International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

Construction Safety Equipment Detection System

Yogesh Kawade[1], Mayur Kasture[2], Yash Mallawat[3], Aniruddh Dubal[4]

[1,2,3]Student, B.Tech., Construction Engineering and Management, Symbiosis Skills and Professional University, Pune, Maharashtra, India
[4]Asst. Prof., Construction Engineering, Symbiosis Skills and Professional University, Pune, Maharashtra, Pune, India

---***---

Abstract – *The construction sector is one of the most unpredictable and hazardous industry sectors. Tens of millions of construction industry accidents occur globally causing damages and accidents to employees every year. This industry makes up one of the primary sectors of the workforce and activity in the international market is taken into consideration as a vital element in operating the financial system of the country. Construction sites remain dynamic and complex structures. The complex motion and interaction of humans, goods, and power traditionally make construction safety management extremely hard. However, many studies have been conducted in the last decade to introduce innovative technologies for the implementation of efficient protection systems within the construction industry. The primary section of solution development involved data training, model selection, model training, and model evaluation. The second segment of the study comprises model optimization; application specific embedded system choice and economic analysis. The current study demonstrates the practicality of deep learning-based object detection solutions for construction situations. Furthermore, the specific understanding, provided in this study, can be employed for numerous functions along with safety monitoring, productiveness assessments and managerial decisions.*

Key Words: **Personal Protective Equipment, Object Detection, Hardhat, Personal Protective Vest, Computer Vision.**

1. INTRODUCTION

Numerous onsite safety policies have been established to ensure construction workers' safety. Within the safety guidelines, an appropriate use of personal protective equipment (PPE) is specified and the contractors need to make sure that the regulations are enforced through the monitoring method. The monitoring of the usage of PPE is usually carried out at the site entry and the onsite construction field. These days, most of the construction industry conduct the monitoring of PPE using manually with the aid of inspectors. This work is tedious, time-consuming and useless because of the excessive number of workers to monitor in the field. Currently, numerous technologies were proposed to enhance the construction safety. A number of the proposed solutions, computer vision has been broadly used [1], [2], [3]. However, most of latest works focus on detecting the usage of hardhat and personal protective vest at the onsite construction field. On this paper, we suggest a fully automatic vision-based PPE detection and monitoring. The proposed system includes two essential additives: PPE detection and recognition. The primary aim of PPE detection system is to determine the presence of required personal protective equipment.

The PPE detection system developed by us employs a camera that monitors the person using object detection in real time. The system detects for the person, hardhat, and the personal protective vest. The system sends a warning message through the speakers installed onsite in the construction field. Despite the warning if the construction worker(s) defies safety rules, the system sends an SMS/email alert to the concerned authority in charge onsite. This framework is illustrated in figure 1.

2. METHODOLOGY

In our system, we appoint the YOLO network for PPE and person detection. The YOLO network has been introduced via Joseph Redmon's team [4] for object detection. When an input of a picture or a video is fed into the system the network does not examine the whole picture, as a substitute; it looks at elements of the picture which have high probabilities of containing the object. YOLO or you only look once, proposed by using Redmon et al. is a singular object detection algorithm a lot unique from the vicinity-based algorithms. In YOLO a single convolutional network predicts the bounding boxes and the class possibilities for these boxes. Whilst the images are fed to the system, the features of human beings are extracted and from that features head localization is completed and identity of hard hat and PPE vest is done by means of drawing bounding boxes and confidence levels are stated at the top of the bounding box. Step one is human identification from the accumulated surveillance photos. A partition is made on each of the applicable construction surveillance images into a set of object regions. In those particular photographs the system detects the workers who're wearing hardhat and PPE vest with their accuracy. The subsequent step is to detect the hard hat and PPE vest detection. This equipment is recognized by the features. Those features are extracted using characteristic extraction techniques. YOLO is implemented on images to predict the objects and in which it is present. It is the faster convolution neural network.

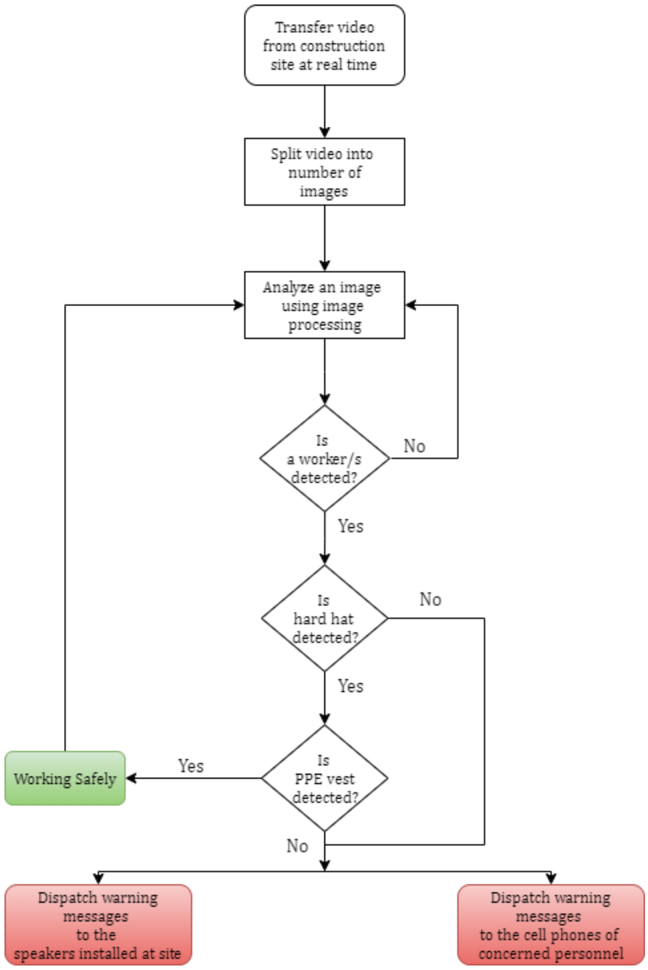

Figure -1: Framework for identifying hard hats and PPE vests with a worker.

Non-maxima Suppression: This system applies Non-Maximal Suppression (NMS) in order that only the fine bounding boxes are saved. The first step in NMS is to delete the so far anticipated bounding boxes which have identity probability that's much less than given NMS threshold. In the current system, we set this NMS threshold to 0.4, i.e., all of the predicted bounding boxes that have a detection probability greater than 0.4 will be stored with the system.

3. EXPERIMENTAL PROCEDURE

3.1. Experimental dataset

To create the training dataset of PPE detection, we accumulated hard hat and PPE vest pictures from two assets: (1) internet photos retrieved using the net crawler and (2) actual images efficiently captured using the webcam. A complete of 3,334 pictures were gathered (table 1) and annotated to train a YOLOv3 version.

Table -1: Framework for identifying hard hats and PPE vests with a worker

	Number of the Internet images	Number of the Real-World Images	Total
Hard hat	856	962	1818
PPE vest	664	852	1516

Areas of decommissioning worker may be captured at distinct resolutions in the images as the surveillance cameras are established at exceptional locations at the decommissioning site and the trajectory of workers is random. For this particular reason, different distance situations (3 m, 5 m, and 7 m) had been taken into consideration in our experiments to validate the robustness of our proposed technique. The impact of individual posture was additionally taken into consideration in current experiments and three common employee postures—standing, sitting, and squatting—had been included inside the testing dataset. To create the testing dataset in such a way that it may validate the overall performance of the trained model, three volunteers had been instructed to carry out different postures while wearing PPEs (or not) at different distances from the digital camera. The information is provided in table 2, wherein advantageous samples refer to those who are wearing PPE properly and poor samples refer to those who are not wearing PPE. Eventually, we randomly selected 500 pictures for each case from the gathered image sequences and created a testing dataset.

3.2. Evaluation Metrics

We adopted precision and recall to assess the overall performance of the proposed method:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Where TP (true positive) is described as the wide variety of precise identification of people who are wearing PPE. FP (false positive) is the number of those who aren't wearing PPE properly however are misidentified as wearing PPE properly, and FN (false negative) represent the variety of not detected ground reality of PPE proper use individuals, as described in the table 2.

Table -3: Identification of outcomes under specific distance.

Category	Distance	TP	FP	FN	Precision (%)	Recall (%)
Hard hat	3m	672	23	14	96.69	97.95
	5m	658	27	87	96.05	88.32
	7m	632	12	151	98.13	80.71
PPE vest	3m	499	0	1	100	99.80
	5m	496	7	4	98.60	99.20
	7m	480	5	18	98.97	96.38

3.2. Implementation Details

We constructed the YOLOv3 version using TensorFlow [5] and initialized it primarily based on pretrained weights on the ImageNet dataset [6]. Training of YOLOv3 was achieved in two degrees: (1) all convolutional layers had been first frozen up to the final convolutional block in Darknet-53, and the version was trained with frozen layers to sustain a solid loss in 50 epochs; (2) all convolutional layers of Darknet-53 proceeded to unfreeze to carry out satisfactory-tuning in 50 epochs. The learning rate agenda is as follows. For the primary level, the model was prepared with a learning rate of 1e − 3; for the second stage, the model was prepared with a learning rate starting out at 1e − 4. An Adam optimizer [7] with a batch length of eight was adopted in the course of training. Figure 2 illustrates the identification results on individuals.

Table -2: Statistics of accumulated testing dataset

	No.	Value	Images	Hard Hat		PPE Vest	
				Positive Samples	Negative Samples	Positive Samples	Negative Samples
Distance	1	3m	500	656	756	567	1124
	2	5m	500	707	633	639	1184
	3	7m	500	857	596	855	1256
Posture	1	Standing	500	1634	1697	976	2227
	2	Sitting	500	354	184	165	272
	3	Squatting	500	515	339	297	309

4. RESULT AND DISCUSSION

4.1. Impact of Distance

The identification results under specific distances are mentioned in table 3. The resolution of the distinctive areas within the image steadily became more limited as the distance between the digital camera and individuals extended. The precision and recall rate for the identity of proper hard hat use steadily decreased with growing distance from the camera, however the precision was above 95% for all measured distances, at the same time as the recall rate was above 90% except for measurements at 7 m. The general overall performance of the identity of proper hard hat use is suitable (Precision rate: 98.07%; recall rate: 93.72%). For the identification of proper PPE vest use, the overall performance of our approach declined only slightly as distance increased; the precision and recall rates remained higher than 97%.

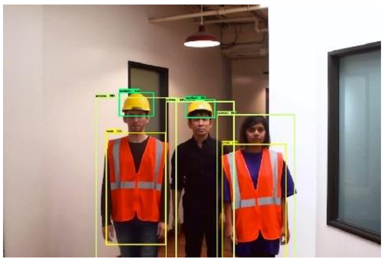

Fig -2: Identification of hard hat and PPE vest on individuals.

4.1. Effect of individual posture

The identification consequences for distinctive individual postures are shown in table 4. The high average precision confirmed in the outcomes represents the excellent overall performance of our model for several individual postures.

For the identification of a proper hard hat use, the recall rate for the squatting position is lower than the others due to failures of the body components detection through OpenPose[8]. However, the recall rate continues to be above 81%. The identification outcomes imply the effect of individual posture has little impact on the identification of a proper PPE vest uses, because the precision and recall rate continue to be sturdy in distinctive individual postures.

The overall precision and recall rates are 98.32% and 92.15%, respectively, which demonstrates the robustness of the proposed technique in the identification of proper PPE use at varying distances and individual postures.

Table -4: Identification of outcomes under different individual postures.

Category	Posture	TP	FP	FN	Precision (%)	Recall (%)
Hard hat	Standing	1508	93	153	94.19	90.78
	Sitting	267	0	47	100	85.03
	Squatting	387	0	82	100	82.51
PPE vest	Standing	978	4	7	99.59	99.28
	Sitting	196	0	4	100	98
	Squatting	326	13	9	96.16	97.31

5. CONCLUSIONS

This paper has provided a unique vision-based method to cope with the difficulties of proper PPE use. Initially, we created a dataset using internet pictures and real-world pictures to improve the YOLOv3 model to apprehend hard hats and PPE vests. Ultimately, we carried out the identification of proper PPE use using geometric relationships of the outputs of OpenPose and YOLOv3. The performance of the proposed method was experimentally evaluated under various distance and individual posture situations. The experimental effects suggest the proposed method became capable of figuring out the decommissioning workers who are not wearing PPE properly with excessive precision (98.32%) and recall rate (92.15%), while making sure real-time performance. The scope for on-site occupational safety monitoring was to persuade a preliminary recognition. This work possesses immense possibilities with an extensive range of applications, e.g., proper PPE use management in a COVID-19 health facility. Future research is to be investigated following the consideration of the augmentation of the training dataset to increase the safety of the monitored objectives, e.g., safety gloves and safety glasses.

ACKNOWLEDGEMENT

This paper could have been impossible without the splendid support of our university, Symbiosis skills and professional university, Pune. The authors of this paper would like to thank all the faculty members who aided us to gain these outcomes. We would also like to sincerely thank the other authors whose research and work we have constructed upon.

REFERENCES

[1]. M. of labour invalids and social affairs. (2018) Notification notice of the situation of labor accidents in 2017. [Online]. Available: http://vnniosh.vn/chitietNCKH/id/7559/T hong − bao − tinh − hinh − tai − nan − lao − dong − nam − 2017

[2]. B. E. Mneymneh, M. Abbas, and H. Khoury, "Automated hardhat detection for construction safety applications," Procedia Engineering, vol. 196, pp. 895 – 902, 2017, creative Construction Conference 2017, CCC 2017, 19-22 June 2017, Primosten, Croatia. [Online]. Available:

http://www.sciencedirect.com/science/article/pii/S1877705817331430

[3]. ——, "Vision-based framework for intelligent monitoring of hardhat wearing on construction sites," Journal of Computing in Civil Engineering, vol. 32, 2019.

[4]. J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," CoRR, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640

[5]. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX symposium on operating systems design and implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016; pp. 265–283.

[6]. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. Int. J. Comput. Vis. 2015, 115, 211–252. [CrossRef]

[7]. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. arXiv 2014, arXiv:1412.6980.

[8]. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.E.; Sheikh, Y. OpenPose: Realtime multi-person 2D pose estimation using Part Affinity Fields. arXiv 2018, arXiv:1812.08008.