

Artificial Intelligence &
Machine Learning
(623.504) - Term Project
Proposal

Personal Prospective
Equipment Detection

Mustafa Tayyip BAYRAM – 12237686

Table of Contents

1. Motivation	1
2. Learning Task.....	1
Performance Measure:.....	2
3. Related Work.....	2
4. Plan	3
5. Risk Management.....	3

1. Motivation

Personal Protective Equipment (PPE) like as helmets (hard hats), safety harnesses, goggles, and other similar items serve a significant role in assuring worker safety in industry, particularly in the manufacturing business. However, numerous accidents continue to occur as a result of workers' and managers' carelessness. Supervisors may make mistakes because such activities are repetitive, and they may not be able to supervise regularly. This initiative attempts to use current camera equipment to help supervisors watch workers more effectively by giving real-time notifications. Violations on this issue can be avoided if it is utilized in this manner in real time.

2. Learning Task

Learning Task will be Object Detection in Real Time. It must me Real Time to running trigger concurrently in the violation situation. Therefore, YOLO (You Look Only Once) architecture is appropriate for this purpose. While choosing the version experientially, versions 5 and 7 will be experimented.



```
1 0.5961538461538461 0.3293269230769231 0.10576923076923077 0.11538461538461539
1 0.31971153846153844 0.3798076923076923 0.07932692307692307 0.09375
1 0.5973557692307693 0.03125 0.10817307692307693 0.0625
```

Suitable data for YOLO is founded on the internet to avoid the hassle of labelling, helmet detection can be handled with it. The data consists of the image and the label that shows the object in which location. This data can be solving the problem, but it would be better to find images close to camera view. Because cameras will be using for the detection, perspective and distance should be similar.

Performance Measure:

Two metrics are critical for identifying PPE. FPS is crucial since it is measured in real time, and Mean Average Precision (mAP) will be used to assess it. It is, in fact, common in Object Detection. The appropriate model (YOLO) version will be chosen based on the FPS appropriateness of each mAP.

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k$$

AP_k = the AP of class k
 n = the number of classes

Figure 1. Mean Average Precision (mAP) for Object Detection

Because it will be Real Time Object Detection, some aspects must be sacrificed, such as accuracy. It is usual to obtain fewer results than with standard Object Detection methods. Essentially, the important criteria will be capturing the optimal value between these two.

According to my study, an aim of 80% mAP and 35FPS is achievable.

3. Related Work

[1] Deep learning for site safety: Real-time detection of personal protective equipment:
<https://www.sciencedirect.com/science/article/pii/S0926580519308325>

[2] PPE Detection using yolo3 and DeepSORT:
https://github.com/AnshulSood11/PPE-Detection-YOLO-Deep_SORT

These two resources are about YOLO PPE Detection. Almost identical architectures are intended to be used in new versions. YOLOv3 is utilized in these resources, while v5 or v7 are better architectural solutions. I believe the data is underfitting, and the data should be more comparable to the camera

perspective, not simply photographs in both of them. The paper's optimal result is 72.3% mAP, which might be increased with the new design and data correction. Both of the results' suggested baselines can be utilized.

4. Plan

First thing to do is to finding proper data in an appropriate amount like more than 1500 images as YOLO format labelled. Priority is Hard-Hat but also a few other PPE equipment data should be found such as google, vest etc.

Learn more about the YOLO architecture and begin selecting versions. There are several types of YOLO versions that are utilized for various purposes. YOLOv7 has less resources than v5, yet it is certainly a superior alternative. They are further subdivided into sub-branches such as v5X and v5s. According to the camera perspective images, v5s will most likely produce the best results in terms of fps and mAP.

5. Risk Management

In most circumstances, I believe the main threat is the data. Is the data we have, correct? Is it sufficient?

Data is gathered, however further labeled data should be discovered from other viewpoints including using different equipment to provide better findings. It might be asked from paper writers in order to avoid dealing with labeling.

For modelling, YOLO architecture seems best option but as a backup option, faster R-CNN and SSD algorithms can be used.