A Computer Vision-Based Approach to Classifying and Storing Image Data for Construction Safety Management

Zhitian Zhang and Hongling Guo

Department of Construction Management, Tsinghua University, Beijing, China

Abstract

The safety issues on construction site have been a critical problem and has received increased attention. The safety management approaches based on computer vision provides more chances to rapidly identify the unsafe hazards and instantly alert. However, related research mainly focused on specific scenarios or tasks and therefore lack of holistic and systematic discussions. On the other hand, the storage and processing mechanism of the collected image data is not timely and efficient enough, so it's impossible for managers to quickly extract the required information. In order to better apply computer vision into practice, a classification and storage approach for site images based on computer vision is proposed. Firstly, this paper provides well-organized descriptions and classifications of the site hazards based on the survey results. Secondly, according to the literature review, sort out the current applications of the computer vision technology in construction safety management, and analyze the various types of site information it requires. Finally, the safety management requirements and on-site technical information are comprehensively processed to establish a framework to classify and store the site image. Then, we assessed the actual effects with this framework in conjunction with a case study. The results imply that the approach proposed promotes the application of computer vision from a holistic view and improves the efficiency in safety management.

Keywords: construction safety management, image data, computer vision, classification, storage

1. Introductoion

In the past 10 years, the total output value of the domestic construction industry has continued to increase, bringing important contributions to the national economic development. However, due to the complex and changeable construction environment, safety accidents have occurred frequently, and the accident rate remains high. According to a report released by the International Labor Organization in 2013, the fatalities of workers in the construction industry are 2 to 3 times higher than in other industries. Therefore, it is necessary to strengthen the safety control measures of construction site to reduce accidents.

In recent years, the rapid development of computer vision technology has made automatic identification of construction hazards possible. This approach mainly collects on-site images, extracts key elements from them, and conducts analysis and training to obtain safety-related information. Compared to traditional safety control measures, computer vision technology is more timely and effective and will not interfere with the working process. Additionally, the surveillance camera is widely used in the construction site and easy to obtain.

However, there are still some difficulties in applying computer vision technology to the practice. The current research results are more focused on specific scenarios, types of work, or operating behaviors, and not general enough. At the same time, due to the large storage capacity of image data, it has an impact on the

efficiency of data analysis. Aiming at the data classification and storage problems in construction site image detection, this paper combines the requirements of safety management with the characteristics of computer vision technology to construct a data storage structure suitable for general site management, thereby supporting timely and effective data analysis.

2. Related work

According to different recognition targets, the application of computer vision technology on the construction site can be divided into action recognition, object detection, and interaction detection.

2.1. Action recognition

Action recognition based on computer technology, mainly through two-dimensional color images of workers [1] or depth images [2,3]. Combined with training or non-training methods to achieve action classification and recognition. Ding et al. [4] used deep learning methods, convolutional neural networks and Long Short-Term Memory (LSTM) to distinguish related actions on the herringbone ladder from two-dimensional images. Han and Lee [5,6] extracted the three-dimensional skeleton of the human body after obtaining the depth image from an infrared camera. Furthermore, they [7] used the stereo camera to simultaneously capture two ordinary images with different angles of view to obtain depth information and achieved the three-dimensional reconstruction of the human skeleton, and finally, classified ladder-related unsafe behavior.

2.2. Object detection

The scope of object detection in this study includes worker wearable entities such as PPE (personal protective equipment) and non-wearable entities such as on-site materials. As the fundamental of image recognition, a number of models have been developed. Much of the literature on PPE concerns safety helmets and safety belts. Wu et al. [8] uses advanced imaging algorithms and a three-dimensional computer aided design perspective view to identify concrete columns from the construction site image. Zhu and Brilakis. [9] utilized shape, color and material information to identify concrete from the image column. Golparvar-Fard M et al. [10,11] proposed a method for classification and recognition based on image excavator behavior with higher fineness.

2.3. Interaction detection

Based on worker action recognition and object detection, we can obtain the interaction information after supplementing the spatial or logical relationship. Fang et al. [12] uses deep learning methods to detect the PPE usage of workers who go to the outer wall through windows to work at height. Chen et al. [13] revealed that the appearance of a person or an objec instance contains informative cues is useful for facilitating interaction prediction. Georgia et al. [14] carried out a model that can predict an action-specific density over target object locations.

Taken together, these studies clearly indicate the importance and feasibility of applying computer vision in the construction site. And object detection has a more mature approach compared with the others, which can play a significant role in this study.

3. Methodology

In order to solve the problems mentioned above, this study proposes a holistic classification and lightweight storage of images approach to support further data analysis and safety management. First, based on the related rules published by Oregon OSHA, we established the classification framework. Second, combined with the image recognition technique, we sort out the elements from image to store. In section 4, a case study was used to show how to conduct this approach.

3.1. Classification of image data on construction site

The Oregon OSHA's safety and health rules for the construction industry has 24 subdivisions. We sort out 14 types of hazards related to workers, three of which are health-related hazards that cannot be captured by the image, which are ①toxic and hazardous substances (e.g. air contaminants), ② noise, and ③X-ray

and other radiation. For the remaining 11 hazards (A-K in Fig.1), they are classified into unsafe conditions and unsafe practices according to the relationship with workers. In addition, we added the unsafe practices (L in Fig.1) such as using mobile phones and smoking, considering that there are some unsafe situations without interactions between workers and hazards.

UNSAFE CONDITIONS

A. Overhead High-voltage Line or Equipment

B. Holes, Wall Openings, Established Floors, Mezzanines, Balconies

C. Rigging

D. Cranes and Derricks

E. Scaffolding

F. Trenching and Excavations

G. Vehicles and Mobile Equipment

UNSAFE PRACTICES

H. Using Tools: Hand Tools, Power-operated Hand Tools, Powder-actuated Tools e.t.

I. Wielding and Cutting

J. Using Ladders and Stairways

K. Wearing or Holding Personal Protective Equipment
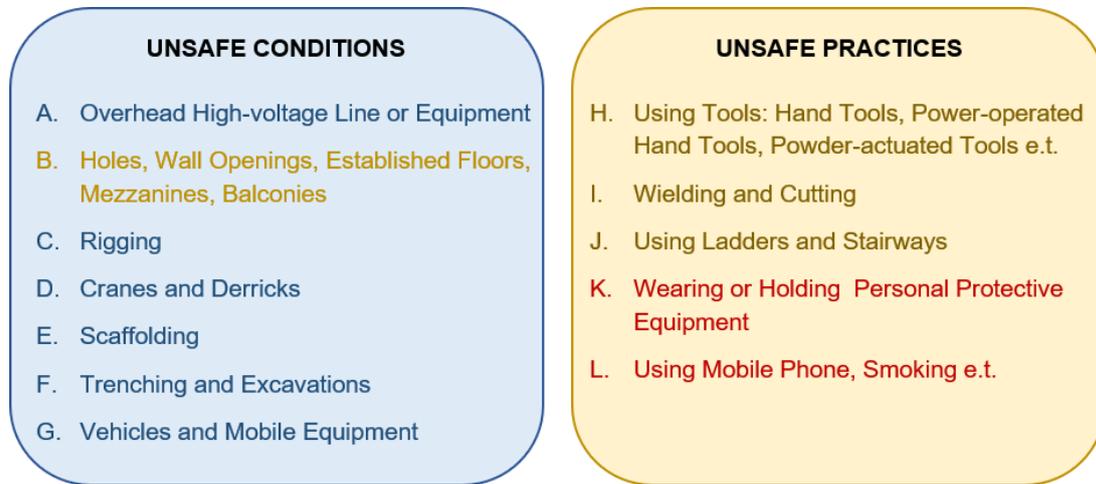
L. Using Mobile Phone, Smoking e.t.

Fig. 1. unsafe conditions and practices in construction site

According to the characteristics of the above hazards and computer vision technique, we classify the images into macro-image and micro-image. Macro-images with longer distance between the camera's location and the object or workers identified, can attain the classification but is difficult to gain the operation information exactly. Such images mainly include outdoor images with wide coverage on the construction site. Micro-images, with shorter distance, can achieve the accuracy of identifying the specific operation and mostly be obtained by the indoor camera.

Overall, the main monitoring target of macro-images is unsafe conditions, and micro-images is mainly for unsafe practices. However, there are two special situations. ① The type B mainly exists in the indoor environment which is supported by micro-images. ② The detection of type K and L are equally important in both images. Therefore, according to the safety management requirements, as shown in Fig.1, the blue font is the main detection target of macro-images, the yellow font is for micro-images, and the red font should be included in both.

3.2. Storage of image data on construction site

Using computer vision to obtain the required safety information is generally divided into two steps. The first is to obtain the detection result from the image, which means identifying the existence of the objects or workers. Then, to further estimate whether the worker is in an unsafe condition or performs an unsafe practice, the logical or spatial relationship from the detection result should be supplied. In this study, by storing the results after detection, the image data is lightened.

Combined with the classification of image data in section 3.1, the storage structure of the data is divided into four situations, as shown in Fig.2.

Unsafe conditions in macro-image
(A, C, D, E, F, G)

Unsafe practices in micro-image
(H, I, J)

Unsafe conditions in micro-image
(B)

Unsafe practices in macro-image and micro-image
(K, L)

Fig 2. four types in storage structure of image data

- Unsafe conditions in macro images: It mainly includes the detection of entities and workers. And entities in construction site can be subdivided into dynamic and static entities. The position of static objects in the image is relatively fixed (such as scaffolding). Dynamic entities and workers positions are more complicated, so time series information needs to be added;

- Unsafe conditions in micro-images: The unsafe conditions in micro-images are also relatively static. Therefore, we only need to store them detection results and there is no need to add time series;

- Unsafe practices in micro-images: The identification of unsafe operations in micro-images requires high accuracy. It is difficult for ordinary two-dimensional images to achieve satisfactory recognition results, so depth image and posture information in detail is necessary;

- Unsafe practices in macro and micro images: the information required of Type K is similar to 'unsafe practices in micro-image', and both involve the accurate detection of objects and posture. Type L only focuses on the worker's posture and movement, and there is no other interaction with the on-site entities.

4. Case study

4.1. Establishment of image database management system

Based on the MySQL5.7 database management system, we utilize the classification and storage approach mentioned above to establish a macro-image and micro-image database. The database storage structure is shown in Fig. 3, where (W, H) is the boundary coordinates of detection results; T represents the time series; (R, G, B) and (X, Y, Z) are the pixel information and space coordinates of workers' joint respectively.

Fig. 3. conceptual model structure of image database

It can be seen from the analysis of construction safety requirements that the gesture recognition of workers is a crucial task. Therefore, this paper takes the storage process of workers' posture as a case study. In order to support the further interactive relationship analysis, the study chooses to use the joint data of workers, which can be collected by depth image and openpose algorithm. After labeling each joint, store the labeled data into the 'construction workers' table of 'macro-image' and 'micro-image' database. The process can be seen in Fig.4.
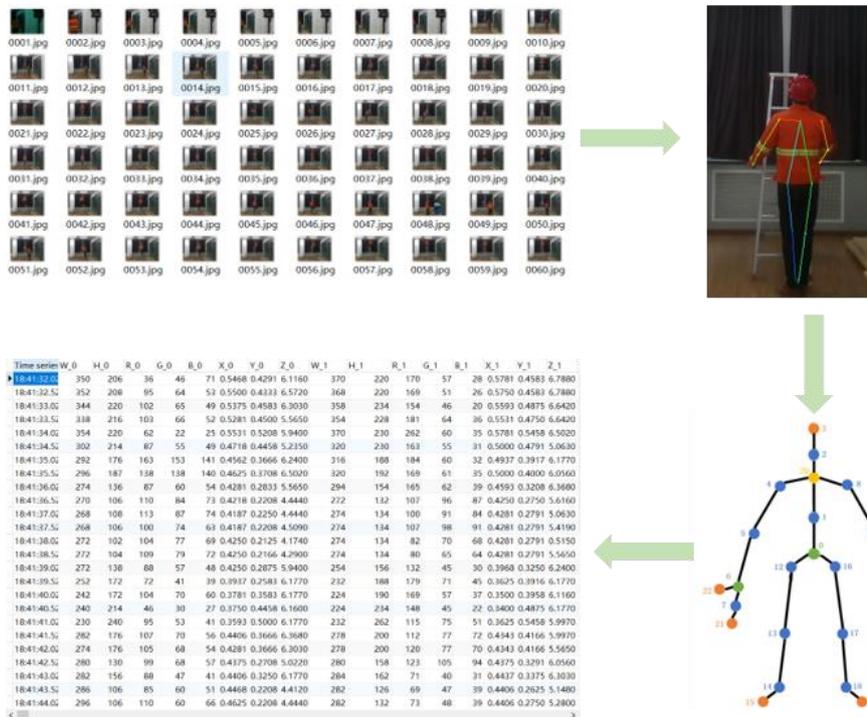

Fig. 4. the storage process of workers' joints

4.2. Findings and discussion

This case presents the storage process of workers' posture data by MySQL5.7. One noticeable finding is the 1.68G 3D point cloud data can be reduced to 0.75M, after extracting the 3D joint data and storing in rules. On the one hand, it greatly saves the storage space of the image data and increases the storage efficiency. On the other hand, the structured storage approach proposed in this study is not limited to special scenarios or types of work, which benefits boarder applications of computer vision in safety management.

5. Conclusion

In this study, the classification and storage approach of construction site image data is proposed. Satisfying the requirements of safety management, the key hazards that need to be identified from images are sorted out, and then classify them into macro and micro-images. Then, think of computer vision techniques, a structured storage method for images is proposed based on the classification result. Finally, the case study further confirmed the feasibility and effectiveness of the approach.

At the same time, there are some important issues for future work. The first is that the division of macro-image and micro-image needs to be clarified. From this point, a more scientific classification method can be proposed. Second, the storage framework of this study is based on the completion of object detection or action recognition, but the accurate identification of small entities still facing some challenges, such as the hooks of seat belts.

6. Acknowledgements

7. References

[1] J. Yang, Z. Shi, Z. Wu, Vision-based action recognition of construction workers using dense trajectories, Advanced Engineering Informatics 30.3 (2016): 327-336. https://doi.org/10.1016/j.aei.2016.04.009

[2] R. Gonsalves, J. Teizer, Human motion analysis using 3D range imaging technology, Int. Symp. on Automation and Robotics in Construction. 2009. https://doi.org/10.22260/ISARC2009/0044

[3] V. Escorcia, M. A. Dávila, M. Golparvar-Fard, J. C. Niebles, Automated vision-based recognition of construction worker actions for building interior construction operations using RGBD cameras, Construction Research Congress 2012: Construction Challenges in a Flat World. 2012. https://doi.org/10.1061/9780784412329.089

[4] L. Ding, W. Fang, H. Luo, P. E. D. Love, B. Zhao, X. Ouyang, A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory, Automation in construction 86 (2018): 118-124. https://doi.org/10.1016/j.autcon.2017.11.002

[5] S.U. Han, S.H. Lee, F. Peña-Mora, Vision-based detection of unsafe actions of a construction worker: Case study of ladder climbing, Journal of Computing in Civil Engineering 27.6 (2013): 635-644. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000279

[6] S.U. Han, S.H. Lee, F. Peña-Mora, Comparative study of motion features for similarity-based modeling and classification of unsafe actions in construction, Journal of Computing in Civil Engineering 28.5 (2014): A4014005. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000339

[7] S.U. Han, S.H. Lee, A vision-based motion capture and recognition framework for behavior-based safety management, Automation in Construction 35 (2013): 131-141. https://doi.org/10.1016/j.autcon.2013.05.001

[8] Y. Wu, H. Kim, C. Kim, S.H. Han, Object recognition in construction-site images using 3D CAD-based filtering, Journal of Computing in Civil Engineering 24.1 (2010): 56-64. https://doi.org/10.1061/(ASCE)0887-3801(2010)24:1(56)

[9] Z. Zhu, I. Brilakis, Concrete column recognition in images and videos, Journal of computing in civil engineering 24.6 (2010): 478-487. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000053

[10] M. Golparvar-Fard, A. Heydarian, J.C. Niebles, Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers, Advanced Engineering Informatics 27.4 (2013): 652-663. https://doi.org/10.1016/j.aei.2013.09.001

[11] R. Bao, M. A. Sadeghi, M. Golparvar-Fard, Characterizing construction equipment activities in long video sequences of earthmoving operations via kinematic features, Construction Research Congress 2016. 2016. https://doi.org/10.1061/9780784479827.086

[12] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, C. Li, Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment, Automation in Construction 93 (2018): 148-164. https://doi.org/10.1016/j.autcon.2018.05.022

[13] C. Gao, Y. Zou, J.B. Huang, ican: Instance-centric attention network for human-object interaction detection, arXiv preprint arXiv:1808.10437 (2018).

[14] G. Gkioxari, R. Girshick, P. Dollár, K. He, Detecting and recognizing human-object interactions, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018. https://doi.org/10.1109/CVPR.2018.00872