IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468

Workers Safety Helmet Wearing Detection on Construction Sites Using deep learning

Arvind Yede [1] Dr. G. N. R. Prasad [2]

MCA VI Semester, Dept. of MCA, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad[1]

Sr. Asst. Professor, Dept. of MCA, Chaitanya Bharathi Institute of Technology(A), Gandipet, Hyderabad[2]

Abstract: Research shows that many workers construction is one of the high-risk industries where construction workers tend to be hurt in the work process. Head injuries are very serious and often fatal. Every worker needs to wear a helmet while working in a factory or any construction site. But many workers are ignored and do work without safety equipment. The management tried to control this problem manually but it is insufficient for the real situation. The ideal solution is to develop an electronic detection system that can be automated recognize this kind of problem without human cost. The motivation for this project is to prevent the death of workers due to head injuries by monitoring real-time if a person is wearing helmet while working using Deep Learning techniques. Here, a robust approach is tried, in which CCTV cameras are used to capture the image of humans. The proposed system uses YOLO (You look only once) v3 model that is the used as the state-of-art method for real-time detection with higher rate of accuracy. The detected humans as objects are utilized for calculating the distance between them the rough Euclidean distance calculation method. The proposed model produces reliable outcome compared with the other prediction systems. We use Convolutional Neural Network (CNN) to identify who are workers are wearing helmet while entering a premise. YOLO Dark net is used to get the dependencies.

Keywords: Convolution neural network, Helmet Detection System, Image processing, real time object detection, Machine learning, YOLO, Deep Learning.

I. INTRODUCTION

Reducing the risk of workers life one of the important task of the work station management but it is also the most difficult task. Working in the construction sites may cause the serious injuries. One of the serious injuries is injury caused on head. So it is essential that worker to wear the helmet while working. But some workers ignore to wear the helmet. By this risk on their life will increase. The injury of head can cause serious mental as well as physical problems and can also lead to loss of their life. This is very serious problem to take care off. As per the current reports many worker losses the life in construction sites or work stations in which many of cases caused due to not wearing helmets.

Tracking weather worker is wearing helmet or not is the one of the major problem of management. The difficulty of management can lead to miss management which can cause the work flow of the construction site or work stations. Manually finding which worker is wearing helmet is also a difficult task and time taking task. Also the human can make mistake to identifying. Hiring a person for this task can also the costly and improper use of the human resource.

In the present days the technology is so advance and also money efficient. Using the CCTV can help the management to overcome this problem. The CCTV are installed with helmet detection system in entrance of construction site or work stations as well as in the various places of construction sites and work stations. Many of the work stations and construction site have pre-installed CCTVs. By upgrading their CCTV system management can effectively check whether the work is wearing helmet or not.

YOLO v3 prediction model which is derived from the YOLO v1 and v2 model. YOLO model is refreshingly simple and precise form of prediction model with lower latency. This model connected with the CCTVs of the work station and construction sites to efficient working. YOLOv3 uses the variant of Dark Net that contains 53 layers stacked with more than 100 layers of fully connected layers. The Sec II of the paper discusses various literature papers that discussed about the same criteria. Sec III discuss the system tools required for the proposed analysis, and Sec IV discuss the methodology adopted here and finally the Sec V discuss the various challenges in concluding the present analysis.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468

II. RELATED WORK

[1]In this paper, the SSD and the neural network modules are used to get the better accuracy and the efficient results for the power station surveillance to check weather workers are wearing helmet or not. The dataset which is used for training is created by the real situation in the power station. The SSD detector gives the good efficiency. The model have achieve the speed of 21fps i.e. real time. This project will work well with the power stations.

[2] In this paper, the neural structure is used to extract the features and prediction is done by passing the video through the original output methods of the YOLO V3 network. Dense net network is used in this process. The better security equipment used monitoring will give better results. GAN network is used to get the better results.

[3] In this project the different methods are used to detect the workers are wearing helmet or not. The method used in this are SSD, R-FCN and YOLO V3. The different versions of these three models are used. The YOLO V3 have given the better results than others. The Darknet-53 is used by YOLO V3 to get the structure for the model.

[4] In this paper, the helmet detection is used to detect the helmet of bikers. The detection is done by the traffic signals cameras. The YOLO V3 as well as CNN is used for detection. The CNN is the light model then the YOLO V3 which gives the effective results. The dataset is created using the different images of the people wearing helmet or not.

[5] In this paper, the work flow of the CNN is explained. It explains about how the CNN model process objects in the different stages. If the image is processed in more stages the model will give more accuracy. Max-pooling layer give the more accuracy as it divides the objects in 4 layers which gives the 4 different values by each layer and at the end the output will be compilation of the 4 layers.

[6] Evaluated a profound learning approach on face identification and face order. Bunching of various countenances are finished utilizing pre-prepared facial dataset. FDDB dataset is applied to prepare and test the proposed model. The proposed model is changed to acquire oddity and high acting in expectation. The precision accomplished on utilizing convolution neural organization is noted high and the future difficulties are announced utilizing constant face pictures and live catching of recordings.

[7] Explained in detail on multi-face recognition model dependent on AI calculations Ada-Boost, SVM (support vector machine), Hear support, and inclination help models. These strategies are obviously disclosed to accomplish high precision with each other. The significant test looked by the scientist is bogus positive rate expansions in certain facial information.

[8] Developed a human location model, an equipment based methodology in which distance estimation utilizing sound wave sensors is utilized. The distance between the two people is resolved through the variety in the sound waves aggregated to shape a distance variable. The sign strength decides the variety somewhere out there. Ordinarily the sign strength is resolved utilizing RSSI (got signal strength esteem). The framework additionally track the situation of the people as for the got estimation of RSSI.

[9] Developed an AI model for picture preparing applications. The framework decides the picture upgrade strategies, picture division, applying channels and assessing the undertaking that accommodating for breaking down different boundaries of the pictures. The introduced framework manages radiation based clinical pictures and dealing with methods. Recreation of pictures with diminished goal additionally decided. For different picture related special advances the diary gives the clues and ideas.

[10] Author on a worldwide conference, evaluated a total survey on AI calculations that suits for the cutting edge specialized registering stages. Many exploration researchers dealing with improving the AI calculations as far as exactness, execution and blunder rate the diary talks about much centered on conventional strategy and versatile focuses made in the customary calculations. Half breed various techniques into new engineering all around planned in numerous applications. The paper unmistakably recognizes another measurement on making new calculations.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

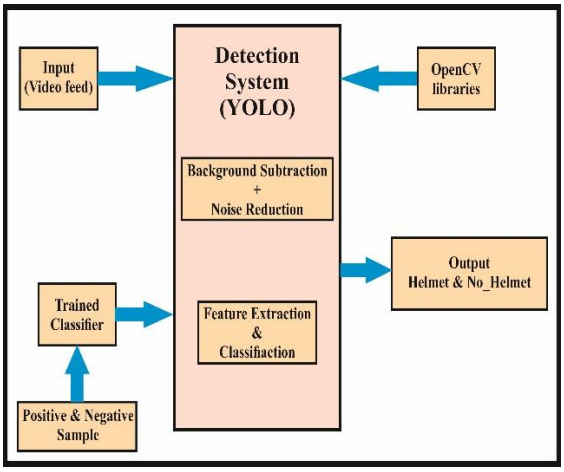
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468

III. PROPOSED SYSTEM ARCHITECTURE

Based on the previous literature considerations discussed in Sec II. The proposed model is evaluated by creating a strong framework that contains improved accuracy and reliable operation on detection of face mask and social distancing. The study is carried out using the basic step to be taken by the society on behalf of avoiding the transit of the virus through social distance maintenance and wearing a face mask. The proposed model uses python as the development platform. YOLO v3 model is proposed herein to focus on achieving accuracy and higher reliability in real-time object detection. YOLO v3 methodology is a kind of one-stage object detection process. Many recent days prediction models need a large GPU to process the batch files of smaller sizes. This process gradually slows down the process during training. Even though YOLO is the stage detector.



IV. METHODOLOGY

1. Data Collection: Data is collected in the form of JPEG files. Downloaded the 2000 images of people with helmets and 2000 images of people without helmets. All the files are stored as Helmet.zip and No_Helmet.zip.

2. Data Pre-processing: We have to give labelling and annotation to every photo. Labelling is the process of tagging the data like text or objects in videos and images to make it detectable and recognizable to computer vision to train the models through machine learning algorithms for the right predictions. For image labelling, We have used a tool called Labelling. Labeling is a free, open-source tool for graphically labeling images. It's written in Python and uses QT for its graphical interface. We have to select the area which we have to annotate using the Draw Rectangle tool. Then give the labeling to the selected area. After saving it will create the two files:-

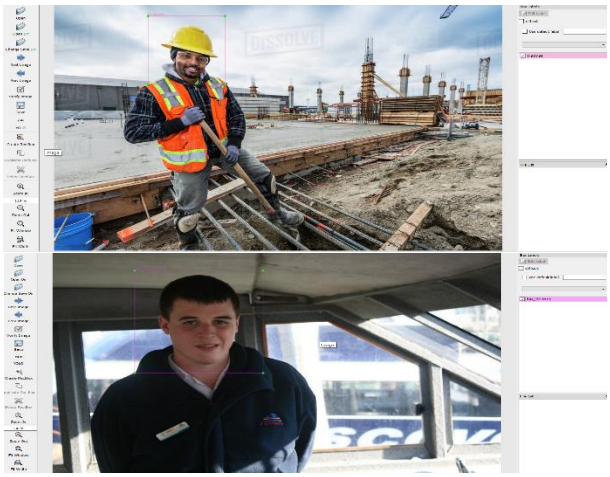
- Image_name.txt
- Classes.txt

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

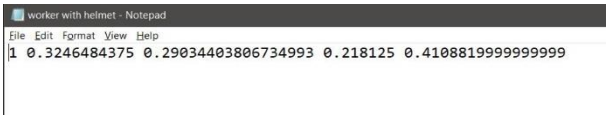
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468



Image_name.txt contains the centre point (X, Y), width and height of labelling.



Classes.txt contains the available class name.

3. Training Model : Yolo V3 Dark net is used to train the model which is developed by the Alexa. Dark net is written using C and CUDA which is open source neural network. The proposed framework uses YOLO v3 prediction model which is derived from the YOLO v1 and v2 model. YOLO model is refreshingly simple and precise form of prediction model with lower latency. YOLOv3 uses the variant of Dark Net that contains 53 layers stacked with more than 100 layers of fully connected layers. The reason behind the latency in YOLO v3 may be the stacked version of layers and followed process. For training we need four files:-

- obj.name file
- obj.data file
- Yolo_custom.cfg file
- Zip of data

obj.name file contains the names of classes required. In our case **Helmet** & **No Helmet**.
obj.data file contains the numbers of classes and location where file will save.
Yolo_custom.cfg file contains the filters and layers the image will be processed by the CNN algorithm.
Zip of data in this we have to make compressed (.zip) file of images and there labeling text file. In our case we have two zip files **Helmet.zip** and **No_Helmet.zip.**
After training we will get **.weights** file which contains the trained model

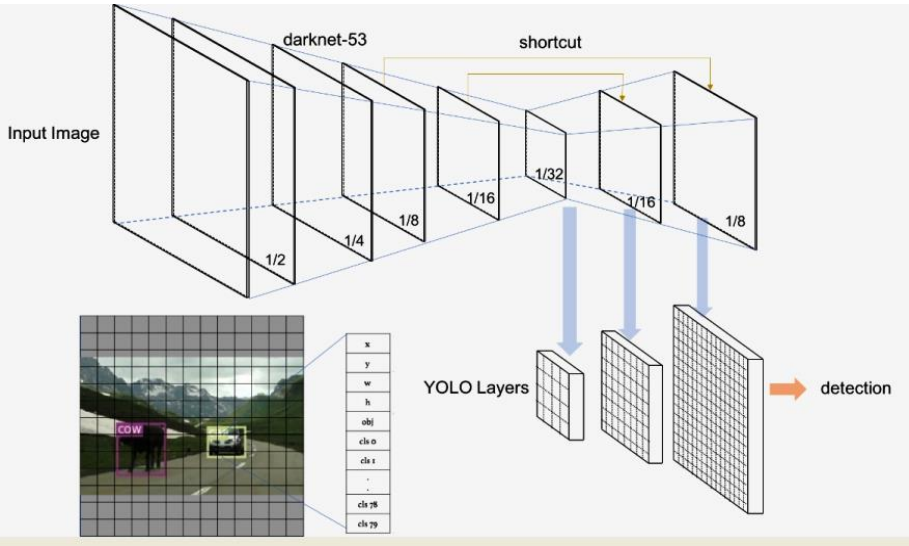
4. Testing Model : For testing the JAVA Script is used to turn on the camera. The script is connected to the Dark net. In real time object is detected and compared with trained model. The system design comprises of three levels of detection process. The first step involves in image acquisition through CCTV camera. The second step is to extract the feature such as human body and detection of face from it. The next step is to extract the trained model of mask structure present with the image. The processed images are stored with the predefined web cloud for anytime access.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468



The architecture of YOLO v3 model in which the uniqueness of the method on handling the feature is clearly depicted. YOLO v3 model is a kind of inspired framework because of the Reset and FPN (Feature Pyramid Network) involved in it. Dark Net 53 act as the feature extractor and it consists of minimum of 53 convolutions. The input test and train images are processed at different spatial domains and then compared.

Feature Pyramid Network (FPN): The various spatial differences of the images are concatenated with the FPN model to form a feature mapping vector that is more unique and precise. The procedure is repeated for various sizes of objects and that could be connected with different detection heads. Yolo v3 is good in detecting large objects compared with other objects. The proposed model is accurate and better in detection of bounding boxes of the detected real time objects. It creates a benchmark model that detects the images of different sizes and different resolutions too.

V. EXPERIMENTAL RESULTS

Results for Pre-Captured Image

Real-Time Results

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468

Iterations	Accuracy
3000	88.13%
6000	92.83%
10000	93.39%

Table [I]

Table [A] shows the accuracy obtained by the iterating the model for training. This results can be differ due to the internet speed and the quality of the camera used.

VI. DISCUSSIONS AND CHALLENGES

From the above discussions, results, and comparisons, the challenges in the proposed works are achieving the accuracy and mapping the unique feature vectors on different sized objects and locations. The variation of the objects in the spatial domain is much helpful for the predictor. The proposed research focused on the detection of violations in social distance in public places and violation of face mask is being concerned. The proposed framework considers the previous research journals and their contributions on detection of social distancing which is discussed in detail in Sec II. From the previous data, the proposed model is created with YOLO v3 architecture that is considerably faster than the previous prediction models and much better in reducing the latency of the same. The challenge in the proposed method lies in the feature mapping process, since the large set of images is considered here, the variation in the sizes and the object differences are considered as challenges. That can be overcome in future works through adopting the global sizes and the bags of specials, bags of information models.

VII. CONCLUSION

Prediction of whether the workers wearing helmets or not is one of the challenges and reducing the risks to workers' life is another task. Most workplaces are installed with CCTV cameras and monitor the people's gatherings more accurately. The coverage of CCTV footage acquires more storage spaces and dealing with the large set of image data requires a fast GPU. The drawbacks of previous methodologies on the same prediction are discussed in Sec II. Provides an idea of the focused work with tenable parameters. The low latency system is the required thing when considering accurate prediction. Many research journals adopted Res Net, dark Net, and VGG Nets to acquire accurate results. The drawbacks of most of the prediction models are high processing time. The single spot detector model or the so-called YOLO version 3 is proposed here. The proposed model considers the features of the images with different scales, different locations, and different angles. The fusions of features are applicable to the YOLO model which added advantage to the accurate prediction of the proposed system. The Table-I shows that if the model has been trained with more iteration the chances of getting the accuracy will be more and also depends upon the quality of the camera used.

REFERENCES

[1] X. Long, W. Cui and Z. Zheng, "Safety Helmet Wearing Detection Based On Deep Learning," 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019, pp. 2495-2499, doi: 10.1109/ITNEC.2019.8729039.

[2] F. Wu, G. Jin, M. Gao, Z. HE and Y. Yang, "Helmet Detection Based On Improved YOLO V3 Deep Model," 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), 2019, pp. 363-368, doe: 10.1109/ICNSC.2019.8743246.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11468

[3] Y. Bo et al., "Helmet Detection under the Power Construction Scene Based on Image Analysis," 2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT), 2019, pp. 67-71, doe: 10.1109/ICCSNT47585.2019.8962495.

[4] J. Hu, X. Gao, H. Wu and S. Gao, "Detection of Workers Without the Helmets in Videos Based on YOLO V3," 2019 12th International Congress on Image and Signal Processing, Biomedical Engineering and Informatics (CISP-BMEI), 2019, pp. 1-4, doe: 10.1109/CISP-BMEI48845.2019.8966045.

[5] M. Disrupt, O. Bandyopadhyay and S. Chatterjee, "Automated Helmet Detection for Multiple Motorcycle Riders using CNN," 2019 IEEE Conference on Information and Communication Technology, 2019, pp. 1-4, doi: 10.1109/CICT48419.2019.9066191.

[6] D. Garg, P. Goel, S. Pandya, A. Ganatra and K. Kotecha, "A Deep Learning Approach for Face Detection using YOLO," 2018 IEEE Punecon, Pune, India, 2018, pp. 1-4, doi: 10.1109/PUNECON.2018.8745376.

[7] H. Filali, J. Riffi, A. M. Mahraz and H. Tairi, "Multiple face detection based on machine learning," 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, 2018, pp. 1-8, doi: 10.1109/ISACV.2018.8354058.

[8] M. Matsumoto, K. Kaneta, M. Naruoka, H. Tanaka and K. Takano, "Tracking Positions of Human Body Parts Based on Distance Measurement with Sound Wave," 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), Taipei, 2017, pp. 514-518, doi: 10.1109/WAINA.2017.135.

[9] M. Hatt, C. Parmar, J. Qi and I. El Naqa, "Machine (Deep) Learning Methods for Image Processing and Radiomics," in IEEE Transactions on Radiation and Plasma Medical Sciences, vol. 3, no. 2, pp. 104-108, March 2019, doi: 10.1109/TRPMS.2019.2899538.

[10] P. P. Shinde and S. Shah, "A Review of Machine Learning and Deep Learning Applications," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697857.

[11] Girshick R. Fast R-CNN [C]// Proc of IEEE International Conference on Computer Vision. 2015:1440-1448.

[12] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]// Proceedings of the 2015 advances in Neural Information Processing Systems. Palais des Congrès de Montréal, Montréal CANADA. 2015:91-99.

[13] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [C]// Proc of European Conference on Computer Vision. Springer, 2016:21-37.

[14] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real time object detection [C]// Computer Vision and Pattern Recognition. 2016:779-786.

[15] Redmon J,Farhadi A.YOLO9000:Better,Faster, Stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2017:6517-6525.

 This work is licensed under a Creative Commons Attribution 4.0 International License