Safety Monitor
(using Deep Learning on the Edge)

W251 Project by

Silas Everett
Andrew Larimer
Sudha Subramanian

Berkeley
UNIVERSITY OF CALIFORNIA

We value workers' safety …

Companies and businesses strive hard to make their workplaces safe for its workers, especially in the Construction Industry. Occupational Safety and Health Administration (OSHA) recommends practices to be followed by employers and among those, the first is '***Prevent workplace injuries and illnesses***'. Our tool provides a way for employers to monitor safety at their site.

Berkeley
UNIVERSITY OF CALIFORNIA

How can businesses ensure that their workers protect themselves on the job?

Tip

Businesses have to do what it takes to ensure **Workers' Safety.**

A simple way is to monitor the workers wear safety equipments, while on the job.

How can we leverage advances in **AI and Deep Learning** to offer a solution to this problem?

Workers' Safety

Personal Protective Equipment (PPE) refers to what is worn on the job by construction workers in order to minimize exposure to hazards, injuries and illnesses. These include:

- **Safety Vests**
  What the worker needs to wear, while on job.
- **Hard Hats**
  Protective headgear to be word on the job.
- Safety Glasses, Boots etc.
  Other protective equipments

Statistics

OSHA) reports that nearly 6.5 million people work at approximately 252,000 construction sites across the nation on any given day.

Safety Monitor is the answer.

(Indigenously developed by Cal Students)

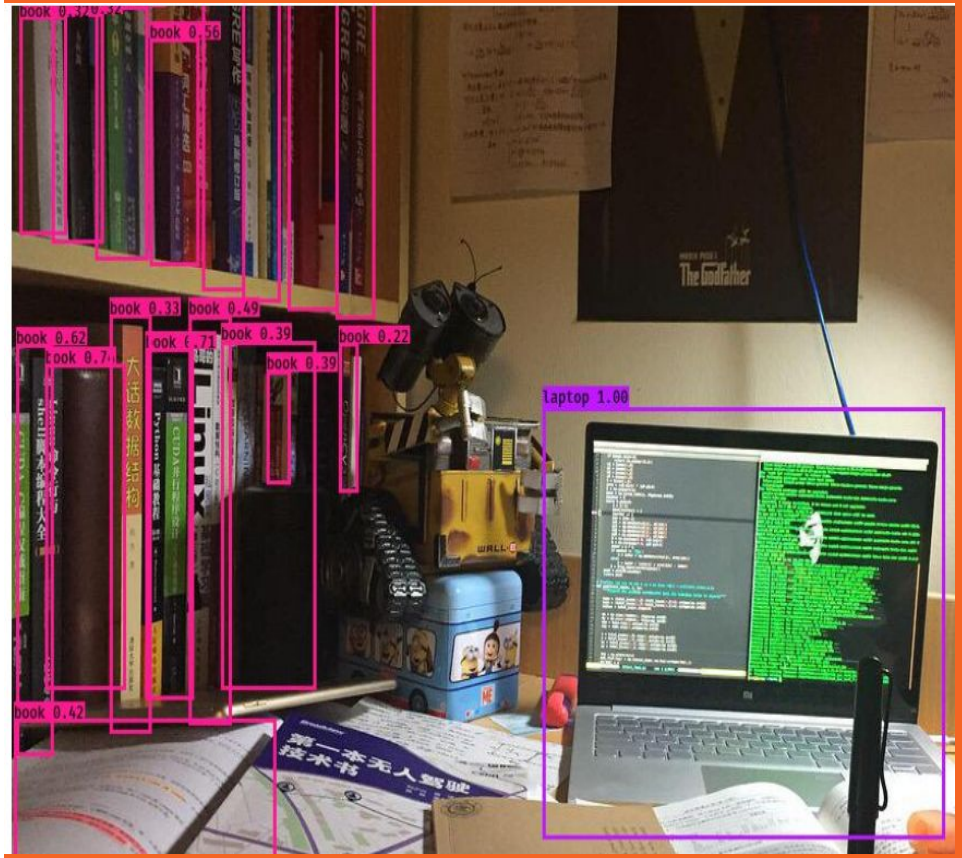
This is a proof-of-concept to show how this tool can be deployed at sites as a way to monitor workers' safety by getting feeds from cameras on sites.

Did you know...

A recent AI-based analysis found 17000 people and 1000 PPE instances of non-compliance due to not wearing safety vests and hard hats.

Link to source

The model is built using **Object Detection using YOLO** and deployed on the EDGE for making predictions.



Dataset

Videos of construction sites used to prepare dataset for training:

- ➔ **Different environments**
  Video recordings of work at multiple construction sites.
- ➔ **Under different lighting**
  Images included under different lighting conditions (day and night).
- ➔ **Distance from camera**
  Data included recordings wherein the images were both close or farther away from camera.

Statistics

Total images: 1009

Objects of interest:

Person, Safety Vest, Hardhat

(included frames with no objects of interest)

Annotation

Images extracted from videos were annotated manually using CVAT

➔ **Run TF Annotation**
Facilitated by automatically annotating persons in all frames, which can further be fine-tuned if needed.

➔ **Manual Annotation**
Annotate manually using CVAT tool; dump annotation details as xml.

➔ **Convert to Pascal VOC Format**
Aggregated annotations from each of our work; converted to PascalVOC format using script.

Statistics

In total, we annotated 1009 images, which included:

Persons: 4421

Safety Vests: 2553

Hardhats: 1908

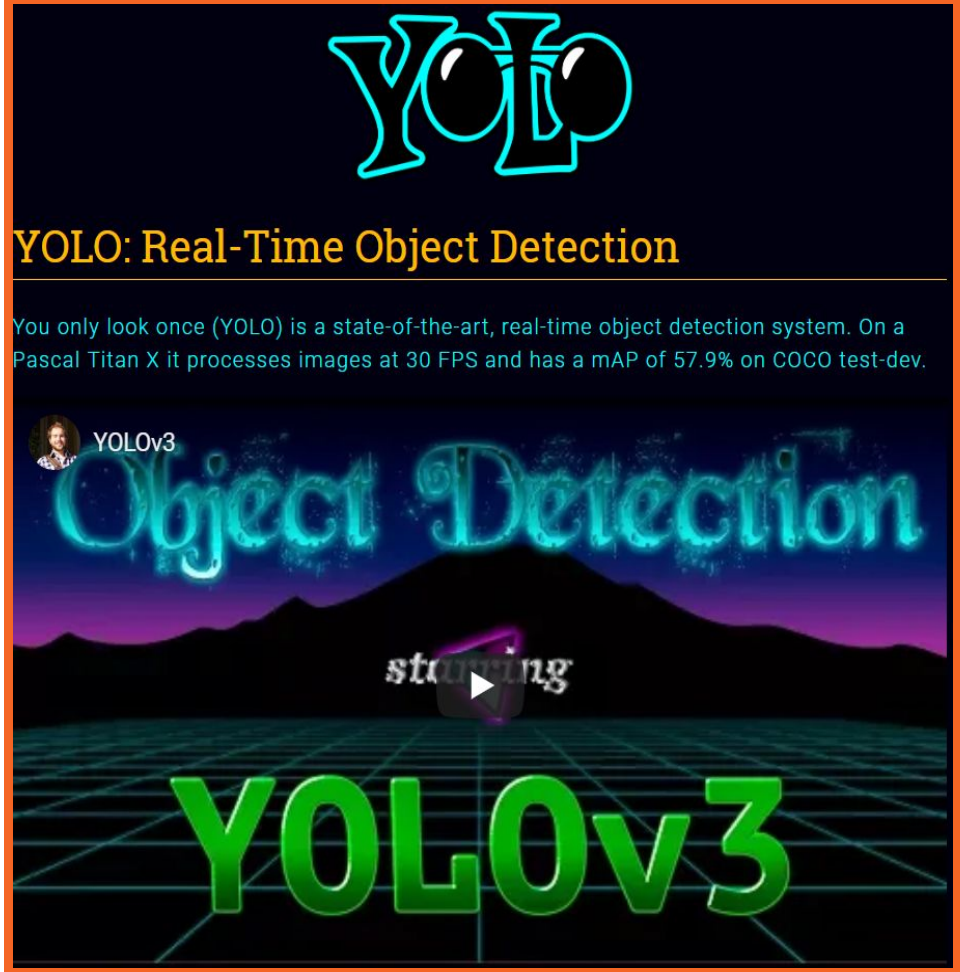
Annotation Dump


```xml
<?xml version="1.0"?>
<annotation>
    <folder>HardHat-Related</folder>
    <filename>Sequence01_0.jpeg</filename>
    <path>C:\Users\sudha\Documents\W251-DL-O
  - <source>
        <database>Unknown</database>
    </source>
  - <size>
        <width>640</width>
        <height>360</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
  - <object>
        <name>person</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
      - <bndbox>
            <xmin>515.40</xmin>
            <ymin>98.90</ymin>
            <xmax>535.20</xmax>
            <ymax>131.58</ymax>
        </bndbox>
    </object>
  - <object>
        <name>safetyvest</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
      - <bndbox>
            <xmin>521.94</xmin>
            <ymin>102.49</ymin>
            <xmax>533.53</xmax>
            <ymax>116.44</ymax>
        </bndbox>
    </object>
</annotation>
```

Train the model

using YOLOv3

(Darknet YOLO → Keras)



Anchor Boxes

Anchor Boxes needed to be defined for better object detection, as the objects to be identified were of different sizes.

➔ **EDA performed**

EDA helped to understand the size of the images of interest.

➔ **Understand objects to detect**

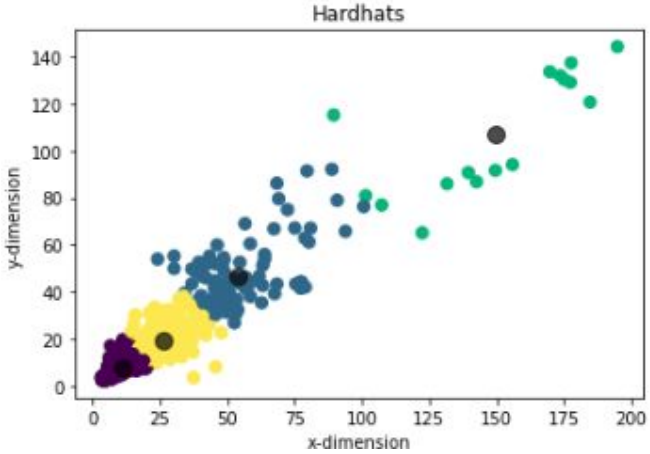
Depends on the size of the objects to detect. Hardhats had the smallest dimensions / sizes.

➔ **KMeans for bounding box**

YOLOv3 uses KMeans to estimate bounding boxes. We performed clustering (image on right) to estimate size of anchor boxes

Statistics

In total, we used 9 anchor boxes, targeted at identifying 3 different objects (Person, Safety Vest and of different sizes



Train on Cloud

Model Training on IBM Cloud by spinning p100 and setting up the environment.

➔ **Image Augmentation**
Images were flipped to get additional training data.

➔ **Padding**
Images were padded to meet size, as required for YOLO training.

➔ **Metric: mAP**
Mean Average Precision computed using tool

Statistics

Trained for:

- 64 epochs
- P100 GPU
- Metric: mAP.

mAP of Model

- Person: 75.85%
- Safety Vest: 67.09%
- Hardhat: 58.92%

- **mAP = 67.28%**

mAP Calculation

Statistics

Generate mAP

Screen record of mAP evaluation using:

https://github.com/Cartucho/mAP

Legend

Green - Ground Truth

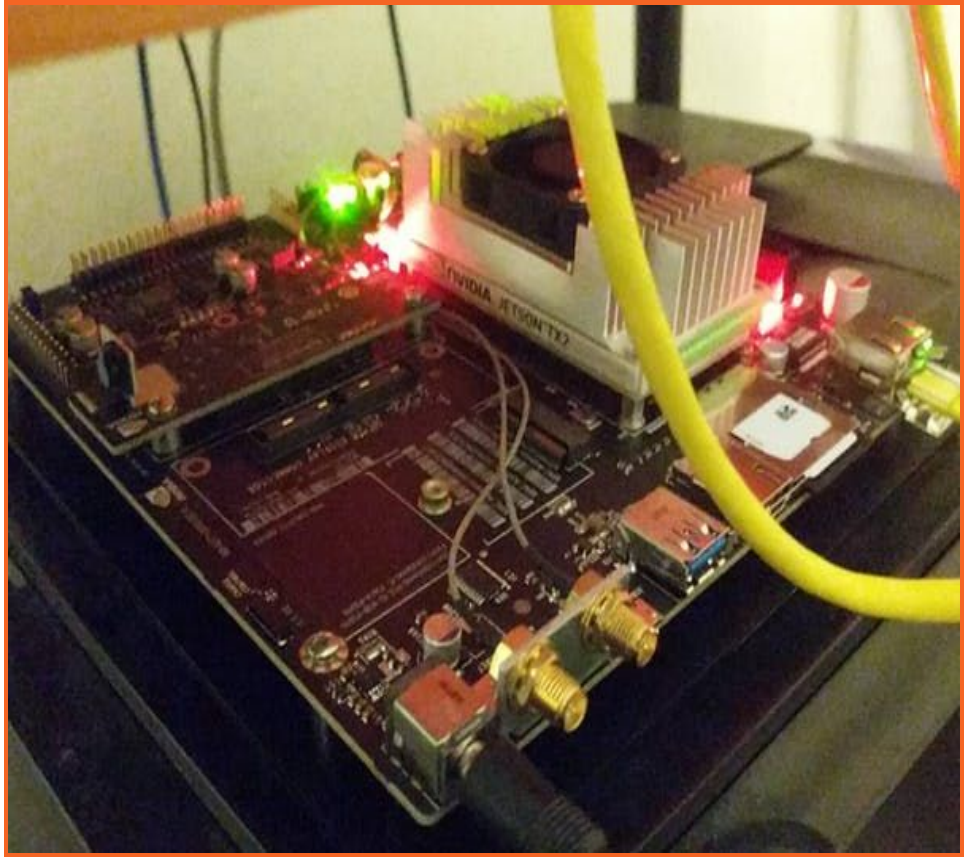
Blue - Model Prediction

Red - Not predicted

Prediction time!

Model on the Edge on TX2

(Get model weights to TX2 and run predictions)



Prediction

Weights downloaded to TX2 and ready to process video stream from webcam to check on PPE compliance!

➔ **Input: Video Stream**

Frames extracted from the stream, say from a webcam and processed to extract images.

➔ **Process / Predict using YOLO Model**

Images are processed to obtain the model predictions and results (bounding box information) written to file.

➔ **Check for PPE Compliance**

Process and send alert if PPE compliance is not met for say, continuous 20 frames

PPE Compliance

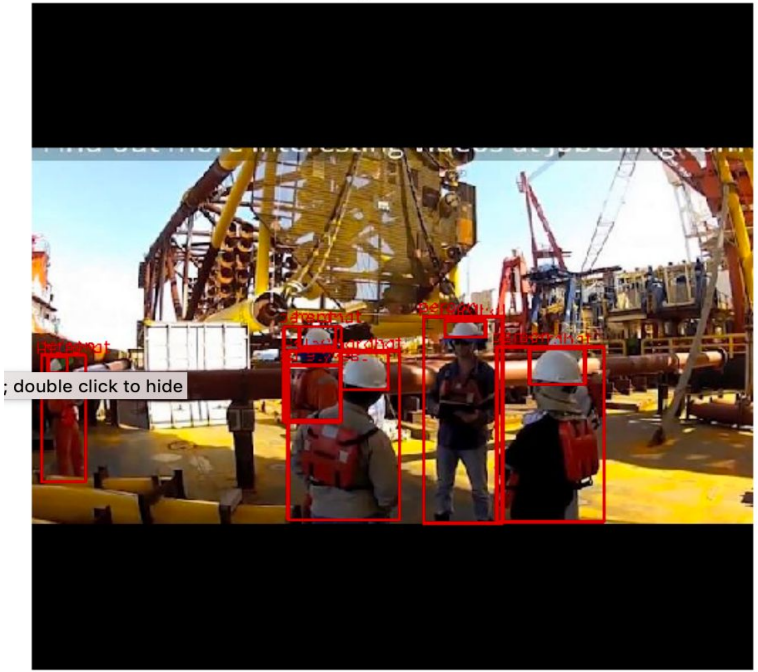
Compliant, if person has both safety vest and hardhat

Partially compliant, if person has either safety vest or hardhat

NOT COMPLIANT, if person has neither safety vest nor hardhat

Ground Truths vs. Model Predictions

Sequence14_732 Ground Truths

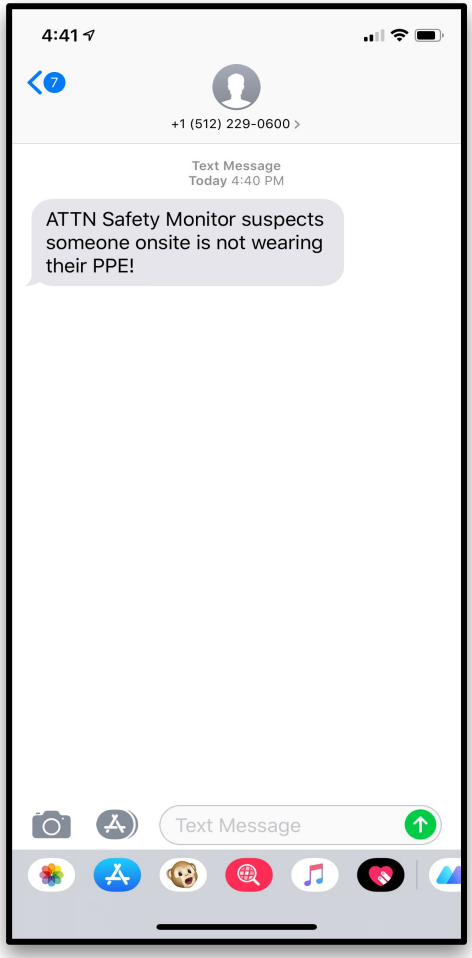


Sequence14_732 Predictions

Alert the Supervisor

Configure to ALERT Supervisor at the field location when it is evident that workers on the field are not PPE compliant. And, we have verified this by sending text messages to our phones.

- ➔ **Input: Video Stream**
  Frames extracted from the stream, say from a webcam and processed to extract images.
- ➔ **Process / Predict using YOLO Model**
  Images are processed to obtain the model predictions and results (bounding box information) written to file.
- ➔ **Check for PPE Compliance**
  Process and send alert if PPE compliance is not met for say, continuous 20 frames



Q&A time!

THANK YOU

Check it out on GitHub

AWESOME COURSE (W251)

DL on the EDGE is cutting-edge and state-of-the-art stuff

Great Exposure to various tools and technologies to explore in real world

GREAT PRODUCT from this PROJECT that can be implemented in fields / sites and there is currently no tool out there that does this!

Thanks, Ryan & Brad