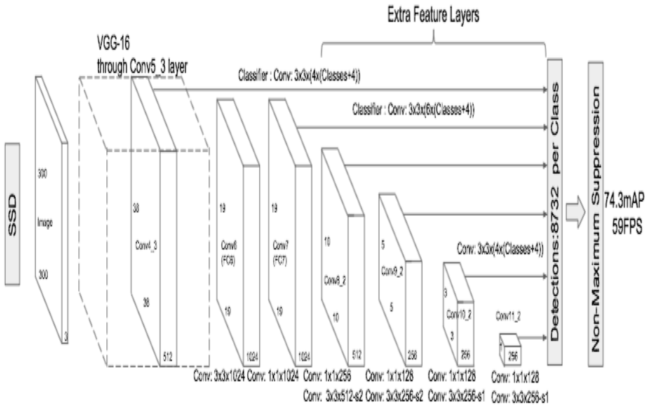


Fig.4. SSD VGG-16 Architecture[5]

The SSD architecture is depicted in the picture and comprises additional layers that are built on top of a base CNN network, such as VGG or MobileNet. The SSD technique extracts information from each grid cell using a sequence of convolutional and pooling layers.Each grid cell is then subjected to a classifier to forecast.

B)Performance Metrics

True Positive Results

This suggests that the model is performing well in recognizing small and blurry objects. The high accuracy on these objects indicates that the model has effectively learned the relevant features and patterns needed to correctly identify them.

Fig.5. Image for True Positive Result

The YOLOv5 model was used to perform object detection on a video footage and the results were evaluated in terms of true positive detections. From the information provided, it seems that the model performed well in detecting small and blurry objects, as evidenced by a high mAP (mean average precision) score of 67% for small objects. This suggests that the YOLOv5 is known for its robustness in accurately detecting small and difficult-to-see objects in video footage. Similarly, the SSD model is known for its ability to perform well in detecting small and blurry objects in video footage,as indicated by the high mAP score of 37% for small objects.the RCNN model performed well in detecting small and blurry objects in video footage, as indicated by the high mAP score of 37% for small objects.

False Positive Results

false positive is a prediction made by the model that an object is present in an image when is reality it is not this can result in incorrect or misleading result and is a common challenge in object detection task a high number of false positive can negatively impact the Precision of the model and reduced its over all performance

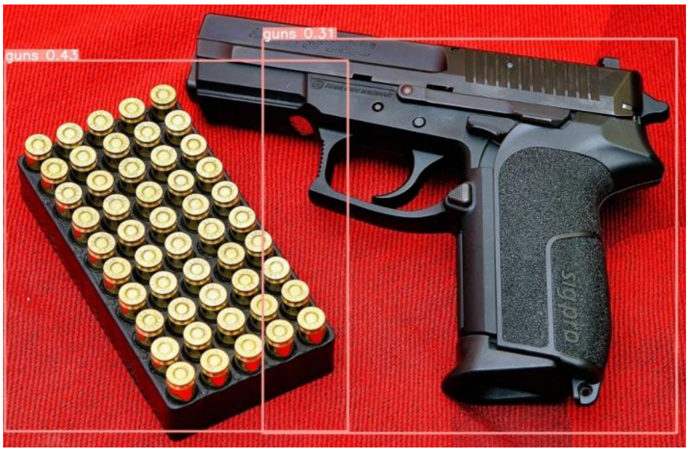


Fig.6. Image for False Positive Result

True Negative

True negative is a Precision made by a model that there is no object in an image when in reality there is not high number of two negative indicates that the model is able to accurately identify when there is no object present in the image which is important for awarding false alarms and improving the precision of model.

Fig.7. Image for True Negative Result

C) Evaluation Metrics

AP (Average Precision) is a commonly used metric in object detection to evaluate the performance of the model. It considers both precision and recall and calculates the average precision for a range of recall values. The higher the AP score, the better the model is at correctly identifying objects in an image.

MODELS	YOLOv5	R-CNN	SSD
mAP value (Mean Precision Value)	56.2	47.1	36.7

Table.1 Comparative analysis of map values

V.CONCLUSION

We used YOLOv5, RCNN, and SSD as different models for image processing and computer vision, and we are comparing their performance on our dataset. As mentioned YOLOv5 has the best prediction accuracy, but it was slower than the other two models.YOLOv5 model included an affine-tuning approach to optimize its performance, and YOLOv5 has improved compared to previous versions and RCNN can be more accurate but slower. Ultimately, the choice between YOLOv5 and RCNN depends on the specific use case and the trade-off between accuracy and speed that is acceptable for the application.

5% Overall Similarity

Top sources found in the following databases:

- 1% Internet database
- 2% Publications database
- Crossref database
- Crossref Posted Content database
- 4% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	**University of Essex on 2021-04-30** Submitted works	1%
2	**American Public University System on 2023-01-02** Submitted works	<1%
3	**"Recent Advances on Soft Computing and Data Mining", Springer Scien...** Crossref	<1%
4	**Liang Huang, Weijian Pan, You Zhang, Liping Qian, Nan Gao, Yuan Wu. ...** Crossref	<1%
5	**Queen Mary and Westfield College on 2023-01-17** Submitted works	<1%
6	**ncbi.nlm.nih.gov** Internet	<1%
7	**Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubba...** Crossref	<1%
8	**Mugla University on 2021-03-30** Submitted works	<1%

9	Sepp Hochreiter, Jürgen Schmidhuber. "Long Short-Term Memory", Ne... Crossref	<1%