Computer Vision-Based Monitoring Method of Non-Wearing Helmet Events Using Face Recognition

Chenrui Liao[1], Hongyan Chen[2]*, Chenxi Liu[3], Ying Yu[4], Pengfei Zhao[5]

[1]Department of Research, Huigong Tech Co. Ltd, Nanchang, China
[2]Department of Life Sciences, Fudan University, Shanghai, China
[3]Department of Civil and Environmental Engineering, University of Washington, Seattle, United States
[4]Department of Computer Science, School of Software, East China Jiaotong University, Nanchang, China
[5]Department of Radiotherapy, Beijing Friendship Hospital, Capital Medical University, Beijing, China

Research Article

Received: 05-Nov-2024, Manuscript No. GRCS-24-151762; **Editor assigned:** 08-Nov-2024, Pre QC No. GRCS-24-151762 (PQ); **Reviewed:** 22-Nov-2024, QC No. GRCS-24-151762; **Revised:** 10-Dec-2025, Manuscript No. GRCS-24-151762 (R); **Published:** 17-Dec-2025, DOI: 10.4172/2229-371X.16.4.001

***For Correspondence:** Hongyan Chen, Department of Life Sciences, Fudan University, Shanghai, China;
E-mail: chenhongyan@jxufe.edu.cn

Citation: Liao C, et al. Computer Vision-Based Monitoring Method of Non-Wearing Helmet Events Using Face Recognition. J Glob Res Comput Sci. 2025;16:001.

ABSTRACT

Wearing helmets is crucial for ensuring the safety of workers in the construction industry because this is the first line of avoiding over 70% of production safety accidents. However, many workers are not willing to wear helmets due to discomfort and reduced work efficiency. To this end, this paper proposes a computer vision-based monitoring method using face recognition to detect and prevent non-wearing helmet events on construction sites. Compared to existing surveillance or monitoring systems, the proposed method has three significant advantages. Firstly, by using a unique structure, the proposed method can achieve up to 97.7% accuracy in detecting workers not wearing helmets. Secondly, the proposed method enables real-time detection, allowing it to prevent dangerous behaviors by stopping them in advance. Finally, the proposed method has been successfully deployed on over 20 real construction sites, and it has detected more than 18,000 related events.

Keywords: Computer vision-based method; Non-wearing helmet event; Face recognition; Progressive sieve analysis; Detection on construction sites; Prevention on construction sites

INTRODUCTION

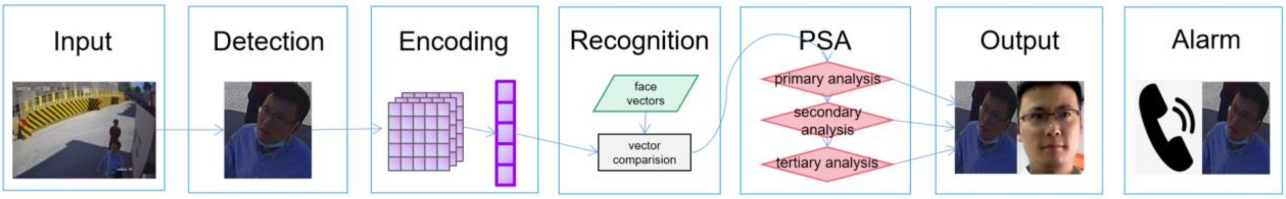
Construction is one of the most dangerous industries in the world [1]. According to statistics, about 7% of the workforce in the world is employed in construction, but the industry accounts for 30-40% of workplace fatalities. In China, 3843 fatal injuries were recorded in 2017 on construction sites, making the industry identified as the most hazardous in the country [2,3]. Meanwhile, the report released by the China Housing and Urban-Rural Development Agency 2020 [4] shows that in 2020, 59.07% of the annual production safety accidents in China were caused by falls from height [5], and 12.05% were caused by object strikes. The two types of events together account for more than 70% of total production safety accidents, and wearing helmets is the first measure to protect the lives of workers under these conditions. However, many workers are not willing to wear helmets due to discomfort and reduced work efficiency [6]. To protect the lives of workers on construction sites, it is significant to perform helmet-wearing detection for workers.

Safety accident warnings and helmet-wearing detection are often realized by methods based on Intelligent Internet of Things (IoT) devices or computer vision. However, these methods fail to address some issues. Some construction sites monitor workers' behaviors and helmet-wearing by purchasing smart IoT devices. Khan et al., used an IoT-based smart hook for preventing falls from height [6]. Zhou et al., developed an IoT-based safety barrier warning system to monitor unsafe behaviors and equipment status [7]. IoT-based items can effectively detect unsafe behaviors, including not wearing helmets, to protect the lives of workers on construction sites, but this practice has a high cost. A construction site needs to procure a complete IoT system, including software and a large number of hardware devices. Also, workers and large machinery require the use of smart devices, bringing a huge burden. Computer vision methods are more efficient and cheaper and are now widely used in construction site safety management, including safety incident warnings and helmet-wearing detection. Fang et al., proposed that deep learning and computer vision-based approaches have the potential to enable managers and engineers to improve the safety performance of their construction operations on-site [6]. Using deep learning to identify unsafe behaviors and work conditions provides a signpost for future research in the emergent area of deep learning within the context of safety. Nash et al., used a CNN-based algorithm to determine whether a worker is wearing a hard hat, vest, or both from an image/video in real-time [3]. Fang et al., employed two CNN-based models to determine if workers are wearing their harnesses when performing tasks while working at heights [8]. Based on the YOLOv5 model, Zhou, Jia, Tan, and Xu's teams made improvements in helmet-wearing detection on construction sites [9,10]. Although these computer vision-based methods realize helmet-wearing detection on construction sites, they can only notify supervisors when detecting unsafe events of not wearing helmets. At this time, the worker is not notified, and supervisors can only respond after this event and conduct critical education or punishment for workers, failing to prevent unsafe behaviors from happening.

To solve this problem, by combining the InsightFace face recognition algorithm with Progressive Screening Analysis (PSA), this paper proposes a new face recognition method that can adapt to real construction sites as shown in Figure 1. The proposed method confirms the identity information of the worker who is not wearing a helmet on the construction site through face recognition, and at the same time, the method directly contacts the worker and his manager by telephone to ask the worker to immediately wear a helmet correctly and discontinue his unsafe behaviour. In this way, this method solves the problem that existing methods only support post-event response, and the behaviour of not wearing helmets is discontinued as soon it is detected, thereby safeguarding the lives of workers.

This paper has two main contributions to methodology and application. First, this paper proposes PSA to achieve high efficiency of face recognition on real construction sites. Second, to improve the recognition efficiency of this method on construction sites, this paper establishes the real site face dataset for model training, which contains 831 construction site workers with 16,620 images. The proposed face recognition method has been applied to more than 20 real construction sites and has discontinued over 18,000 potential incidents. The statistical results on the test set with 1662 events indicate that this algorithm achieves an accuracy of 0.977, a recall of 0.884, an F1 score of 0.928, a top-1 accuracy of 0.698, and a top-5 accuracy of 0.892 for face recognition on real construction sites. To sum up, the proposed computer vision-based monitoring method provides an effective solution to address the issue of not wearing helmets on construction sites. Its outstanding features, including high accuracy, real-time detection, and successful deployment on multiple construction sites, make it a practical and valuable tool for improving workers' safety in the construction industry.

Figure 1. The workflow of the detecting and alarming system.



Related work

Deep learning and computer vision in construction: Artificial intelligence technologies such as deep learning and computer vision are widely used in the construction industry, especially in the areas of schedule management, quality management, safety management, cost management, and assisted construction.

In safety management, using artificial intelligence and deep learning is the mainstream approach. Bangaru et al., used an Artificial Neural Network (ANN) model to guide and detect earplug-wearing steps based on construction workers' gestures. Pistolesi adopted artificial intelligence to analyse data fed by wearable sensors to determine whether construction workers were handling goods and whether safety codes were met. Aythan et al., performed artificial intelligence analysis to predict possible accidents at construction sites and prevent them. Computer vision methods are critical safety management tools based on camera capture footage. Fang et al., proposed a novel framework to check whether a site worker is working within the constraints of their certification. Ding and Fang et al., developed different models to detect unsafe behaviors and falling from height. Liu et al., used a Region-based Convolutional Neural Network (RCNN) for object detection and improving the safety of the built environment. Foong et al., employed ANN to estimate the safety factor of slope stability during construction. Brunetti et al., used computer vision and deep learning methods for pedestrian detection and tracking. Son et al., detected workers under different poses by using very deep residual networks. Li et al., proposed a deep learning method for detecting non-hardhat use. Li et al., developed a new method of monitoring helmet-wearing based on SSD improvement, which achieved accurate and real-time helmet-wearing detection and replaced manual inspection with automatic detection. Xu et al., performed target detection of helmet-wearing based on the yolov3 algorithm. They combined the algorithm with natural language processing technology to describe the helmet-wearing situation of construction personnel to reduce the accident rate due to construction workers not wearing helmets. The above studies demonstrate essential applications of deep learning, especially computer vision techniques, in safety management. However, these studies only focus on identifying unsafe events after they happen. To identify the unsafe actors and stop unsafe behaviors such as not wearing helmets, face recognition must be performed on construction sites.

Face recognition algorithm: Face recognition is a technology that can match a human face from a digital image or a video frame against a database of faces (add). Simonyan et al., proposed the deep learning backbone network VGG-16, which opened a new era of computer vision and face recognition. Schroff et al., proposed the FaceNet system, which obtained feature vectors from face images for face similarity measurement. Liu et al., studied the angular softmax (A-softmax) loss problem and designed the SphereFace algorithm that approximates the ideal feature criterion. Deng et al., further improved it based on softmax and proposed the ArcFace loss function, which improved the efficiency of face recognition. The face detection model SCRFD proposed by the team of InsightFace in 2021 introduces sample redistribution and computational redistribution methods. The team of InsightFace also proposed SDU-Nets, which greatly improves the accuracy of face alignment. Partial FC is a model training strategy that enables model training under memory-limited conditions.

Face recognition algorithms mainly include face detection, face calibration, face recognition, and other steps. The face recognition method proposed in this paper is mainly based on the InsightFace face analysis project, a mainstream face recognition solution improved from the MXNet framework. It runs 2.82 times faster than the original framework in data parallelism, 2.45 times faster in model parallelism, and 1.38 times faster in hybrid parallelism+Partial FC. Currently, the project has been fully open-sourced, with no restrictions on academic research and commercial use, and it has collected 4.1 k forks and 13.3 k stars on Git Hub.

Insightface pioneered ArcFace, a softmax-based improved loss function shown in the following equation.

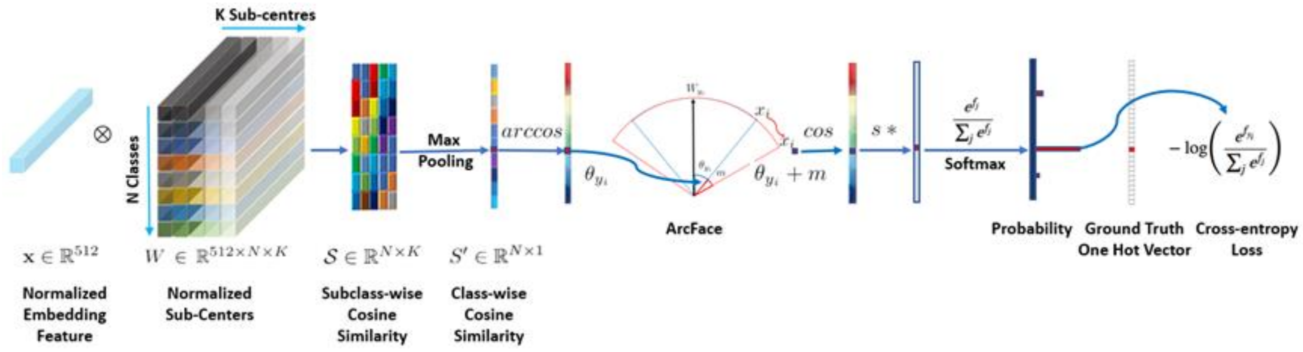
$$L = -\log \frac{e^{s\cos(\theta_y + m)}}{e^{s\cos(\theta_y + m)} + \sum_{j=1, j \neq y}^{N} e^{s\cos\theta_j}} \qquad (1)$$

The decision boundary is

$$\cos(\theta_1 + m) = \cos\theta_2 \quad (2)$$

With the ArcFace loss function shown in Figure 2, the InsightFace algorithm maximizes the classification in the angle space, and its impact is more direct than that in the cosine space. Meanwhile, the algorithm has a significantly better effect than FaceNet, SphereFace, and other algorithms. Its performance reaches 99.83% on the LFW dataset and 98.02% on the YTF dataset, which is remarkable. Besides, InsightFace also performs better on construction sites against FaceNet and other competitors.

Figure 2. The flow of training a deep neural network with the ArcFace loss function.



The InsightFace project also proposed the face detection model SCRFD. The model introduces sample redistribution and computational redistribution methods. Its performance is 3.86% higher than that of TinaFace, the strongest competitor in the field of face detection. Meanwhile, SCRFD runs more than three times faster than the TinaFace algorithm on GPUs with VGA-resolution images, achieving accurate and efficient face detection.

SDU-Nets is a high-accuracy face alignment algorithm. It is spatially invariant to arbitrary input face images and has excellent robustness to extreme poses, exaggerated expressions, as well as severe occlusions. The algorithm achieved State-of-the-Art (SOTA) in the field, with a mean error of 5.55% on the cofw dataset and an optimal accuracy of 98.51% on the CFP-FP dataset. The use of this algorithm greatly improves face alignment accuracy for images taken on construction sites.

For face recognition, this paper adopts the partial FC face recognition training strategy, which greatly improves the training efficiency of face recognition networks. As a load-balanced sparse distributed classification training method, partial FC enables model training under memory-limited conditions. With this strategy, the training of face recognition models can be conducted on low-configuration computers on the construction site without the need to purchase additional high-configuration computers.

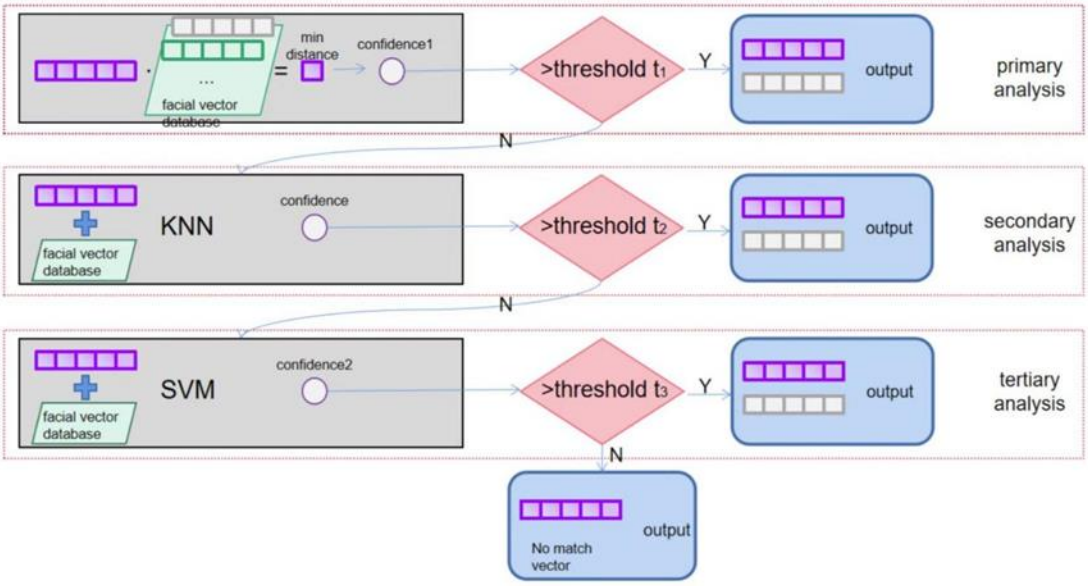
These face recognition networks and algorithms perform well on public datasets, but tests in scenarios of construction sites show that they perform poorly and need further improvement to meet the needs of face recognition on construction sites.

MATERIALS AND METHODS

This paper proposes Progressive Sieve Analysis (PSA) to solve the problem that the face recognition algorithms do not meet the demand for face recognition on construction sites. The accuracy of algorithms such as VGG-16 and InsightFace decreases in actual construction scenarios. When face recognition was performed on the collected images of workers not wearing helmets in an actual construction project, the precision of the InsightFace algorithm dropped to 0.893, and the recall dropped to 0.562, which differs greatly from the performance on public datasets.

As shown in Figure 3, PSA performs face recognition on real construction sites at three levels: Primary analysis, secondary analysis, and tertiary analysis. Specifically, the primary analysis is based on the InsightFace algorithm. In the secondary analysis, the K-Nearest Neighbor (KNN) algorithm is used to recognize events with low credibility in the primary analysis. If primary analysis and secondary analysis both fail to obtain a result, tertiary analysis based on the Support Vector Machine (SVM) algorithm is used for the final level.

Figure 3. The working principle of PSA.



The purple vector is the facial vector to be tested. The facial vector database is the Milvus vector database that contains all face vectors of workers on a real construction site. Confidence is the reliability of the face recognition, KNN, and SVM methods.

The primary analysis is performed based on the InsightFace algorithm. A 512-dimension facial vector called a tested vector is obtained from videos on construction sites after face detection and encoding. By comparing the Euclidean distances between the tested vector and all vectors in the facial vector database, a vector with the minimum distance in the facial vector database and confidence (called confidence 1) related to the distance can be obtained. If the minimum distance is extremely low so that confidence 1 is higher than the threshold t_1, the tested worker in the video is matched to the identification in the facial vector database. The result of face recognition is delivered to the front end without secondary and tertiary analysis. If confidence 1 is lower than the threshold t_1, the secondary analysis will be performed to obtain a more precise face recognition result.

The secondary analysis is to run the KNN algorithm when confidence 1 is lower than the threshold t_1 in the primary analysis. In the secondary analysis, the Euclidean distances between the tested vector and all vectors in the facial vector database are still compared. The vectors with top-k minimum distances are picked out. Meanwhile, the frequency of each worker among the top-k vectors is counted, and the worker with the highest frequency is the most likely result of face recognition. The confidence is n/k, where n is the number of workers most likely to appear in the top-k vectors. If the confidence is higher than the threshold t_2, the tested worker in the video is matched to the identification of the most likely worker; otherwise, the tertiary analysis will be performed.

The tertiary analysis is based on the SVM algorithm when secondary analysis cannot determine a precise result of face recognition. In this condition, the confidence in the secondary analysis is lower than the threshold t_2. The top two workers are picked up from the top-k vectors, and then the SVM classifier is used to fit if the tested vector is more likely to be one of the top two workers. The Gaussian function is selected as the core function in the classifier. The confidence of fitting is called confidence 2. If confidence 2 is higher than the threshold t_3, the tested worker in the video is matched to the more likely worker between the top two workers; otherwise, the output turns out to be none, indicating that no vector in the facial vector database matches the tested vector.

PSA performs PSA on low-trustworthiness face recognition results to accurately recognize low-trustworthiness recognition events, which significantly improves the recognition accuracy. The thresholds t_1, t_2, and t_3 for primary, secondary, and tertiary analyses are obtained by tests using the test dataset.

RESULTS AND DISCUSSION

Experiments

Datasets: The Glint360K public dataset and RealSiteFace local dataset were used as the training dataset of this experiment. Glint360K is an extremely clean and large dataset that achieves SOTA performance by using only 10% of the classes for training.

The RealSiteFace dataset contains 16620 images on construction sites as is shown in Figure 4. To solve the problem that the models trained on mainstream public datasets do not perform well in face recognition on construction sites, this paper collects 831 construction site workers with 16,620 images to establish the dataset RealSiteFace. RealSiteFace collects attendance images from construction sites, and each data sample contains workers' images, ID, gender, age, shooting angle, lighting condition, and whether they are wearing helmets or not. All workers are of East Asian ethnicity, 20 images were selected from the daily attendance images, with no less than five images of both helmets and non-helmets. These images were authorized by staff and managers and were used only for construction site safety management and related research.

Figure 4. Data samples of RealSiteFace dataset.

The RealSiteFace dataset has significant advantages over public datasets such as Glint360K. All faces in the RealSiteFace dataset are of East Asian ethnicity and are collected from the attendance records of construction sites, which is very close to the real application scenario. The data samples contain the key concerns of face recognition on construction sites, and they are labeled for issues such as lighting conditions, shooting angles, and whether a helmet is worn. To ensure the reliability of the training results, 20 different images were selected for each worker in the dataset, with at least 5 images with helmets and 5 images without helmets. The dataset was divided into a learning set, a validation set, and a test set in the ratio of 8:1:1. The comparison of the face recognition performance of the models trained on RealSiteFace with the models trained on other datasets in the construction site scenario is presented in Table 1.

Table 1. Comparison of face recognition performance.

	Top-1 accuracy	Top-5 accuracy
Glint360K	0.54	0.85
VGG Face2	0.49	0.78
RealSiteFace	0.69	0.89

Then, the models trained on multiple datasets were used to perform face recognition on the test set to verify their performance on the construction sites. The model trained on the RealSiteFace dataset achieves a top-1 accuracy of 0.69 and a top-5 accuracy of 0.89, showing better performance that the model trained on the large datasets Glint360K and VGG

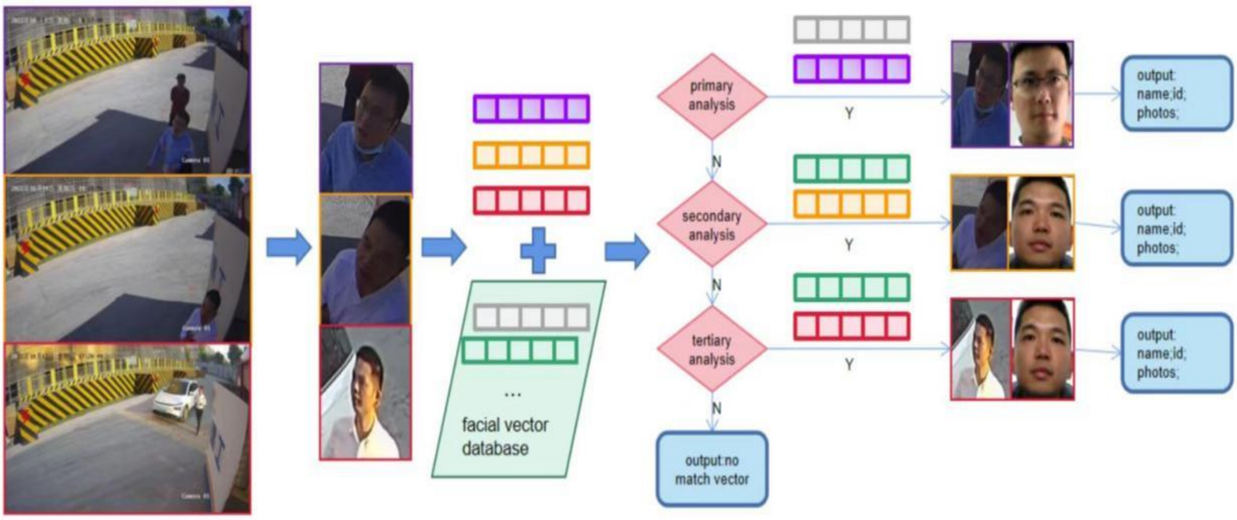
Face2.

Project flow: Before the training and testing of the face recognition network, the data needs to be pre-processed. In the neural network training process, the SCRFD model is used for face detection, and the SDU-Nets algorithm is used for face alignment. The Glint360K public dataset and the RealSiteFace local dataset were used as the training dataset of this experiment.

After completing the data pre-processing, 19 attendance images and 1 real-name registration image of each worker were used to form a facial vector database on each construction site. The faces in these images are encoded as 512-dimensional vectors and saved in the Milvus vector database with the ID and other information of each worker for unified management and operation. These real name and attendance images are authorized by workers for safety management and personnel management on construction sites.

Figure 5 shows how the face recognition system works on a real construction site. A video with non-wearing helmet events was provided by surveillance cameras on construction sites. After face detection and encoding, 512-dimensional facial vectors (tested vectors) are collected from the video. Then, the PSA method is used to obtain the result of face recognition. If the tested vectors are matched to facial vectors in the database, screenshots of the video on the construction site and real-name registration images are sent to the front end. Detailed information is shown, and the worker is called immediately.

Figure 5. The working principle of face recognition on a real construction site.



Experiments:

Threshold t_1, t_2, and t_3: The thresholds t_1, t_2, and t_3 are obtained by testing on the test dataset. There are 1662 images in the test dataset. With other threshold parameters fixed, different thresholds, t_1, t_2, and t_3, are taken for testing to obtain the top-1 accuracy of the face identification, as shown in Tables 2-4. Then, the threshold t with the largest top-1 accuracy is taken (Tables 2-4).

Table 2. Variation of top-1 accuracy with t_1.

t_1	Top 1 accuracy
0.6	0.564
0.7	0.64
0.8	0.693
0.9	0.698
0.95	0.582

Table 3. Variation of top-1 accuracy with t_2.

t_2	Top 1 accuracy
0.6	0.567
0.7	0.676
0.8	0.698
0.9	0.578
0.95	0.372

Table 4. Variation of top-1 accuracy with t_3.

t_3	Top 1 accuracy
0.6	0.651
0.7	0.651
0.8	0.664
0.9	0.698
0.95	0.604

The top-1 accuracy reaches its maximum when t_1=0.9, t_2=0.8, and t_3=0.9. The thresholds t_1, t_2, and t_3 are set to these values in practical applications on construction sites.

Face recognition algorithms with or without PSA:

Various face recognition algorithms and the proposed InsightFace+PSA method are used to recognize faces on construction sites, and the recognition results are presented in Tables 5 and 6.

Table 5. Face recognition results.

	Accuracy	Precision	Recall	F1-score
VGG16	0.577	0.802	0.407	0.54
InsightFace	0.716	0.893	0.562	0.69
InsightFace+PSA	0.892	0.977	0.884	0.928

Table 6. Top-1 and top-5 accuracy.

	Top-1 accuracy	Top-5 accuracy
VGG16	0.248	0.577
InsightFace	0.316	0.716
InsightFace+PSA	0.698	0.892

Three different methods, VGG16, InsightFace, and InsightFace+PSA method (*i.e.,* the method used in this paper), were used for face recognition for 1662 images in the test set. The results of face recognition show that the accuracy of the InsightFace+PSA method is 0.892, the precision is 0.977, the recall is 0.884, and the F1-score is 0.928. Compared with the InsightFace algorithm without the PSA method, the accuracy is 1.24 times, the precision is 1.09 times, the recall is 1.57 times, and the F1-score is 1.34 times.

The face recognition results show that the top-1 accuracy and top-5 accuracy of the InsightFace+PSA method are also the highest among the three methods, where the top-1 and top-5 accuracy is 0.698 and 0.892, which is 2.21 and 1.25 times of the InsightFace algorithm, respectively.

To sum up, compared with the InsightFacealgorithm, the InsightFace+PSA method in construction site scenes has greatly improved the face recognition efficiency, especially in the indexes of recall rate, F1-score, and top-1 accuracy, which can meet the needs of face recognition on construction sites.

Analysis of experimental results: The classical confusion matrix reflects how PSA successfully achieves face recognition in construction site scenarios. The experimental results show that using the RealSiteFace dataset for model training and performing face recognition with the InsightFace+PSA method successfully achieves face recognition under construction site scenes. Whether the PSA method is used or not has a great impact on the efficiency of face recognition in construction site scenes. Compared with the InsightFace algorithm without the PSA method, this method has a relatively small improvement in the accuracy of face recognition and a great improvement in recall. This is determined by the nature of the present method. The classical confusion matrix consists of TP (True Positive), FN (False Negative), FP (False Positive), and TN (True Negative). The precision and recall are calculated by the following equations.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

In this method, T means the face recognition result is correct, F means the face recognition result is wrong, P means the face recognition result is credible (the system outputs the recognition result), and N means the face recognition result is not credible (the system outputs a null result).

The essence of the PSA method is to perform secondary analysis and tertiary analysis on the face recognition result when the output of the InsightFace algorithm is N so that the ratio of P/N is substantially increased. That is, if the result of a certain event is FP using only the InsightFace algorithm, then the result of the event is still FP after using PSA. If the result of a certain event is N (TN or FN) using only the InsightFace algorithm, the result of the event is likely to be recognized as TP or FP after using PSA. Therefore, after using the InsightFace+PSA method, the FN in the event is reduced substantially, while the proportion of TP and FP does not very much in relative terms. This explains why the use of the PSA method can significantly improve the recall of face recognition, and thus the efficiency of face recognition (F1-score).

Filed projects

The system has been put into use in more than 20 construction sites, and it has discontinued more than 18,000 potential incidents. Here, the actual under-construction project "xx Apartment" is taken as an example. This project covers an area of 54,210 square meters, with a total construction area of 156,427 square meters, and it is under construction at present. In this example, the surveillance camera is placed four meters high and faces the construction site, which is an area with a large flow of people and is also the critical monitoring area for safety alarm events such as not wearing helmets. Before the application of this face recognition method, the event of not wearing a helmet is usually achieved by manual monitoring. The high position and low clarity of the surveillance camera cause significant difficulties for the safety management of the staff. Meanwhile, direct sunlight and wall shadows on the captured image are also severe.

The original image of the event reports for not wearing a helmet is shown in Figure 6. The three images correspond to the results output under three different levels of PSA. Specifically, (a) shows the result that can be directly discriminated by primary analysis, (b) shows the high-confidence result output after secondary analysis using the KNN method, and (c) shows the high-confidence result output after tertiary analysis using the SVM method after using the KNN method.

Figure 6. Three images reported on a construction site.

(a) (b) (c)

In Figure 6(a), the detected person without wearing a helmet is closer to the camera and has a head-up behavior, and a more positive portrait is collected. The algorithm identifies that the confidence of the face is more than 0.98, and the accuracy of face recognition is 0.97, which can be directly completed by the PSA method once analyzed. In Figure 6(b), the detected person without wearing a helmet is closer to the camera, the angle between the portrait and the camera is about 30 degrees, and it is in a shadow area of the construction site. Since the brightness and clarity of the collected face are insufficient, the accuracy of the algorithm for face recognition is less than 0.9, and the reliability of this recognition result could not be confirmed. After the secondary analysis of PSA, the face recognition of this event was completed by comparing k results in the face database. The reliability of this recognition result was confirmed to be 0.96. In Figure 6(c), the detected person without wearing a helmet is far from the camera, the shooting angle between the portrait and the camera is greater than 60 degrees, and the captured face is not clear. Even if the face recognition accuracy is only 0.57 after the primary analysis, the face recognition of the event is completed after the secondary and tertiary analysis by the PSA method. The reliability of the face recognition result is 0.93.

After completing face recognition, Table 7 presents the information of the workers that are called automatically. The workers are required to wear a helmet as soon as possible, and project management is also notified. Then, a non-wearing helmet event is discontinued, and the safety of the worker is ensured.

Table 7. Front-end display of face recognition results.

Project	Name	Time	Id	Confidence 1	Confidence 2	Tel
P00016	Fu	12-10-2022 14:57	11816	0.97	Null	136xxxxxxxx
P00016	Yan	11-10-2022 17:35	104616	0.89	0.96	139xxxxxxxx
P00016	Yan	13-10-2022 17:20	104616	0.57	0.93	139xxxxxxxx

CONCLUSION

This paper proposes progressive sieve analysis for face recognition of certain special events on construction sites (such as non-wearing helmet events). A new face recognition method that can adapt to construction site scenarios is developed by combining theInsightFace algorithm with PSA, to improve face recognition precision and recall. Then, the method proposed in this paper is applied to the real construction site, the face recognition precision and recall of workers not wearing a helmet reach 0.977 and 0.884, respectively. The F1 score is 0.928, and the top-1 accuracy is 0.698. This method has been applied to monitor unsafe behaviors on more than 20 real construction sites, and it has warned and successfully discontinued over 18,000 events of not wearing helmets, effectively safeguarding the lives of workers.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AUTHORS CONTRIBUTION STATEMENT

Chenrui Liao: Data curation, formal analysis, methodology, software, validation, writing - original draft.
Hongyan Chen: Conceptualization, investigation, project administration, supervision.
Chenxi Liu: Software, visualization, writing - review and editing
Ying Yu: Funding acquisition, investigation, supervision.
Pengfei Zhao: Funding acquisition, resource, supervision.

ACKNOWLEDGEMENTS

This research was supported by Huigong Tech Co.Ltd.

DATA AVAILABILITY AND ACCESS

The data that has been used is confidential. By the way, it can only have been used on worker's safety on construction sites. Any one wants to request data please contact Chenrui Liao.

REFERENCES

1. Tixier AJP, et al. Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports. Autom Constr. 2016;62:45-56.
2. Fang W, et al. Computer vision applications in construction safety assurance. Autom Constr. 2020;110:103013.
3. Nath ND, et al. Deep learning for site safety: Real-time detection of personal protective equipment. Autom Constr. 2020;112:103085.
4. The State Council. Ministry of Housing and Urban-Rural Development of the Peoples of Republic of China. State on work safety reported by the Ministry of Housing and Urban-Rural Development in 2020. The peoples republic of China. 2021.
5. Aguwa EN, et al. A Review of Sir Thomas Legge's Aphorisms and Workplace Personal Protective Equipments. Is There Gap in Knowledge, Attitude and Utilization. Occup Med Health Aff. 2013;1:134.
6. Khan M, et al. Tag and IoT based safety hook monitoring for prevention of falls from height. Autom Constr. 2022;136:104153.
7. Zhou C, et al. Safety barrier warning system for underground construction sites using Internet-of-Things technologies. Autom Constr. 2017;83:372-389.
8. Fang W, et al. Falls from heights: A computer vision-based approach for safety harness detection. Autom constr. 2018;91:53-61.
9. Zhou F, et al. Safety helmet detection based on YOLOv5. In2021 IEEE International conference on power electronics, computer applications (ICPECA). IEEE. 2021:6-11.
10. Jia W, et al. Real-time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector. IET Image Process. 2021;15:3623-3637.