**Project Overview - Job Site Safety**

**Authors: Jason Raimondi,** [Jeremy Cryer](mailto:jcryer@sandiego.edu)**,** [Maimuna Bashir](mailto:mbashir@sandiego.edu)

**Business Name: PPE Vision Solutions**

**Publication Date: 1/20/2024**

**Github Project Link:** [**https://github.com/jerm914/Job-Site-Safety---AAI-540**](https://github.com/jerm914/Job-Site-Safety---AAI-540)

* **1 Notebook for setting up data lake**
  + [**https://github.com/jerm914/Job-Site-Safety---AAI-540/blob/main/Jason/00\_S3\_Datalake\_Jason.ipynb**](https://github.com/jerm914/Job-Site-Safety---AAI-540/blob/main/Jason/00_S3_Datalake_Jason.ipynb)
* **1 Notebook for setting up feature store**
  + [**https://github.com/jerm914/Job-Site-Safety---AAI-540/blob/main/Jason/01\_FeatureStore\_Jason.ipynb**](https://github.com/jerm914/Job-Site-Safety---AAI-540/blob/main/Jason/01_FeatureStore_Jason.ipynb)

**Asana Board Link:** [**https://app.asana.com/0/1206390072539638/1206390072539638**](https://app.asana.com/0/1206390072539638/1206390072539638)

**Project Background:**

1-2 paragraphs describing, at a high level, the problem you are trying to solve. This should provide the reader context for the technical solution they will review. This should be a quick elevator pitch for your project. Be sure to answer the following:

* What is the model’s objective?
* What type of Machine Learning problem will you be solving?

The ultimate objective of the project is to increase safety compliance, specifically adherence to wearing required Personal Protective Equipment (PPE) and reduce occurrences of serious injury on job sites. The objective of the model will be to detect, classify, and report instances of ‘Helmet’, ‘Vest’, and ‘Head’ from a provided input image. Instances of ‘Head’ may indicate absence of a required helmet being worn.

This is an object detection and classification Machine Learning problem to solve. First, the model will need to be able to sufficiently identify when these three types of objects are present in the image. Second, the model will need to be able to accurately classify the detected objects as the correct type or belonging to the correct category.

**Technical Background:**

1-2 paragraphs describing the technical details of the problem you are trying to solve. This should help the reader understand the project constraints. Be sure to answer the following:

* How will you evaluate your model?
* What is your data source?
  + How will you need to prepare your data?
  + How will you explore your data?
  + What do you hypothesize your main features will be?
* What type of model do you want to use?

To evaluate the model, a few different business metrics can be considered. First, the business can monitor and track instances of serious injury over time to see if the model implementation may be improving safety records (i.e., assuming business is intervening with employees when unsafe conditions are investigated). Second, the business can track instances of ‘Head’ being detected in a given time period, work shift, etc. to see if their interventions are having a positive impact (i.e., instances of ‘Head’ reduced over time in favor of instances of ‘Helmet’). Important model metrics include box loss (i.e., reduce error in object detection) and classification loss (i.e., reduce errors in classification).

The data source is from Kaggle:

HardHat-Vest Dataset

<https://www.kaggle.com/datasets/muhammetzahitaydn/hardhat-vest-dataset-v3/data>

The dataset contains 23,673 images, and there are at least 10,000 instances of each class available. Per Kaggle, the images should be consistently sized at 640x640 pixels, so resizing is not expected unless required for the model architecture. However, we plan to still validate the image sizing as claimed. Preparation should be minimal, assuming image annotations are of sufficient quality (expected to be, as the Kaggle usability rating is a 10.00). The dataset should be explored, minimally, by viewing sample images to gain an understanding of contents, considerations, and annotation quality. Bounding boxes will need to be applied to the images to inspect the quality and accuracy. Model features are hypothesized to be constructed via multiple Convolutional Neural Network (CNN) layers to obtain lower-level features and then higher-level features to ultimately support the classification task after detection. An initial suggestion would be to use a pre-trained YOLOv8 model architecture for this project, with fine-tuning for the specific dataset to improve the performance initially and to keep maintainability and adaptability in mind for ongoing use.

**Goals vs Non-Goals:**

Write a bulleted list of (3-5 points each) of goals and non-goals. This should help the reader understand the context that would factor into solution selections and trade-offs. Goals will help the reader understand what a successful outcome looks like. Non-goals will help limit the scope of your project and prevent scope creep.

Goals:

* Reduced instances of serious injury on construction sites (Business Metric)
* Increased safety compliance (wearing required PPE) (Business Metric)
* Low box loss (specific amount to be determined w/ company) (Model Metric)
* Low classification loss (focus on ‘Helmet’ and ‘Head’) (Model Metric)

Non-Goals:

* Resolving annotation issues / bounding boxes (if issues identified during exploration, discard from dataset)
* Model metrics for ‘Vest’ category (not the primary focus of this project to maintain simplicity)
* Use of multiple model architectures (also for simplicity)

**Solution Overview**

1-2 paragraphs summarizing your ML System. Include a system architecture diagram containing the components you use to store data, pre-process data, engineer features, build/train/debug your model, and deploy your model. Also add notes on what you will monitor and what you will test prior to releasing a new model.

**Data Sources:**

What is your data source?

Kaggle - HardHat-Vest Dataset

<https://www.kaggle.com/datasets/muhammetzahitaydn/hardhat-vest-dataset-v3/data>

What is your data volume?

23,673 images.

Why did you select this data set?

Our team has interest in the business problem that this dataset can help solve, and this dataset appears to be sufficient to support development of a solution to this problem. There are at least 10,000 instances of each class, including ‘Helmet’, ‘Vest’, and ‘Head’, which should provide a sufficient number of samples for our computer vision task.

Any risks (bias, sensitive features, etc)?

We are not aware of any bias or sensitive features present at this time. However, we will, minimally, inspect image samples and predictions to determine if there could be any considerations. For instance, we would want to make sure that we understand if performance may vary when detecting and classifying objects when considering different genders, race, skin tones, hairstyles, etc. Another risk we need to keep in mind is if there are any types of occlusions that could impact performance. If a worker is wearing sunglasses, a respirator, or if there is any other type of occlusion, does this affect the object being detected and/or classified correctly? This could also include other items such as ball caps and other items that are similar and may conflict with the PPE being detected in this problem. We need to be aware of any limitations such as this so that the solution can be advertised and implemented in suitable environments while providing us the necessary information to continually improve the solution.

**Data Engineering:**

How will you store this data?

What data pre-processing do you need to do before you feed it into your ML system?

We are anticipating minimal pre-processing of our data. The images are consistently sized at 640x640 pixels, so resizing is not expected unless required for the model architecture. Preparation should be minimal, assuming image annotations are of sufficient quality (expected to be, as the Kaggle usability rating is a 10.00). We are planning to store the raw data into S3 initially, and may look into loading image/label data into the feature store later on if time permits.

**Training Data:**

How will you split your data into training, test and validation?

We are planning to split the data using the recommendations in the final project instructions:

* ~40% Train
* ~10% Validation
* ~10% Test
* ~40% “Production Data”

Will you use any data labeling techniques?

We are not expecting to need to use any data labeling techniques. As mentioned, the Kaggle usability rating is high, and the images come with annotations. We will inspect some random samples of images, bounding boxes, and labels, however, we are not anticipating needing to make any adjustments for this dataset.

**Feature Engineering:**

Which fields from your data will you use or exclude?

We will be using class instances of ‘Helmet’ and ‘Head’ but excluding instances of ‘Vest’ in an effort to simplify the implementation and put more focus on the overall ML system.

Which fields will be combined or bucketed?

We are not planning to combine or bucket any fields. We are working on a computer vision problem using raw data and not a dataset with predefined fields.

What other data transformations will you apply to your data?

We are not performing any data transformations but are instead engineering new features derived from properties and contents of the raw data files.

**Model Training & Evaluation:**

How will you train your model?

We are currently in the process of developing and training our benchmark model.

What algorithm will you use?

We are currently looking into creating a PyTorch estimator to provide compatibility with our intended YOLO model deployment.

What parameters will you use?

We do not know which parameters we will use yet, however, we plan to use a minimal/baseline set of parameters for our benchmark then will adjust for further model iterations.

How will you evaluate your model?

We are planning to use the following metrics for evaluation and comparison between models:

* Box loss
* Classification loss (focus on ‘Helmet’ and ‘Head’)

**Model Deployment:**

What instance size will you use?

We plan to use a small instance size, as the AWS student environment does not have access to GPUs. We are prioritizing simplicity and computational efficiency to work within these constraints and focus efforts on the overall ML system.

Will your model function as a batch or real time model? Why?

Our model will function as a batch. This will increase computational efficiency to process in batches. It is also not a requirement of our ML solution to process and produce results in or near real-time. It is instead to provide insights on trends in the data.