Journal of Advanced College of Engineering and Management

Comparison of YOLO Algorithms for Sitting Posture Assessment of Office-Based Workers in Construction

[1]S. Sivanraj, [2]D.N.L.S Uduwage, [3*]R.K. Shiwakoti, [4]M. Thripath
[1]*School of Architecture and Built Environment, Deakin University, Australia*
[2]*Deptartment of Building Economics, University of Moratuwa, Sri Lanka*
[3]*Institute of Engineering, Pulchowk Campus, Tribhuvan University, Nepal*
[4]*School of ICT, Thammasat University, Thailand*
*Corresponding email: *ranju.shiwakoti@ioe.edu.np*

DOI: 10.3126/jacem.v12i01.93906

Abstract
Improper sitting posture is a significantly contributes to physical and psychological damages to office-based workers particularly where there is a lack of continuous ergonomic monitoring. This study analyses four automated sitting posture detection models based on recent versions of YOLO object detection algorithm. The models demonstrated reliable sitting posture classification by accurately detecting sitting posture variation. The proposed models provided an efficient solution for office-based workers siting posture monitoring compared to existing sensor based and computer vision-based models. The models were evaluated based on key performance matrices such as true positive values, precision score, recall score, mAP values and inference speed. Among the four models, YOLOv11-s was selected as the optimal model due to its performance in accuracy and computational efficiency. The model can be further developed by integrating real-time feedback mechanisms, such as automated posture alerts to improve workplace ergonomics, worker wellbeing, and overall safety practices.
Keywords*—Construction, Ergonomics, Object detection, Pose detection, Sitting Posture, YOLO Algorithms*

1. INTRODUCTION

Ergonomics is the science of designing workplaces, tools and tasks to align with the capabilities and limitations of the workers within the working environment [1]. In modern office settings, sitting posture represents a fundamental aspect of ergonomic design because office-based workers are often required to remain seated for long hours while performing computer-based tasks [2]. This is particularly significant in construction offices where engineers, project managers, quantity surveyors, architects and administrative staff regularly engage in work involving documentation, digital modelling, and project coordination. Despite the significance of maintaining proper sitting posture, office-based construction workers frequently overlook ergonomic practices during prolonged work activities.

Prolonged improper sitting has led to a wide range of health impacts, including physical and psychological damage [3]. As a result, monitoring and assessing sitting posture has become a significant consideration in occupational ergonomics. Various approaches have been developed to assess and monitor sitting posture including traditional ergonomic assessment methods which rely on observational techniques and worker surveys which often require continuous human involvement. As a result, technological solutions such as pressure sensor systems [4], wearable sensors [5], radio-frequency sensing [6], and computer vision-based monitoring systems [7] have been introduced to automate sitting posture detection. While sensor-based systems can provide accurate measurements of body orientation and pressure distribution, they often require specialized hardware installations, which increase implementation costs.

Pose estimation models have been widely used for sitting posture monitoring by identifying key points of human body joints and analysing body alignment [7]. Although these techniques can provide detailed skeletal information for sitting posture assessment, they often rely on complex deep learning architectures that require substantial computational resources. Furthermore, many existing studies utilize webcam-based monitoring systems within controlled environments, which may raise privacy concerns and create practical limitations for real-world workplace implementation.

To address these challenges, object detection algorithms such as You Only Look Once (YOLO) offer an alternative computer vision approach. YOLO has gained considerable attention in construction [8], [9] due to its real-time detection capability, lower computational requirements, and simpler training process compared to many other computer vision models [9]. Over time, the algorithm has undergone several advancements, leading to improved versions such as YOLOv9, YOLOv11, YOLOv12, and YOLOv26, each providing enhancements in detection accuracy, efficiency, and processing speed. However, despite the application of YOLO algorithms in various construction-related computer vision tasks, its application for sitting posture detection in construction office environments remains unexplored.

2. LITERATURE REVIEW

A. Impacts of Improper Sitting Posture

Sitting posture is a critical ergonomic factor that impacts the health and wellbeing of office-based construction workers. According to a study conducted by Markova et al. [3], when workers maintain improper sitting posture over a prolonged time it has resulted in a range of physical and psychological damages [10]. One of the most significant physical damages relates to the spinal and musculoskeletal system as improper sitting posture can alter the natural spinal alignment causing chronic back pain [11]. A survey-based study revealed that between 60% and 90% of individuals experiencing prolonged sitting postures have reported musculoskeletal discomfort [3].

In addition to musculoskeletal impacts, improper sitting posture has also influenced respiratory and functional performance. Slump sitting positions restrict the normal diaphragm movement and impacts respiratory efficiency. Furthermore, in occupation that involves extensive use of digital devices, posture-related musculoskeletal discomfort has also been linked to fatigue, sleep disturbances, and psychological effects such as anxiety and reduced concentration, which have negatively impacted the overall work productivity [2].

B. Existing technological advancement on sitting posture detection

Various sensing technologies and computational models have been developed to monitor sitting posture and identify improper ergonomic behaviours. These approaches generally rely on pressure sensor systems [4], wearable sensors [5], radio-frequency sensing [6], and computer vision-based monitoring [7]. The pressure sensor-based system uses pressure sensors which are embedded in seats to measure pressure distribution across the chair surface [4], [12]. The systems analyse the pressure exerted during different sitting postures and classifies them. Although these methods have been successful, these systems incur high hardware costs and have higher system complexity [13]. Similarly, wearable sensor-based systems are also widely used in sitting posture monitoring studies [14]. These systems commonly use sensors attached to different parts of the body to capture orientation and movement data. Although these can provide precise measurements of body posture and movement patterns, studies highlighted several practical challenges including worker discomfort and hesitations [5], [14]. Studies have also explored radio frequency and radar-based sensing techniques, which use wireless signal reflections to analyse worker sitting posture. These methods preserve the worker privacy as they do not capture visual information. However, radio frequency-based systems require hardware with specific capabilities to receive and process the signals which results in a higher implementation cost [14]. Additionally, these systems are also impacted by other interfering signals which affect their reliability.

Within computer vision-based approaches for ergonomic assessment, one of the most commonly used techniques is posture monitoring through pose estimation models [7], [15]. These systems identify the key points of human body joints, such as the shoulders, neck, hips and spine, and analyse their angles to determine whether a person is maintaining a proper or improper sitting posture. Pose estimation-based systems provide detailed information about body alignment and have therefore been widely explored in sitting posture assessment applications. However, these approaches often rely on complex deep learning architectures and continuous real-time skeletal tracking, which can be computationally intensive and require significant processing resources. In addition, many of the current studies have focused on sitting posture detection using webcam-based video recordings, where the camera is positioned directly in front of the user in a controlled workspace. This raises concerns regarding the privacy and compliance of the user.

C. You Only Look Once (YOLO) Algorithm

In addition to pose estimation methods, another widely used technique within computer vision is object detection which has gained considerable attention in the construction industry [8], [16], [17]. Among the various object detection algorithms, the YOLO algorithms are widely used [18]. YOLO is designed to perform object detection efficiently by processing the entire image in a single pass, enabling real-time performance [19]. Due to its relatively lower computational complexity and ease of training and implementation, YOLO has been widely adopted in construction-related computer vision applications [8].

Over time, the algorithm has undergone several advancements providing improved accuracy, computational efficiency and real-time performances. These developments have resulted in several versions of the model including YOLOv9, YOLOv11 and YOLOv12, followed by newer releases such as YOLOv26. In addition, each model version consists of multiple variants allowing users to select a configuration based on the requirements of the application. These variants range from nano (n) variant, which are suitable for application with faster but limited computational capabilities, to extra-large (x) variants, which prioritise detection accuracy and high computational capabilities. Among these the small variants provide a balance between detection accuracy, computation capabilities and processing speed making them ideal for medium scale real-time application.

D. Significance and Transferability

Several studies have addressed sitting posture monitoring using pressure sensors, wearable devices, radio frequency sensing and pose estimation-based computer vision systems. However, studies have highlighted that many of these approaches use high computational models which require specialized hardware and incur high cost. In addition, most ergonomic monitoring studies conducted within the construction industry focused on site-based workers comparatively limited attention on office-based construction workers. Despite significant advancements, existing approaches fail to simultaneously achieve low-cost deployment, minimal computational requirements and real-time performance in construction office environments. This highlights a critical research gap in developing a practical posture monitoring model suitable for office-based construction workers. Recent advancements in real-time object detection, particularly YOLO-based architectures, provide an alternative to traditional pose estimation models. Unlike skeletal tracking approaches, YOLO-based models can directly classify posture states from visual data with significantly lower computational overhead, making them suitable for real-time applications. Therefore, this study evaluates and compares the performances of four YOLO based models which developed using the recent versions of YOLO including YOLOv9, YOLOv11, YOLOv12 and YOLOv26 and identifies the most suitable model for assessing the sitting posture of office-based construction workers in construction. The models were trained, tested and validated using an online dataset during the development phase.

3. METHODOLOGY

This study trained four versions of YOLO algorithms, YOLOv9-s, YOLOv11-s, YOLOv12-s, and YOLOv26-s, for office-based worker sitting posture detection. Each model was trained to detect and classify sitting postures from image data. The models were developed using an online training platform and their performance was evaluated using standard object detection metrics specified by Sivanraj et al. [9]. Figure 1 illustrates the overall architectural diagram for the pose detection model.

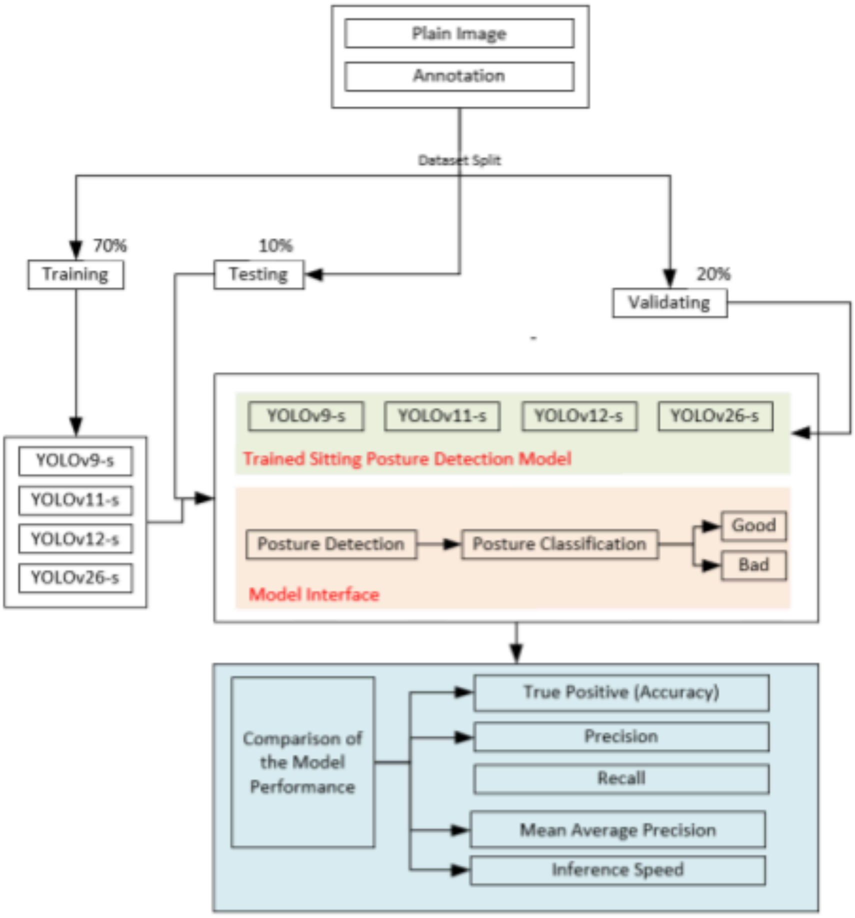


Figure 1: Overall architectural diagram for PPE detection models

A. Dataset

Obtaining a dataset with adequate number of images is a critical step required for training YOLO based object detection models. In this study, a dataset consisting of 3,438 images with annotations were obtained from an online database called RoboFlow. The dataset consisted of various sitting posture, body orientation and sitting positions taken from different camera angles. The dataset was split into 70% for training, 20% for internal validation, and 10% for testing as shown in Figure 1. This is a commonly adopted practice in computer vision literature [20], [21]. The dataset from Roboflow was already provided in this configuration, ensuring consistency. If the split is changed, the results may vary due to differences in training data size and evaluation reliability; however, for sufficiently large datasets, performance differences are generally limited.

All the images were pre-processed to ensure consistent training, testing and validation process. The images were resized to a 640x640x3 pixel resolution using the RGB colour format which is a common input configuration for YOLO based models. Each image consisted of annotations, which provided the coordinates of the bounding box required by the model for localization, as described by Sivanraj et al. [9]

B. Model Training

This section describes the training process involved in developing the sitting posture detection models. Identical training parameters and procedures were maintained to ensure fair assessment and identification of the optimal YOLO model for sitting posture detection. Prior to training the model, it is significant to configure the hyperparameters of the YOLO models. Hyperparameters are significant to control the performance and computational efficiency of the model. Several key hyperparameters were used in this study to balance training efficiency and model performance. Accordingly, the values highlighted in Table 1 were set for the key hyperparameters highlighted by Sivanraj et al. [9]. The values for the hyperparameters were selected based on commonly used YOLO training configurations.

Table 1: Hyperparameters values of the YOLO models

Hyperparameters	Values
Batch Size	64 images
Learning Rate	0.01
Momentum	0.937
Weight Decay	0.0005
Training Epochs	300

Since the primary objective of this study is to evaluate the performance of different YOLO models rather than optimize individual configurations, most hyperparameters were kept fixed. These hyperparameters follow the standard training configurations reported in the original YOLO implementations, ensuring consistency with prior work and enabling a fair comparison across different model variants without introducing additional tuning bias. Only the number of training epochs and early stopping patience were adjusted. Specifically, training was conducted for up to 300 epochs with a patience of 20 epochs. This step stopped the model training when no improvements in performance were observed over 20 consecutive epochs. This allows sufficient time for convergence while preventing overfitting, as training stops if no improvement is observed for 20 consecutive epochs.

After splitting the dataset, the four models were trained independently using 2,419 images from the dataset to detect and classify the sitting postures as good (proper) and bad (improper) sitting postures. The models received the plain images and their corresponding annotations. With each training epoch, the model underwent validation with the 682 images form the dataset. Through this process, trained posture detection models were obtained. Once the training was completed, the performance of each sitting posture detection models was compared. Table 2 provides the training time and the total number of epochs completed by each model.

Table 2: Training time and epochs summary of the models

Model	YOLOv9-s	YOLOv11-S	YOLOv12-s	YOLOv26-s
Training time	2.56	2.37	3.49	2.86
Total Epochs	103	149	142	150

4. RESULTS AND PERFORMANCE EVALUATION

The performance of the trained sitting posture detection models was evaluated using several standard object detection metrics, including true positive classification values, precision-recall values, mean Average Precision (mAP), and inference speed.

A. True Positive Values

According to Caelen [22], a confusion matrix is used to evaluate the classification performance of a model by comparing predicted classifications against ground truth labels. The confusion matrix enables the calculation of true positive values, which represent the instances where the model correctly identifies the classes. In this study, true positive classification represents the correct identification of good sitting posture and bad sitting posture by the trained posture detection models. The overall accuracy of each model in correctly classifying good and bad sitting postures is summarized in Table 3.

Table 3: Accuracy of trained posture detection models

Model	**YOLOv9-s**	**YOLOv11-S**	**YOLOv12-s**	**YOLOv26-s**
Good Posture	0.76	0.77	0.74	0.78
Bad Posture	0.87	0.88	0.88	0.85

The results indicate that all models were able to identify bad posture with relatively higher accuracy compared to good posture. Among the models, YOLOv26-s achieved the highest true positive rate for good posture detection, while YOLOv11-s and YOLOv12-s demonstrated the highest accuracy for bad posture classification.

B. Precision Recall Values

Precision and recall are significant evaluation metrics used to measure the classification performance of object detection models. Precision represents the proportion of correctly predicted positive classifications relative to the total number of positive predictions made by the trained posture detection model. A higher precision value indicates that the model produces fewer false positive predictions. Similarly, recall represents the proportion of correctly predicted positive classifications relative to the total number of actual positive instances in the dataset. Table 4 summarizes the precision and recall values for the four trained sitting posture detection models.

The results indicate that YOLOv26-s achieved the highest precision value, demonstrating its ability to correctly classify the sitting posture detections with fewer false positives. Conversely, YOLOv11-s achieved the highest recall value, indicating a stronger capability in identifying the majority of actual positive sitting posture instances.

Table 4: Precision and recall values for the trained sitting posture detection models

Model	**YOLOv9-s**	**YOLOv11-S**	**YOLOv12-s**	**YOLOv26-s**
Precision	0.819	0.792	0.816	0.838
Recall	0.764	0.817	0.782	0.797

However, a model should maintain a balance between precision and recall achieving reliable detection performance. The precision-recall values can be calculated using precision-recall curves, which illustrate the trade-off between the two metrics across different detection thresholds. Figure 2 presents the precision-recall curves generated for the trained sitting posture detection models.

Journal of Advanced College of Engineering and Management

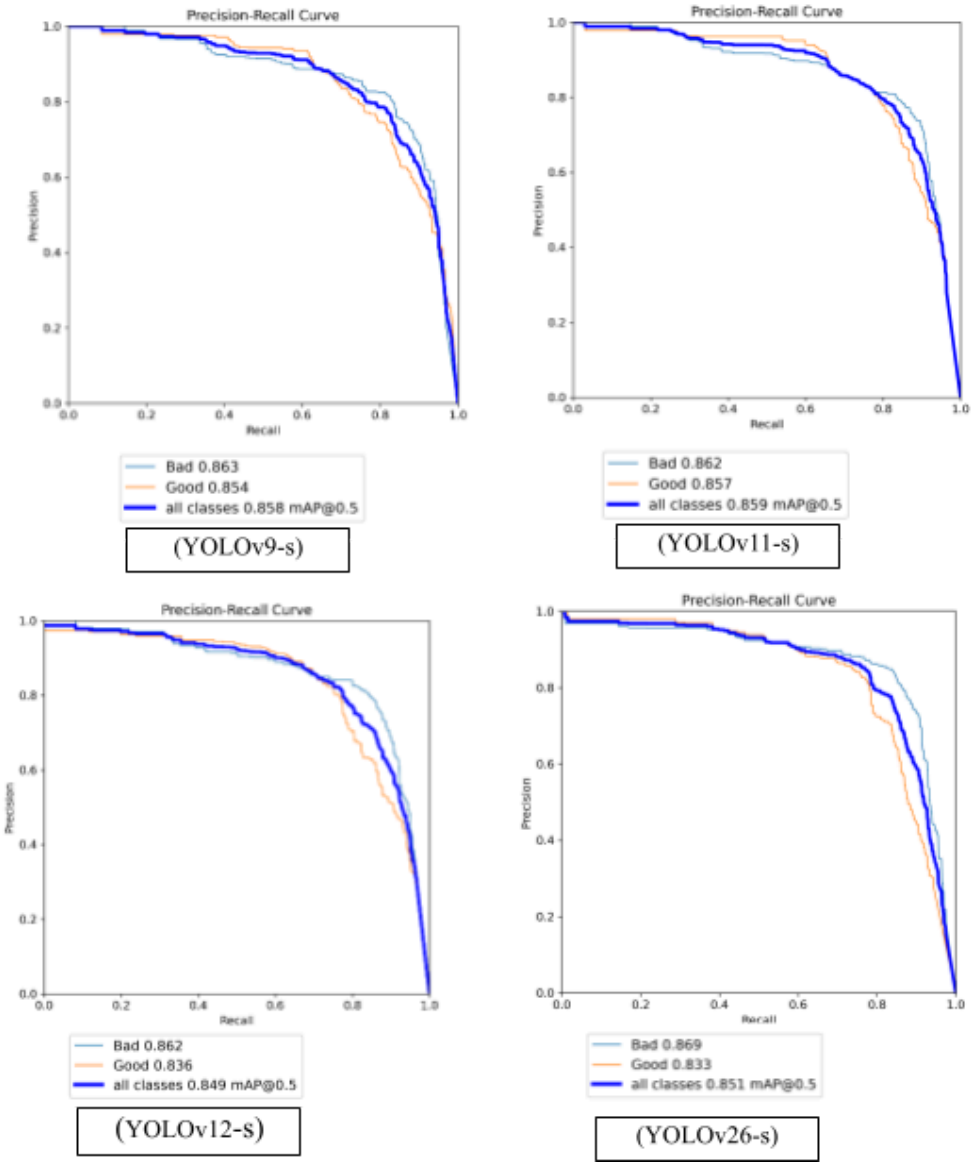


Figure 2: Precision-recall curves of trained models

C. Mean Average Precision (mAP)

Mean Average Precision (mAP) is another matrix used for evaluating the performance of object detection models based on the balance between the precision and recall values [23]. The mAP value is calculated by calculating the mean of the average precision values across all the detection classes. The mAP is evaluated at different Intersection over Union (IoU) thresholds, which measure the overlap between the predicted bounding box and the ground truth bounding box as highlighted by Sivanraj et al. [9]. In this study, mAP@0.5 and mAP@0.5:0.95 were calculated for each model as summarized in Table 5.

Table 5: mAP values of the trained sitting posture detection models

Model	YOLOv9-s	YOLOv11-S	YOLOv12-s	YOLOv26-s
mAP@0.5	0.858	0.859	0.849	0.851
mAP0.5:0.95	0.610	0.624	0.621	0.622

According to the results, YOLOv11-s achieved the highest mAP values at both IoU thresholds, indicating its strong overall detection performance in sitting posture classification compared to the other three trained sitting posture detection models.

D. Inference Speed

In addition to detection accuracy, the computational efficiency of the models was evaluated using inference speed, which represents the time required for the model to process an image and produce a prediction. The total inference time consists of pre-processing, model inference, and post-processing stages. The speed at each stage was calculated in the unit milliseconds per image (ms/image) as given in Table 6.

Table 6: Inference speed of the trained sitting posture detection models

Model	YOLOv9-s	YOLOv11-S	YOLOv12-s	YOLOv26-s
Preprocessing	0.2	0.2	0.2	0.2
Inference	5.7	4.3	6.8	4.3
Post processing	1.0	1	1.1	0.1
Total Inference speed	6.9	5.5	8.1	4.6

The results show that all models recorded similar pre-processing speeds of approximately 0.2 ms/image. However, differences were observed in the inference and post-processing stages. The inference time ranged from 4.3 ms/image image to 6.8 ms/image, with YOLOv11-s and YOLOv26-s demonstrating the fastest inference speeds. Additionally, YOLOv26-s demonstrated a significantly lower post-processing time compared to the other models. As a result, when the total inference time is considered, YOLOv26-s achieved the fastest overall processing speed.

E. Overall comparison of the trained sitting posture detection models

Table 7 presents the overall comparison of the four trained sitting posture detection models based on the key evaluation metrics considered in this study. The results indicate that YOLOv11-s demonstrated the highest accuracy in terms of true positive classification, indicating its strong capability in correctly identifying both good and bad sitting postures. Additionally, YOLOv11-s also demonstrated the highest recall and mAP values. In contrast, YOLOv26-s achieved the highest precision value, showing its effectiveness in reducing false positive detections during sitting posture classification. The model also outperformed the other models by achieving the fastest inference speed, making it the most efficient model for real-time sitting posture detection. Overall, among the five key evaluation metrics considered in this study, YOLOv11-s demonstrated outperformed the other trained sitting posture detection models in three of the metrics, indicating its strong overall capability for accurate sitting posture detection.

Table 7: Overall comparison of the trained sitting posture detection models

Model	YOLOv9-s	YOLOv11-s	YOLOv12-s	YOLOv26-s
Accuracy		✔		
Precision				✔
Recall		✔		
mAP		✔		
Inference speed				✔

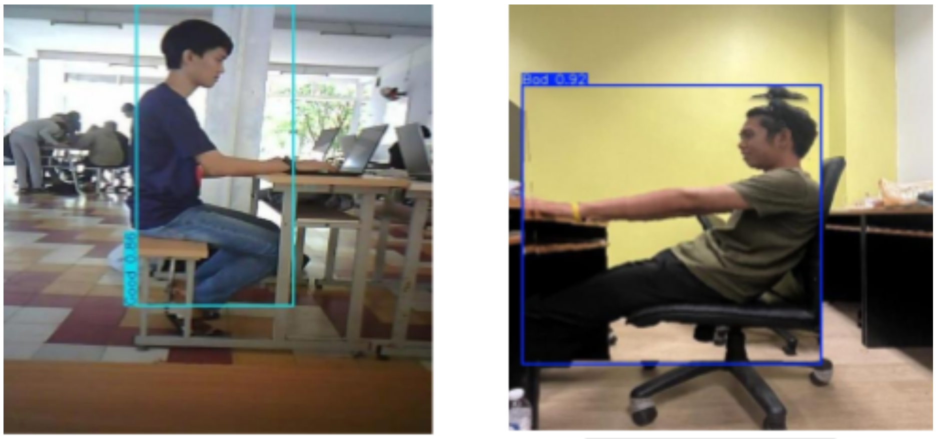
The observed performance difference between YOLOv11-s and YOLOv26-s can be attributed to their architectural and training design choices. YOLOv11-s achieved a higher accuracy, recall, and mAP due to its use of Distribution Focal Loss (DFL), which enhanced bounding box localisation though modelling spatial uncertainty more effectively. This results in improved detection of challenging and small objects, thereby increasing recall and overall mAP. Additionally, its deeper feature representation contributes to more robust object discrimination.

In contrast, YOLOv26-s demonstrated higher precision and faster inference speed primarily due to the removal of DFL and simplification of the detection head [24]. This reduces computational overhead and minimises latency, making it more suitable for real-time and edge-device applications. The streamlined architecture also leads to fewer false positive detections, thereby improving precision. Consequently, YOLOv26-s offers a favorable trade-off between speed and precision, while YOLOv11-s remains more effective for comprehensive detection performance.

Compared to the existing pressure and wearable sensor-based systems, the proposed YOLO-based approach eliminates the need for additional hardware and body-attached devices, thereby reducing implementation cost and improving user comfort. While these traditional methods often achieve high accuracy, their scalability and practical deployment in office environments remain limited. Additionally, the proposed method utilises standard RGB imaging and can be deployed using existing camera infrastructure, making it more accessible and cost-effective compared to radio frequency-based approaches. Furthermore, compared to pose estimation-based approaches, which rely on computationally intensive key point detection and analysis which is often controlled by camera positioning, the YOLOv11-s model provided a simpler and more efficient alternative.

F. Model Output

After evaluation, YOLOv11-s was selected as the model with the best performances among the four models. This model was validated using office-based worker images. The validation results are presented in Figure 3, demonstrating one instance of good sitting posture detection (Figure 3a) and one instance of bad sitting posture detection (Figure 3b). The model successfully classified both sitting postures with bounding boxes and confidence scores.



3a 3b

Figure 3: Model output

5. CONCLUSION AND RECOMMENDATION

Although ergonomic research in the construction industry has focused on site-based construction workers, office-based construction workers such as engineers, project managers, quantity surveyors, architects, and administrative staff are also exposed to significant ergonomic risks due to prolonged sitting and extensive computer-based work. However, office-based construction workers have received comparatively limited focus in ergonomic research. To address this gap, and to overcome challenges associated with existing sitting posture monitoring approaches, such as the need for high computational resources and specialised hardware, this study proposed an alternative computer vision-based approach using YOLO object detection model to identify improper sitting postures.

Accordingly, this study developed four sitting posture detection models based on recent YOLO versions, enabling a comparative evaluation of their performance in detecting the sitting posture. The models were assessed using several key evaluation metrics, including true positive classification, precision, recall, mAP values, and inference speed. The results of the model evaluation indicated that YOLOv11-s demonstrated better performance in terms of accuracy, recall, and mAP, while YOLOv26-s achieved the highest precision and fastest inference speed. Considering the overall balance between detection accuracy and reliability, YOLOv11-s was identified as the most suitable model for sitting posture detection in construction office environments.

Finally, future research could explore the integration of real-time feedback mechanisms, such as automated sitting posture alerts or ergonomic guidance systems, to encourage behavioural improvements and promote healthier sitting habits among office-based construction workers. By incorporating such advancements, sitting posture monitoring technologies could contribute significantly to improving workplace ergonomics, worker wellbeing, and overall safety practices within the construction industry, particularly for office-based construction workers who are often overlooked in traditional safety management frameworks.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the technical support provided by Mr.A.A.D.W.V Udayashantha for his assistance in managing the software and hardware related to data analysis of this study.

REFERENCES

[1] A. Jirapongsuwan, P. Klainin-Yobas, W. Songkham, S. Somboon, N. Pumsopa, and P. Bhatarasakoon, "The effectiveness of ergonomic intervention for preventing work-related musculoskeletal disorders in agricultural workers: A systematic review protocol," *PLoS One*, vol. 18, no. 7, p. e0288131, Jul. 2023, Doi: 10.1371/journal.pone.0288131.

[2] B. Toko *et al.*, "The modern epidemic—digital era bad posture and its musculoskeletal consequences: A questionnaire-based study," *J. Family Med. Prim. Care*, vol. 14, no. 10, pp. 4187–4192, Oct. 2025, Doi: 10.4103/jfmpc.jfmpc_299_25.

[3] V. Markova, M. Markov, Z. Petrova, and S. Filkova, "Assessing the Impact of Prolonged Sitting and Poor Posture on Lower Back Pain: A Photogrammetric and Machine Learning Approach," *Computers*, vol. 13, no. 9, p. 231, Sep. 2024, doi: 10.3390/computers13090231.

[4] M.-C. Tsai, E. T.-H. Chu, and C.-R. Lee, "An Automated Sitting Posture Recognition System Utilizing Pressure Sensors," *Sensors*, vol. 23, no. 13, p. 5894, Jun. 2023, doi: 10.3390/s23135894.

[5] P. Vermander, A. Mancisidor, I. Cabanes, and N. Perez, "Intelligent systems for sitting posture monitoring and anomaly detection: an overview," *J. Neuroeng. Rehabil.*, vol. 21, no. 1, p. 28, Feb. 2024, doi: 10.1186/s12984-024-01322-z.

[6] G. Liu, X. Li, C. Xu, L. Ma, and H. Li, “FMCW Radar-Based Human Sitting Posture Detection,” *IEEE Access*, vol. 11, pp. 102746–102756, 2023, doi: 10.1109/ACCESS.2023.3312328.

[7] H. Jin, X. He, L. Wang, Y. Zhu, W. Jiang, and X. Zhou, “SitPose: Real-Time Detection of Sitting Posture and Sedentary Behavior Using Ensemble Learning With Depth Sensor,” *IEEE Sens. J.*, vol. 25, no. 7, pp. 12444–12454, Apr. 2025, doi: 10.1109/JSEN.2025.3541821.

[8] S. Sivanraj, D. N. L. S. Uduwage, and M. Tripathi, “Real-time safety detection on construction sites using a vision-language and NLP-based model,” *Advanced Engineering Informatics*, vol. 69, p. 103889, Jan. 2026, doi: 10.1016/j.aei.2025.103889.

[9] S. Sivanraj, D. N. L. S. Uduwage, R. K. Shiwakoti, and M. Thripathi, “YOLOv11-s-Based Pothole Detection Model with Integrated Traffic Data for Maintenance Decision Support,” *Journal of Advanced College of Engineering and Management*, vol. 11, pp. 77–92, Sep. 2025, doi: 10.3126/jacem.v11i1.84529.

[10] M. Orczyk *et al.*, “THE EFFECT OF SITTING POSTURE ON THE INCIDENCE AND SEVERITY OF LOW BACK PAIN IN ADULT POPULATION – SURVEY STUDY,” *Archiv Euromedica*, vol. 15, no. 1, Feb. 2025, doi: 10.35630/2025/15/1.109.

[11] R. K. R. Guduru *et al.*, “The Impact of Hunched Posture on the Elasticity of the Upper and Middle Trapezius Muscles During Sedentary Activity,” *Biotechnol. Appl. Biochem.*, vol. 73, no. 1, pp. 113–122, Feb. 2026, doi: 10.1002/bab.70008.

[12] D. F. Odesola, J. Kulon, S. Verghese, A. Partlow, and C. Gibson, “A Smart System for Continuous Sitting Posture Monitoring, Assessment, and Personalized Feedback,” *Sensors*, vol. 25, no. 18, p. 5610, Sep. 2025, doi: 10.3390/s25185610.

[13] D. F. Odesola, J. Kulon, S. Verghese, A. Partlow, and C. Gibson, “Smart Sensing Chairs for Sitting Posture Detection, Classification, and Monitoring: A Comprehensive Review,” *Sensors*, vol. 24, no. 9, p. 2940, May 2024, doi: 10.3390/s24092940.

[14] A. Kulikajevas, R. Maskeliunas, and R. Damaševičius, “Detection of sitting posture using hierarchical image composition and deep learning,” *PeerJ Comput. Sci.*, vol. 7, p. e442, Mar. 2021, doi: 10.7717/peerj-cs.442.

[15] Z. Cao *et al.*, “KeypointNet: An Efficient Deep Learning Model with Multi-View Recognition Capability for Sitting Posture Recognition,” *Electronics (Basel).*, vol. 14, no. 4, p. 718, Feb. 2025, doi: 10.3390/electronics14040718.

[16] Z. He, K. Wang, T. Fang, L. Su, R. Chen, and X. Fei, “Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment,” in *2025 37th Chinese Control and Decision Conference (CCDC)*, IEEE, May 2025, pp. 1281–1286. doi: 10.1109/CCDC65474.2025.11090973.

[17] F. Y. A’la, M. A. Safi’ie, and A. Supriyadi, “YOLOv8 vs. YOLOv9: Safety Helmet Detection Performance,” in *2024 7th International Conference of Computer and Informatics Engineering (IC2IE)*, IEEE, Sep. 2024, pp. 1–5. doi: 10.1109/IC2IE63342.2024.10748076.

[18] G. Jocher and J. Qiu, “Ultralytics YOLO11 (Version 11.0.0).” Accessed: Aug. 27, 2025. [Online]. Available: https://docs.ultralytics.com/tasks/pose/#how-do-i-validate-a-trained-yolo11-pose-model

[19] M. C. García, J. T. Mateo, P. L. Benítez, and J. G. Gutiérrez, “On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data,” *Remote Sens. (Basel).*, vol. 13, no. 1, pp. 1–23, Jan. 2021, doi: 10.3390/rs13010089.

[20] O. Caelen, “A Bayesian Interpretation of the Confusion Matrix,” *Ann. Math. Artif. Intell.*, vol. 81, pp. 429–450, 2017, doi: https://doi.org/10.1007/s10472-017-9564-8.

[21] P. Henderson and V. Ferrari, “End-to-End training of object class detectors for mean average precision,” 2017, pp. 198–213. doi: 10.1007/978-3-319-54193-8_13.