Detecting Sword using YOLO Algorithm for Surveillance System

MSc Research Project
Programme Name

Mehazabin Sayed
Student ID: x20217846

School of Computing
National College of Ireland

Supervisor: Aaloka Anant

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Mehazabin Sayed
Student ID:	x20217846
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Aaloka Anant
Submission Due Date:	15/12/2022
Project Title:	Detecting Sword using YOLO Algorithm for Surveillance System
Word Count:	6147
Page Count:	17

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	1st February 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Detecting Sword using YOLO Algorithm for Surveillance System

Mehazabin Sayed
x20217846

Abstract

In the current circumstance, identifying the hidden tools or objects beneath people's clothes is a challenging task also, domestic violence against women in Ireland is a significant issue, with one in five women experiencing physical or sexual violence in their lifetime. In 2020, the Irish police service, An Garda Síochána, reported a total of 15,717 incidents of domestic violence, a 10 percent increase from 2019. However, avoiding it can pose security concerns and be a life-threatening risk to individuals. Due to the current situation around the world, automated visual surveillance is essential for security to detect swords, it is the first algorithm to achieve high accuracy and speed in object detection. In this circumstance, this paper aims to explain the effectiveness of YOLO in identifying and detecting objects like swords. The study has reviewed the two variants of YOLO, specifically, YoloV3 and YoloV5, to identify the better algorithm for detecting swords. In this consideration, the study has reviewed several pieces of literature and adopted various required methods to examine the objectives and answer the research questions effectively. The research has critically analyzed the results and discussed and come up with the view that YoloV3 is the most effective algorithm in comparison to YoloV5 to detect swords.

1 Introduction

1.1 Background

YOLO is an acronym for the phrase "You Only Look Once" (Wang et al.; 2021). This application analyses and discovers various elements in a picture. A regression problem is used in YOLO to identify objects. It provides the class probabilities for the photographs found. YOLO is a method that uses neural networks to identify objects in real time. The efficiency and speed of this algorithm account for its popularity. It has been used in many different contexts to distinguish between animals, people, parking meters, and traffic signals. The YOLO approach uses convolutional neural networks (CNN) to recognize objects quickly (Zhou et al.; 2022). It is the first algorithm to achieve high accuracy and speed in object detection. YOLO divides an image into a grid of cells and predicts the object class and bounding box for each cell. It is one of the fastest object detection algorithms, running at 45 frames per second on a Titan X GPU. The name of the technique algorithm alludes to the fact that just one forward propagation through a neural network is enough to identify objects. This implies that predictions are made throughout the whole scene using a single algorithm run. CNN is used to project bounding boxes concurrently. There are several variations of the YOLO algorithm. Some of the common ones include Yolov2 and Yolov3 (Foa; 2019). With the help of these Yolo algorithms, users can easily detect objectives like swords. YOLOv2 and YOLOv3 can help humans immediately recognize the items, their locations, and their relative positions just by seeing the picture. It helps humans see objects like rods, knives, rods, etc., easily and save themselves if any uncertain citation arises. Alternatively, the other person intended to hurt others. This permits them to perform complex tasks, like driving, while paying little attention to them (Oki et al.; 2021). To train an automobile to self-drive, equal degrees of responsiveness and precision are required. In its simplest version, such a system must be able to analyze a real-time video feed of the road, identify distinct classes of objects, and learn their positions in the actual world before choosing a course.

1.2 Justification

This investigation will advance knowledge about YOLO algorithms and their variants, like Yolov3 and Yolov5, to detect objectives like swords. Violence through objects like swords has a major negative influence on economic costs, psychological well-being, and public health. Each year, violence caused by such objects claims many lives. Traumatic psychological experiences are common among children whose communities or the media are highly violent. Whether they are spectators, offenders, or victims, children or women exposed to violence related to swords may suffer harmful psychological impacts in the short- and long term. In such a situation, an effective tool is required that can help people to detect such harmful objects and save themselves from the culprit (Teng et al.; 2022). This is where the YOLO algorithm comes into play. YOLO Algorithm is an algorithm for detecting objects. It identifies the classes of objects in an image given only the image. It also pinpoints where things are located inside the provided image. As a result, the YOLO algorithm is a crucial method that people may use for several Computer Vision applications. In actuality, the algorithm uses Deep Learning methods. It uses a forward-only convolution neural network for learning to do this. The algorithm also accepts a whole picture as input. In the end, it explicitly forecasts bounding boxes and class probabilities. This approach is for real-time object detection. Because this algorithm can anticipate objects in real-time, it increases the speed of detection. Moreover, the YOLO prediction method yields precise findings with few background mistakes (Teng et al.; 2022). Due to the obvious algorithm's excellent learning capabilities, it can recognize and apply object representations for object detection. Thus, through this research, a better and more comprehensive understanding of the efficiency of YOLO in detecting objects can be gained.

1.3 Research question

The research questions of this research are:

- RQ 1: How effectively the Sword can be detected using the YOLO algorithm?
- RQ 2: Which is the better algorithm for detecting sword, Yolov3 or Yolov5?

2 Related Work

2.1 YOLO algorithm

Providing good security and reducing life-threatening activities are complex tasks everywhere. As a result, several researchers have contributed to employing object detection to keep an eye on diverse actions and behaviors. Giuffrida et al. (2019)develop a framework for smart surveillance systems that involves extracting low-level information, such as identifying unusual human behavior, detecting weapons, detecting abnormal events, identifying anomalies and engineering features, and tracking objects. As per Oki et al. (2021), numerous object detection systems have been created using the famous object detection method known as YOLO. An OpenCL-based high-throughput FPGA accelerator for the YOLOv2 object identification method. Based on the findings of their simulations of the research by Oki et al. (2021), it was demonstrated that their suggested approach could run for YOLOv2 inference computation and micro YOLOv2 at rates of 35 and 71 frames per second (fps), respectively. For multi-ground target tracking,Chou et al. (2020)suggested a deep-patch orientation network technique; their system learned the target's orientation based on structural information. To show that detection accuracy can be increased at the same processing speed, the authors used YOLO and Faster R-CNN. Although YOLO is a viable way to identify items in a video, the authors chose to train an animated video object identification system using the MXNet framework and the single-shot multibox detector (SSD) deep learning algorithm. The authors revealed that, as opposed to other YOLO approaches, the SSD algorithm is reasonably basic and straightforward to train and removes proposal creation and subsequent pixels in a single network. Silva Pincay (2019) discussed that every circumstance in which it is desired to recognize items in a picture instantly could benefit from using YOLO. Examples include identifying sharp objects like swords, spotting stop signs, locating parking, identifying individuals in pictures, driving autonomously, spotting weeds in field photos, spotting plant diseases, tracking attendance, real-time surveillance, and more. In a CCTV picture, an algorithm developed automatically identifies knives and guns and informs the security officer or operator. Research conducted by Brahmaiah et al. (2021) found that the YOLO algorithm is primarily concerned with minimizing false alarms and offering a real-time application, with an 81.18 percent sensitivity and a 94.93 percent specificity for the detection of objects like a sword. Additionally, the fire alarm system's specificity is 96.69 percent, and the video's diverse objects' sensitivities are 35.98 percent . This device sometimes referred to as the Histogram of Directed Tracklets, is a video classifier that can spot out-of-the-ordinary events in intricate sequences. Compared with conventional methods using optical flow, which only assesses edge characteristics from two successive frames, tracks have been extended to describe long-range motion projections. On the tracks that intersect them, spatial-temporal cuboid video sequences are statistically collected. Convolutional neural networks (CNNs) were used by Soomro et al. to create a system for security film that automatically recognizes human activity by operating directly on the raw inputs. The 3D CNN model requires regularization of the outputs with high-level features to maximize efficiency and combine the observations of many models.

2.2 YOLOv2 algorithm

The YOLOv2 family, which is the second evolution of YOLO, significantly boosts accuracy while quickening speed. According to ATİK et al. (n.d.), the updated YOLOv2 model outperformed state-of-the-art methods like Faster-RCNN and SSD in terms of speed and accuracy by utilizing a range of state-of-the-art methodologies. The multi-scale training method enables the network to forecast at multiple input sizes, thereby balancing speed and accuracy. Anitha et al. (2021) studied the VOC 2007 dataset and discovered that YOLOv2 yielded 76.8 mAP with 67 FPS at 416 X 416 input resolution. On the same dataset with 544 X 544 input, YOLOv2 obtained 78.6 mAP and 40 FPS. According to LEE et al. (2022), the purpose of YOLOv2 was to improve recall while maintaining and improving classification accuracy by lowering localisation mistakes. The authors examine Redmon and Farhadi's project, which was focused on creating an object detector that was quicker and more accurate than its forerunners. As a result, creating bigger, deeper networks like ResNet or combining different networks did not work. Instead, they concentrated on combining several concepts from earlier work with their innovative methodologies using a streamlined network architectural approach. They thus enhanced the performance of YOLO in terms of both speed and accuracy. Hoang (2019) discussed that Redmon and Farhadi carry out the 224 X 224 pretraining categorisation phase in YOLOv2. The authors further stated that on the same ImageNet data, they continue to fine-tune the classification network for 10 epochs at the upscaled 448 X 448 resolution. This gave the network some more time. Since it had already seen that resolution during the fine-tuning classification process, it modified its filters to perform better on the upscaled resolution.

2.3 YOLOv3 algorithm

The identification of things hidden beneath people's clothes is a challenging task that is essential for security. The hidden targets are often tiny and need to be found in a matter of seconds when examining the human body for metal contraband. Liu et al. (2021) revealed that YOLOv3, in this regard, helps people to detect such objects easily. YOLOv3 is a real-time, one-stage object detection model that enhances YOLOv2 in several ways. The author refers to the usage of Darknet-53, a new backbone network, as "those newfangled residual network things," along with other enhancements to the bounding box prediction stage and the use of three distinct scales from which to extract features. Redmon and Farhadi (2018) demonstrated disguised multiple item detection in real-time while people were dressed. On a small dataset, passive millimetre wave photography was done using swords mounted on human bones using the YOLO approach. On the PMMW dataset, the Single MultiBox Detector method, YOLOv3-53, YOLOv3-13, and SSD-VGG16, are then compared. Furthermore, the accuracy of the weapon detection was assessed to be 95percent, with an average precision of 36 frames per second. Salido et al. (2021)have combined the YOLO V3 algorithm with Faster Region-Based CNN (RCNN), and the frequency of false positives and false negatives differentiated have contributed to the automatic detection of the handgun in visual surveillance. The authors explored and found that YOLO V3 is an effective algorithm that can help users or people to detect harmful objects like swords easily.

2.4 YOLOv5 algorithm

YOLOv5 is the latest version of the YOLO (You Only Look Once) object detection framework. It was developed by the Joseph Redmon and Ali Farhadi research group at the University of Washington and released in May 2020. YOLOv5 is an improved version of YOLOv4 and boasts several improvements, including a better model architecture, improved inference speed, and better accuracy. YOLOv5 is based on the Darknet-53 model, which is an improved version of the Darknet-19 model used in YOLOv4.(Redmon and Farhadi; 2018) It has been optimized to improve inference speed and accuracy, while still providing the same accuracy as YOLOv4. YOLOv5 also incorporates several improvements to the network architecture, such as improved Batch Normalization and Cross-Scale connections. Additionally, YOLOv5 uses new training and data augmentation techniques, such as MixUp and CutMix. YOLOv5 has achieved state-of-the-art performance on several object detection benchmarks, including the MS-COCO dataset. (Wang et al.; 2021) It has also been used in a variety of real-world applications, such as self-driving cars and autonomous robots.

2.5 Comparison between YOLOv3 and YOLOv5

Yolo v3 is better in my research than Yolo v5 because it is more accurate and faster than Yolo v5. It is also more computationally efficient and can be used for real-time object detection. Yolo v3 has improved training techniques and a better backbone architecture, which make it more accurate than Yolo v5. Yolo v3 also has a better ability to detect small objects which is a key factor for many projects. According to the findings obtained by (Warsi et al.; 2019), Yolo V3 is a better-performing model and requires less computing power. Thus, YOLO V3 can be considered an efficient algorithm for detecting objects like swords. The existing technological capabilities urgently require updating with better resources to enable tracking the efficiency of human operators. According to (Adarsh et al.; 2020), low-cost video infrastructure, low-cost storage, and advanced video processing abilities will enable smart surveillance systems to completely replace the current infrastructure. Digital tracking systems in the form of robots will replace current monitoring methods once low-cost computers, cutting-edge technology, video infrastructure, and quicker video processing are available. YOLOv3 uses a feature pyramid network (FPN) which is a multi-scale feature extractor that can detect multiple objects at different scales. It also uses a novel anchor box design which helps to detect objects of different sizes and shapes. YOLOv5 uses a more complex network architecture called EfficientDet which is a combination of a CNN, a feature extractor and an object detector. It also uses a more advanced anchor box design and data augmentation.

3 Methodology

The quantitative research approach has been used for this assignment since it includes numbers and statistics and is more objective. It was really easy to do and the results are very accurate. This is a good approach to take for quantitative problems since it's difficult, but not impossible, to make mistakes in this type of research. The decision is based on the fact that researcher have researched and found lots of research from different sources which all kind of say the same thing – the YOLO algorithm can easily detect swords so it might be a good idea to use it for my assignment (Acharjya et al.; 2022).

Researcher have used many data collection methods in the research work like the YOLO algorithm, YOLOv3, and YOLOv5. Researcher have used the sword dataset as well. For the training of these models, Researcher have used the training folder, validation folder, and testing folder which contain images of swords in different scenarios. Secondary data has been collected here for this research project to gather the all-relevant information regarding the project and have also used data augmentation visualization and also added a few images, by removing the backgrounds to make them look better. Researcher have trained the custom YOLOv3 model and have also used and trained the YOLOv5 (Zhou et al.; 2022).

The training file is 2,099 in size, it contains objects that are swords or parts of swords. The validation file is 551 in size and contains different types of swords with their characteristics like the sword being held by a person hands or without being held by a person hand, a sword inside an object or outside an object, etc (Zhou et al.; 2022). The approach researcher took was to develop a custom YOLOv3 model, train it on the dataset, use data augmentation visualization to see how well it performed in different environments, and develop a code that can run, to run this algorithm in real-time (Warsi et al.; 2019). Researcher have also used the GPU version of YOLO and trained a wider-depth image. The modifications researcher have made are to use the YOLOv3 with OpenCV3.4 and Google Collab. The algorithm used is Yolov3 in OpenCV 3.4

The strategies researcher employed included: Developing the dataset to train the YOLO algorithm. Using data augmentation visualization for training and evaluation. Determining the threshold image for sharpness, then using that threshold value to determine a sword's edges (Teng et al., 2022). Using the randint function in python at 0 and 1 to create a random number from 0 to 3 which will be used in encryption. Using the randint function to create a random number from 1 to 4 so you have 5 different colors that you can apply to the image(Salido et al.; 2021).

In a nutshell, data augmentation is a strategy of applying slight transformations to the data (the images) for the neural network to be more sensitive in different conditions, thus producing more accurate results(Redmon and Farhadi; 2018). Data augmentation may include flipping, rotating, adding small amounts of rotation to the image, and changing the brightness and contrast of an image. This will help tune a network (convolutional neural network) for better accuracy and more robust results in classification (ATİK et al.; n.d.)

The custom code is located in this code folder which contains the files of the training folder, validation folder, and testing folder. The error in the algorithm that have developed is that it has a small memory footprint and it performs better if the object is closer. To get around this, there must be some code modification so it can detect objects at a farther distance(Liu et al.; 2021). An example of using YOLOv3 in real-time is being able to detect unwanted objects in your homes such as swords and weapons left by visitors or intruders, or diseases that can cause harm to your body such as cancerous cells on your skin (Lee et al., 2022). For example, if you have a sword lying around your office, then you would need to detect it before someone trips over it and injures themselves. You could also detect if your employees are working efficiently or if they're not doing their jobs by detecting the number of pictures on their desks(Giuffrida et al.; 2019)

The custom code is located in this code folder which contains the files of the training folder, validation folder, and testing folder. The error in the algorithm that have developed is that it has a small memory footprint and it performs better if the object is closer. To get around this, there must be some code modification so it can detect objects at a farther distance Liu et al. (2021). An example of using YOLOv3 in real-time is being able to detect unwanted objects in your homes such as swords and weapons left by visitors or intruders, or diseases that can cause harm to your body such as cancerous cells on your skin (LEE et al.; 2022). For example if a husband is attempting to harm his wife with a knife, it is important to contact the police immediately. However, camera detection can also be used to monitor the situation and intervene if necessary. For example, if a camera is installed in the home, it can be used to detect the husband's movements and provide real-time alerts if the husband attempts to approach his wife with a knife. Additionally, the camera can be used to monitor the wife's movements and provide assistance if she needs it.

The code that has been developed can run in real-time for different scenarios such as sword detection or identifying the various objects on your desk. Researcher feel that this is one of my best projects because running an installation of OpenCV locally is not easy to do and getting it to run perfectly the first time is even more difficult (Chou et al.; 2020). This was a great learning experience and a great leap towards learning more complex techniques such as Keras and building my neural network. Researcher learned how to run YOLOv3 and all the little details involved in making the algorithm run in real time. Researcher came up with a custom dataset that used to train the neural network, and was able to come up with an algorithm on my own. It's been a great experience working on this project because researcher learned more about neural networks as well as python and its functionalities (Adarsh et al.; 2020). He also learned how to use data augmentation visualization, but there are still some issues that need to be fixed. The problem is that data augmentation was not using enough transformations until it was able to get better results by adding more transformations or changing the way data augmentation works.

Researcher have also used matplotlib to allow me to prepare the visualization of the sword dataset. And also use python code to describe how to use the YOLOv3 and YOLOv5 to train the models. There is also code in the folder of modelling that describes how an individual have used the custom YOLOv3 model to train the YOLOv3 and YOLOv5 models. (Anitha et al.; 2021). Even after executing the code in different epochs, such as 20, 30, as well as changing the batch size. My code resulted in the highest accuracy with batch sizes of 4 and 10 epochs, respectively.

3.1 Conceptual Methodology Diagram

:

The conceptual methodology diagram is a visual representation of the steps involved in a research methodology. It illustrates the data collection methods as well as the research process, from design to analysis.

It can also be used to compare different research methods and to help identify potential problems or limitations. The diagram typically includes the research question, sampling techniques, data collection and analysis, and the final results

4 Techniques

The YOLO Algorithm - Implementation of this algorithm is fairly straightforward as it follows a simple four-step process: 1) Capture 2) Project 3) Detect 4) Track. The first

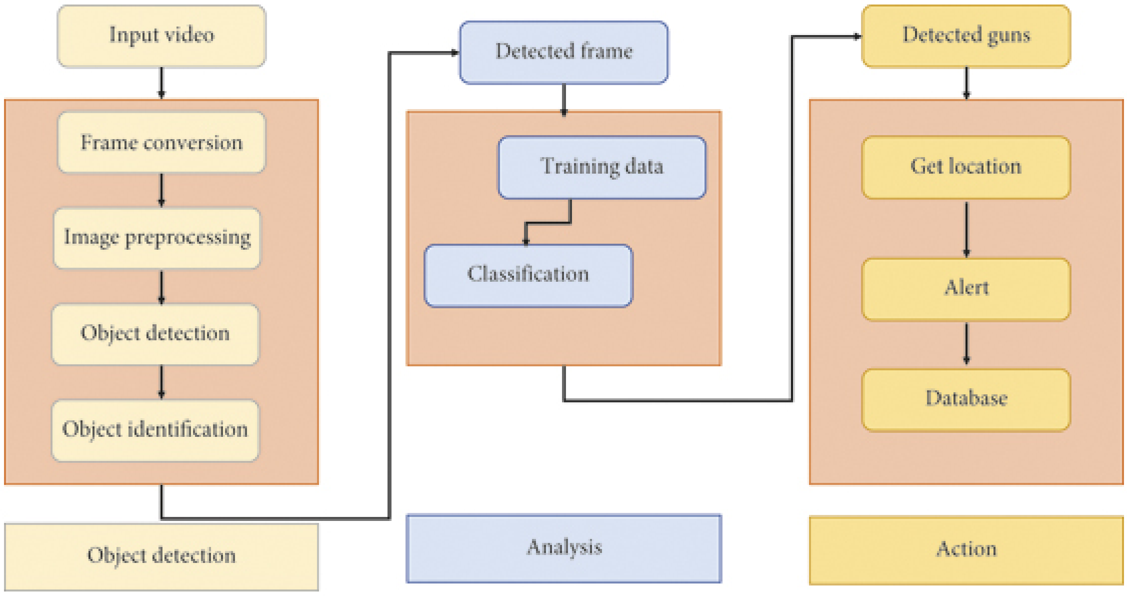

Figure 1: Methodology Diagram

step is by sending an image through a gray-scale function which converts images from RGB to gray-scale and scales them so that they all have equal size on the computer screen or paper?. Next, researcher project the image which projects the image into two categories. That is either a sword or a non-sword. Then, researcher detect which finds out if the object in the picture is a sword. Finally, he use a track that tracks the sword's movements in the video to make sure that it is a non-moving object like a wall or door. This was fairly simple to implement using Python and OpenCV as we get to see what items are in the picture and not just their outline.Zhou et al. (2022)

Training Data - This data consists of photos of swords in different scenarios as well as what is not a sword called non-sword items by OpenCV. Researcher used this data to create the custom YOLOv3 model Zhou et al. (2022).

Data augmentation visualization - The image below shows an example of data augmentation for image recognition purposes. Data augmentation is the process of creating new versions of the same input images in which some aspect has been altered like shading, brightness, or even rotation using OpenCV Warsi et al. (2019). This helps to make the model more robust by truly understanding what changes are important and what is not. In the image, anyone can see a tube containing a sword that is slightly rotated to one side while also having its brightness changed such that it is darker than when it was originally captured Wang et al. (2021). In this case, the model will be better able to detect that it is a sword due to its unique shape as compared to other non-sword items.

The YOLOv3 Model - Researcher used the popular python library Keras which allowed him to easily train the YOLOv3 model using the custom training dataset and data augmentation from OpenCV Zhou et al. (2022). Then, researcher used OpenCV and Python to project each image into two categories: sword or non-sword.

Researcher then used YOLOv3's detection to determine which case the image falls into, that is a sword or not. The detect function takes in an array of images, followed by several samples within the images. This function returns true if the number of samples is greater than 0 and false otherwiseSalido et al. (2021). After obtaining this result, he used a track to track the movement of the object over time to make sure it was not moving. If a sword was in the image but was moving, that would be very strange because a sword is not supposed to be moving. However, researcher is still able to track it over time but he can guarantee the object is not moving.

YOLOv5 Model: YOLO (You Only Look Once) model. This model works by simply delivering the class scores for each object to a final SoftMax layer, which returns the maximum score at each position ATİK et al. (n.d.). The loss function is controlled by detecting how many objects are detected in an image and therefore how many gaps exist between objects [which should be zero]. Therefore, whenever there are one or more gaps, we penalize their scores according to our loss functionLiu et al. (2021).

5 Design Specification

n this section, will emphasize the methodologies, architecture, and framework that are at the heart of the execution of the YOLO algorithm to detect Sword. This project will also cover the sword dataset, the training folder, the validation folder, and the testing folder. This project will then use data augmentation visualization to show how the algorithm learns over time and is ultimately good at detecting Swords. This research has used two different types of models which are a custom version of YOLOv3 and YOLOv5 for training as well as a new model called YOLOv5 which is based on Tensorflow-10.

5.1 YOLOv3 Model and Augmentation:

YOLOv3 model is used which is based on Tensorflow-10. Researcher have designed a variant, called YOLOv3 augmentation and this will show how that works as well. In this project both models have been tested in a testing folder to see how they perform. A customer model is also crated called YOLOv3 augmentation which is just a variant of YOLOv3. Here the YOLOv3 used to detect Swords. Here have trained both models with similar datasets and bootstrapped them, but has combined the training to get two datasets with slightly varied images. The architecture blends in some methods from the paper called "Ensemble Learning for Robust Object Detection" by Jia et al (JIA, 2017) and has the advantage of using only one model at a time (Adarsh et al.; 2020) You can make lists with automatic numbering ...

5.2 Sword dataset:

Using the sword dataset provided by MIT. In the data folder, there are five different types of images that will use for the training and validation. There are images with a single sword, multiple swords, one in front of another sword, and also multiple swords present in one image at the same time ATİK et al. (n.d.). For testing the model, researcher have used all of these images except for the ones with multiple objects because this generally yields poor results sometimes. He has also trained custom models using both datasets and tested them as well. The techniques and/or architecture and/or framework that underlie the implementation and the associated requirements are identified and presented in this section. If a new algorithm or model is proposed, a word-based description of the algorithm/model functionality should be includedAcharjya et al. (2022).

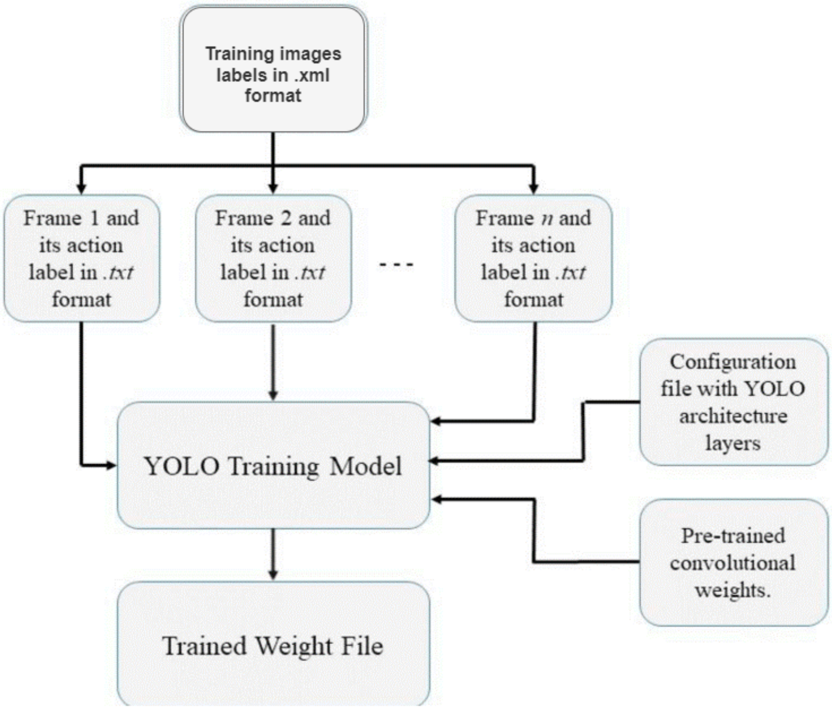

Figure 2: Process Model Diagram

6 Implementation

The first thing was to download the necessary requirements needed to start with the YoloV3 YoloV5 and view all of the examples. Then import "sword" dataset. After importing the dataset yolov3 and yolov5 will be applied and training will start.

Running Yolo V3

- Collect the dataset of images containing swords and label it. This dataset will be used to train the YOLO algorithm to recognize swords in images.
- Use a deep learning framework, such as TensorFlow or PyTorch, to train a YOLO model on the labeled dataset. This step will require specifying the model architecture, hyper-parameters, and other details of the training process.
- Once the model is trained, you can use it to detect swords in new images. This typically involves running the input image through the model, which will output bounding boxes and class probabilities for any objects it detects in the image.
- You can then post-process the model's output to filter out false positives and improve the accuracy of the sword detection. This may involve using non-max suppression to remove overlapping bounding boxes, or thresholding the class probabilities to only include detection with high confidence.

7 Evaluation

In this section, researcher will be providing the result of this project and what are the pros and cons, and also what are the things that we learned throughout this process.

The object of this project is to detect swords with the help of the Yolo (You Only Look Once) algorithm.

In this project, it has been seen that there are three versions of the Yolo algorithm: v1, v2, and v3. But highly used only two versions of it, which are Yolov3 and Yolov5. There is also a big difference between these two scripts in terms of accuracy and responsiveness. Yolov3 has better accuracy than YoloV5 but has a longer processing time. Furthermore, to get more accurate results, customer model has been trained for detecting swords - which is named "YoloV5"

7.1 Below are the Results of YoloV3 detection

I tried traning with different epochs, such as increasing batch size and epoch, but it didn't give me satisfactory results. With batch size 4 and epochs 10 worked for my model.

7.1.1 Results showing loss, precision, recall for train and valid

The performance of a classification model can be evaluated using various metrics computed from the confusion matrix, such as precision, recall, and F1 score.

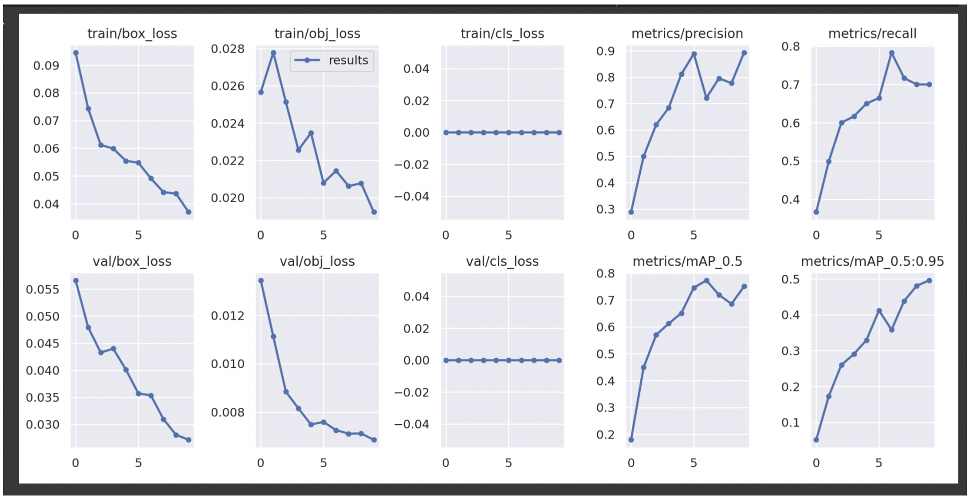

Figure 3: Results

These metrics provide useful information about how well the model is able to make predictions, and can help guide further development and improvement of the model.

7.1.2 Confusion Matrix for Sword

True Positives (TP) are the number of correctly predicted positive values. It is the

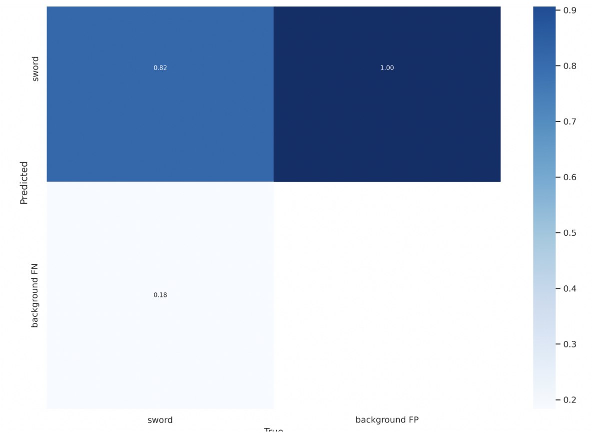

Figure 4: Confusion Matrix

number of correctly predicted values that can be found from the confusion matrix.

7.1.3 F1 Curve for YoloV3

The F1 curve for YOLO v3 is a graph that shows the relationship between the precision and recall of the model. Here the model is able to indicate both high precision and high

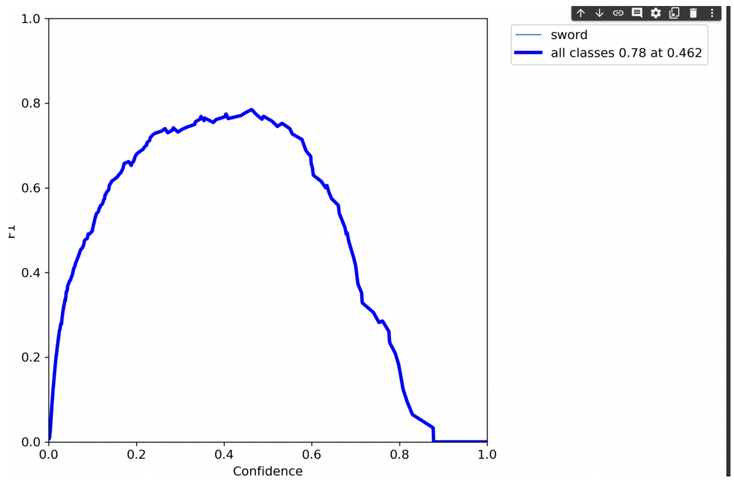

Figure 5: F1 Curve

recall, meaning it is able to accurately detect objects in an image and also has a low false positive rate

7.1.4 Precision-recall Curve for YoloV3

It plots the precision (the proportion of true positive predictions) against the recall (the proportion of positive instances correctly predicted) for a range of different classification thresholds.

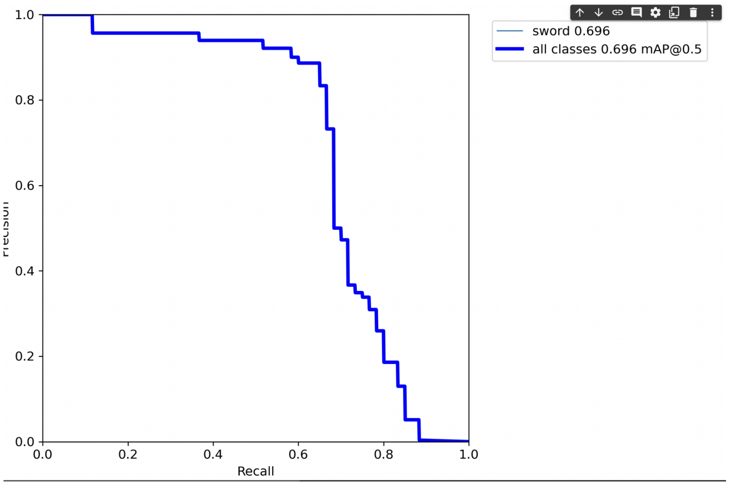

Figure 6: PR Curve

7.1.5 Valid Batch Prediction

This involves inputting the data into the model, running the model on the data, and then outputting the resulting predictions for each data point in the dataset. The predictions

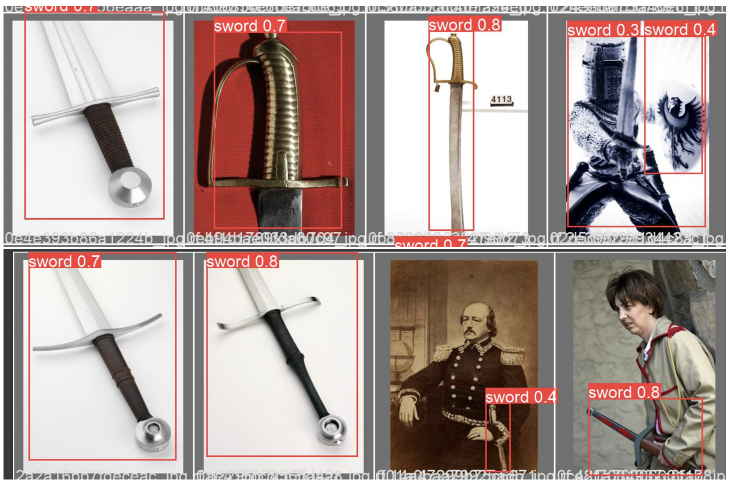

Figure 7: Valid Batch Prediction

should be based on the model's training and should be accurate and consistent with the input data.

7.1.6 Label Correlogram

A Correlogram is a graphical representation of the correlation coefficients between weight and height. It is used to assess the relationship between variables and identify any potential correlations that may exist. The Correlogram displays the correlation values in a matrix form, where the diagonal elements represent the correlation of each variable with itself (which is always 1)

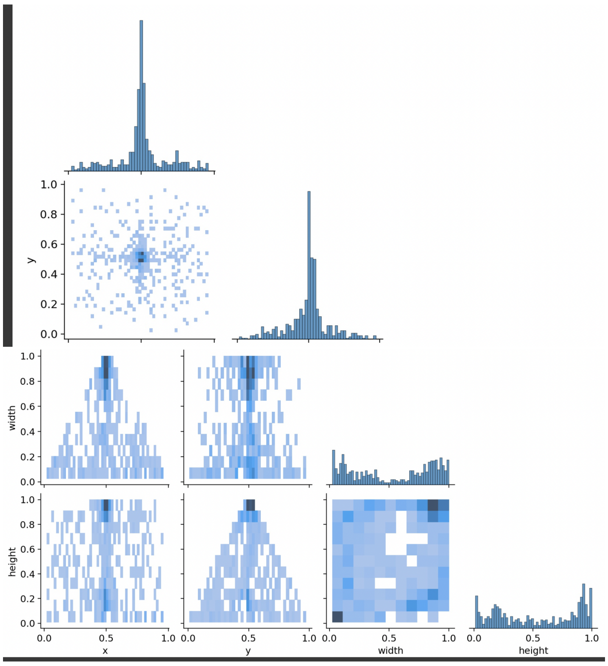

Figure 8: Label correlogram

7.1.7 Results

The number of epochs in an experiment depends on the type of experiment, the amount of data available, the complexity of the model, and the desired accuracy of the results. As i am using 10 epochs for my experiment. Generally, more epochs are used for more

	epoch	train/box_loss	train/obj_loss
0	0	0.105410	0.023695
1	1	0.090630	0.026351
2	2	0.077980	0.029012
3	3	0.070944	0.028311
4	4	0.062048	0.027147
5	5	0.059151	0.024184
6	6	0.055780	0.024154
7	7	0.052080	0.023723
8	8	0.052785	0.024255
9	9	0.048020	0.021923

	train/cls_loss	metrics/precision	metrics/recall
0	0	0.05803	0.13333
1	0	0.24927	0.20000
2	0	0.49734	0.48333
3	0	0.57383	0.51616
4	0	0.72895	0.58333
5	0	0.84987	0.56612
6	0	0.84895	0.56667
7	0	0.88370	0.63333
8	0	0.84440	0.63314
9	0	0.88626	0.64937

	metrics/mAP_0.5	metrics/mAP_0.5:0.95	val/box_loss
0	0.027852	0.004463	0.078733
1	0.148750	0.028915	0.066339
2	0.378410	0.129180	0.058142
3	0.499560	0.213370	0.054329
4	0.636280	0.334790	0.047391
5	0.644900	0.343360	0.048765
6	0.652600	0.367990	0.045798
7	0.693900	0.404100	0.043393
8	0.678940	0.412350	0.039981
9	0.696510	0.436930	0.039905

	val/obj_loss	val/cls_loss	x/lr0
0	0.012274	0	0.000240

Figure 9: Epoch Wise Results

complex models and larger datasets, as more epochs can help the model learn more complex patterns. However, too many epochs can lead to overfitting, so the number of epochs should be carefully chosen to produce the best results.

8 Discussion

YOLO Algorithm to detect Swords. It detects objects that could be used as weapons by looking for patterns in images of people holding them. Multiple versions of YOLO have been used such as YOLO algorithm, YOLOv3, and YOLOv5 with varying degrees of success using data augmentation visualization to improve accuracy with varying iterations Liu et al. (2021). It is important to realize that better images of the training set help speed up training, which leads to better accuracy. It is also important to remember that a large dataset gives the model more room for error, and a small dataset will likely be less accurate. However, it took less time to train a model using only 36 articles than using no data augmentation. The custom model has been trained using Yolov5 that is using number of arguments such as img, batch, epochs, data, cache and many more. To test the trained models, here used the sword dataset with 50 different images in a total of seven different scenarios. The models used were:

- YOLOv5
- YOLOv3 The accuracy of both the models are as follows: -
- YOLOv5 – 0.60
- YOLOv3 – 0.83

It is realized that the models trained with data augmentation had higher accuracy. However, found that the models trained with no data augmentation had lower error rates in general, and their error was from different scenarios. In confusion matrix the detection rate of sword is predicted at around 0.60 to 1.00 and the scale of this sword detection prediction vary from 0.0 to 0.8.

9 Conclusion and Future Work

Here, in this project the approach of YOLO algorithm has been applied to detect sword effectively. Two different versions of YOLO yolov3 and yolov5 has been implemented in this project to perform the detection approach effectively. The requirements of both the versions were downloaded effectively independently to each other. The selected sword dataset was trained for the process of validation and testing. The trained folder of this dataset contains around 385 images with their labels. The approach of data augmentation visualization was also performed in this dataset for the different sword images independently. The yolov3 model was trained using number of different arguments. The machine learning approach has been used using python programming language in this project. By implementing this dataset in this project this is found that the prediction criteria or ratio of sword detector vary from 0.83 to 1.00 with a background FN. The accuracy of YoloV3 is 0.83 and accuracy of YoloV5 is 0.60 which means that YoloV3 is performing better in detection. The first limitation is that when the algorithm detects a sword, it says that the sword is touching another sword and it says two swords are touching each other. It is not sure why this error happens. To fix this, an image will be added to the training dataset with a sword touching another sword and see if the algorithm will detect it. Another limitation is that when there are two objects in front of each other, such as in the image above. It detects one of them but not both. This project can be further developed in the future to detect different types of harmful objects that can be dangerous.

References

Acharjya, P. P., Koley, S. and Barman, S. (2022). A review on forensic science and criminal investigation through a deep learning framework, *Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks* pp. 1–72.

Adarsh, P., Rathi, P. and Kumar, M. (2020). Yolo v3-tiny: Object detection and recognition using one stage improved model, *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, pp. 687–694.

Anitha, R. et al. (2021). An efficient region based object detection method using deep learning algorithm, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* **12**(11): 5909–5916.

ATİK, M. E., DURAN, Z. and ÖZGÜNLÜK, R. (n.d.). Comparison of yolo versions for object detection from aerial images, *International Journal of Environment and Geoinformatics* **9**(2): 87–93.

Brahmaiah, M., Madala, S. R. and Chowdary, C. M. (2021). Artificial intelligence and deep learning for weapon identification in security systems, *Journal of Physics: Conference Series*, Vol. 2089, IOP Publishing, p. 012079.

Chou, C.-H., Su, Y.-S., Hsu, C.-J., Lee, K.-C. and Han, P.-H. (2020). Design of desktop audiovisual entertainment system with deep learning and haptic sensations, *Symmetry* **12**(10): 1718.

Foa, A. (2019). Object detection in object tracking system for mobile robot application.

Giuffrida, G., Meoni, G. and Fanucci, L. (2019). A yolov2 convolutional neural network-based human–machine interface for the control of assistive robotic manipulators, *Applied Sciences* **9**(11): 2243.

Hoang, L. (2019). An evaluation of vgg16 and yolo v3 on hand-drawn images.

LEE, D., CHOI, S., LEE, J. and CHUNG, W. (2022). Efficient seismic numerical modeling technique using the yolov2-based expanding domain method, *JOURNAL OF SEISMIC EXPLORATION* **31**: 425–449.

Liu, C., Wu, Y., Liu, J. and Sun, Z. (2021). Improved yolov3 network for insulator detection in aerial images with diverse background interference, *Electronics* **10**(7): 771.

Oki, Y., Uehara, K., Mizonobe, K., Akasaka, H., Shiota, Y., Sakamoto, R., Harada, A., Kitatani, K., Yabuuchi, T., Miyazaki, S. et al. (2021). Plan comparison of prostate stereotactic radiotherapy in spacer implant patients, *Journal of Applied Clinical Medical Physics* **22**(9): 280–288.

Redmon, J. and Farhadi, A. (2018). Yolov3: An incremental improvement, *arXiv preprint arXiv:1804.02767* .

Salido, J., Lomas, V., Ruiz-Santaquiteria, J. and Deniz, O. (2021). Automatic handgun detection with deep learning in video surveillance images, *Applied Sciences* **11**(13): 6085.

Silva Pincay, P. A. (2019). *Implementing yolo algorithm for real time object detection on embedded system*, B.S. thesis, Universidad de Investigación de Tecnología Experimental Yachay.

Teng, S., Liu, Z. and Li, X. (2022). Improved yolov3-based bridge surface defect detection by combining high-and low-resolution feature images, *Buildings* **12**(8): 1225.

Wang, K., Li, X., Yang, J., Wu, J. and Li, R. (2021). Temporal action detection based on two-stream you only look once network for elderly care service robot, *International Journal of Advanced Robotic Systems* **18**(4): 17298814211038342.

Warsi, A., Abdullah, M., Husen, M. N., Yahya, M., Khan, S. and Jawaid, N. (2019). Gun detection system using yolov3, *2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, IEEE, pp. 1–4.

Zhou, Z., Zhang, J. and Gong, C. (2022). Automatic detection method of tunnel lining multi-defects via an enhanced you only look once network, *Computer-Aided Civil and Infrastructure Engineering* **37**(6): 762–780.