CONSTRUCTION SITE SAFETY MONITORING

By : Ashwanth V

PROBLEM STATEMENT

- Construction workers have to be continuously monitored to make sure that they adopt the usage of proper Personal Protective Equipment (PPE), such as hardhats, vests ,etc. But monitoring them manually through CCTV footages is a tedious task, hence there is a need for introducing an automate system.

LITERATURE SURVEY

I. **“Pose guided anchoring for detecting proper use of personal protective equipment”**

– *Ruoxin Xiong, Pingbo Tang, Automation in Construction,Volume 130,2021*.

- Uses a pose estimator to detect worker body parts as spatial anchors and localize the part attention regions.
- CNN based classifier used to recognize both PPE and non-PPE classes within the attention regions.
- Uses a new CPPE dataset – worker, hardhat, vest.
- Higher precision and recall.
- Dataset not construction domain specific.
- The cropped body part attention regions are typically in low resolution.
- Used a lightweight MobileNet rather than a deep network.

II. “Dataset and benchmark for detecting moving objects in construction sites”

– *AnXuehui, et.al, Automation in Construction,Volume 122,2021.*

- In this study, the Moving Objects in Construction Sites (MOCS) image dataset is presented.
- Images collected from 174 construction sites.
- Thirteen categories of moving objects found in construction sites were annotated.
- A benchmark containing 15 different DNN-based detectors was made using the MOCS dataset.
- Faster R-CNN and Mask R-CNN with three different backbones were trained and evaluated in this study.
- Comparison was done with COCO dataset.

III. **“SINGLE- AND MULTI-LABEL CLASSIFICATION OF CONSTRUCTION OBJECTS USING DEEP TRANSFER LEARNING METHODS”** – *Nipun D.Nath ,et.al., Virtual, Augmented and Mixed: New Realities in Construction, Journal of Information Technology in Construction 2019.*

- Deep learning (transfer learning) algorithms to annotate construction imagery from unconstrained real-world settings with high fidelity.
- VGG-16 model - pre-trained on the ImageNet dataset, is trained on construction images retrieved with web mining techniques and labeled by human annotators.
- Two fully-connected layers added to the existing VGG-16’s convolutional layers and weights were updated through training and fine-tuning the model.
- Model tends to classify an image as an object that has a larger visual footprint on the image.
- Single-label classification : 91.1% accuracy
- Multi-label classification : 86.0% accuracy

IV. “Development of an Image Data Set of Construction Machines for Deep Learning Object Detection” – *Bo Xiao,et.al., Journal of Computing in Civil Engineering, ASCE, 2020.*

- Presents a case study on developing an image data set specifically for construction machines named the Alberta Construction Image Data Set (ACID).
- Images collected through online collection (web crawler) and onsite collection (UAVs, cctv cameras, manual clicks).
- To validate the feasibility of ACID, trained the data set using four existing deep learning object detection algorithms- YOLO-v3, Inception-SSD, R-FCN-ResNet101, and Faster-RCNN-ResNet101.
- The mean average precision (mAP) is 83.0% for Inception-SSD, 87.8% for YOLO-v3, 88.8% for R-FCN-ResNet101, and 89.2% for Faster-RCNN-ResNet101.
- The average detection speed of the four algorithms is 16.7 frames per second (fps), which satisfies the needs of most studies in the field of automation in construction.
- One-stage detectors are faster than two stage detectors with the loss of accuracy.

V. **“Deep Convolutional Networks for Construction Object Detection Under Different Visual Conditions”** – *Nipun D.Nath ,et.al., Frontiers in Built Environment, 2020.*

- The paper investigates YOLO-based CNN models in fast detection of construction objects.
- Pictor-v2 dataset used as the transfer learning dataset to train the model YOLOv2 and YOLOv3.
- YOLO-v3 outperforms YOLO-v2 by focusing on smaller, harder-to-detect objects.
- YOLO-v3 model has a 78.2%mAP
- YOLO model integrated with a NN model - predict the relative distances of objects in synthetic images from a single camera view with high accuracy.

VI. "Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset" – *Jixiu Wu, et.al., Automation in Construction, 2019.*

- A one-stage system based on convolutional neural network is proposed to automatically monitor whether construction personnel are wearing hardhats and identify the corresponding colors.
- Constructs a new and publicly available hardhat wearing detection benchmark dataset, GDUT-HWD, which consists of 3174 images covering various on-site conditions.
- Proposed reverse progressive attention to generate a new feature pyramid, which will be fed into the single Shot Multibox Detector (SSD) to predict the final detection results.
- The experimental results demonstrate that the proposed system is effective under all kinds of on-site conditions, which can achieve 83.89% mAP (mean average precision) with the input size 512×512.

VII. "Deep learning for site safety: Real-time detection of personal protective equipment" – *Nipun D. Nath, et.al., Automation in Construction, 2020.*

- This paper presents three deep learning (DL) models built on You-Only-Look-Once (YOLO) architecture to verify PPE compliance of workers; i.e., if a worker is wearing hard hat, vest, or both, from image/video in real-time.
- Trained on Pictor v3 dataset.
- First approach - detects workers, hats, and vests and then, a machine learning model (e.g., neural network and decision tree) verifies if each detected worker is properly wearing hat or vest.
- Second approach - algorithm simultaneously detects individual workers and verifies PPE compliance with a single convolutional neural network (CNN) framework.
- Third approach - the algorithm first detects only the workers in the input image which are then cropped and classified by CNN-based classifiers (i.e., VGG-16, ResNet-50, and Xception) according to the presence of PPE attire.
- Second approach achieves the best performance - 72.3% mean average precision (mAP), in real-world settings,- can process 11 frames per second (FPS) on a laptop computer which makes it suitable for real-time detection, as well as a good candidate for running on light-weight mobile devices.

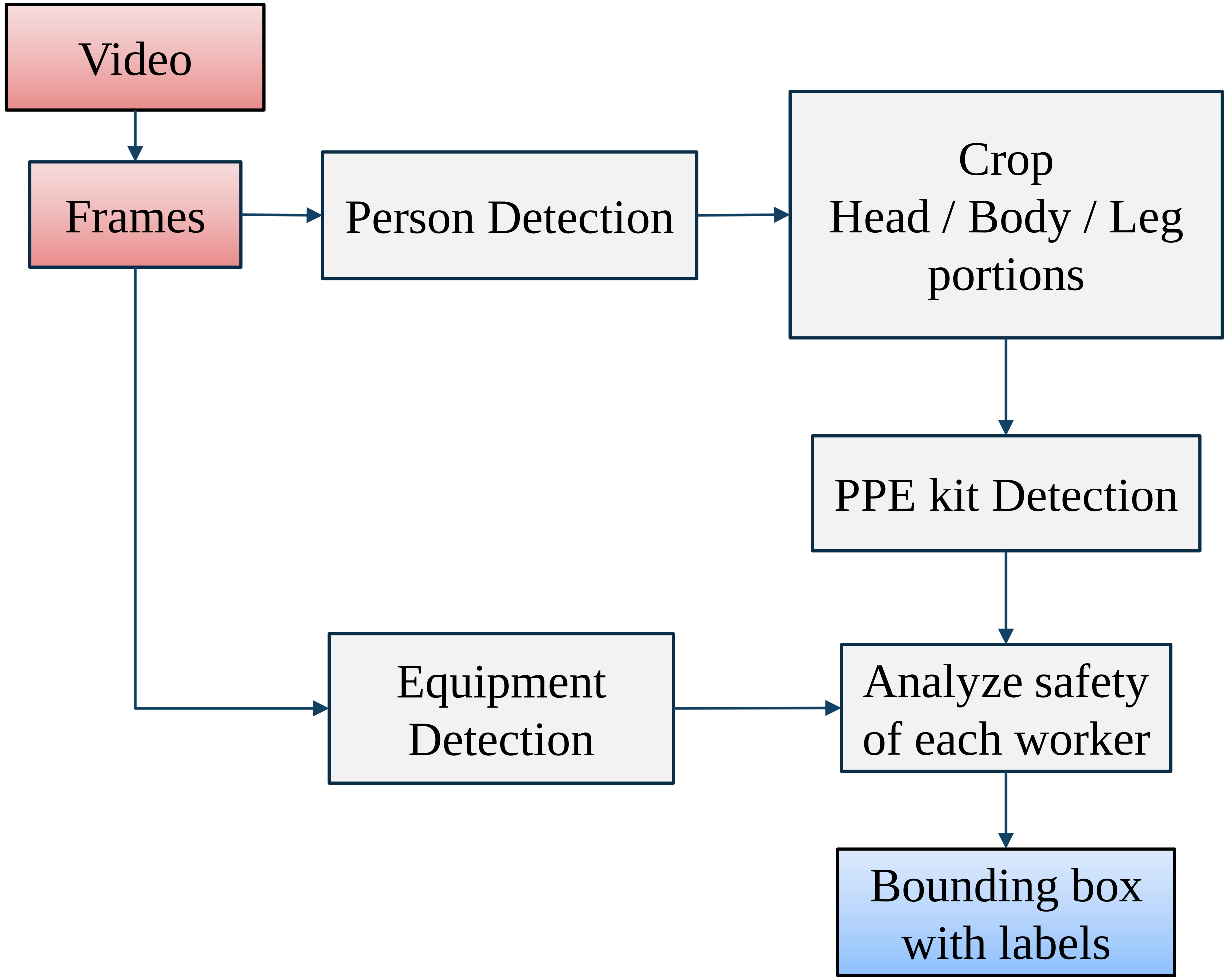
VIII.“YOLOv4 algorithm for the real-time detection of fire and personal protective equipment at construction sites” – *Saurav Kumar, et.al, Multimed Tools Appl (2021).*

- This article presents an approach towards the detection of fire and PPEs to assist in the monitoring and evacuation tasks.
- Open images dataset V6 by Google (10k images).
- Six classes: Fire, Person With Helmet, Person, Safety Vest, Fire Extinguisher and Safety Glass.
- Augmentation techniques used for making the dataset are as follows: Flipping, Rotation, Shearing, Cropping, Zoom in, Zoom out, and Changing brightness or Contrast.
- Utilizes the YOLOv4 and YOLOv4-tiny algorithms based on deep learning for carrying out the detection task.
- YOLOv4 offers mAP of 76.86% over multiple classes.

ARCHITECTURE AND SYSTEM DESIGN

- **Input :** Video
- PPE kit detection using **EfficientNet-B5.**
- Equipment detection using **YOLOv4 .**
- **Safety of worker :** Distance of Centroids from the worker and the equipment. (Euclidean distance)

Work-Flow Diagram



Dataset Preparation

PPE kit:

- CHV, Pictor v3 and other web mined images.
- Identify people from the images using YOLO v4
- Bounding box divided into 4 halves – Head : $1/4^{th}$ part, Body – $2/4^{th}$ and $3/4^{th}$ part, Leg- final part

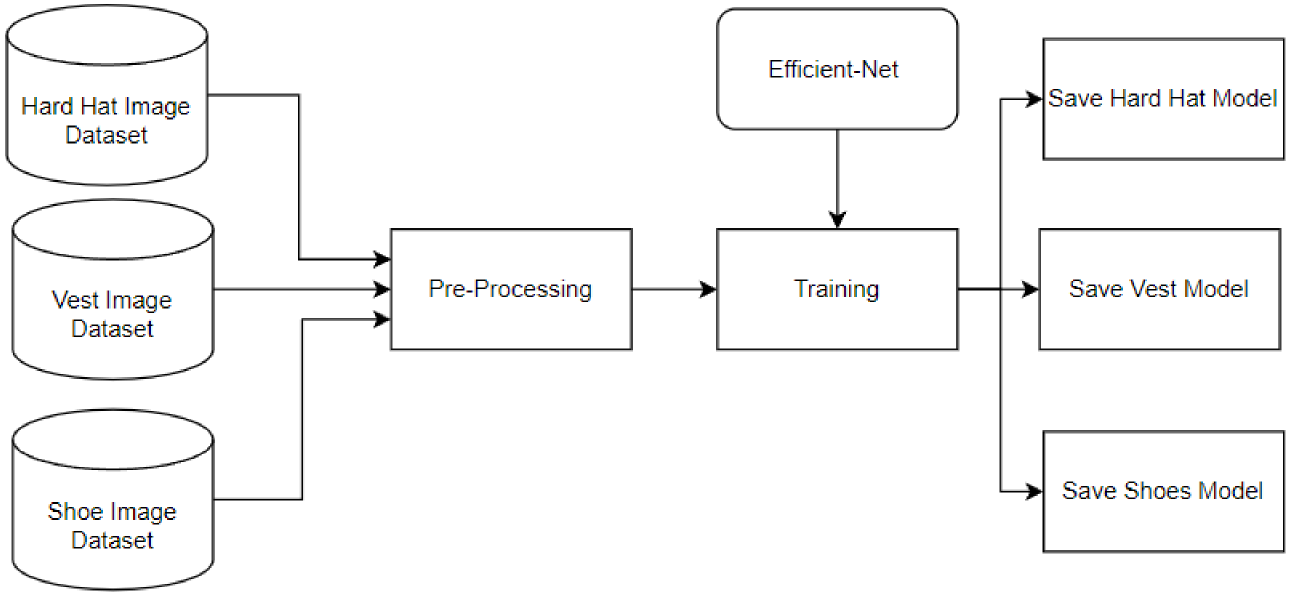
Construction Equipment / Tools :

- ACID, MOCS, Google v6 tools and web mined tools images.
- Custom labelled used CVAT tool. – ‘Equipment’ and ‘Tool’

DESIGN DIAGRAMS

PPE Kit Detection

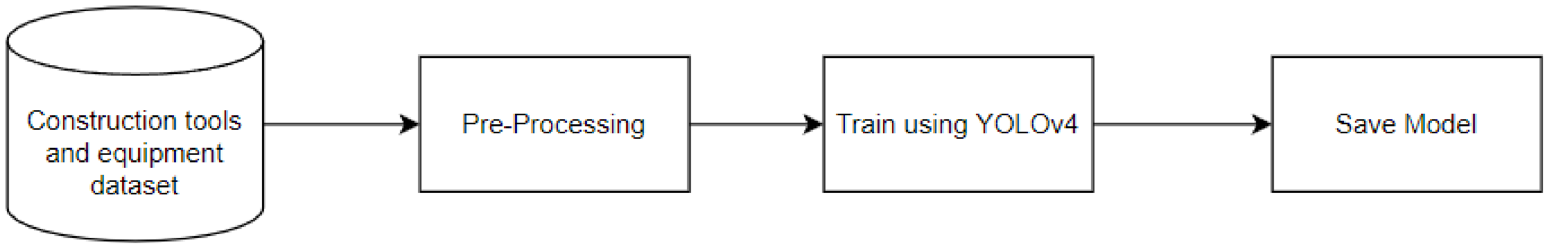
- **Pre-Processing** : Image Resizing- 456x456x3, Image Augmentation – flipping, rotation.
- **Efficient-Net B5** pre-trained on ImageNet dataset.



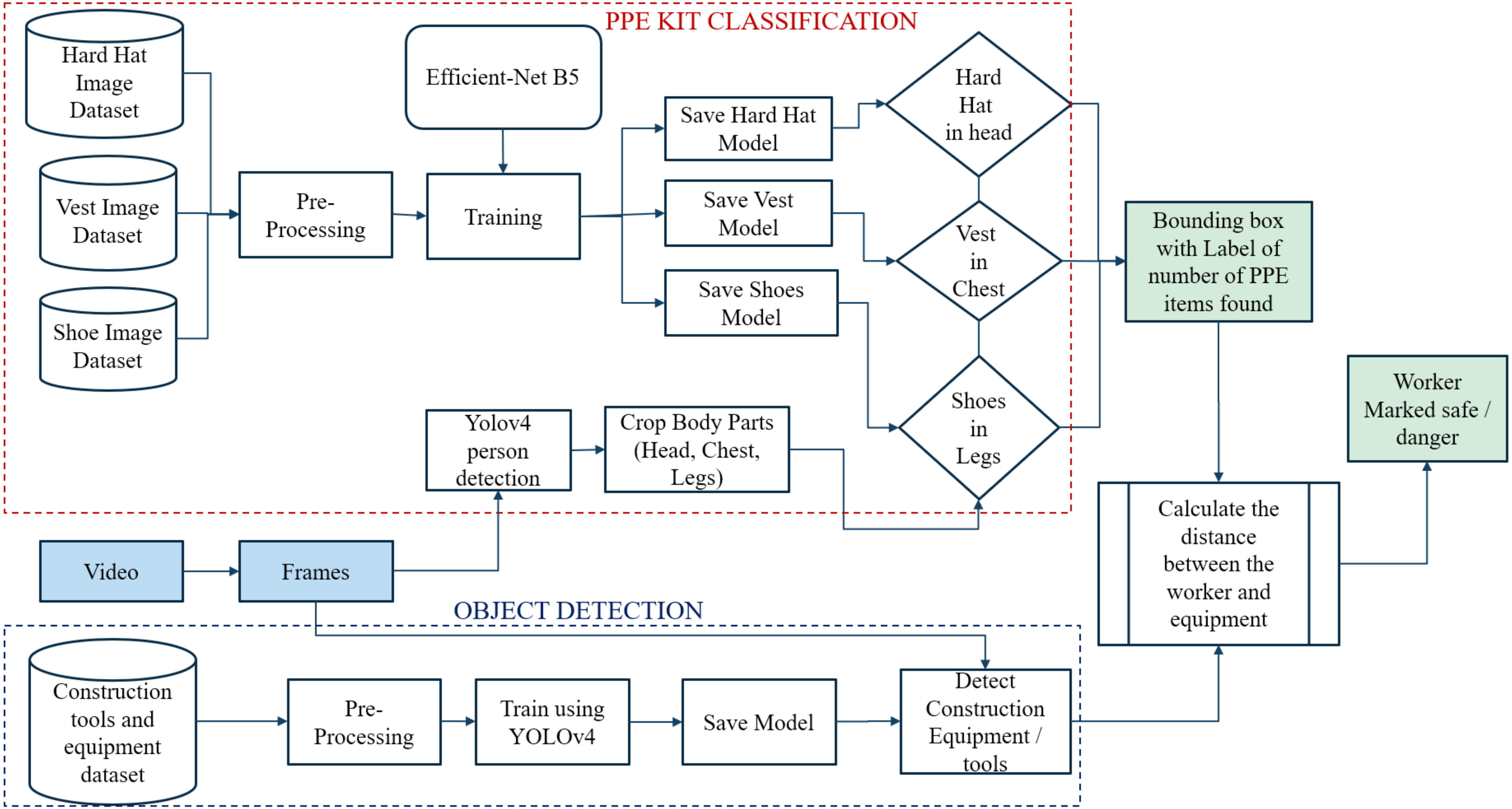
DESIGN DIAGRAMS (Contd.)

Construction Object Detection

- **Pre-Processing :** Image Resizing – 416x416x3
- Pre-trained YOLOv4 model trained on COCO dataset.



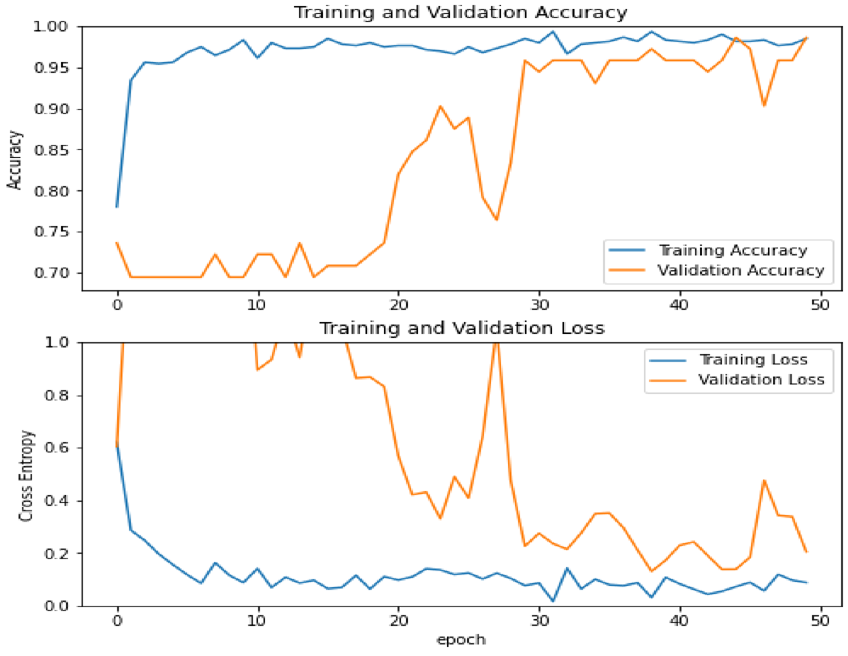
ARCHITECTURE



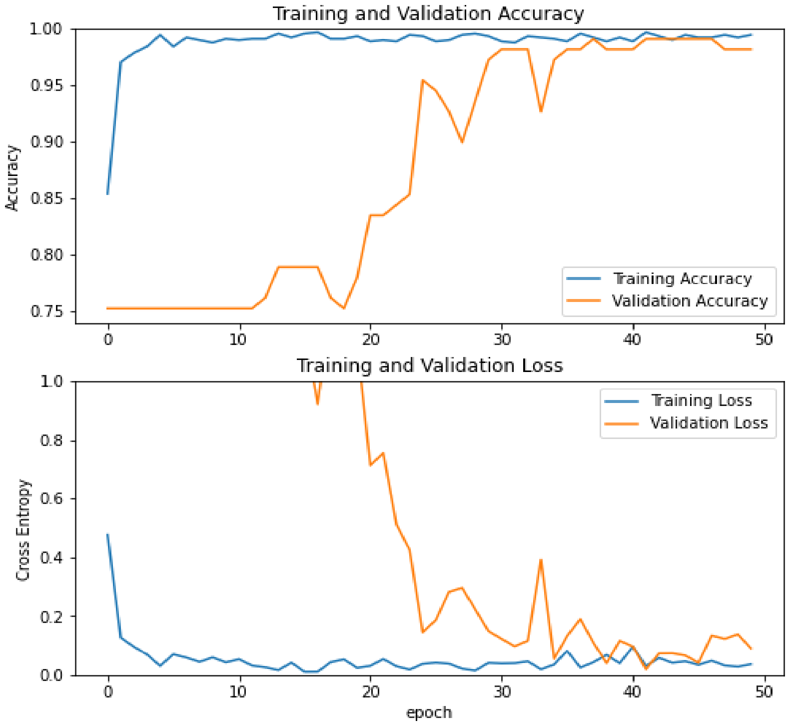
EXPERIMENTAL RESULTS

PPE kit detection

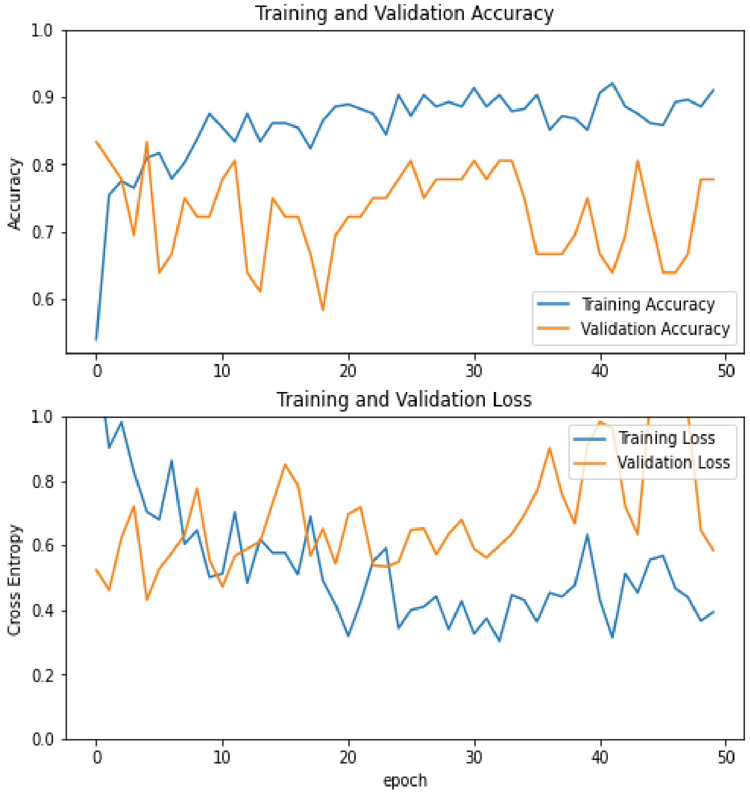
98.48% for hard-hat, 99.43% for vest and 91.00% for shoes.



Hard-Hat



Vest

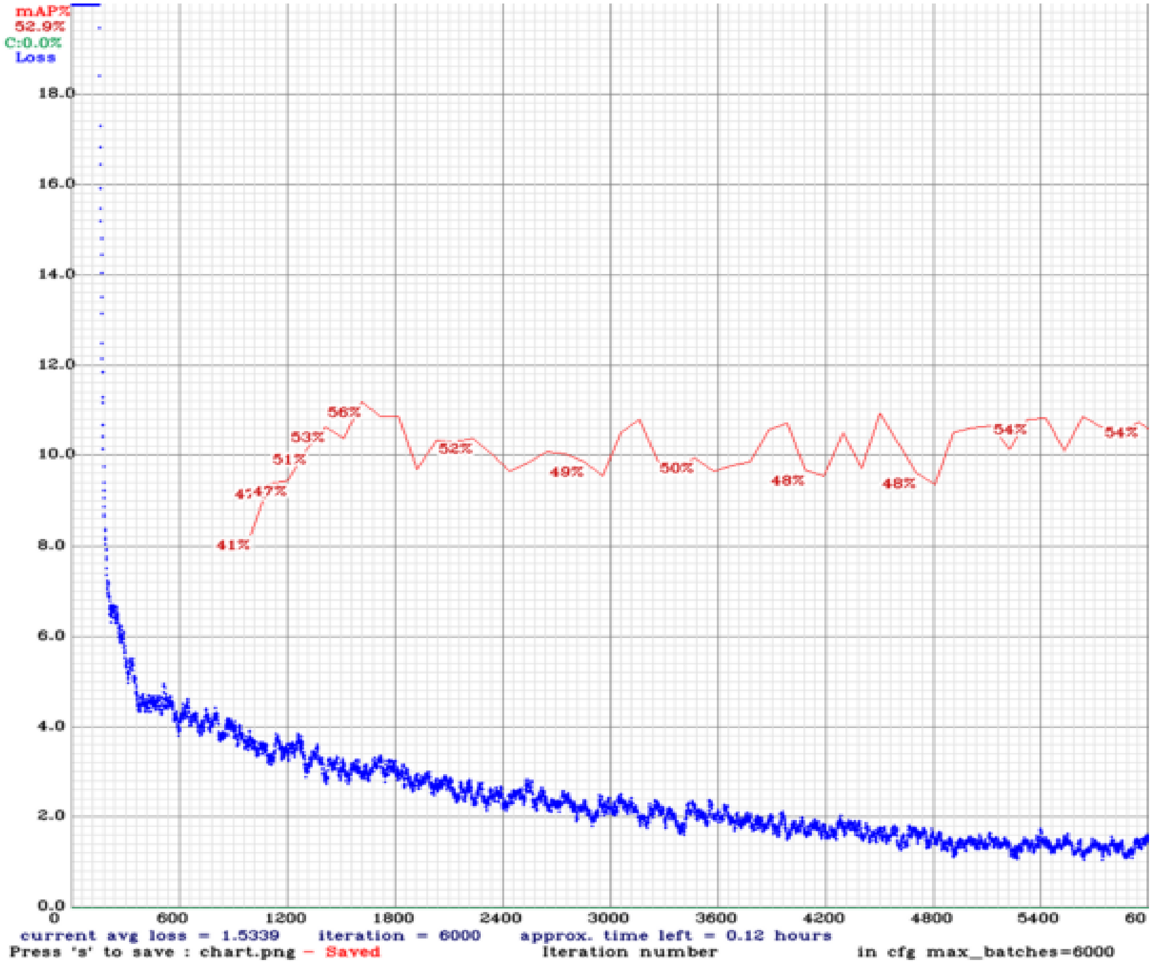


Shoes

Equipment Detection:

- About 1000 images of construction equipment and 100 images of tools were used for the training

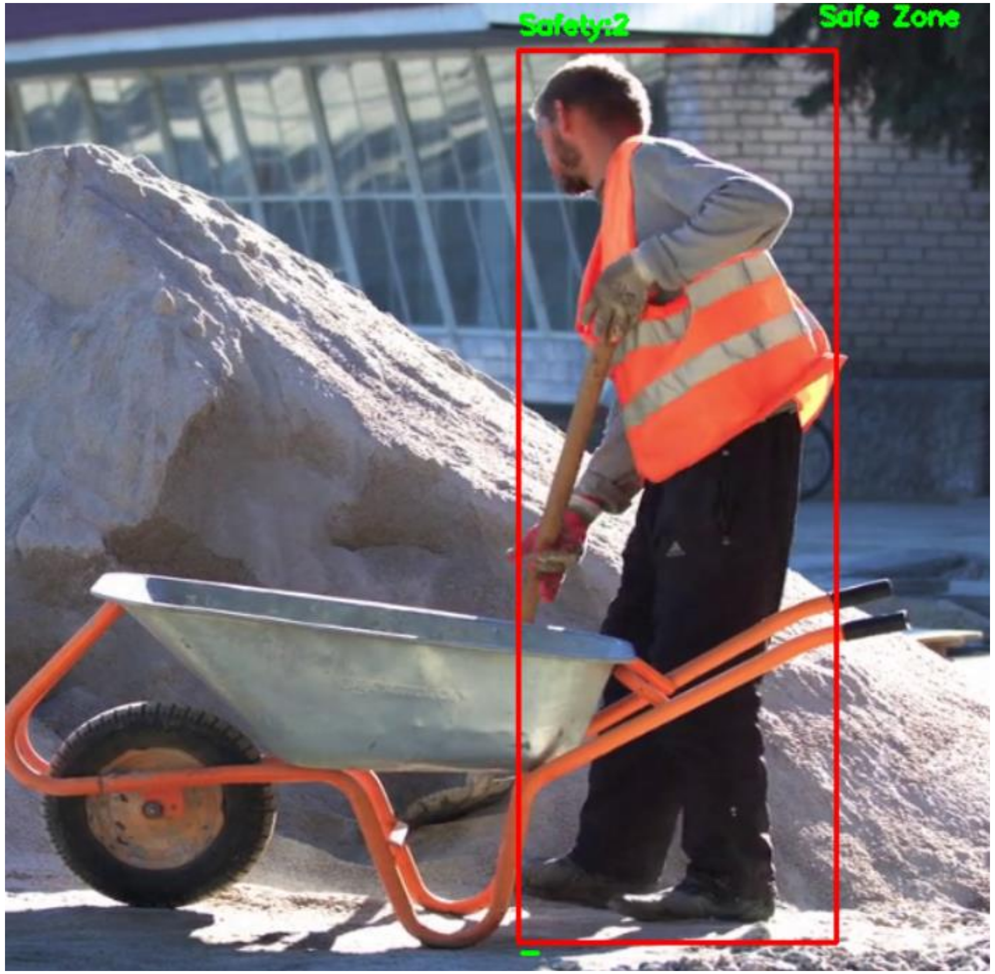
```
detections_count = 4839, unique_truth_count = 912
class_id = 0, name = equipment, ap = 73.81%     (TP = 655, FP = 371)
class_id = 1, name = tool, ap = 38.05%          (TP = 29, FP = 31)

for conf_thresh = 0.25, precision = 0.63, recall = 0.75, F1-score = 0.68
for conf_thresh = 0.25, TP = 684, FP = 402, FN = 228, average IoU = 45.84 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 0.559294, or 55.93 %
```

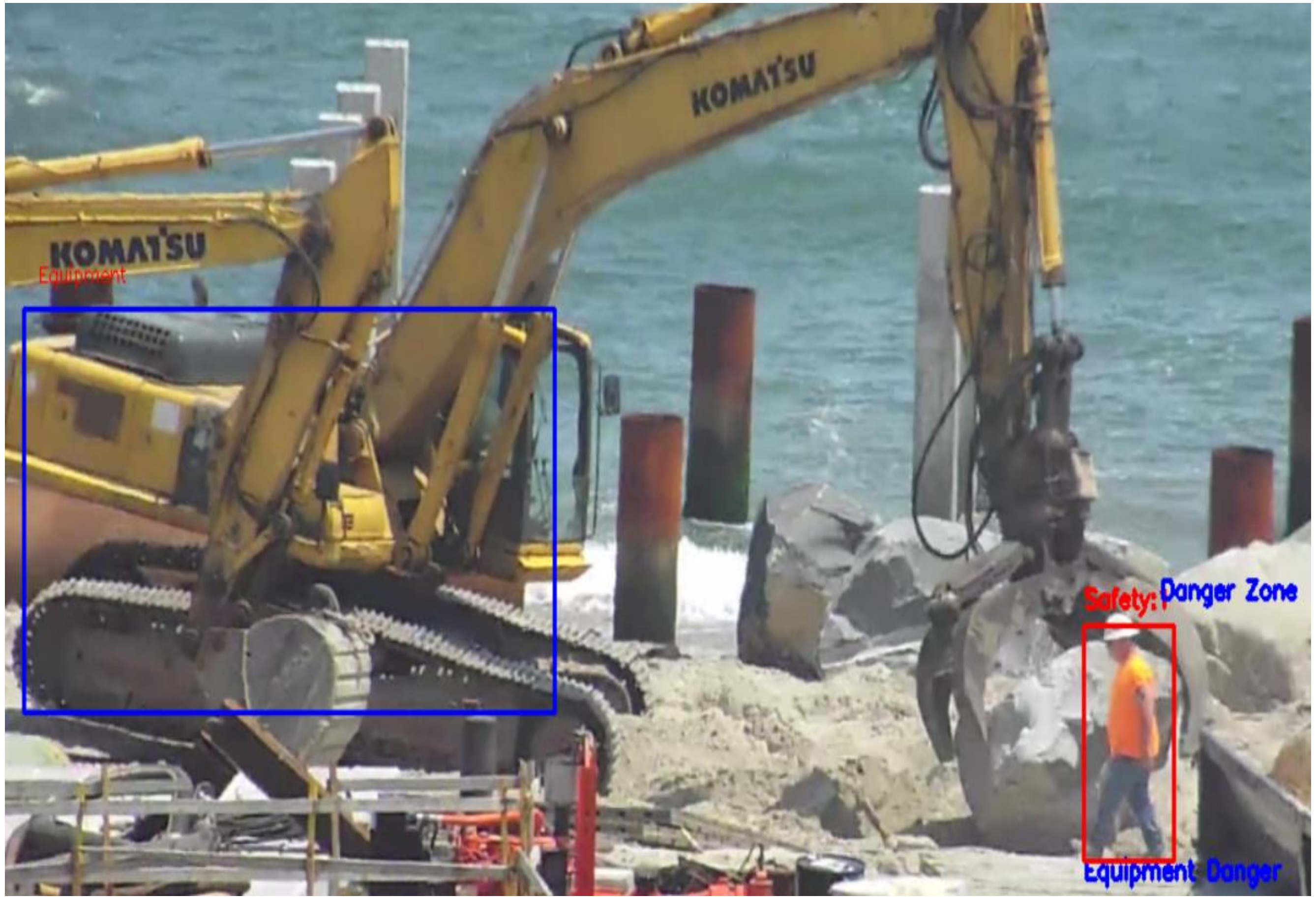


Equipment Detection Output

Tools Detection Output

FINAL OUTPUT





Limitations

- The problem of occlusion occurs in certain frames which leads to wrong recognitions.
- The distance calculation between the equipment and worker can be improved, since the camera angle would play a major role in detection of the objects and its bounding boxes.

CONCLUSION

- Using the concept of body part localization, the accuracy of the PPE detection can be increased.
- As an added safety check, the risk factor of each worker is calculated.
- An automatic construction site monitoring system can be created using the mentioned method of body part localization and risk factor calculation.

Future work

- The quality of the images used to train the EfficientNet classifier could be improved by adopting super resolution (SR) techniques.
- Output has the problem of occlusion in some frames, which could be improved probably by detecting the workers using techniques like pose estimation at the cost of FPS and other resources.
- More types of PPE components could be detected, such as gloves, goggles, and also domain-specific types of PPE kits can also be detected.

REFERENCES

[1] Ruoxin Xiong, Pingbo Tang, “Pose guided anchoring for detecting proper use of personal protective equipment”, Automation in Construction, Volume 130, 2021.

[2] Neha Sharma, Vibhor Jain, Anju Mishra, “An Analysis Of Convolutional Neural Networks For Image Classification”, Procedia Computer Science, Volume 132, 2018, Pages 377-384, ISSN 1877-0509, 2018.

[3] Nath Nipun D., Behzadan Amir H, “Deep Convolutional Networks for Construction Object Detection Under Different Visual Conditions”, Frontiers in Built Environment, volume 6, 2020.

[4] Yu Fan and Xianqiao Chen and Jinguang Xie and Zhenjie Fu, “An Algorithm for Detecting the Integrity of Outer Frame Protection Net on Construction Site Based on Improved {SSD}”, Journal of Physics: Conference Series, IOP Publishing, 2021.

[5] Jixiu Wu, Nian Cai, Wenjie Chen, Huiheng Wang, Guotian Wang, “Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset”, Automation in Construction, Volume 106, 2019.

[6] Nipun D. Nath, Amir H. Behzadan, Stephanie G. Paal, “Deep learning for site safety: Real-time detection of personal protective equipment”, Automation in Construction, Volume 112, 2020.

[7] Yang, Zhen, Yongbo Yuan, Mingyuan Zhang, Xuefeng Zhao, Yang Zhang, and Boquan Tian, "Safety Distance Identification for Crane Drivers Based on Mask R-CNN" Sensors 19, no. 12: 2789 . 2019.

REFERENCES

[8] Kumar, S., Gupta, H., Yadav, D. *et al.* YOLOv4 algorithm for the real-time detection of fire and personal protective equipments at construction sites. *Multimed Tools Appl* (2021).

[9] Weili Fang, Lieyun Ding, Hanbin Luo, Peter E.D. Love,"Falls from heights: A computer vision-based approach for safety harness detection",Automation in Construction,Volume 91, 2018.

[10] An Xuehui, Zhou Li, Liu Zuguang, Wang Chengzhi, Li Pengfei, Li Zhiwei,"Dataset and benchmark for detecting moving objects in construction sites",Automation in Construction,Volume 122, 2021.

[11] Nath N D, Chaspari T, Behzadan A H, "Single- and multi-label classification of construction objects using deep transfer learning methods", ITcon Vol. 24, Special issue Virtual, Augmented and Mixed: New Realities in Construction, pg. 511-526,2019.

[12] Bo Xiao and Shih-Chung Kang , "Development of an Image Data Set of Construction Machines for Deep Learning Object Detection", Journal of Computing in Civil Engineering, volume -35, 202.

[13] Bin Dai, Yuhu Nie, Wenpeng Cui, Rui Liu, and Zhe Zheng, "Real-time Safety Helmet Detection System based on Improved SSD". In Proceedings of the 2nd International Conference on Artificial Intelligence and Advanced Manufacture AIAM2020.

THANK YOU