SAFETY MONITORING AT A CONSTRUCTION SITE

ARNAV SATISH[1], SWASTIKA SHARMA[1],VEDANTIKA SHARMA[2] , HARSHITA CHHAPARIA[2], SAHANA DEVI KJ[2]

[1]Department of Computer Science, PES University, Bengaluru, Karnataka, India.

[2]Department of Computer Science in AI&ML, PES University, Bengaluru, Karnataka, India

Corresponding author: Sahana Devi KJ <email:sahanadevikj@pes.edu>

***Abstract*— Our IEEE project focuses on making construction sites safer by detecting hazards at a construction site. This model watches out for unworn Personal Protective Equipment (PPE). It also addresses issues such as compromised safety nets, the proximity of people to Heavy Earth Moving Machinary (HEMM) and fire risks. Our specialty lies in our ability to spot damages to safety nets, detecting people near HEMMs and flammable substances near fire sources that are one of the most neglected hazards in a construction site. When it detects a problem, it sends alerts to on-site supervisors and workers so they can quickly do the needful. This timely intervention helps prevent accidents and ensures that everyone on site is better protected. By providing real-time updates, our model helps create a safer working environment.**

***Keywords*— Safety Monitoring, Construction Sites, Computer Vision, Object Detection, Deep Learning and YoloV5.**

I.INTRODUCTION

The construction industry of India stands as the fastest growing construction industry in the world standing as the second largest contributor to the country's GDP. While being the key industry, it's also one of the most dangerous industries with up to 38 fatal accidents daily. A recent study for the International Labor Organisation (ILO) says that India currently has the highest number of fatal accidents on a construction sites in the world.

In 2023 alone, 51 laborers have died in 24 mishaps as compared to 19 deaths in 2021 and 39 deaths in 2022. The current year is marked with the tragic incident of the death of 17 people after a crane collapses at a viaduct site. Ideally safety officers must be present for a daily routine of supervision. However, it is rarely the case.

These mishaps occur due to a range of factors including poor maintenance of makeshift life, failure of equipment like cranes and scaffolding, malfunctioning of safety belts.

Given these challenges, prioritizing the study and identification of critical causes behind construction accidents assumes paramount importance. By understanding and ranking these factors, the supervisors and the managers can implement targeted interventions to foster safer working conditions and uphold the industry's growth sustainably. The aim of this project is to analyse construction safety hazards and help prevent it.

II. LITERATURE SURVEY

The papers that were referred to so far allowed us to arrive at the following conclusions:

[1] The research on how deep learning algorithms are currently being implemented on construction sites has mostly been limited towards the implementation of a single feature.

[2] The image augmentations and feature extractions that would be ideal for such a model which would monitor safety on a construction site

[3] Fire hazard detection technique exists but the major drawback here being that it does not go beyond detection of sparks and inflammable material. The usage of Euclidean distance in calculating the distances between two detected objects in a image however the drawback being the distance between objects in a image does not usually correspond to the actual distance on the ground.

III. DATASETS

Since our project involves safety monitoring at a construction site, the dataset that we required were images and videos from a construction site of various environments that the workers can be subjected to Datasets were collected from Roboflow which is a computer vision platform that simplifies the process of building computer vision models by providing pre-annotated datasets.

1. The first set of data for our very first and the basic model includes images of workers on the construction site with and without their Personal Protective Equipment’s.
2. The second model is about detection of damaged safety nets and harness on the construction site, so the dataset that we have collected for this model includes the images of damaged safety nets and harness
3. The third set of the data for our fire hazard detection model includes images of sparks due to welding and inflammable materials such as Styrofoam, plastic and gas cylinders.
4. The fourth model which is about danger zone detection. The dataset that we collected for this model were images of heavy earth moving machines like bulldozer, crane, roller and excavators and workers.

So overall our dataset consists of 16872 images across 16 classes

A) Sample Dataset:

Figure 1 Sample Dataset

IV. MODEL

YOLOv5 offered by Ultralytics is a sated of the art object detection algorithm built on a PyTorch framework as opposed to the previous versions of YOLO which were implemented on the Darknet framework thus offering a greater control over the encoded operations. The CNN model forms the backbone of YOLO. Architecturally the YOLOv5 model consists of three layers:

1) Backbone:

It's the main body of the network which is designed using the NEW CSP - Darknet 53 framework

2) Neck:

Connects the backbone and the head and utilises the SPPF and the new CSP-PAN structures

3) Head:

The head is the final stage of the architecture of the which has been constructed using the YOLOv3 head.

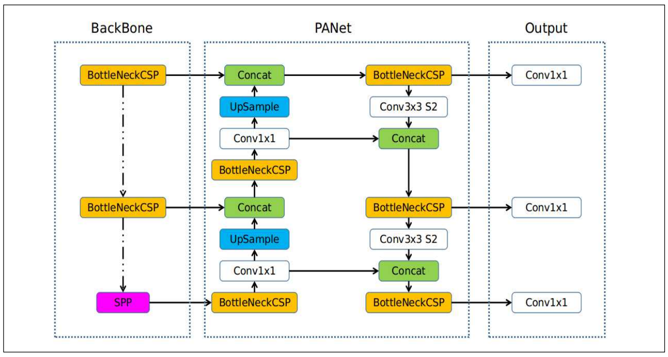

Figure 2 Yolo V5 Architecture

A) The reasons why the YOLOv5 model is being used:

YOLOv5 retains more features in the feature extraction network and often optimizes the algorithm details thus combining semantic and detailed information to improve the detection effect of small objects which results in the ability of YOLOv5 to deal with complex problems.

The usage if a more complex architecture in YOLO v5 allows it to achieve higher accuracy and better generalization to a wider range of object categories.

It was hard for the previous versions of YOLO to detect bounding boxes on image corners mainly due to the equations used to predict the bounding boxes, but YOLOv5 helped in solving this problem by expanding the range of the centre point offset from (0-1) to (-0.5,1.5) therefore the offset can be easily 1 or 0 (coordinates can be in the image's edge). Also, the height and width scaling ratios were unbounded in the previous equations which may lead to training instabilities but now this problem has been reduced.

B) How yolov5 works

YOLO v5 introduces the concept of "spatial pyramid pooling" (SPP), a type of pooling layer used to reduce the spatial resolution of the feature maps. SPP is used to improve the detection performance on small objects, as it allows the model to see the objects at multiple scales.

It is composed from three convolution layers that predicts the location of the bounding boxes (x, y, height, width), the scores and the objects classes.

The equations to compute the target coordinates for the bounding boxes is shown in the figure below.

$$b_x = (2 \cdot \sigma(t_x) - 0.5) + c_x$$
$$b_y = (2 \cdot \sigma(t_y) - 0.5) + c_y$$
$$b_w = p_w \cdot (2 \cdot \sigma(t_w))^2$$
$$b_h = p_h \cdot (2 \cdot \sigma(t_h))^2$$

Figure 3 Bounding Box

YOLOv5 returns three outputs: the classes of the detected objects, their bounding boxes and the objectness scores. Thus, it uses BCE (Binary Cross Entropy) to compute the classes loss and the objectness loss. While CIoU (Complete Intersection over Union) loss to compute the location loss. The formula for the final loss is given by the following equation.

$$Loss = \lambda_1 L_{cls} + \lambda_2 L_{obj} + \lambda_3 L_{loc}$$

Figure 4 Loss Function

$$Euclidean\,(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Figure 5 Euclidean Formula

C) Limitations of yolov5:

The major limitations of YOLOv5 include low accuracy and a high rate of false detection.

- The YOLOv5 algorithm may not achieve the desired results when applied to specific tasks.
- YOLOv5 does not output angular predictions, which are crucial for reflecting attitudes and shapes of targets in aerial images.

V. WORKFLOW

Our project focuses on safety monitoring at a construction site. Our goal is to reduce and avoid accidents that occur at construction sites due to negligence through continuous monitoring. Our model aims to reduce the risk of accidents by being able to detect various objects and provide the necessary information. We have done this using the Yolov5 CNN framework for its flexible and predefined object detection capabilities.

A) Dataset and augmentation:

The dataset for all our features was sourced from Robofow. Our classes include gloves, no gloves, vest, no vest, helmet, no helmet, sparks, people, etc. The dataset we have is pre-annotated. We performed further augmentations on them, such as grayscale, noise, exposure, shear, and blur. For this, we used the 'opencv' library and 'os' module. The next step done was performing a train, test, and validate split on our complete dataset, i.e., the original one and the augmented dataset.

B) Model training:

After downloading and cloning Yolov5 requirements, we defined the number classes in the data.yaml file and trained the model for 90 epochs.

C) Features:

1) Basic PPE (Personal Protective Equipment) detection:

This feature ensures that the workers are wearing their PPE accurately. There are many cases where the workers in elevators meet with accidents. Therefore, this feature also detects whether the workers working in elevators are wearing safety harnesses. For this, we have given the dataset as workers with harnesses and workers without harnesses.

Furthermore, accidents often occur during bar bending or other construction activities performed at heights where safety nets are provided, but they are mostly damaged. Typically, there's no one to monitor and check the condition of these nest. So, the last aspect of this feature is the detection of damaged safety nets.

2) Prevent fire accidents caused by sparks:

The thought behind this feature is to ensure there are no inflammable objects in the vicinity of welding or any other task involving sparks. The dataset for this feature includes sparks and inflammable materials. We have focused on three types of inflammable materials: cylinders, Styrofoam, and polythene.

We made two assumptions for this feature. The first one is that our camera would be placed at eye level and the other one being, the spark reach would be 35 feet (10.66 m). However, since we could not capture an image over such a long distance, we have assumed a threshold distance of 7 metres. To implement this feature, we placed two objects 7 metres apart in real life, and calculated their pixel distance, which came out to be 3338.079 pixels. Using the real-life distance and pixel distance, we calculated the scale factor, i.e., 7 / 3338.079 as 0.002097. Now that we have the scale factor, given any image, we calculate the euclidian distance between sparks and inflammable objects and multiply it by the scale factor to obtain the real-life distance between the sparks and any inflammable object. If this distance is less than or equal to 7 metres, an alert is raised. We have shown the distance and marked it with suitable colours in our testing video.

3) Preventing accidents due to Heavy Earth Moving Machines (HEMM):

Many accidents occur due to these heavy, moving construction vehicles because the driver cannot spot any person in front of their vehicle. The dataset for this feature includes people and all types of construction vehicles.

The first step is to detect whether the vehicle is moving using, CSRT (Discriminative Correlation Filter with Channel and Spatial Reliability) tracker. This tracker can be used to track the position of a vehicle across frames, helping determine whether the vehicle is moving. If it is moving, there are two cases:

1: If the bounding boxes of the vehicle and the person intersect, an immediate alert is raised.

2: The euclidian distance between the person and the vehicle is calculated for all frames. If this distance is found to be reduced to half or less, an alert is raised, thereby warning them.

These features collectively enhance site safety by ensuring proper PPE use, preventing fire hazards, and avoiding accidents with heavy machinery.

VI. RESULTS

Accuracy of the first model which is concerned with a detection of PPE equipment is at 86%.

The second model which is concerned with detection of damaged safety nets and the donning of harness is at 93% as indicated by the PR curves in Fig6.

The third model which is concerned with the detection of fire hazards works provided a few parameters are satisfied. The parameters being: The camera should be placed at an eye level, the distance between the welding station and the camera should be known and that the sparks reach up to a distance of 7m.

The fourth model which is concerned with the estimation of a danger zone near HEMs. Proper testing could not be implemented as a good video in order to test the model could not be obtained, hence right now, the overall accuracy and proper working could be not figured out.

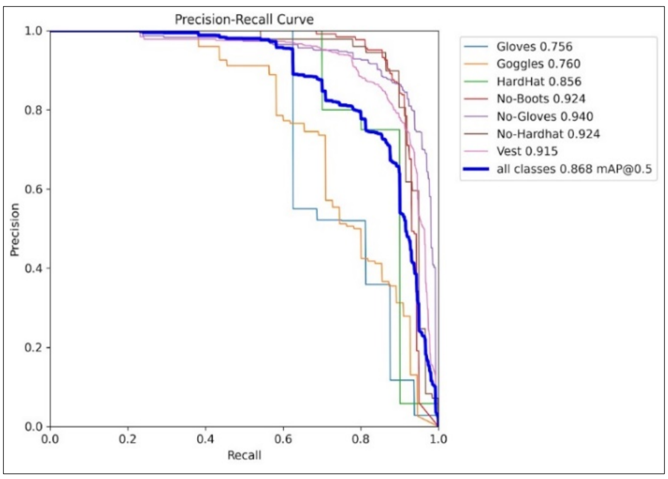

Figure 6a PPE Detection

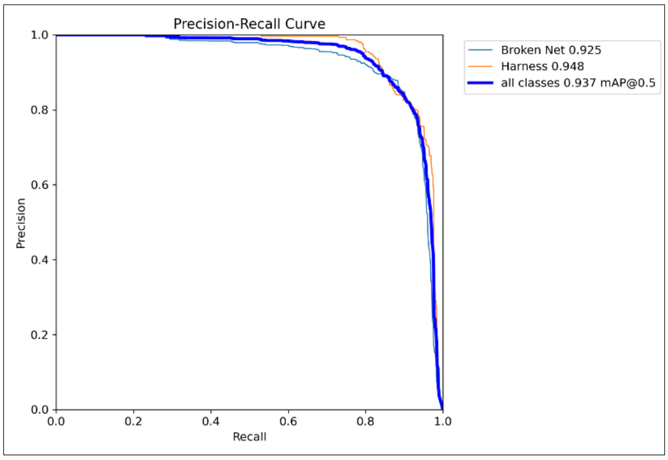

Figure 6b Danger Zone

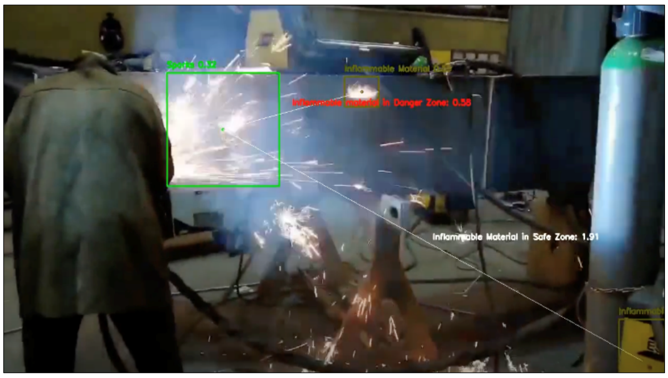

Figure 7 Test Sample for Fire hazard Detection

VII. CONCLUSION

Our revolutionary model offers diverse abilities to monitor safety in varied environments in a construction site. We are able to provide our model with an ability to not just detect objects but go a step further and detect hazards by establishing danger zones and monitoring safe distances. This is a significant improvement over previous models that were limited to single feature or were focused solely on object detection.

VIII. REFERENCES

[1] J. Lee and S. Lee, "Construction site safety Management: A computer vision and deep learning approach," Sensors, vol. 23, no. 2, p. 944, 2023. [Online]. Available: https://doi.org/10.3390/s23020944

[2] Q. Wu, W. Wang, H. Chen, L. Zhou, Y. Lu, and X. Qian, "A Safety Detection Method on Construction Sites under Fewer Samples," Electronics, vol. 12, no. 8, p. 1933, 2023. [Online]. Available: https://doi.org/10.3390/electronics12081933

[3] H. Ann and K. Y. Koo, "Deep learning-based fire risk

detection on construction sites," Sensors, vol. 23, no. 22, p. 9095, 2023. [Online]. Available: https://doi.org/10.3390/s23229095

[4] M. M. Alateeq, F. R. PP, and M. A. S. Ali, "Construction site hazards identification using deep learning and computer vision," Sustainability, vol. 15, no. 3, p. 2358, 2023. [Online]. Available: https://doi.org/10.3390/su15032358

[5] I. Journal, "Safety helmet detection in engineering and management," SlideShare, Oct. 22, 2022. [Online]. Available: https://www.slideshare.net/slideshow/safety-helmet-detection-in-engineering-and-management/253755868#1

[6] N. Khan, S. F. A. Zaidi, J. Yang, C. Park, and D. Lee, "Construction work-stage-based rule compliance monitoring framework using computer vision (CV) technology," Buildings, vol. 13, no. 8, p. 2093, 2023. [Online]. Available: https://doi.org/10.3390/buildings13082093