International Journal of Autonomous Robotics and Intelligent Systems
Vol. 2, No. 1, June 2026 [31-41]

Human Pose Estimation and Action Classification in Cluttered Industrial Environment for Safety Monitoring

Farid Azhari Norzaidi[1,3], Latifah Munirah Kamarudin[1,3,4,6*], Ammar Zakaria[1,5,6], Ahmad Shakaff Ali Yeon[1,5] and Syed Muhammad Mamduh Syed Zakaria [1,2]

[1]Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia
[2] Faculty of Electronic Engineering & Technology (FKTEN), Universiti Malaysia Perlis (UniMAP), 02600, Arau, Perlis, Malaysia
[3]Faculty of Intelligent Computing (FKC), Universiti Malaysia Perlis (UniMAP), 02600, Arau, Perlis, Malaysia
[4]Faculty of Engineering, Graduate Faculty of Interdisciplinary Research, University of Yamanashi, Kofu, Japan
[5]Faculty of Electrical Engineering & Technology (FKTE), Universiti Malaysia Perlis (UniMAP), 02600, Arau, Perlis, Malaysia
[6]Centre for Applied Intelligence, Management and Science University (MSU), Selangor, Malaysia

Received 9 May 2026, Revised 20 May 2026, Accepted 20 June 2026

ABSTRACT

This paper presents an automated worker postural classification system using deep learning for real-time activity monitoring at Malaysian industrial sites, particularly fabrication yards and construction sites. The system runs raw CCTV video frames through the YOLOv8x-pose architecture to pull out 17 skeletal keypoints per detected worker. From there, a rule-based geometric classification logic is applied to the lower-body keypoints, specifically the hip, knee and ankle joints, to tell apart standing and sitting postural states without needing the extra computational load that comes with temporal neural networks. The classification logic works through four scale-invariant geometric metrics: a vertical height ratio, a knee flexion angle worked out using the law of cosines, a torso lean and tilt angle and a normalised knee-to-torso proximity measure. All four metrics together build up a solid geometric profile for each action class that stays consistent across different worker body sizes and camera distances. Validation is done using a custom dataset of real-world industrial site footage annotated in YOLO format. The system came out with an overall accuracy of 79.80% and an F1-score of 86.70%. Standing classification came in at 93.00% precision and 90.00% recall, while sitting classification recorded a perfect precision of 100.00% with a recall of 63.20%. The lower sitting recall is mainly put down to lower-body occlusion from on-site equipment and the steep mounting angles of industrial CCTV cameras, both of which block out the lower-body keypoints the geometric logic needs to work with.

Keywords: Action Classification, Deep Learning, Human Pose Estimation, Industrial Safety

1. INTRODUCTION

Worker safety in industrial environments has long been a serious concern in Malaysia. Facilities such as fabrication yards and active construction sites operate in environments where workers are regularly exposed to risks that can lead to serious injuries or fatalities if proper safety practices are not followed [1][2]. Traditional safety monitoring at these sites relies heavily on manual supervision, which is not only labour-intensive but also inconsistent [3][4], especially across large site areas with many workers moving around at the same time. Actions within these

*latifahmunirah@unimap.edu.my

high-risk environments has high mortality rates and any actions that are not related to work can not only effect safety, but it will also affect productivity.

The growing availability of CCTV infrastructure across industrial sites in Malaysia has opened the possibility of automating safety monitoring using computer vision and deep learning [5]. By processing existing camera feeds in real time, an automated system can flag unsafe worker behaviour without needing additional hardware or extra manpower on the ground. Among the behaviours worth monitoring, worker posture is one of the more informative signals [6]. Whether a worker is standing or sitting at a given moment can indicate a lot about what task they are carrying out and whether they are in a vulnerable or non-compliant position [2]. Human pose estimation has seen significant progress in recent years [7], with deep learning models now capable of extracting skeletal keypoints from video frames with high accuracy even in cluttered environments. Models such as YOLOv8x-pose [8][9] have made it practical to deploy pose estimation in real-time settings, offering a good balance between detection accuracy and processing speed.

This paper addresses that gap by proposing an automated postural classification system developed specifically for Malaysian industrial environments. The system makes use of YOLOv8x-pose to extract lower-body skeletal keypoints directly from CCTV footage and applies rule-based geometric logic to classify each detected worker as either standing or sitting in real time. Rather than depending on computationally heavy temporal networks, the proposed approach relies on four scale-invariant geometric metrics derived from joint coordinates, keeping the overall system lightweight and practical enough for actual on-site deployment. Validation is carried out using real-world footage captured on site, ensuring that the system's performance is measured against conditions that closely reflect what is encountered on Malaysian industrial worksites.

2. MATERIAL AND METHODS

This research focuses on building an automated postural classification system, specifically for standing and sitting, using deep learning for real-time monitoring at Malaysian industrial sites, mainly fabrication yards and construction sites. The system takes raw video frames and runs them through YOLOv8x-pose, which pulls out 17 skeletal keypoints from each worker detected in the frame [10]. That skeletal data is then used by a rule-based geometric logic to identify specific postures, without needing to pile on extra temporal networks that would add unnecessary computational load. All models are validated using a custom dataset of site footage annotated in YOLO format.

The system architecture is built around the YOLOv8x-pose model, which serves as the backbone for extracting 17 skeletal keypoints, covering Joint IDs 0 through 16, from each detected worker in the video frame [11]. These keypoints provide high-accuracy spatial coordinates for joints across the full body, from the nose and eyes down to the ankles, giving the classification logic the raw geometric data it needs to work with. The keypoints most relevant to postural classification are those covering the lower body, specifically Joint IDs 11 through 16, which map to the left and right hips, knees and ankles respectively. For standing classification, the system checks whether the S ratio exceeds 0.45, where a large enough vertical distance between the hip and ankle joints relative to the bounding box height is taken as confirmation of an upright posture. For sitting classification, the S ratio needs to drop below 0.45 and the knee angle θ needs to fall below 100°, which together indicate that the lower body is compressed and the knee is bent. Either way, the approach is scale-invariant, so the classification holds up whether the worker is standing close to the camera or far in the background.

System performance is measured using a standard set of evaluation metrics. Precision tells us how many of the model's positive detections were genuinely correct, while recall tells us how many of the actual positive cases the model managed to pick up. The F1-score brings both together into one balanced figure, and overall accuracy gives a general sense of how well the system performed across all tested instances. These metrics are then applied across a multi-class confusion matrix covering both the sitting and standing classes, which helps paint a clearer picture of where the system is holding up well and where there is still room for improvement. The classification logic is built to tell apart two basic postural states, that is standing and sitting, by computing the spatial relationships and joint angles from the extracted skeletal keypoint coordinates:

- Standing Posture: A worker is classified as standing when the S ratio exceeds 0.45, indicating that the vertical distance between the hip and ankle joints is sufficiently large relative to the bounding box height to confirm an extended, upright leg configuration.

- Sitting Posture: This state gets triggered when the S ratio drops below 0.45 and the knee angle θ falls below 100°, indicating that the lower body is compressed and the knee is in a flexed position consistent with a seated posture. On top of that, the system also checks that the torso tilt φ and normalised knee proximity D norm are consistent with a seated configuration before the classification is confirmed.

Running this check frame by frame keeps the monitoring at high frequency, so any postural changes happening on a busy, cluttered site floor can be caught quickly.

2.1 Model and Data

The research framework uses YOLOv8x-pose as the main deep learning model for human pose estimation and postural classification. This model was chosen because of its high representational capacity, meaning it can reliably extract 17 skeletal keypoints as labelled in Table 1 from each worker detected in the video frame. These keypoints give the spatial coordinates (x, y) for joints like the hips, knees and ankles, and from there, the data gets passed through a rule-based geometric logic to determine what posture the worker is actually in.

Table 1 The skeletal keypoints

Keypoint ID	Body Part
0	Nose
1,2	Left/Right Eye
3,4	Left/Right Ear
5,6	Left/Right Shoulder
7,8	Left/Right Elbow
9,10	Left/Right Wrist
11,12	Left/Right Hip
13,14	Left/Right Knee
15,16	Left/Right Ankle

For validating the action classification results, a custom dataset was put together from actual industrial site visits, covering locations like Kasawari and a few construction areas. The dataset is a combination of images manually annotated and taken from real site conditions to ensure variances such as inconsistent lighting and high-mounted camera angles are all represented in there. The images are labelled under two activity classes, standing and sitting, and these are used specifically to measure the accuracy, precision and recall of the postural classification module.

On the data preparation side, the LabelImg [1][12] tool was used to go through each validation image and manually annotate the workers together with their postural states. Everything is saved in YOLO format (.txt) to keep the annotations consistent, so the validation script can properly match up what the model predicted against the ground truth labels. This dataset is kept strictly for validating the action classification side of things, and it gives a proper real-world benchmark to check whether the system can reliably pick up worker postures in the kind of messy, cluttered conditions that are common at fabrication yards.

2.1.1 Model Calibration

The postural classification system is calibrated using a set of rule-based geometric constraints derived from the 17 skeletal keypoints that YOLOv8x-pose pulls out. The approach uses the spatial coordinates of joint as labelled in Figure 1. Keypoints 11 through 16, which cover the hips, knees and ankles, able to tell apart standing and sitting states without needing the extra computational load that comes with temporal neural networks. The logic is also built to be scale-invariant, so postural detection stays accurate no matter how far the worker is standing from the fixed CCTV camera.

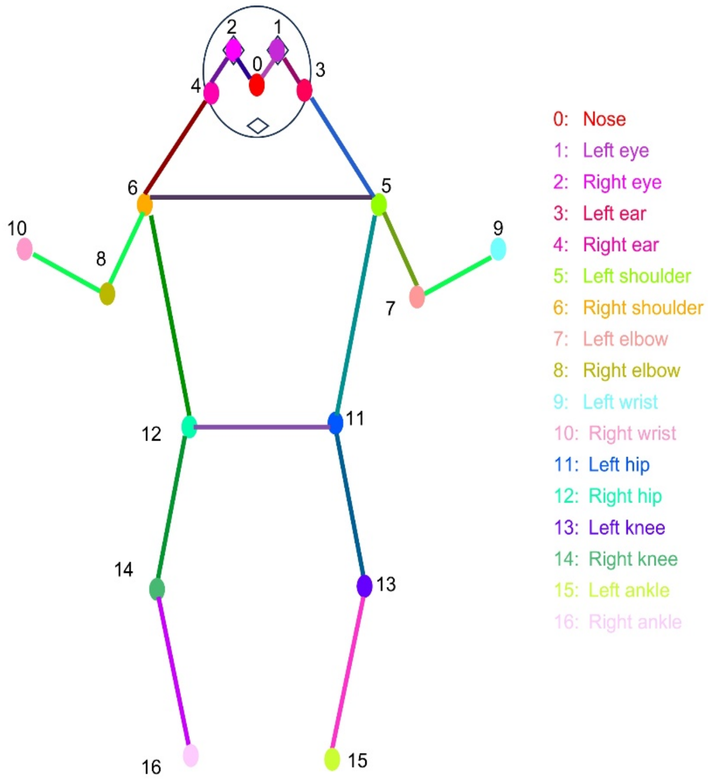

Figure 1. Keypoints positions on the human skeleton [13]

The main metric used for lower-body state detection is the vertical separation and height ratio, S ratio, as defined in Equation (1). For a worker who is standing, the vertical distance between the average ankle coordinates and the hip coordinates is expected to be relatively large compared to the total bounding box height (bh). On the flip side, a sitting state is picked up when this ratio drops below an empirically determined threshold, which tells the system that the hips have come down closer to ankle level. This vertical check is then refined further by calculating the knee angle, θ knee, using Equation (2). When someone is standing, applying the law of cosines across the thigh and calf segments should give an angle close to full extension. When someone is sitting, the knee angle comes out noticeably flexed instead. To identify vertical separation and height ratios

(S ratio), the vertical distance between the average ankle and hip coordinates is normalized by the person's bounding box height (bh) as presented by Equation (1):

$$S = \frac{avg(y_{ankles}) - avg(y_{hips})}{bh} \tag{1}$$

To handle camera tilt and off-angle views common in fabrication yards, torso orientation and joint proximity are included as supporting metrics alongside the primary height ratio and knee angle. As shown in Figure 2, the system first extracts lower-body keypoints (Joint IDs 11–16) from the YOLOv8x-pose output, then computes four scale-invariant geometric metrics before applying the classification logic. The height ratio S and knee angle θ serve as the primary decision layer. Standing is confirmed when S exceeds 0.45 and the knee is near full extension, indicating the legs are sufficiently straightened. Sitting is flagged when S drops below 0.45 and θ falls below 100°, reflecting a compressed lower body with a flexed knee. These thresholds were set based on observed joint geometry across the validation dataset a standing worker typically produces S values above 0.45, while a seated worker commonly records knee angles between 80° and 100°. When the height ratio alone is not conclusive, typically due to perspective foreshortening from high-mounted cameras, the system checks the torso tilt φ torso and normalized knee proximity D norm as a secondary confirmation. A torso tilt beyond 20° indicates forward lean consistent with a seated posture, while a D norm above 0.50 confirms the knee has risen close to the torso center. Only when all four conditions align is the sitting classification confirmed, which is what produces the zero false positive result seen in the validation.

The flexion of the lower limbs is determined by calculating the knee angle (θ knee) using the law of cosines based on the lengths of the thigh (a), calf (c), and the distance between the hip and ankle (b) as presented by Equation (2):

$$\theta = \arccos\left(\frac{a^2 + c^2 - b^2}{2(ac)}\right) \times \frac{180}{\pi} \tag{2}$$

The angle of the torso relative to a vertical axis (ϕ torso) is calculated to assess leaning or tilting postures) as presented by Equation (3):

$$\emptyset_{\mathrm{torso}} = \arccos\left(\frac{\mathrm{v}.\,|0,-1|}{||\mathrm{v}||}\right) \times \frac{180}{\pi} \tag{3}$$

Additionally, the proximity of specific joints to the torso center (D) is normalized to ensure scale invariance as presented by Equation (4):

$$D_{norm} = \frac{|y_{knee} - TC_y|}{bh} \tag{4}$$

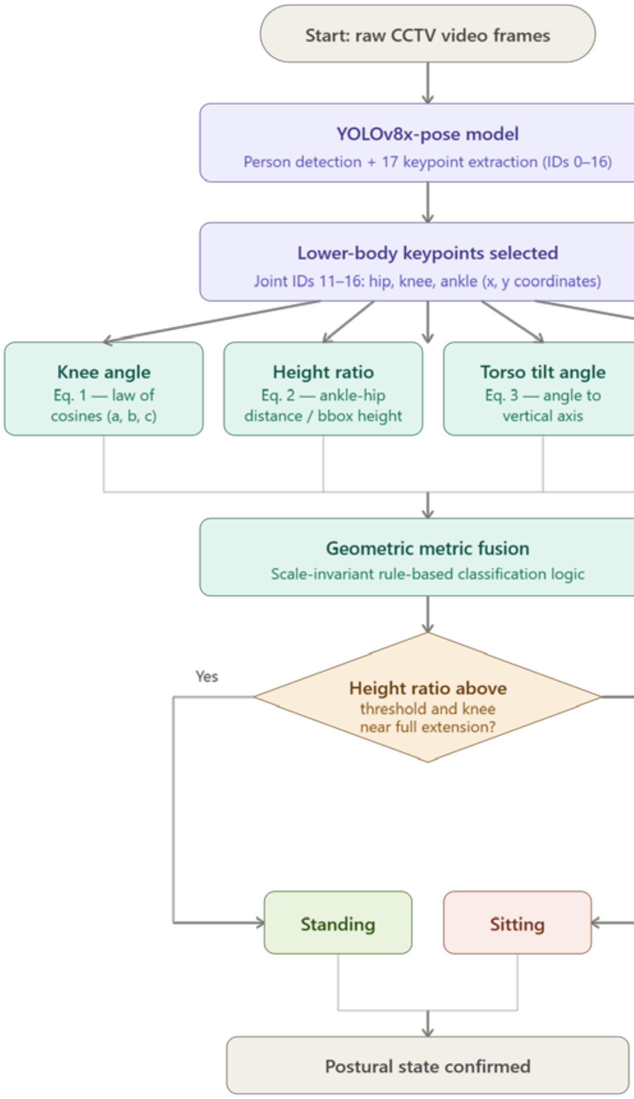

Figure 2. Flowchart of the geometric-based pose classification pipeline

2.1.3 Evaluation Metrics

Once the postural state of each detected worker is confirmed by the classification logic, the system begins outputting results on a frame-by-frame basis, streaming the classified posture labels in real time to the safety monitoring interface. This continuous output allows safety officers to track worker activity as happens across the site without any manual review of footage. To evaluate how well the system performs, a dedicated validation dataset consisting of real-world industrial site footage was prepared using LabelImg and saved in YOLO format, serving as the ground truth reference for all performance measurements as further visualized in Figure 3.

The metric result of precision as given in Equation (5) measures the proportion of correctly identified positive instances out of all instances identified as positive by the model, reflecting the accuracy of the detections:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall, assesses the proportion of actual positive instances that the model correctly identifies, indicating its ability to detect all relevant objects as shown in Equation (6):

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Mean Average Precision (mAP) exemplified in Equation (7) combines precision and recall across different confidence thresholds to provide a single performance metric, offering a comprehensive view of the model's accuracy:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{7}$$

mAP 0.5:0.95 shown in Equation (8) extends this concept by averaging the mAP values across multiple Intersections over Union (IoU) thresholds, typically from 0.5 to 0.95 in steps of 0.05, thus evaluating the model's performance across varying levels of localization precision:

$$mAP@[.5:.95] = \frac{1}{N}\sum_{i=1}^{N} AP@IoU_i \tag{8}$$

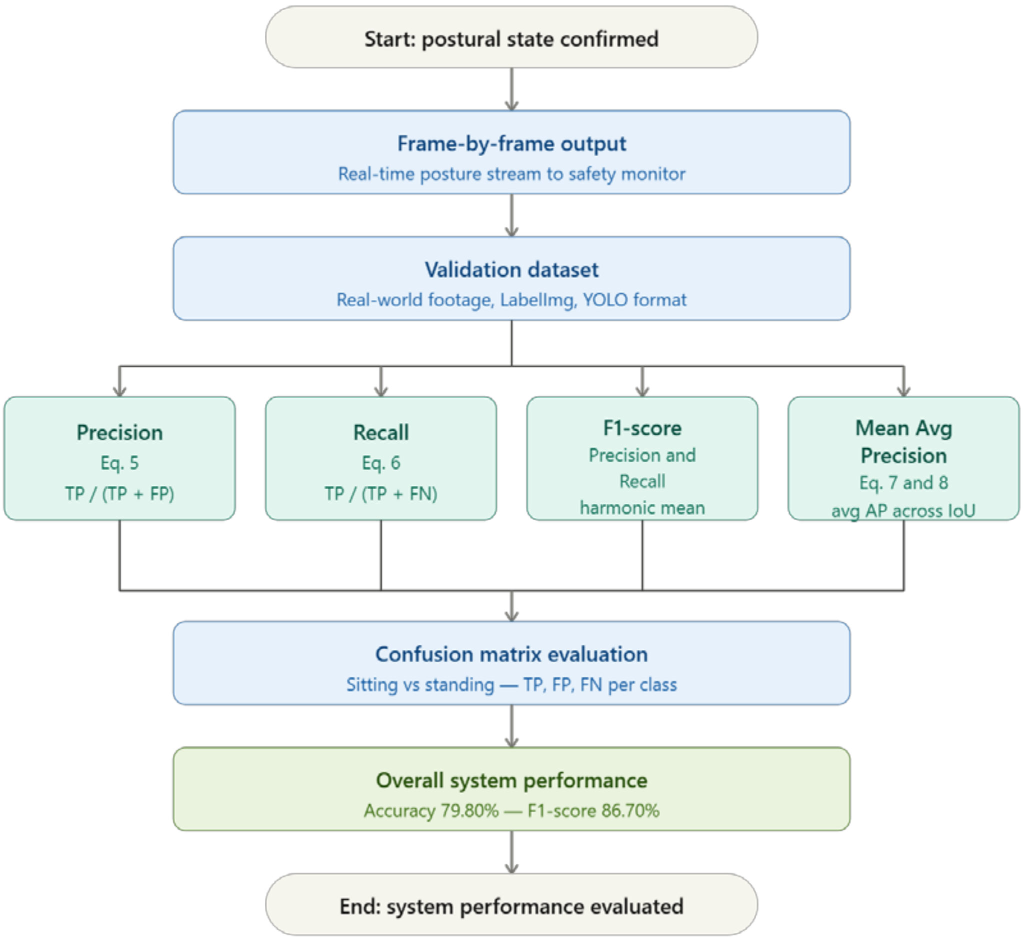

Figure 3. System performance evaluation flowchart

3. RESULTS AND DISCUSSION

This section discusses the performance of the pose-based action classification module, which categorises workers into standing and sitting postures. The system uses geometric rules based on the 17 keypoints extracted by the YOLOv8x-pose model. Based on the validation using the manually annotated site dataset, the overall system achieved an accuracy of 79.80% and an F1-score of 86.70% across 641 tested instances as further shown in Table 2.

Table 2 Overall performance metrics

Class	TP	FP	FN	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Standing	343	26	38	93.00	90.00	91.50	89.60
Sitting	148	0	86	100.00	63.20	77.50	86.00
Overall	491	26	124	95.00	79.80	86.70	79.80

The findings show that the standing posture was detected with 93.00% precision, 90.00% recall, and an F1-score of 91.50%. During testing, the system recorded 343 true positives (TP) and only 26 false positives (FP). This high success rate is expected because a standing worker generally exposes their full body to the camera without much blockage. In an upright position, the distance between the hip and knee joints is fully extended. Due to these keypoints are clearly separated, the geometric rules can compute the measurements easily to confirm the standing pose.

Sitting classification achieved a perfect precision of 100.00%, meaning every instance the system flagged as a seated worker was a correct detection with zero FP. This confirms that the geometric thresholds calibrated for the sitting posture are strict and well-grounded. For the system to confirm a sitting state, the S ratio must drop below the empirical threshold, the knee angle must show clear flexion, and the supporting metrics for torso tilt and knee proximity must also be consistent with a seated configuration. The fact that no FP were recorded across the entire validation set shows that this combination of conditions is specific enough to sitting that no other posture, whether crouching, leaning or bending, was able to satisfy all four criteria at the same time. This is an important result for real industrial deployments, where a false alarm flagging a standing worker as non-compliant could undermine the confidence of safety officers in the system.

However, the sitting recall of 63.20% and the 86 recorded false negatives point to a genuine limitation that is tied closely to the conditions of real industrial footage rather than any fundamental flaw in the classification logic itself. The most significant contributing factor is the presence of workers in the background of frames who were included in the ground truth annotations but were simply too far from the camera for YOLOv8x-pose to reliably localise their keypoints. At greater distances, the worker's body occupies a smaller pixel region in the frame, and the model either fails to detect the person entirely or produces keypoint confidence scores that fall below the threshold needed for the geometric logic to run [11][14]. These instances are still counted as false negatives in the validation results, which brings the recall figure down even though the system was never realistically given enough visual information to make the classification in the first place.

Partial occlusion of seated workers is another major factor [15]. In a busy fabrication yard or construction site, workers who are sitting are often doing so behind heavy equipment, workpieces, scaffolding or other workers standing in the foreground. Since the sitting classification logic specifically depends on having reliable and simultaneous coordinates for the hip, knee and ankle joints, any occlusion that removes or degrades one of these keypoints breaks the geometric chain. Unlike standing, where the joints tend to be spread across a larger vertical range and are harder to fully obscure, a seated worker compacts their lower body into a smaller spatial area, which makes it far more likely that a single piece of equipment or a foreground obstruction will block multiple critical joints at once. When the keypoints are missing or flagged as low confidence, the system conservatively skips the classification rather than risk a FP, which is the right design decision for precision but directly costs sitting recall.

The steep downward mounting angle of industrial CCTV cameras introduces an additional geometric challenge that is harder to fully compensate for through normalisation alone. When a camera is mounted high above the work floor and angled downward, the resulting perspective foreshortens the vertical distance between joints in a seated worker. The hip and knee, which should appear well-separated vertically in a side-on or level view, appear much closer together when viewed from above. This compresses the S ratio reading and pushes it toward the standing threshold, making it harder for the rule-based logic to confidently register the posture as sitting even when the worker clearly is. While the torso tilt metric ϕ torso and the normalised knee proximity D norm help to partially offset this effect by capturing postural information that is less sensitive to vertical foreshortening, they cannot fully overcome the perspective distortion introduced by extreme camera angles. Expanding the calibration dataset to include more samples taken from steep angles and retuning the empirical thresholds specifically for high-mounted camera configurations would likely produce a meaningful improvement in sitting recall for these cases.

Taken together, these results show that the system is highly reliable when it does make a classification, as evidenced by the 95.00% overall precision and the perfect sitting precision, but that the recall is constrained by real-world data conditions that are common across Malaysian industrial sites. Addressing these constraints through better handling of low-resolution background detections, partial keypoint imputation strategies, and angle-aware threshold calibration would bring the sitting recall closer to the level already achieved for standing and further strengthen the system's overall suitability for deployment as a real-time safety monitoring tool as exemplified in Figure 4.

Figure 4. YOLOv8x-pose keypoint extraction and activity prediction result

4. CONCLUSION

Overall, the system achieved an accuracy of 79.80% and an F1-score of 86.70%, with standing classification reaching 93.00% precision and 90.00% recall, and sitting classification achieving a perfect precision of 100.00%, meaning every seated worker that was flagged by the system was a genuine case.

The sitting recall of 63.20% is admittedly on the lower side, but the perfect sitting precision is arguably more valuable for this kind of safety monitoring application. On an actual industrial site, a false alarm is not just an inconvenience as it erodes trust in the system over time and makes safety officers less likely to act on alerts. Future work can look at fine-tuning the joint angle and distance thresholds to handle cases where a worker's lower body is partially blocked by site equipment, as well as incorporating partial keypoint logic and collecting more varied footage that better represents how workers sit on the ground.

Looking further ahead, the keypoint-based foundation built in this work can be extended into a broader action classification system capable of recognising more specific industrial tasks beyond just sitting and standing. By feeding temporal keypoint sequences into a classification model, the system can learn to tell apart activities like welding, grinding and general movement, each of which has its own distinct body posture and motion signature. Once the system knows what a worker is doing, it becomes possible to enforce task-specific PPE requirements in real time, for example alerting when someone is welding without a proper welding mask or grinding without face protection. This kind of capability would make the system genuinely useful not just as a posture monitor but as a full safety enforcement and manpower tracking tool for Malaysian fabrication yards and construction sites.

DISCLAIMER

The dataset utilised in this study is proprietary in nature and is not made available for public access. Parties who are interested in obtaining the dataset for academic or research purposes are kindly requested to reach out directly to Dr. Latifah Munirah through the appropriate institutional channels. Any subsequent use of the dataset shall be subject to the terms and conditions as stipulated by the data owner, and access will only be granted upon receiving proper authorisation.

REFERENCES

[1] S. Madiahlagan *et al.*, 2024. Implementation of Object Tracking using YOLOv7 Pose Estimation for Human Activity Recognition in Fabrication Yard via CCTV Videos, in 5th International Conference on Smart Sensors and Application: Shaping the Future of Intelligent Innovation, ICSSA 2024, Institute of Electrical and Electronics Engineers Inc., doi: 10.1109/ICSSA62312.2024.10788582.

[2] M. Imam *et al.*, 2025. Integrating real-time pose estimation and PPE detection with cutting-edge deep learning for enhanced safety and rescue operations in the mining industry, Neurocomputing, vol. 618, Feb. 2025, doi: 10.1016/j.neucom.2024.129080.

[3] N. A. A. Bakar *et al.*, 2024. Feasibility Analysis of PPE Detection via YOLOv7 & YOLOv7-Pose to Enhance Workplace Safety, in 5th International Conference on Smart Sensors and Application: Shaping the Future of Intelligent Innovation, ICSSA 2024, Institute of Electrical and Electronics Engineers Inc., doi: 10.1109/ICSSA62312.2024.10788619.

[4] S. Sivanraj, D. N. L. S. Uduwage, M. Tripathi, 2026. Real-time safety detection on construction sites using a vision-language and NLP-based model, Advanced Engineering Informatics, vol. 69, Jan. 2026, doi: 10.1016/j.aei.2025.103889.

[5] N. Azmi, L. M. Kamarudin, A. Zakaria, S. Muhammad, and M. Syed Zakaria, 2023. AI-Based Analytics for Hawkers Identification in Video Surveillance for Smart Community, Journal of Techno-Social, vol. 15, no. 2, pp. 77–87, doi: 10.30880/jts.2023.15.02.008.
[6] Q. Zhao, C. Zheng, M. Liu, P. Wang, C. Chen, 2023. PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation, [Online]. Available: http://arxiv.org/abs/2303.17472.
[7] D. Bermuth, A. Poeppel, W. Reif, 2025. SimpleDepthPose: Fast and Reliable Human Pose Estimation with RGBD-Images, [Online]. Available: http://arxiv.org/abs/2501.18478.
[8] S. Wang, 2025. Automated non-PPE detection on construction sites using YOLOv10 and transformer architectures for surveillance and body worn cameras with benchmark datasets, Sci. Rep., vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-12468-8.
[9] N. Malik, A. Nayak, 2025. Artificial Intelligence System Based Personal Protective Equipment Detection for Construction Site Safety using YOLOv8, Int. J. Res. Appl. Sci. Eng. Technol., vol. 13, no. 3, pp. 2675–2683, Mar. 2025, doi: 10.22214/ijraset.2025.67829.
[10] R. Sapkota, M. Karkee, 2026. Ultralytics YOLO Evolution: An Overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 Object Detectors for Computer Vision and Pattern Recognition, [Online]. Available: http://arxiv.org/abs/2510.09653.
[11] P. Kadam, G. Fang, F. Amirabdollahian, J. J. Zou, P. Holthaus, 2024. Hand Pose Detection Using YOLOv8-pose, in 2024 IEEE Conference on Engineering Informatics, ICEI 2024, Institute of Electrical and Electronics Engineers Inc., doi: 10.1109/ICEI64305.2024.10912185.
[12] H. M. Ahmad, A. Rahimi, 2025. SH17: A dataset for human safety and personal protective equipment detection in manufacturing industry, Journal of Safety Science and Resilience, vol. 6, no. 2, pp. 175–185, Jun. 2025, doi: 10.1016/j.jnlssr.2024.09.002.
[13] M. Qixiang, G. Zhilin, W. Jintao, K. Qiqi, B. Fanliang, 2025. Collaborative detection network for sensitive targets and abnormal human behaviour in public places, Guangdian Gongcheng/Opto-Electronic Engineering, vol. 52, no. 8, doi: 10.12086/oee.2025.250131.
[14] Y. Chen, L. Zhao, Y. Xu, H. Zu, X. An, G. Li, 2024. Domain adaptive pose estimation via multi-level alignment, [Online]. Available: http://arxiv.org/abs/2404.14885.
[15] M. Iqbal, S. Z. Shah, Z. Ullah, 2026. Graph-Based Pattern Restoration for Occlusion-Robust Human Pose Estimation in Crowded Scenes, Algorithms, vol. 19, no. 2, Feb. 2026, doi: 10.3390/a19020142.