Truck Dumper Control System with Safety Unit Using Real-Time Deep Learning

Nittaya Kerdprasop[1,*], Apirak Worrakantapon[1], Paradee Chuaybamroong[2], and Kittisak Kerdprasop[1]

[1]School of Computer Engineering, Suranaree University of Technology, Thailand
[2]Department of Environmental Science, Thammasat University, Thailand
Email: nittaya@sut.ac.th (N.K.); apirak.wor@gmail.com (A.W.); paradee@tu.ac.th (P.C.); kerdpras@sut.ac.th (K.K.)
[*]Corresponding author

Manuscript received December 30, 2023; revised March 4, 2024; accepted April 15, 2024; published August 9, 2024

Abstract—**Presently, industrial factories have moved toward the era of intelligent automation systems, especially, animal feed industries. They contained several automation machines such as a mixer, pellet mill, and truck dumper. The truck dumper has been used to lift the whole truck and dump the raw material into an intake hopper. However, during its process, neighbor areas are identified as a risk area. Thus, if there is a person standing on the truck dumper platform at the time, this may cause fatal injury or death. In this study, a new proposal to develop a model of automatic human detection using a convolutional neural network that learns and recognizes the important characteristics of the target objects has been introduced. The objective of the study was to develop a system of human detection and risk-level identification in the truck dumper control system to prevent accidents. Herein, several experiments have been conducted in order to select the effective and optimal object detection architecture to apply to the truck dumper control system. In the results, the YOLOv4 model outperformed the Faster Region-Convolution Neural Network (R-CNN) model in both precision and processing speed. The average precision was 99.93% on a day-time dataset and 94.25% on a night-time dataset. The overall accuracy of risk identification was 94.18%. An average processing speed was 31.96 frames per second.**

Keywords—**deep learning, You Only Look Once (YOLO), truck dumper control system, human detection**

I. Introduction

Nowadays, automation technology and robots have played an important role in the production process of industrial plants, which are considered the key to creating productivity throughout the process in order to achieve the most efficiency. In animal feed production plants, production process involves receiving raw materials, storing raw materials, mixing the food pellets, packaging food, and delivering the food to customer vehicles. In order to increase efficiency and quality of food production, machinery is required and involved throughout the process. For this reason, while the machine is working, the surrounding area is considered a danger zone and no person is allowed to enter the vicinity because the operation of the machine may cause accidents to employees or persons. In the process of receiving raw materials, a general factory will receive raw materials from the seller's truck. Customer trucks are divided into 2 types: those that can lift the back of the truck by themselves and those that cannot lift the back of the truck. For the second type, once the seller has parked the truck in the raw materials dumping area, the driver must get out of the truck and walk away from the area called truck dump. In this area, the factory has a machine for lifting the entire truck to tilt it up so that the raw materials inside the truck can flow down to the raw material receiving pit. This machine is called a truck dumper which is controlled by employees in the control room through the Truck Dumper Control System.

While the machine is working, no person will be allowed to enter the vicinity of the truck platform, especially on areas where the dump truck is raised. This is because while the floor is raised and tilted up, if someone stands on it, that person may slip and fall into the raw material receiving pit resulting in injury or possibly death. Preventing such accidents in general industrial plants still requires humans to observe. However, there may be a high chance of human error from the blind spots that cannot be seen from many angles from inside the control room. Computer vision technology and deep learning are thus applied to improve safety in the dangerous areas.

In this work, we adopt a real-time deep learning technique to automatically detect human in the danger zones. A system has been developed to prevent accidents that may occur to people in the truck dump work area. A one-stage detection network called You Only Look Once (YOLO) is used for detecting living objects on the three-dimensional camera frames. When the people within the medium-risk area is detected, the system will send an alarm signal to the screen in the control room. When the system detects a person within a high-risk area, the system will send a signal to stop the machine immediately. In the experiment to test the accuracy of the system, we use video footage installed at real locations from various events during the day and at night.

II. Literature Review

Within the past, computer vision based on different statistical and numerical strategies has been a well known procedure for identifying people in still pictures and video stream and for recognizing activity [1–5]. Nevertheless, recognition accuracy is still an important aspect waiting for constant improvement since light and locate perspectives influence the execution of the computer vision strategies. One possible solution to solve the recognition accuracy problem is by applying the background subtraction techniques. But for recognizing the moving objects, this technique is difficult to apply. Utilizing the three dimensional camera may solve this problem, but such camera is quite expensive [6, 7]. Applying advanced software techniques is an inexpensive way to help recognizing moving objects with a satisfied accuracy rate.

The complex deep learning techniques such as Convolutional Neural Network (CNN) have thus been extensively applied to recognize moving things such as human and vehicles [8–12]. The extensiveness of CNN application is based on the high performances of the Central

Processing Unit (CPU) and Graphics Processing Unit (GPU) in modern computer machines, as well as the advancement of the deep learning algorithms. The success of CNN in detecting objects is due to the multi-layer architecture of this kind of neural network that makes it excellent in extracting representative features from images [13, 14]. Common architecture of most CNN contains two main parts, that are, the convolution part and the classification part. The convolution part is responsible for extracting only representative features that are useful for recognizing objects. The classification part is for classifying objects. When applying CNN to recognize moving objects in the real-time applications, the speed in recognizing objects is still unsatisfactory because the learning phase takes a lot of time.

To solve the time-consuming problem during the learning phase of CNN, Redmon and colleagues proposed in [15] 2015 a faster image recognition scheme called You Only Look Once, or YOLO. This new method has such name because it performs just one round of image scan and then transforming an image classification problem to be a logistic regression learning. A wide range of real-time machine vision applications have shown successful results [16–22].

YOLO recognizing moving objects faster than regular CNNs because it locates an object in the image and classifies that object at the same time. An image is firstly divided into S×S grids, then the position and size of interested object has been predicted with a confidence score for possible object type. The high confidence score represents the high probability of the target. The object detection occurs on each grid simultaneously [15, 20], therefore, object detection within the image takes drastically less time than ordinary CNN techniques. In this research work, we adopt YOLO for real-time human detection to prevent harm in the restricted area of truck dumper.

III. Materials and Methods

A. Research Framework

The research concept for developing a truck dumper control system with safety unit is shown in Fig. 1. Real-time data are obtained from the Internet Protocol (IP) camera. The image data are then sent to the proposed system for detecting human in the truck dumper area. If human is detected, the system is also evaluating risk level of the human. In this research, we define two levels of risk: medium-risk and high-risk. Medium-risk results in showing alarm to the staff in the control room, whereas high-risk can send a signal to the Programmable Logic Control (PLC) to immediately stop the truck dumper machine. Communication between the proposed system and the truck dumper control is through the software known as Open Platform Communications (OPC). The OPC server is responsible for direct communication between the result obtained from the human detection and risk-level identification part to the PLC to control the truck dumper machine.

In the proposed system, YOLO architecture is adopted to automatically detect human in the truck dumper area. We also develop the risk-level identification module to evaluate and identify degree of risk as either medium or high.

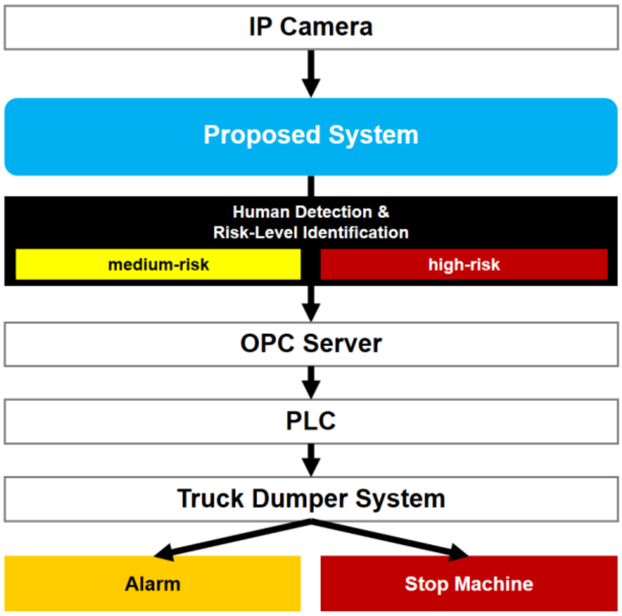

Fig. 1. Research framework for truck dumper control system.

B. Data

The video streaming data used for training and testing the YOLO human detection model are from the real situation of truck dumper areas at both day time and night time. Images are prepared to be of two sizes: 1920×1080 and 963×1080 pixels. Situation at day time contains 5,043 images, whereas the night time has 4,068 images. Examples of images are displayed in Fig. 2.

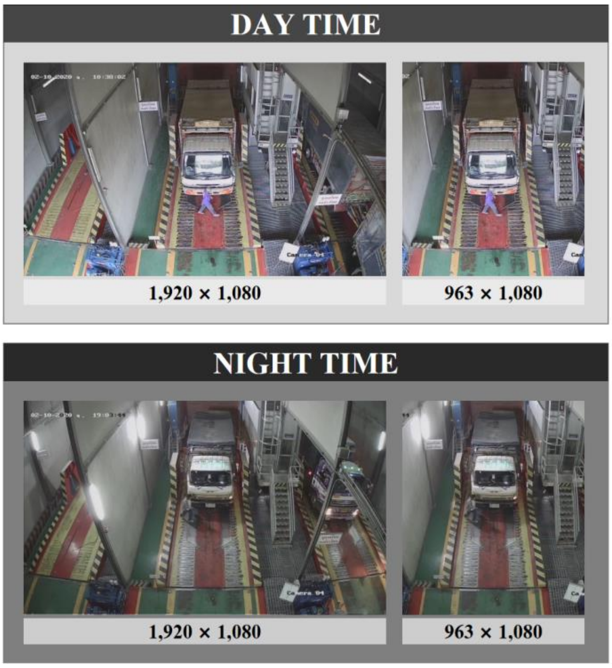

Fig. 2. Examples of truck dumper area images at day and night time.

To train the automatic human detection process, boundary of human in the picture has to be identified. The software LabelImg is applied to identify human object. Its example is shown in Fig. 3. Position of a human to be classified as either medium-risk or high-risk is also shown in Fig. 4.

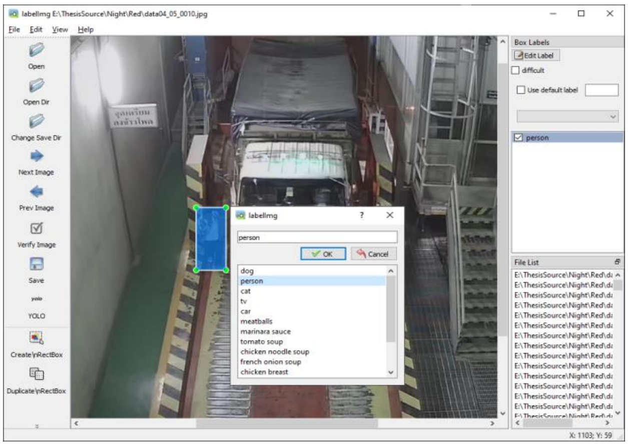

Fig. 3. Bounding human object in the truck dumper area image.

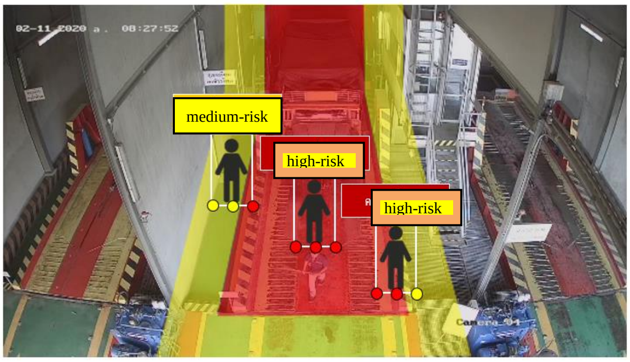

Fig. 4. Risk-level identification of human position in the truck dumper area as either medium-risk or high-risk.

C. Performance Evaluation

The important module of the proposed truck dumper control system is the human detection and risk-level identification. On performance evaluation of the system, we thus assess it on two aspects: efficiency of human detection and correctness of risk-level identification.

Efficiency of human detection is assessed with the two main metrics: Average Precision (AP) and Frame per Second (FPS). AP is measured from the recall and precision values. Its computation is shown in Eqs. (1) and (2).

$$AP = \sum(r_{n+1} - r_n)Pinterp(r_{n+1}) \quad (1)$$

where $Pinterp(r_{n+1})$ is the precision of object detection at position r_{n+1} such that r_n is the object recall value at current position and r_{n+1} is the object recall value at the next position.

$$Pinterp(r_{n+1}) = \max_{r':r' \geq r_n} P(r') \quad (2)$$

where max $P(r')$ is any precision value of object detection in the range r_n to r' such that r' is the position of any recall value that is greater than or equal to r_n.

FPS is the metric to measure speed in detecting object of the model. The higher FPS is the faster a system can operate in real-time environment. FPS can be computer as in Eq. (3).

$$FPS = \frac{1000}{Processing\ Time\ (ms)} \quad (3)$$

To compare performance of human detection, we apply two deep learning algorithms: YOLO (version 4) and Faster R-CNN. Correctness of risk-level identification is measure from the model result compared against the ground truth.

IV. Experimentation and Result

On evaluating performance in the aspect of precision of the model to detect human in the truck dumper area, faster R-CNN is compared against YOLO. The results are shown in Table 1. It can be seen from the results that YOLO performs better than Faster R-CNN in terms of average precision and number of frame per second. However, model performance during the night time is slightly lower than the day time.

Table 1. Performance on human detection of YOLO versus faster R-CNN

Model	Time	Image size	AP	FPS
Faster R-CNN	Day	963×1080	63.0854	17.0958
		1920×1080	51.5875	14.2366
	Night	963×1080	63.3128	17.0958
		1920×1080	50.6053	14.2366
YOLO	Day	963×1080	**68.1349**	**31.9644**
		1920×1080	54.3163	27.1791
	Night	963×1080	**63.8033**	**31.9644**
		1920×1080	51.8455	27.1791

From better performance of YOLO, this architecture is thus selected to be used in our system. The next step of evaluation is the correctness on identifying risk level. Experimental results are illustrated in Table 2.

Table 2. Correctness on identifying risk-level of YOLO models

Risk level	Time	Image size	Correctness (%)
High	Day	963×1080	91.53
		1920×1080	86.42
	Night	963×1080	96.58
		1920×1080	95.03
Medium	Day	963×1080	99.64
		1920×1080	95.03
	Night	963×1080	85.76
		1920×1080	89.76

It is noticeable that correctness on classifying risk level of small size images is a little higher than large size images. During day time operation, risk-level identification correctness is as high as 99%. The lowest correctness at 85% is still acceptable. Result of our system development is also shown in Fig. 5.

Fig. 5. The truck dumper control system with YOLO as a module for human detection and the developed risk-level identification

V. Conclusion

This research presents the development of system to improve safety in the truck dumper area of the feed-mill industry. Truck dumper is the strictly control area during its operation. To prevent harm to humans working in this area, staff in the control room is responsible for monitoring safety in the area. This traditional process is however unsatisfied due to the operation spots existing in the surrounding.

We thus design and develop the automatic system to help improving safety. Our system design has adopted the IP camera to replace human eyes and include deep learning technique in our design. The real-time deep learning architecture named YOLO is used for automatic human detection, then the detected human is to be identified the level of risk based on standing position. The medium-risk can result in alarm signal to alert staff in the control room, whereas the high-risk level will result in the immediate stop of the truck dump machine. This action can be done through the communication between OPC server and PLC that control the truck dumper.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Nittaya Kerdprasop is responsible for manuscript preparation and submission. Apirak Worrakantapon is the main contributor in data collection and experimentation. Paradee Chuaybamroong helps revising the manuscript. Kittisak Kerdprasop contributes the research idea and research design improvement. All authors had approved the final version.

References

[1] Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," in *Proc. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2006, vol. 2, pp. 1491–1498.

[2] N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance," in *Proc. European Conference on Computer Vision*, 2006, pp. 428–441.

[3] C. Raghavachari, V. Aparna, S. Chithira, and V. Balasubramanian, "A comparative study of vision based human detection techniques in people counting applications," *Procedia Computer Science*, vol. 58, pp. 461–469, August 2015.

[4] H. Zhang, Y. Zhang, B. Zhong, Q. Lei, L. Yang, J. Du, and D. Chen, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, 1005, February 2019.

[5] J. K. Agrawal and M. S. Ryoo, "Human activity analysis: A review," *ACM Computing Survey*, vol. 43, no. 3, 16, April 2011.

[6] X. Yang and Y. Tian, "Effective 3D action recognition using EigenJoints," *Journal of Visual Coomunication and Image Reoresentation*, vol. 25, no. 1, pp. 2–11, 2014.

[7] D. I. B. Hagebeuker and P. Marketing, "A 3D time of flight camera for object detection," *PMD Technol. GmbH, Siegen*, 2007.

[8] J. Zhao, G. Zhang, L. Tian, and Y. Q. Chen, "Real-time human detection with depth camera via a physical radius-depth detector and a CNN descriptor," in *Proc. 2017 IEEE International Conference on Multimedia and Expo (ICME)*, 2017, pp. 1536–1541.

[9] J. H. Kim, H. G. Hong, and K. R. Park, "Convolutional neural network-based human detection in nighttime images using visible light camera sensors," *Sensors*, vol. 17, no. 5, 1065, May 2017.

[10] T. Liu and T. Stathaki, "Faster R-CNN for robust pedestrian detection using semantic segmentation network," *Fromtiers in Neurorobotics*, vol. 12, 64, October 2018.

[11] N. AlDahoul, A. Q. M. Sabri, and A. M. Mansoor, "Real-time human detection for aerial captured video sequence via deep models," *Computational Intelligence and Neuroscience*, vol. 2018, 1639561, February 2018.

[12] X. Wang and S. Hosseinyalamdary, "Human detection based on a sequence of thermal images using deep learning," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W13, pp. 127–132, 2019.

[13] D. Strigl, K. Kofler, and S. Podlipnig, "Performance and scalability of GPU-based convolutional neural networks," in *Proc. 2010 18th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP)*, 2010, pp. 317–332.

[14] H. Jiang and E. Learned-Miller, "Face detection with the faster R-CNN," in *Proc. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 650–657.

[15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[16] R. Laroca *et al.*, "A robust real-time automatic license plate recognition based on the YOLO detector," in *Proc. 2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–10.

[17] Q. Peng *et al.*, "Pedestrian detection for transformer substation based on gaussian mixture model and YOLO," in *Proc. 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2016, vol. 2, pp. 562–565.

[18] D. Shen, X. Chen, M. Nguyen, and W. Q. Yan, "Flame detection using deep learning," in *Proc. 2018 4th International Conference on Control, Automation and Robotics (ICCAR)*, 2018, pp. 416–420.

[19] C. Tang, Y. Ling, X. Yang, W. Jin, and C. Zheng, "Multi-view object detection based on deep learning," *Applied Sciences*, vol. 8, no. 9, article 1423, 2018.

[20] J. Li, J. Gu, Z. Huang, and J. Wen, "Application research of improved YOLO v3 algorithm in PCB electronic component detection," *Applied Sciences*, vol. 9, 3750, September 2019.

[21] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, "Apple detection during different growth stages in orchards using the improved YOLO-V3 model," *Computers and Electronics in Agriculture*, vol. 157, pp. 417–426, February 2019.

[22] X. Yin, Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, and L. Kurach, "A deep learning-based framework for an automated defect detection system for sewer pipes," *Automation in Construction*, vol. 109, 102967, January 2020.