Journal of Science and Technology in Civil Engineering, HUCE, 2024, 18 (2): 163–174

AUTOMATED EVALUATION OF UNSAFE WORKING POSTURES IN LIFTING AND CARRYING HEAVY OBJECTS IN CONSTRUCTION USING A CNN DEEP LEARNING MODEL

Quangdung Tran[a,*]

[a]*Faculty of Building and Industrial Construction, Hanoi University of Civil Engineering, 55 Giai Phong road, Hai Ba Trung district, Hanoi, Vietnam*

Article history:
Received 02/04/2024, Revised 20/05/2024, Accepted 07/06/2024

Abstract

Lifting and carrying heavy objects takes place very commonly and with high intensity on construction sites. Wrong working posture likely leads to musculoskeletal disorders, then risks of health damage and illness for workers. This study seeks to propose a method to automatically assess unsafe postures of lifting and carrying heavy objects by combining the RTMPose deep learning model to detect people from videos and a convolutional neural network (CNN) model to automatically extract, evaluate and classify the worker's posture skeleton frames into two states "safe posture" and "unsafe posture". The study also provided two datasets of the worker's skeleton posture skeleton frames of lifting and carrying heavy objects for further research and applications. The proposed method has been experimentally tested with good results. Finally, some ways to apply this method in managing and controlling occupational health risks in practice have also been discussed.

Keywords: lifting and carrying heavy objects; unsafe posture skeleton frame; occupational safe and health risks; CNN model; construction site.

https://doi.org/10.31814/stce.huce2024-18(2)-12

1. Introduction

Automation in construction is being promoted in many aspects from design, estimation, site supervision, building operations, bidding, and procurement to contract management activities [1–3]. Especially, the application of information technology and automation in occupational safety and health management on construction sites has recently attracted a lot of attention from research to practical implementation [4–7].

The construction industry is a very dangerous and toxic industry with very high annual rates of fatal accidents, injuries and illnesses [8, 9]. In general, accidents and illness in construction sector are caused by many reasons in which two main root factors being unsafe working conditions and unsafe working behavior [10–12]. Construction activities require workers to move body parts with great intensity and frequencies; and in many cases with very uncomfortable positions, especially lifting, carrying, and moving heavy objects. These lead to risks of musculoskeletal disorders, health damage and disease [6, 9]. Overall, health problems related to musculoskeletal disorders can cause workers to miss an average of four days of work-off, and companies spend a significant amount in compensation and health care costs [13, 14]. Therefore, monitoring the workers' working posture and assessing the risk of musculoskeletal disorders is very important, from which to have plans and solutions to effectively control the risks of health damage and disease for workers in construction. Currently, monitoring onsite working conditions, unsafe behaviors and postures of workers is mainly

*Corresponding author. *E-mail address:* dungtq@huce.edu.vn (Tran, Q.)

done by using manual techniques and is considered to be highly subjective. Furthermore, this work is often performed according to one planned schedule and not continuously in real time [15].

Besides, construction site is often large and complex with many uncertainty and unpredictable factors; especially with the presence of many workers, machines, and equipment operating at the same time. As a result, assessing and monitoring the workers' unsafe working behavior and posture, thereby controlling the risk of occupational health damage and illness on construction site becomes a very difficult and poor effective task.

Working postural assessment based on observation has been widely adopted to identify safety and health risks in construction sector. In order to overcome the limitation of time consuming and human errors, the recent studies have proposed the advantaged approaches to assess automatically real-time working postures using both wearable sensors and computer vision techniques. Both these approaches have shown promising results in terms of awkward posture identification, the computer vision-based approach is considered to be more suitable, effective without any interference with on-going works. This present study aims to improve the effectiveness of controlling the risk of health damage and disease due to musculoskeletal disorders by developing an artificial intelligence model to automatically real-time assess and classify unsafe postures of workers on construction sites. The study has the following two objectives:

(1) Create a data set of the workers' working posture skeleton frames for lifting, carrying, and moving heavy objects;

(2) Build a convolutional neural network (CNN) model to automatically evaluate and classify the workers' working postures when lifting, carrying, and moving heavy objects into two groups: "safe posture" and "unsafe posture".

2. Literature review

Behavior-based safety is an effective approach to modify the workers' unsafe behavior to undertake works more safely and healthily. In construction, there are the three key groups of unsafe behavior, namely (1) failure to use personal protective equipment (PPE); (2) exposure to a hazardous area; and (3) failure to follow safety procedures (including unsafe working posture). As mentioned, controlling the people's unsafe behavior has been still reliant mainly on the ability of onsite engineers and safety staffs; this process is criticized to be very time-consuming, discontinuous, less effective and subjective [16]. In order to improve such tasks on construction sites, modern and advantaged technologies have been applied to automatically real-time perform such tasks.

Firstly, with the development of measurement sensoring technology, many previous studies have applied this technology to measure and record unsafe human movements and postures in real time [17, 18]. For instance, the studies by Zhao and Obonyo [19] and Yan, Li [20] proposed a solution for workers to wear measurement sensor devices to record movement in real time and develop an early warning system to help workers be aware of the level of inappropriate posture and make timely adjustments to prevent the risk of musculoskeletal disorders. The study of Antwi-Afari, Li [21] has used a sensor to detect abnormal foot pressure patterns in order to identify ergonomically hazardous postures caused by over-exertion during construction activities. However, this technology is less likely to be applied in practice at construction sites in near future due to the high cost and inconvenience of wearing sensor devices. Nearly, there is a shift toward the computer vision-based approaches to solve problems of unsafe behavior in construction [7, 22] (Table 1).

For evaluation of the working posture, previous studies have used many various algorithms, such as object detection algorithms, object tracking algorithms, image classification algorithms, and activity recognition algorithms to identify automatically unsafe behavior. For example, Seo and Lee

Table 1. Previous studies using computer vision-based approach on unsafe behavior in construction

Categories of unsafe behavior	Objectives	Approaches	References
Failure to use personal protective equipment	Detect the workers not wearing safety harness	R-CNN and classification deep network	Fang, Ding [23]
	Detect various PPEs	Yolo5 and seahorse optimization (SHO) algorithm	Nguyen, Tran [5]
	Detect the workers not wearing safety hardhat	R-CNN using 2D images	Fang, Li [24]
	Safety harness hanging does not meet requirements (e.g. hook to fixed objective)	Deep neural network architecture based on the MobileNet using 3D images	Zhu, Zhu [25]
Exposure to a hazardous area	Detect the workers who enter dangerous area	Yolo2	Luo, Liu [26]
	Struck or be close to bulldozer/excavator that is backing up	Yolo3 using UAV images to measure the distance among objects	Kim, Liu [27]
Failure to follow safety procedures (including unsafe working posture)	Detect the workers traverse supporting beams on a deep foundation pit	R-CNN	Fang, Zhong [28]
	Transporting works by hanging objects	Using deep learning and ontologies with 3D images to detect heavy equipment and people	Qiao, Qie [29]
	Detecting and monitoring the movement of construction equipment	Using spatio-temporal features and support vector machine classifiers	Golparvar-Fard, Heydarian [30]
	Evaluate workers' working posture	Using Support Vector Machine (SVM) to classify 2D images	Seo and Lee [16]; Martínez-Rojas, Marín [31]
	Identify the workers' unsafe postures	Combined deep learning models and Bayesian models to identify unsafe postures and behaviors of workers from 2D images	Luo, Li [32]

[16] and Martínez-Rojas, Marín [31] applied the computer vision technology to classify and evaluate workers' working posture in real time based on body silhouettes using two dimension (2D) images from cameras. Luo, Li [32] combined deep learning models and Bayesian models to identify unsafe postures and behaviors of workers from 2D images. Besides, some studies have applied human posture estimation models to extract 2D or 3D skeleton from camera data to evaluate the ergonomic posture of workers [19, 20, 33]. Such studies have tried to improve the unsafe working posture assessment at construction sites by automatically classifying risky postures while tasks are ongoing. However, this 2D image-based deep learning approach faces a considerable challenge of having large, comprehensive training datasets. Such datasets of training images must be large enough to include any possible variations on images. This is a more significant issue for vision-based working posture classification in construction sector as a human body consists of the head, torso and limbs, creating a very huge range of postures depending on different tasks. Moreover, there is still a limitation of research on specific tasks on construction sites, such as lifting and carrying heavy objects.

To fill the gap, the current study proposes a solution to apply a deep learning model trained by the COCO-pose dataset to identify the worker's image and posture combining with a proposed CNN model to automatically assess and classify the posture of workers in lifting and carrying heavy objects on construction site into two states: "safe" and "unsafe" in real time.

3. Research methodology

3.1. Research framework

In order to detect unsafe working postures of workers when lifting, carrying heavy objects on the construction site, the study proposes an implementation process shown in Fig. 1. Accordingly,

- Step 1: Collect an original database of images or videos of safe and unsafe working postures in lifting and carrying heavy objects;

- Step 2: Using a deep learning model to detect people in the original images and videos, then identify key points belonging to body parts and draw a skeleton frame of that person's posture;

- Step 3: Developing an althorism to extract the skeleton frames from the images, then collect them into two datasets including "safe posture skeleton frames" and "unsafe posture skeleton frames"; and

- Step 4: Developing and training a CNN model to automatically evaluate and classify the worker's posture skeleton frame in lifting and carrying heavy objects.

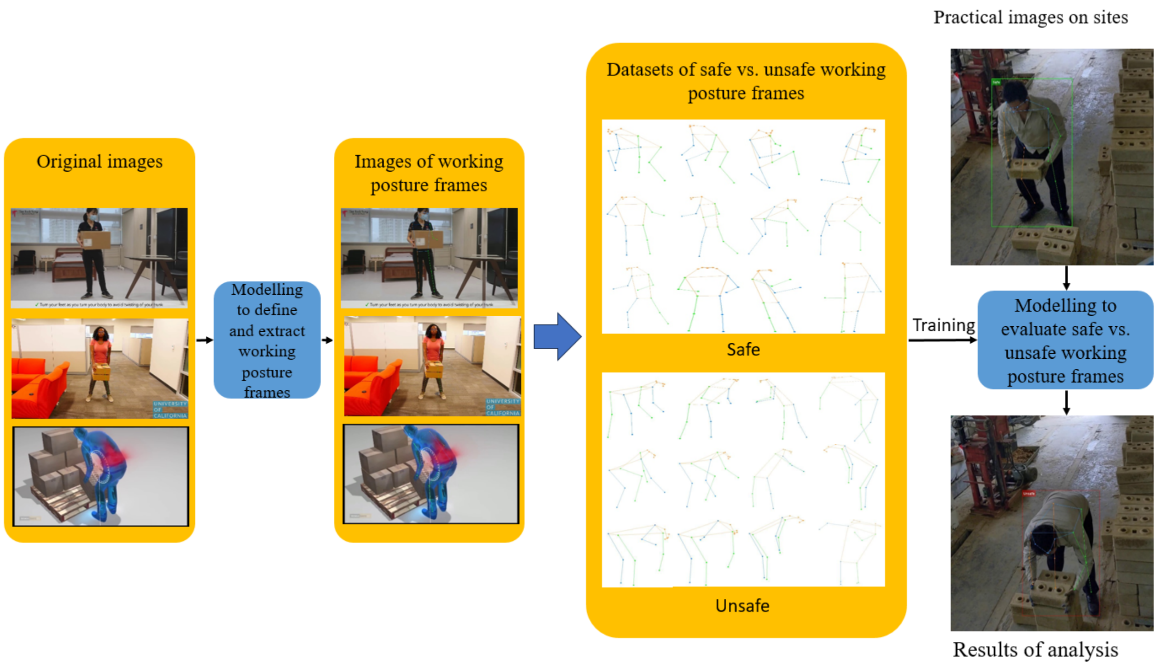

Figure 1. Research framework

3.2. Developing datasets of worker's posture skeleton frames in lifting and carrying heavy objects

To build two datasets of safe vs. unsafe human working posture skeleton frames in lifting and carrying heavy objects, this study uses a RTMPose deep learning model that has been developed, trained with the COCO-pose dataset and widely shared on the Mmpose open library [34, 35]. This is a powerful model to detect the human posture skeleton frames at high speed, suitable for tasks that need to be performed in real time [36]. The COCO-pose dataset includes 200 thousand of images labeled with 17 key points on the body to support assessment of workers' working postures. These points are connected together to represent the human postural skeleton frame in the image (Fig. 2).

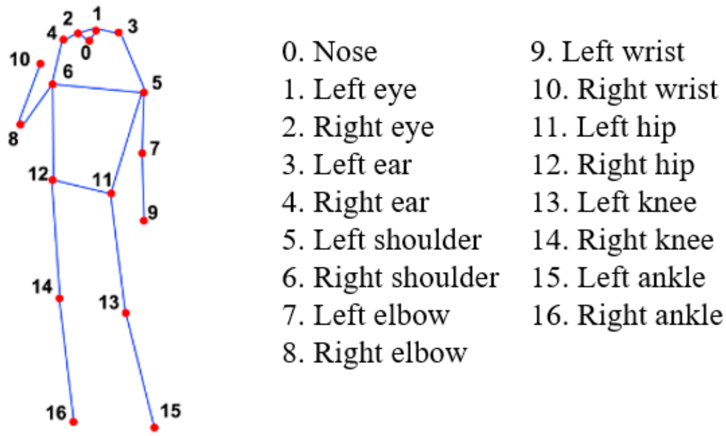

Figure 2. Human skeleton frame

In this study, the initial data on the posture of lifting, carrying, and moving heavy objects were extracted from the guideline videos provided by ergonomic experts from Internet sources [37, 38]. In these guidelines, the various postures of lifting and carrying heavy objects have been recommended by experts as safe (should be followed) or unsafe (should be avoided). Then, the human posture skeleton frames will be extracted from these instructional videos and compiled into 2 datasets including 2220 images of "safe working posture skeleton frames" and 2584 images of "unsafe working posture skeleton frames". For classification in deep learning, the proportion of classification labels is often set to be approximately equal; this is to avoid the circumstance where the model prioritizes identifying one type of label that dominates the rest, making it difficult to train the model to achieve the desired results. Figs. 3 and 4 show an overview of the process of creating the two datasets of working postures in lifting, carrying, and moving heavy objects.

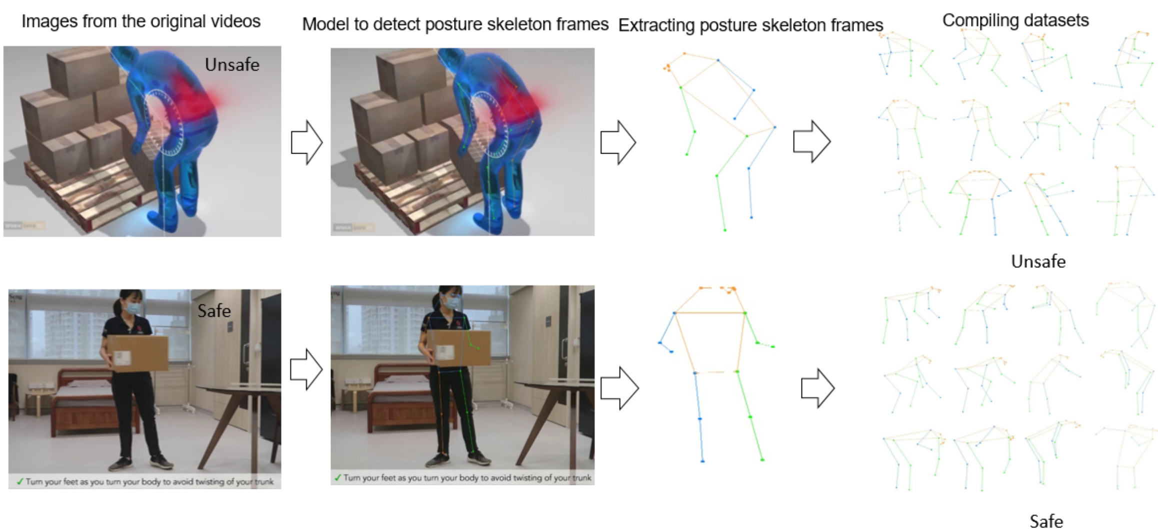

Figure 3. An overall process to build the two datasets of workers' posture skeleton frames

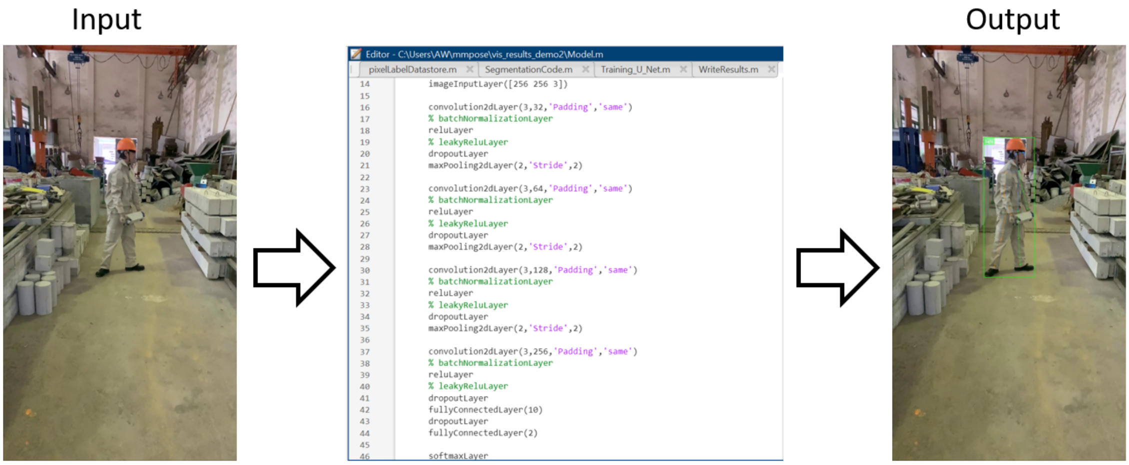

Figure 4. An althorism to extract the human posture skeleton frames

3.3. Developing and training a CNN model to evaluate and classify the working posture skeleton frames

This study develops a CNN model to automatically evaluate and classify the posture of lifting and carrying heavy objects as safe or unsafe. Fig. 5 shows the architecture of the proposed CNN model; and the details of the size of each layer are shown in Table 2. The model includes convolution layers to determine image features and max-pooling layers to highlight identified features. The Relu and dropout activate functions are intended to reduce the linearity of the model, allowing the model to learn and detect complex cases without sticking too closely to the original data. Next, all the values that are the results of the convolutional and max-pooling layers in the previous stage will be spread out in vector form and calculated through two fully connected layers as in the artificial neural network. A softmax function is placed in the last layer, this function will use the result vector from the previous layer to calculate the probability that the input data belongs to the label "safe" or "unsafe". The formula of the softmax function is calculated as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{1}$$

where σ is the value of the softmax probability; $\vec{z}$ is the input vector, the result of data from previous layers in the model; e^{z_i} is the standard exponential function for the i^{th} z value; K is the number of predicted labels, in this study $K = 2$ corresponding to 2 labels "safe" and "unsafe".

Once the softmax layer determines which label the input image belongs to, so that the model is updated with weights after each training process, a cross-entropy loss function is added in the final layer calculated using the formula:

$$\text{Loss} = -\left(y\log(p) + (1-y)\log(1-p)\right) \tag{2}$$

where y is the binary indicator value (0 or 1) corresponding to the two labels that need to be classified and p is the prediction probability calculated from the softmax function. The value of this loss function represents the difference between the model's prediction and the desired result. From there, the model has a basis to update its weights and make an assessment of the model's ability to accurately detect itself.

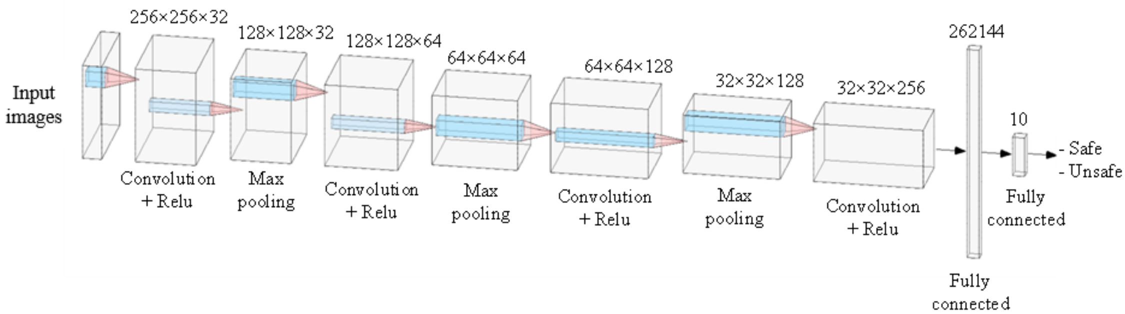

Figure 5. Architecture of the proposed CNN model

Table 2. Architectural details of the layers in the proposed model

No	Layers	Input data	Size	Learning parameters
1	Input images	$256 \times 256 \times 3$	-	
2	Convolution	$256 \times 256 \times 3$	32 filters size $(3 \times 3 \times 3)$	$3 \times 3 \times 3 \times 32 + 1 \times 1 \times 32$
3	Relu + dropout	$256 \times 256 \times 32$	-	
4	Max pooling	$256 \times 256 \times 32$	2×2	
5	Convolution	$128 \times 128 \times 32$	64 filters size $(3 \times 3 \times 32)$	$3 \times 3 \times 32 \times 64 + 1 \times 1 \times 64$
6	Relu + dropout	$128 \times 128 \times 64$	-	
7	Max pooling	$128 \times 128 \times 64$	2×2	
8	Convolution	$64 \times 64 \times 64$	128 filters size $(3 \times 3 \times 64)$	$3 \times 3 \times 64 \times 128 + 1 \times 1 \times 128$
9	Relu + dropout	$64 \times 64 \times 128$	-	
10	Max pooling	$64 \times 64 \times 128$	2×2	
11	Convolution	$32 \times 32 \times 128$	256 filters size $(3 \times 3 \times 128)$	$3 \times 3 \times 128 \times 256 + 1 \times 1 \times 256$
12	Relu + dropout	$32 \times 32 \times 256$	-	
13	Fully connected	$32 \times 32 \times 256$	- $10 \times 262144 + 10 \times 1$	
14	Dropout	$1 \times 1 \times 10$	-	
15	Fully connected	$1 \times 1 \times 10$	- $2 \times 10 + 2 \times 1$	
16	Softmax	$1 \times 1 \times 2$	-	
17	Outputs	$1 \times 1 \times 2$	-	
	Total:			3,009,888

3.4. Training and experimental evaluating environment of the proposed CNN model

The process of training and experimental evaluating the model is performed by Python 3.8.18 language. The study was performed on an Alienware 17R4 laptop with Intel(R) Core (TM) i7-7700HQ @2.80GHz CPU configuration, NVIDIA GeForce GTX 1070 graphics card and CUDA Toolkit 9.0 to accelerate computation.

4. Results and discussion

4.1. Training the CNN model and evaluation parameters

Fig. 6 shows the training graph of the CNN model after 30 training times. It can be seen that with the existing posture frame dataset, the model quickly achieves high accuracy. At the last training time (30th), the model achieved the highest accuracy of 99.58%. Then, the model's loss function graph approaches the value 0 after the 5th training time. After that, the model's loss function value fluctuates stably above the value 0 and ends at the 30th time with the result = 0.0021. The results from the two graphs show that the training achieved good results and the model was suitable for the purpose of the study.

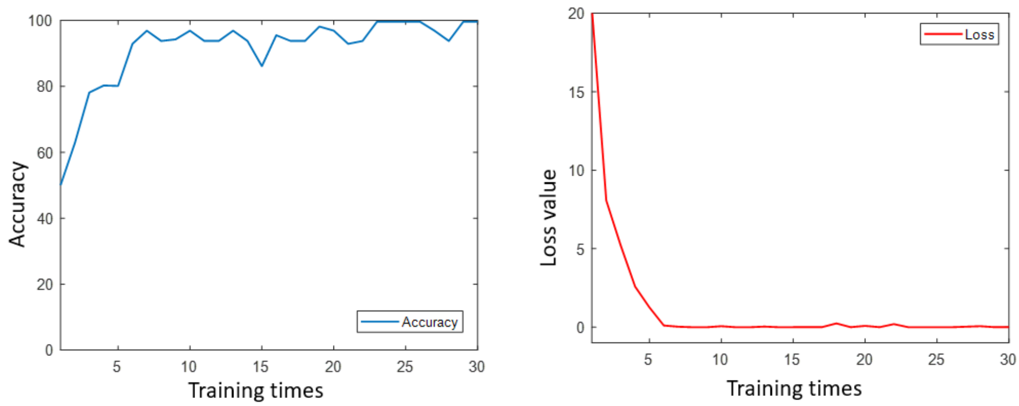

Figure 6. Results of the training process

The proposed method of assessing the workers' working posture in this study is tested by the following indicators: accuracy (ACC), true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV) and negative predictive value (NPV). In order to support controlling disease risks due to musculoskeletal disorders, detecting unsafe working postures should be given priority over safe ones; therefore, the evaluation parameters of the model consider the "unsafe" label as a positive label and the "safe" label as a negative label. The results of the evaluation parameters are shown in Table 3.

Table 3. Results of the evaluation parameters of the proposed method

No	Parameters	Results
1	ACC	99.58%
2	TPR	97.81%
3	TNR	97.32%
4	PPV	96.42%
5	NPV	96.78%

4.2. Experimental testing the proposed method

The proposed method has been experimentally tested for its ability to accurately work in real conditions on complex construction sites by evaluating and identifying unsafe working postures in two videos of workers performing lifting and carrying formwork and big bricks (Fig. 7). The results show that the proposed method is applicable in practical environments.

Figure 7. Results of experimental testing of the proposed model

4.3. Discussion

The indicators calculated in Table 3 showed that the proposed method has a good capacity to evaluate and classify the workers' working postures into the two states, namely "safe" and "unsafe". These results are better than that of previous studies. For example, Vukicevic, Petrovic [39] developed a method to automatically determine the safe and unsafe positions of three types of pushing and pulling, drilling, and polishing collaborative operations with an accuracy of 84.67%, 92%, and 98%, respectively. In another study, Hossain, Azam [14] used the Human 3.6 dataset to detect 3D human pose, then detect labor risks divided into 5 levels, the model achieved an overall accuracy of 89.01%. The result of experimental test also indicated the applicability of the proposed method into the practical environment.

In fact, monitoring camera systems are now quite low cost and are being installed on very many construction sites in Vietnam, so the proposed method becomes more feasible for practical application and potentially effective in terms of economic. This technology can be applied to manage and control occupational safety and health risks on construction sites in many different ways. For example, each worker will wear a small-size sensor attached to their protective gear to receive signals from the working posture assessment system. When detecting a worker performing unsafe working postures within a certain period of time, the system will automatically send a warning signal directly to the worker in real time. Furthermore, the system can also send signals and provide information to site managers and supervisors about the code and location of the workers performing unsafe working postures for timely reminders. The system can also record personal data of each worker to support evaluation, reward/punishment or allocate construction tasks reasonably during the management process on the construction site. The proposed method will be a useful means of quickly identifying and classify risky activities that need immediate intervention. Finally, it is worth noting that the proposed method is not limited in construction sector but also can be adapted to apply in other industrial sectors.

However, from a technical point of views, the computer vision-based approach is facing a challenge to collect training images from a real-world. The performance of vision-based working posture classification would depend mainly on the quantity of training images that should be large enough to not miss any possible critical variations on images. This challenge is more significant for vision-based working posture classification in construction sector.

This study has some notable limitations. First, this newly established method just focuses on identifying, evaluating and classifying safe vs. unsafe working postures; they are the first steps of a musculoskeletal disorders-related risk assessment process. The risk level of musculoskeletal disorders depends on very many factors beside the wrong working posture, such as the lifted object's weight, frequency of wrong working posture, duration of doing wrong postures, lifting height of objects... Therefore, this study should be further developed to quantify risks of musculoskeletal disorders with the whole consideration of these factors. Another limitation is related to obstructed vision problems. The CNN model in the current study only extracts the working posture skeleton in 2-dimensional (2D) format from the video data of a single camera. Construction sites are often large and covered by many obstacles, so data collected from one camera will be difficult to be complete and continuous. Future research efforts should combine the data from both images, videos and wearable sensors to not only enhance the accuracy of ergonomic risk assessment, but also provide additional information (e.g., locations of workers, distance between objects...) that would be needed for more effective intervention. The next study also should consider building models to extract and evaluate the working posture skeleton frames in 3D format from data collected by multiple cameras to increase the accuracy of results in practice. Finally, the proposed method in this study can only utilized to assess human posture in lifting, carrying heavy objects but not other construction works. In fact, construction work is a labor intensive work, there are very many other ergonomic issues on construction sites that should also be more automatically monitored, assessed and managed to better control the musculoskeletal disorders-related risks. It is necessary to study further on these issues in the future.

5. Conclusions

Monitoring the workers'working posture is one of the important requirements to control disease risks due to musculoskeletal disorders in construction environment. On construction sites, manual lifting, carrying and moving of heavy objects is very common and with great intensity; furthermore, there is still the subjective psychology of the workers themselves and the management team in controlling disease risks related to these jobs. For the aim to improve the monitoring activities on these issues, the study achieved the following main results:

- Having compiled two large datasets of safe vs. unsafe working posture 2D skeleton frames in lifting, carrying, and moving heavy objects in construction.

These datasets can be used in a range of ergonomic posture analysis studies in construction. The computer vision community have published many datasets, such as PASCAL and MS COCO, but it can be said that no database of unsafe working posture skeleton frames is available specific for use in construction. The datasets can be accessed at link provided in Appendix.

- An established method for automatically assessing and classifying the workers' posture in lifting and carrying heavy objects by combining the RTMPose deep learning model to detect people in the videos collected from cameras placed on the construction site and a CNN model to automatically evaluate and classify the human posture skeleton of workers into two states: "safe posture" and "unsafe posture".

The established database and proposed method provide a springboard to aid in developing more advanced unsafe working posture detection techniques in construction.

Acknowledgement

This research is funded by the Vietnamese Ministry of Education and Training under grant number B2023 – XDA - 08.

References

[1] Anh Nguyen, D., Nguyen, N.-T., Tran, Q., Tran, D.-H. (2023). Tradeoff different construction project goals in using a novel multi-objective sea horse algorithm. *Alexandria Engineering Journal*, 82:55–68.

[2] Nguyen, D. A., Tran, D. Q., Nguyen, T. N., Tran, H. H. (2023). Modeling labor productivity in high-rise building construction projects using neural networks. *Archives of Civil Engineering*, 675–692.

[3] Tran, Q., Drew, S., Stewart, R. A. (2021). Evolutionary Model of e-Procurement Adoption: A Case of the Vietnam Construction Industry. *International Journal of Sustainable Construction Engineering and Technology*, 12(3).

[4] Loan, L. T. P., Dung, T. Q., Hai, N. V. (2020). Automatically determining safe spaces for pedestrians on floor under construction using BIM-4D. *Journal of Science and Technology in Civil Engieering (STCE) - HUCE*, 14(5V):144–155. (in Vietnamese).

[5] Nguyen, N.-T., Tran, Q., Dao, C.-H., Nguyen, D. A., Tran, D.-H. (2024). Automatic Detection of Personal Protective Equipment in Construction Sites Using Metaheuristic Optimized YOLOv5. *Arabian Journal for Science and Engineering*.

[6] Johansen, K. W., Schultz, C., Teizer, J. (2023). Hazard ontology and 4D benchmark model for facilitation of automated construction safety requirement analysis. *Computer-Aided Civil and Infrastructure Engineering*, 38(15):2128–2144.

[7] Fang, W., Ding, L., Love, P. E. D., Luo, H., Li, H., Peña-Mora, F., Zhong, B., Zhou, C. (2020). Computer vision applications in construction safety assurance. *Automation in Construction*, 110:103013.

[8] Thoan, N. N., Phuong Loan, L. T., Dung, T. Q., Duc, N. A., Hoa, L. T., Hieu, D. (2023). Occupational risk ranking in the construction of superstructure of high-rise buildings. *Journal of Science and Technology in Civil Engieering (STCE) - HUCE*, 17(3V):152–165. (in Vietnamese).

[9] Dung, T. Q. (2023). *Occupational Safety and Health in Construction*. Construction Publisher, Hanoi, Vietnam.

[10] OSHA (2003). *Scaffolding eTool*.

[11] Yuan, X., Anumba, C. J. (2020). Cyber-Physical Systems for Temporary Structures Monitoring. In *Cyber-Physical Systems in the Built Environment*, Springer International Publishing, 107–138.

[12] Sooksil, N., Benjaoran, V. (2021). Non-linear modelling of construction workers' behaviors for accident prediction. *Songklanakarin Journal of Science & Technology*, 43(2):596–602.

[13] Jonsson, R., Dellve, L., Halleröd, B. (2019). Work despite poor health? A 14-year follow-up of how individual work accommodations are extending the time to retirement for workers with poor health conditions. *SSM - Population Health*, 9:100514.

[14] Hossain, M. S., Azam, S., Karim, A., Montaha, S., Quadir, R., De Boer, F., Altaf-Ul-Amin, M. (2023). Ergonomic Risk Prediction for Awkward Postures From 3D Keypoints Using Deep Learning. *IEEE Access*, 11:114497–114508.

[15] Cho, C., Kim, K., Park, J., Cho, Y. K. (2018). Data-Driven Monitoring System for Preventing the Collapse of Scaffolding Structures. *Journal of Construction Engineering and Management*, 144(8).

[16] Seo, J., Lee, S. (2021). Automated postural ergonomic risk assessment using vision-based posture classification. *Automation in Construction*, 128:103725.

[17] Chen, J., Qiu, J., Ahn, C. (2017). Construction worker's awkward posture recognition through supervised motion tensor decomposition. *Automation in Construction*, 77:67–81.

[18] Wang, D., Dai, F., Ning, X. (2015). Risk Assessment of Work-Related Musculoskeletal Disorders in Construction: State-of-the-Art Review. *Journal of Construction Engineering and Management*, 141(6).

[19] Zhao, J., Obonyo, E. (2020). Convolutional long short-term memory model for recognizing construction workers' postures from wearable inertial measurement units. *Advanced Engineering Informatics*, 46: 101177.

[20] Yan, X., Li, H., Li, A. R., Zhang, H. (2017). Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention. *Automation in Construction*, 74:2–11.

[21] Antwi-Afari, M. F., Li, H., Umer, W., Yu, Y., Xing, X. (2020). Construction Activity Recognition and Ergonomic Risk Assessment Using a Wearable Insole Pressure System. *Journal of Construction Engineering and Management*, 146(7).

[22] Guo, H., Lin, J.-R., Yu, Y. (2023). Intelligent and Computer Technologies' Application in Construction. *Buildings*, 13(3):641.

[23] Fang, W., Ding, L., Luo, H., Love, P. E. D. (2018). Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction*, 91:53–61.

[24] Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, 85:1–9.

[25] Zhu, J., Zhu, J., Wan, X., Wu, C., Xu, C. (2019). Object detection and localization in 3D environment by fusing raw fisheye image and attitude data. *Journal of Visual Communication and Image Representation*, 59:128–139.

[26] Luo, H., Liu, J., Fang, W., Love, P. E. D., Yu, Q., Lu, Z. (2020). Real-time smart video surveillance to manage safety: A case study of a transport mega-project. *Advanced Engineering Informatics*, 45:101100.

[27] Kim, D., Liu, M., Lee, S., Kamat, V. R. (2019). Remote proximity monitoring between mobile construction resources using camera-mounted UAVs. *Automation in Construction*, 99:168–182.

[28] Fang, W., Zhong, B., Zhao, N., Love, P. E. D., Luo, H., Xue, J., Xu, S. (2019). A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network. *Advanced Engineering Informatics*, 39:170–177.

[29] Qiao, L., Qie, Y., Zhu, Z., Zhu, Y., Zaman, U. K. u., Anwer, N. (2017). An ontology-based modelling and reasoning framework for assembly sequence planning. *The International Journal of Advanced Manufacturing Technology*, 94(9–12):4187–4197.

[30] Golparvar-Fard, M., Heydarian, A., Niebles, J. C. (2013). Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Advanced Engineering Informatics*, 27(4):652–663.

[31] Martínez-Rojas, M., Marín, N., Vila, M. A. (2015). An Approach for the Automatic Classification of Work Descriptions in Construction Projects. *Computer-Aided Civil and Infrastructure Engineering*, 30 (12):919–934.

[32] Luo, X., Li, H., Yu, Y., Zhou, C., Cao, D. (2020). Combining deep features and activity context to improve recognition of activities of workers in groups. *Computer-Aided Civil and Infrastructure Engineering*, 35 (9):965–978.

[33] Roberts, D., Torres Calderon, W., Tang, S., Golparvar-Fard, M. (2020). Vision-Based Construction Worker Activity Analysis Informed by Body Posture. *Journal of Computing in Civil Engineering*, 34 (4).

[34] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference*, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, Springer, 740–755.

[35] MMPose_Contributors (2020). *OpenMMLab Pose Estimation Toolbox and Benchmark*. Visited on 15/3/2024.

[36] Jiang, T., Lu, P., Zhang, L., Ma, N., Han, R., Lyu, C., Li, Y., Chen, K. (2023). Rtmpose: Real-time multi-person pose estimation based on mmpose. *arXiv preprint arXiv:2303.07399*.

[37] Tan Tock Seng Hospital. *Back Care Tips*. Visited on 15/3/2024.

[38] iHasco. *Manual handling*. Visited on 15/3/2024.

[39] Vukicevic, A. M., Petrovic, M. N., Knezevic, N. M., Jovanovic, K. M. (2023). Deep Learning-Based Recognition of Unsafe Acts in Manufacturing Industry. *IEEE Access*, 11:103406–103418.