Research Proposal: Pre-Shot Sniper Detection

Jan Meijers
Studentnummer : 836897315

Arne Lescrauwaet
Studentnummer : 852617312

Abstract

Enemy snipers are a growing threat to army forces worldwide. Currently no anti-sniper system exist that provides a seamless response to a so-called pre-shot sniper (i.e. detection of a sniper before a sniper weapon is fired) detection. This research contributes to the design of a prototype YOLOv8 based Pre-Shot Sniper Detector and the development of a non-existent high quality custom sniper dataset to assist Ministry of Defence in saving more lives of soldiers and civilians.

Introduction

Background

Enemy snipers are a growing threat to army forces worldwide. Snipers, armed only with a rifle, have rapidly become one of the deadliest threats (NBC news, 2006). As the war in Ukraine continues, sniper teams hunting down and killing Russian soldiers at a time from a place that can't be seen and with a sound that can't be heard. These so-called "sniper ghosts" do have psychological effect on their enemy (BBC news, 2023).

A sniper is a military/paramilitary marksman who engages targets from positions of concealment or at distances exceeding the target's detection capabilities. Snipers generally have specialized training and are equipped with high-precision rifles and high-magnification optics, and often also serve as scouts/observers feeding tactical information back to their units or command headquarters (Wikipedia.org).

Snipers had played an important role, especially in (urban) warfare, where the challenges for snipers and counter-snipers factored heavily in battles. The use of snipers has changed with technological development since the end of World War II. Sharp shooting in contemporary technology-intensive wars requires pro-activeness and has also commenced the development of systems for finding the counter-snipers positioned against snipers (Barişik & Baltacıoğlu, 2014).

Problem Statement

Several initiatives by Military of Defence (MOD) have led to the design of many (also commercially) available anti-sniper systems, merely based on radar, infrared (thermal) and sound signaling processing. The majority of these systems however detect snipers after a sniper weapon is fired. Currently no anti-sniper system exist that provides a seamless response to a so-called pre-shot sniper i.e. detection of a sniper before a sniper weapon is fired. In accordance to (US Ministry of Defense Patent, 2015), MOD strives to equip soldiers with effective and overall anti-sniper detection capabilities. For this purpose and application, MOD decided to study the feasibility of the latest version of YOLO (You Only Look Once) algorithm, which is in general a single neural network that predicts bounding boxes and class probabilities directly from full images in one evaluation (Redmon, Divvala, Girshick, & Farhadi, 2016).

Literature review

Because of security reasons little is published by MOD concerning anti-sniper systems. A research field which is most closely related to the anti-sniper detection domain - using YOLO computer vision technology - is weapon detection in civil surveillance domain.

Weapon detection in civil surveillance, using Artificial Intelligence (AI), has gained significant attention in recent years, and numerous research studies have been conducted in this area. Computer vision technology has been used for real-time weapon detection, and YOLO object detection algorithm has emerged as a popular and efficient technique for this purpose, but further research and development are needed to improve the accuracy and reliability of weapon detection systems to ensure their responsible deployment in real-world settings Murugan et al. (2023).

An another important and crucial part of any application is to have a desired and suitable datasets in order to train the machine learning models (Narejo, Pandey, Esenarro Vargas, Rodriguez, & Anjum, 2021). Yadav, Gupta, and Sharma (2023) evaluated 8 different public available weapon detection datasets originated from 2016 to 2021 that can be used to classify and weapons. They conclude there are still certain challenges in the field of weapons detection that need to be addressed, such as a lack of (large and well-balanced) datasets, the detection of weapons in a variety of lighting conditions as well as that only a few researchers have tackled the subject of partial occlusion of weapon. Another significant problem is the capability of detecting different kinds of weapons and the detection of small sized weapons.

Murugan et al. (2023) remarks in addition that there is no standard dataset for weapons detection and recognition. They do emphasize the importance of quantity and quality of the training data to achieve high accuracy. They trained a YOLO model using a dataset of 17k images annotated weapon images with bounding boxes containing more than 8 classes i.e. firearms, knives and other weapons. They experienced to achieve higher accuracy to remove irrelevant object in each weapon image. Also, to use these techniques for weapon detection, it is necessary to obtain a dataset of images that includes a variety of different type of weapons in various contexts and in different lightning conditions and backgrounds according to the content. The dataset should also include images of non-weapons to help the algorithm learn to distinguish weapons and other objects.

Performance metrics are key tools to evaluate the accuracy and efficiency of pre-shot sniper detection models. Some metrics that are not only important to YOLOv8 but are broadly applicable across different object detection models as noted by Ultralytics Performance Metrics (2023) i.e.:

1. Intersection over Union (IoU): IoU is a measure that quantifies the overlap between a predicted bounding box and a ground truth bounding box. It plays a fundamental role in evaluating the accuracy of object localization.

2. Average Precision (AP): AP computes the area under the precision-recall curve, providing a single value that encapsulates the model's precision and recall performance.

3. Mean Average Precision (mAP): mAP extends the concept of AP by calculating the average AP values across multiple object classes. This is useful in multi-class object detection scenarios to provide a comprehensive evaluation of the model's performance.

4. Precision and Recall: Precision quantifies the proportion of true positives among all positive predictions, assessing the model's capability to avoid false positives. On the other hand, Recall calculates the proportion of true positives among all actual positives, measuring the model's ability to detect all instances of a class.

5. F1 Score: The F1 Score is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance while considering both false positives and false negatives.

The most common metric used for object detection models like YOLO is the mean average precision (mAP). This metric is suitable for a broad assessment of model performance. This metric considers the precision and recall of the model's predictions and provides a single numerical score for the model's performance (Terven & Cordova-Esparza, 2023). In addition to evaluating the performance of the model on the test set, it is important to analyze the types of errors the model makes to understand where improvements can be made. This can be done by analyzing the confusion matrix, which shows the number of true positives, false positives, true negatives, and false negatives for each class (Bhatti, Khan, Aslam, & Fiaz, 2021).

Knowlegde Gaps

While there is a growing body of knowledge in the field that is described in the literature review section, there are some knowledge gaps that this research can contribute to.

The literature review reveals that much research on detecting weapons uses public (labelled) datasets with images of all kinds of fire weapons (i.e. guns, pistols, rifles), which are -in general- not very suitable for the pre-shot detection problem. Mainly because they do not represent typical and substantial sniper rifles and - most importantly - they do not represent parts (partial occlusion) of a sniper rifle. This, however, is a very important aspect of an anti-sniper detection system, because sniper and/or sniper rifles are mostly not visible as a whole.

The main problem with snipers is the gap or challenge to effectively detect real-time enemy snipers before they (can) fire their deadly weapons and kill soldiers and/or civilians instantly. This is especially true in rural environments but also in urban environments.

Snipers' most important equipment is the sniper weapon i.e. sniper rifle. The most obvious pre-shot detectable target(s) snipers is most probably in detecting uncamouflaged (parts of a) sniper equipment i.e. the weapon itself, in particular part of a sniper rifle i.e. the rifle barrel of the weapon, any rifle barrel muzzle break, rifle suppressors/silencers and/or in most cases any (part of a) riflescope. Sniper rifles almost naturally come installed with the attributes like barrel, muzzle break, suppressors and/or riflescope. A sniper rifle without a riflescope is not less a sniper rifle, but less effective. Also, detecting a sniper rifle in military or hostile circumstances almost inherently detects or expects a sniper presence and/or sniper location as well.

Detecting pre-shot snipers based on the presents of sniper rifles is challenging. There is a lack of useful and purposefully labelled images, which can contribute to the detection of pre-shot snipers using computer vision methods. Although many public images datasets do exist, none of them consist of specific and/or applicable labelled images in detecting (parts of) sniper rifles such as barrel, muzzle break, suppressors and/or riflescopes.

Research Design and Methods

Research Goal

In this research, MOD aims to construct a prototype pre-shot sniper detection and a high quality custom sniper dataset, to extend their knowledge of how to effectively detect snipers before a sniper rifle is fired using state of the art detection algorithms, with a minimum goal of > 85% mean average accuracy (mAP) in testing. The results of this research (including performance measurements) are reported in so-called MOD protocol for future anti-sniper system studies and developments. For this purpose we formulate the following main research question:

How to construct a YOLOv8 based pre-shot sniper detection prototype and a high quality custom sniper image dataset that satisfies a fast and accurate detection (i.e. classification and localization) of snipers before a sniper rifle is fired, so that army forces can save more lives of soldiers and/or civilians in the growing threat of enemy snipers in different rural and urban backgrounds and/or battle field conditions?

Research Objectives

To investigate the main research question the following sub research questions are relevant:

1. What is the minimum number of images required for training the pre-shot anti-sniper detection system to a predetermined threshold?

2. What is the influence of the variability of the background? Is the sniper easier to notice in certain backgrounds?

3. Does it matter how tight the bounding box is? Should one spend the extra time to get the bounding boxes just right, or is it better to annotate more images, but care less about the bounding box?

Data Collection

Selecting Pre-Shot Sniper (Rifle) Images

We concluded that custom sniper datasets for our research purposes are not publicly available. Because of this we need to build our own custom sniper dataset. Images to extend (i.e. fine tuning) pre-trained YOLO models will be collected from Internet. Useful and purposeful images must be collected and selected by hand. Selecting these images will be time-consuming. For our experiments about 100-600 images will be initially selected for labeling and answering research sub questions purposes. Depending on the performance measurements results of the model more images will be selected and labelled.

Labeling/Annotate Pre-shot Sniper (Rifle) Images

Selected images will be labelled by hand using internet available free tools like Roboflow and annotating these images with the corresponding object classes will be an intensive and time-consuming part, but is an important step in reaching quality of the custom pre-shot sniper dataset. For this labeling purpose we are planning to use only one classification class i.e. "Snipers", incorporating any comprehensive aspect like distance, environment conditions (surroundings, weather, lightning), occultation, sniper position, scale size and/or camouflage.

Training and Validating

The constructed labelled custom sniper dataset will be divided into training (80%), validation(10%) and test (10%) datasets. For specific and practical YOLOv8 issues and training settings we will make use of publicly available guidelines, best practices, GitHub sources and/or YouTube sources to assist our experiments as wel as build our knowledge.

Methodology

This section presents the methods that will be used in order to answer the research (sub) questions and to reach the goal of this research.

Design Science Research

This section presents the methods that will be used in order to answer the research (sub) questions as listed in section 1.4 and to reach the goal of this research.

The core idea is to propose a prototype Pre-Shot Sniper Detector based on a newly designed high quality custom pre-shot sniper detection dataset. For this, empirical evidence is needed of the working of different methods to undertake high accurate detection of pre-shot snipers in rural and urban circumstances. This calls for the design of an artefact that embodies the knowledge needed by MOD and can be used in practice by MOD. For this science research it is chosen to use the design science research methodology, as this methodology allows for the design of a MOD protocol for training Pre-Shot Sniper Detection models and datasets suited for application in practice. In the process of designing and evaluating artefact, knowledge is gained on what works and why, thus contributing to theory (Hevner, MS, & Ram, 2004). The relation between context, problem, artefact, knowledge gap and knowledge base is schematically depicted in figure 1.

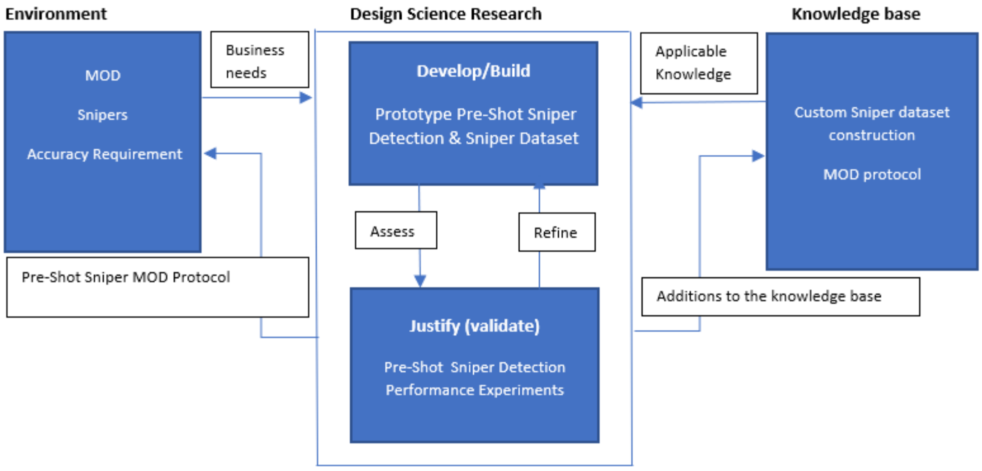


Figure 1. Detecting pre-shot snipers applied to the design science framework of Hevner et al. (2004).

In this research we refer to this artefact as a Pre-Shot Sniper Detector. This research will deliver a construction of a high quality custom sniper dataset, and the knowledge that was required to come up with the construction and its evaluation. Peffers, Tuunanen, Rothenberger, and Chatterjee (2007) describes six elements that in sequence form a design research methodology:

1. Identify the problem to be solved
2. Define the objectives for an artefact that helps solving the problem
3. Design and develop an artefact
4. Demonstrate the working of the designed artefact
5. Evaluate how well the artefact contributes to solving the problem
6. Communicate about the first five steps.

The six and final action is to communicate about the first five steps, thus sharing the knowledge derived from the design process. This communication will be done by the research report (a.k.a. MOD protocol).

Training and Expirements

There are five models in each category of YOLOv8 models for detection, segmentation, and classification, namely YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l and YOLOv8x. YOLOv8n (Nano) is the fastest and smallest, while YOLOv8 Extra Large (YOLOv8x) is the most accurate yet the slowest among them. Also for object detection YOLOv8 comes bundled with Common Objects in Context (COCO) dataset pre-trained models. The COCO dataset consists of 80 object categories (*https://learnopencv.com/ultralytics-yolov8/*, 2023).

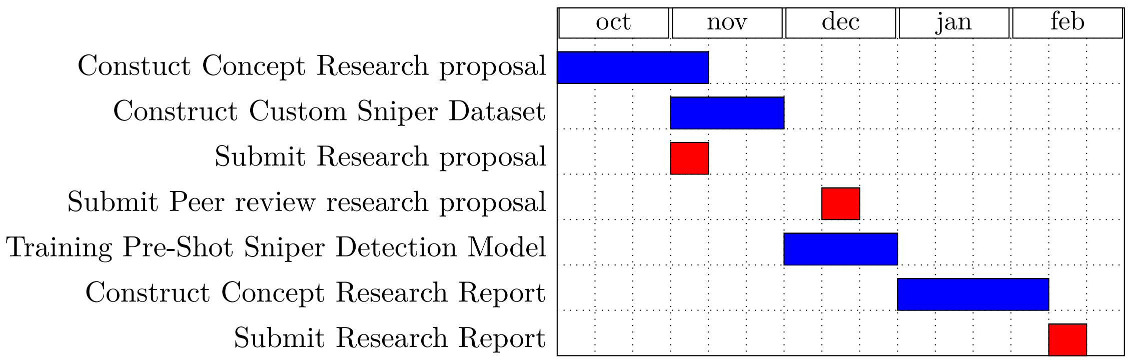
In order to speed up the process of learning to detect snipers, fine-tuning will be applied to the YOLOv8m(edium) weights that are pretrained on the COCO dataset. The model will be trained with the standard values according to the documentation Ultralytics Documentation (2023) i.e.: 100 epochs with a confidence of 0.001 and image size of 640. If time allows we will compare performance measures with the small version of pre-trained YOLOv8. Other version are not applicable because of limited capabilities of our local GPU machine for training purposes.

The model will first be fine-tuned on 100 images, this amount will be incremented by 50 images until we reach MOD goal of $> 85\%$ mean average accuracy (mAP) in testing. With every updated working dataset, we will start finetuning from the baseline YOLOv8medium weights. Sources and results of training, validation, testing and experiments will be archived at GitHub for demonstration purposes, evidence and further research.

Inference

After training, if project time allows, we will also run inference on videos to check the real-world performance of these models. This will give us a better idea of the best model. To test our model in a scenario as close as we can get to the real context, we will run inference on a YouTube video about snipers. The Yolov8 model can run inference on online media by passing the url to the 'source' argument.

Research Project Plan 2023-2024 Time Frame



References

Barişik, E., & Baltacıoğlu, G. (2014). The employment of sniper in modern battlefield. *Journal of Management and Information Science*, *2*(1), 13–19.

BBC news. (2023). `https://https://www.bbc.com/news/world-europe-66354363`.

Bhatti, M. T., Khan, M. G., Aslam, M., & Fiaz, M. J. (2021). Weapon detection in real-time cctv videos using deep learning. *IEEE Access*, *9*, 34366–34382.

Hevner, A., MS, P. J., & Ram, S. (2004). Design science in information system research. *MIS Quarterly*, *28*(1), 75–105.

Murugan, P., Merry ida, A., Aashika, R., Amshiga Brillian, S., Seals Ancy, M., & Sneha, S. (2023). Weapon detection system using deep learning. *https://ijirt.org/master/publishedpaper/IJIRT160929_PAPER.pdf*.

Narejo, S., Pandey, B., Esenarro Vargas, D., Rodriguez, C., & Anjum, M. R. (2021). Weapon detection using yolo v3 for smart surveillance system. *Mathematical Problems in Engineering*, *2021*, 1–9.

NBC news. (2006). `https://www.nbcnews.com/id/wbna15336981`.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, *24*(3), 45–77.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 779–788).

Terven, J., & Cordova-Esparza, D.-M. (2023, 04). A comprehensive review of yolo: From yolov1 to yolov8 and beyond.

Ultralytics Documentation. (2023). `https://https://docs.ultralytics.com/usage/python/`.

Ultralytics Performance Metrics. (2023). `https://docs.ultralytics.com/guides/yolo-performance-metrics/`.

(2023).

US Ministry of Defense Patent. (2015). `https://patents.google.com/patent/US9488442B2/en`.

Yadav, P., Gupta, N., & Sharma, P. K. (2023). A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods. *Expert Systems with Applications*, *212*, 118698.