Scientific Journal of Impact Factor (SJIF): 5.71
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 8, Issue 04, April-2021

Construction site SOP on Hard Hat Detection with Deep Learning Techniques

Nilam Mehta[1], Dr. Gopi Sanghani[2]

[1]*M.E. Scholar, Computer Engineering, Darshan Institute of Engineering & Technology, Rajkot*
[2]*Professor, Computer Engineering, Darshan Institute of Engineering & Technology, Rajkot*

Abstract — *This Health and safety of the worker is the main concern to save them from the dangerous situation while working at the construction site. In India, many workers are ignoring safety precautions like a hard hat, personal protective equipment, etc which results in fatal injuries, and in some cases, they are losing their lives too. The life of the worker should not be that cheap to let go by ignoring safety majors. Therefore, the protection of the workers from injuries and accidents is an important issue that should be taken care of. As the construction area is too vital, It is impossible to be discovered by one supervisor or contractor thus model/ tool needs to be developed to monitor workers by installing cameras at the construction site. The model will detect hard hats automatically and gives a notification if it finds a worker without a safety hat.*

Several Deep learning algorithms are being used to detect the hard hat of the worker. computer vision and image processing techniques are used to get better outcomes. R-CNN family, YOLO versions, SSD, and many more algorithms are on priority to detect an object. Many machine learning algorithms like the histogram of oriented gradients, haar-like features were used previously to detect the safety hat of the worker.

Keywords*- You only look once (YOLO), Convolutional Neural network (CNN), Region-based Convolutional Neural network (R-CNN), Hard hat detection, Global positioning system (GPS)*

I. INTRODUCTION

The Deep Learning approach is largely applied for object detection. Various computer vision techniques are used to detect the object from images and videos and represent the outcome with a bounding box drawn on the identified object. The object is classified in different categories like human, umbrella, car, laptop, cell phone and so on which is known as the object class, and the procedure is known as object classification. The next important task is to find an exact location of the object class from an image which is known as object localization. Object detection is the combination of object identification, object classification, and object localization. Here safety hat of the worker at the construction site is the main object to be detected. Several CCTV cameras were installed at the construction site with different angles and views. Images and video frames are being monitored at the back office to detect the wearing of the safety hat. If the worker is found breaking the safety protocols then the notification will be generated and alerted to the supervisor. Computer vision and image processing techniques are applied with machine learning and deep learning algorithms to get better efficiency. Region-based convolutional neural network, you only look once, histogram of oriented gradient, single shot multibox detector like algorithms are used with distinct frameworks like PyTorch, TensorFlow, Keras to identify the hard hat. The framework is the software used by programmers to make tasks easier as it contains predefined classes, built-in functions, code blocks, and pre-tested functionalities. Deep learning algorithms are applied with frameworks like TensorFlow and PyTorch to detect the safety hat of the worker with better computation speed and higher accuracy.

II. LITERATURE SURVEY

Object detection is the process of identifying the object from images and videos and detected objects are displayed with a rectangular box which is known as the bounding box. The process of object identification, object recognization, and object localization is known as object detection. The object detection approach is divided into two basic approaches for example deep learning approach and the machine learning approach.

YOLO (you only look once), SSD (single-shot multibox detector), R-CNN (region-based convolutional neural network) families are the algorithms that belong to the deep learning approach.

Haar-like features of viola-jones detection, HOG (histogram of oriented gradients) are the algorithm belongs to the machine learning approach.

Previously, machine learning algorithms were being used widely to detect an object. Presently deep learning approach is on-trend for object detection as it provides higher computation speed. Hence, the Deep learning approach is again divided into two types. The first is the one-stage detector and the next is the two-stage detector.

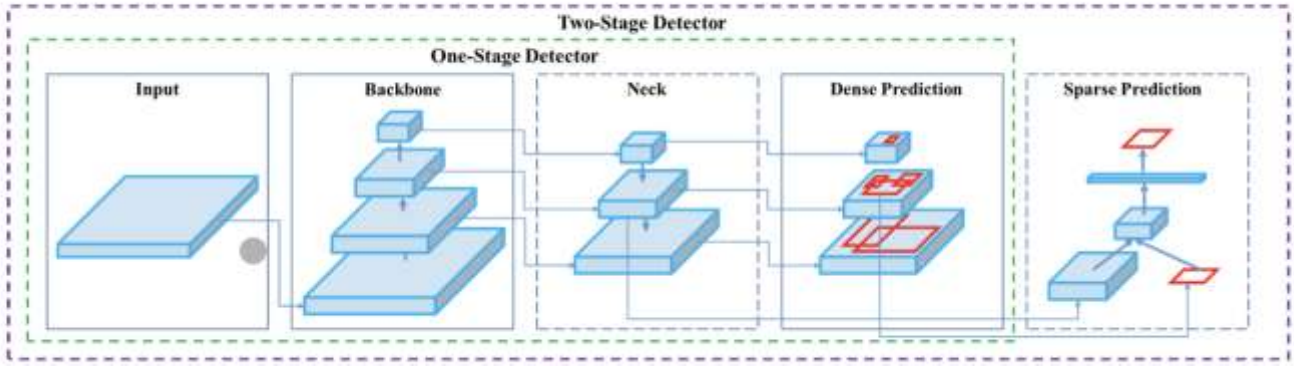

Figure 1. Object detection architecture [10]

Two-stage detector

Region-based Convolutional Neural Network family-like R-CNN, Fast R-CNN, Faster R-CNN uses two-stage detection. The two-stage detector uses the classification-based method. Region of Interest (ROI) is identified from the image in the first stage of detection. ROI is the portion of the image where there is a chance of the object being is maximum. In the second stage, an object detection algorithm is applied to each Region of Interest which leads to computation overhead. Thus two-stage detector takes more computation time which makes it slower but increases accuracy.

One-stage detector

You Only Look Once (YOLO), Single-shot Multibox Detector (SSD) algorithms use the one-stage detector. The one-stage detector uses the regression-based method. As its name suggests, the image is divided into grids, and for each grid, it finds the class of the object and shows the outcome with a bounding box having an object in it. The one-stage detector consumes fewer computations thus it provides a high operating speed.

According to [1], the convolutional neural network-based algorithm is used to detect hard-hat. Computer vision techniques are used to check the proper wearing of the hard hat. The object detection algorithm is implemented in two-phase. In the first phase, safety hat identification is performed using a convolutional neural network. In the second phase, a safety hat-wearing identification algorithm is performed to check whether the worker is appropriately wearing a safety hat or not. The CNN-based approach towards hard-hat detection (CAHD) consists of five steps to train a model. In the first step, a three-channel RGB image with a standard size is used as the input. In the second step, to maintain consistency of the input and output, the processed image is transferred to the residual layer. In the third step, feature extraction is performed on the base of the Yolo algorithm and darknet framework. In the fourth step, the feature map is generated. In the last step, the number of channels of the feature map is being modified to get better accuracy for object recognization. The optimization of the convolutional neural network is responsible for reducing complexity and getting high precision.

According to [2], improved faster R-CNN is used to identify the safety hat of the worker at the substation. Dataset of the hard hat is constructed first and then Retinex image enhancement techniques are applied to improve the image quality of the complex outdoor images. These images are fed to the feature extraction framework and a feature map is formed which is given to the Regional proposed network. To detect smaller size safety helmets, the K-means++ algorithm is applied. According to the experimental result, the mAP of the model is about 94.3% and it detects 11.62 images per second. Thus algorithm is having good detection ability but poor detection speed.

According to [3], the deep learning algorithm and computer vision techniques are applied to detect safety hats and personal protective equipment (PPE). The SqueezeDet neural network is used with the MobileNet classifier to improve object detection by 9% without compromising operating speed. The overall system has accomplished 0.75 F1 scores on the test dataset. The SqueezeDet is a tiny neural network with limited functionalities and parameters. The outcome of the model is compared with tiny-Yolo and RetinaNet where SqueezeDet with the MobileNet classifier gives better results.

According to [4], authors have used the Internet of Things (IoT) to save workers from disaster. They have used Raspberry 3 and various sensors like MQ2 toxic gas detector with Global Positioning system to get the view using the Responsive Web Design (RWD). RWD is responsible to get the best visual effects by connecting it with several devices. Real-time audio-video streaming, voice and face recognization were used. Hakka’s speech recognization is used to communicate between the backside server and the client. They have installed 21 instant alarms to get instant notifications

about future disasters. The only disadvantage of this paper is that almost all the techniques are depended on the network. Maintenance and management of the network should be done precisely.
According to [5], Improved faster R-CNN is used with multi-scale training, increasing anchor strategy, and online hard example mining (OHEM) to detect the safety helmet of the worker at the construction site. Authors have used improved faster R-CNN over the faster R-CNN because of these disadvantages like a problem to detect smaller size objects, occlusion and to overcome the large negative sample in faster R-CNN. Faster R-CNN consists of two modules. First is the Region proposal network module (RPN) and object detection module. The Multi-scale strategy is applied at the training phase to increase the robustness of the original faster R-CNN. Anchor size is increased from 9 to 12 to detect smaller size objects as the original faster R-CNN uses 9 anchors. Online hard example mining technology is applied to reduce negative samples and to keep the balance between positive and negative samples. The multipart combination method is used to identify the missed object and to improve detection accuracy. Improved faster R-CNN is applied over faster R-CNN to increase the detection accuracy by 7%.

According to [6], Safety helmet detection of perambulatory workers in power substations is implemented using machine learning, computer vision, and image processing techniques. To detect the object in motion, the Visual background extractor (ViBe) modeling algorithm is applied. ViBe model is mainly used to detect moving objects as quickly as possible and the process is known as moving object segmentation. Pediatrician classification is applied after the result of the moving object segmentation to extract the feature of images and to classify the pediatrician at the power station. C4 pediatrician classification is used to get the exact location of the worker. The ultimate step is hard hat detection. the color space transformation and color feature discrimination and the head location are used to detect the safety helmet of the worker more accurately.

According to [7], Computer vision and machine learning techniques were used to detect the safety hat of the worker. The object of interest means the construction worker was detected and then the safety hat wearing of the worker was detected. First of all image segmentation was applied to the dataset of the collected images. Histogram of oriented gradient (HOG) and Discrete cosine transform (DCT) features were extracted and supplied to the Support vector machine (SVM) for the detection of the object of interest. SVM uses non-linear mapping to transform training data into higher dimensions. In this model linear SVM was used to classify two classes like human and non-human. Then after Hough transform and color-based feature extraction is applied combined to detect the safety hat of the worker.

According to [8], machine learning, computer vision and, pattern recognization were used on Video-based surveillance to detect the safety hat of the worker at the construction area. The algorithm is divided into three phases. First is person's face detection, second is motion detection and skin color detection and the third is hard hat detection. At first, the motion was detected on the video sequences. If there is motion in the video frame then face detection was applied only on those regions of the image rather than scanning the whole image. Thus it reduces computation time and reduces false alarms. The haar-like feature is the machine learning technique applied to detect the face region of the person. After face detection, skin color detection was applied to reduce false prediction of the faces and non-faces. At last, hard hat detection was applied using the color-based information above the face regions. The main disadvantage of this process is that it takes lots of computations on motion detection, face detection, skin color detection which makes it slower.

According to [9], computer vision, the Yolo v3 algorithm, and deep learning techniques were used to detect the safety hat of the worker. Yolo algorithm is the faster convolutional neural network. The video sequence is divided into image frames. The image frame was scanned at once, divided into grids, and predicts the class of the object i.e. construction worker using the bounding box. The hard hat is identified by features. These features are extracted from feature extraction techniques and a feature map was generated and passed to the convolutional layers for the safety hat detection. At last non-maxima suppression was applied to select the best bounding box with the predicted object.

Many algorithms like YOLO (you only look once), SSD (single-shot multibox detector), R-CNN (region-based convolutional neural network) families, Haar-like features of viola-jones detection, and HOG (histogram of oriented gradients) were used to detect the safety hat of the worker. Haar-like features, HOG is the machine learning algorithm that has more computations and less operating speed therefore they have been used formerly. The main comparison is between YOLO and the R-CNN family. Both the algorithm belongs to the deep learning approaches.

R-CNN (Region-based convolutional neural network) family is upgraded from R-CNN, Fast R-CNN to Faster R-CNN. YOLO algorithm family increases from Yolo V1, Yolo V2 or Yolo 900, Tiny-Yolo, Yolo V3, and Yolo V4. The latest versions from both are the Faster R-CNN and Yolo V4 thus the comparative analysis between Faster R-CNN and Yolo V4 is given below.

R-CNN and Fast R-CNN uses selective search to find out the region proposals which is time-consuming therefore Faster R-CNN comes up with a separate network to predict the region proposals called the Region Proposal Network (RPN). RPN is used to identify the Region of Interest (ROI). ROI is the portion of the image where the chances of the presence of the object are higher. Objects are being identified at the region proposal network and then for each ROI detection is performed to classify the object and represented with the bounding box. [11]

YOLO V4 divides the image into a grid of AxA size and each grid predicts N bounding box, classification score, and confidence. Classification score is the probability of an object being present in the image. confidence returns the accuracy of the bounding box. YOLO scan image at once and classifies the objects. Thus AxAxN bounding boxes were predicted and non-maxima suppression was applied to eliminate bounding boxes with a low confidence score. [11]

In the Faster R-CNN Region of Interest is detected and for each ROI detection is performed which makes lots of computations and decreases operating speed. There is no such ROI was detected, the image is scanned at once in YOLO V4. Thus YOlO V4 is faster than the Faster R-CNN. Alexey et al., YOLO V4 processes 33 frames per second with 43.5% of the average precision where Faster R-CNN processes 9.4 frames per second with 39.8% of the average precision [10].

CONCLUSION

The safety hat of the worker is being detected by the main approaches like machine learning approach and deep learning approach. The machine learning approach is slower with limited functionalities. The deep learning approach contains many algorithms like the R-CNN family, YOLO versions, SSD, and so on. R-CNN has a lot number of computations which reduces its operating speed and YOLO V4 is faster than the Faster R-CNN.

REFERENCES

[1] Xie Z., Liu H., Li Z., & He Y., “A convolutional neural network-based approach towards real-time hard hat detection.”, IEEE International Conference on Progress in Informatics and Computing (PIC), 2018, pp. 430-434.

[2] Chen S., et al., “Detection of Safety Helmet Wearing Based on Improved Faster R-CNN”, International Joint Conference on Neural Networks (IJCNN), 2020, pp. 1-7

[3] Filatov N., Maltseva N., & Bakhshiev A., “Development of Hard Hat Wearing Monitoring System Using Deep Neural Networks with High Inference Speed”, International Russian Automation Conference (RusAutoCon), 2020, pp. 459-463.

[4] Huang F. L., Liao Z. Z., Wang T. H., Chen Q. M., Wu T. H., & Chang C. H., “Intelligent and Disaster Prevention Hard Hat Based on AIOT and Speeches Recognition”, International Conference on Machine Learning and Cybernetics (ICMLC), 2019, pp. 1-5.

[5] Gu Y., Xu S., Wang Y., & Shi L., “An Advanced Deep Learning Approach for Safety Helmet Wearing Detection”, International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 669-674.

[6] Li K., Zhao X., Bian J., & Tan M., “Automatic Safety Helmet Wearing Detection”, IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), 2017, pp. 617-622.

[7] Rubaiyat A. H., Toma T. T., Kalantari-Khandani M., Rahman S. A., Chen L., Ye Y., & Pan C. S., “Automatic detection of helmet uses for construction safety”, IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW), 2016, pp. 135-142.

[8] Du S., Shehata M., & Badawy W., “Hard hat detection in video sequences based on face features, motion and color information”, 3rd International Conference on Computer Research and Development, 2011, Vol. 4, pp. 25-29.

[9] Jagadeeswari C., Nagamani G., Sneha B., and Dr. G. NagaSatish, “Hard Hat Detection Using Deep Learning Techniques.”, International Journal of Advanced Science and Technology, 2020, Vol. 29, No. 11s, pp. 1292-1298.

[10] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M., “YOLOv4: Optimal Speed and Accuracy of Object Detection”, 2020, arXiv preprint arXiv:2004.10934.

[11] Gandhi R., “R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms”, accessed on 11 April 2021, https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e.

