Automated Helmet Detection in Construction Sites using UWB and Machine Learning

Amir Shahbazi Ojghaz[1], Sayeh Bayat[1], and Farnaz Sadeghpour[1]

[1]Schulich School of Engineering, University of Calgary, Canada
amir.ojghaz@ucalgary.ca, sayeh.bayat@ucalgary.ca, farnaz@ucalgary.ca

Abstract

Ensuring the proper use of personal protective equipment (PPE), particularly helmets, is crucial for enhancing safety on construction sites. This study proposes a novel approach for detecting helmet usage and worker states using ultra-wideband (UWB) localization sensors combined with machine learning algorithms. Unlike traditional sensor-based or image-based systems, the proposed method integrates helmet detection into existing proximity warning systems, offering a cost-effective solution without the need for additional hardware. Data was collected from 12 participants in a controlled environment using UWB sensors, and a machine learning model was developed to classify four worker states: standing with a helmet, standing while holding a helmet, walking with a helmet, and walking while holding a helmet. The Gradient Boosting Decision Trees (GBDT) algorithm was selected for model development due to its superior performance. The model achieved an overall accuracy of 74.46% and performed well in detecting unsafe conditions. However, variability was observed in identifying safe states. Additionally, the study explored the impact of worker height on Z-axis localization error, revealing a correlation that suggests the need for height-adjusted safety monitoring systems. This research demonstrates that UWB-based systems can enhance PPE monitoring, reduce computational costs, and address privacy concerns, while also highlighting areas for future improvement, such as expanding the model to detect other PPE and refining its ability to differentiate between worker postures.

Keywords

Personal Protective Equipment (PPE); Construction Safety; Ultra-wideband Technology (UWB); Machine Learning.

1 Introduction

Construction sites are among the most hazardous environments for workers, with many sustaining injuries or losing their lives in workplace accidents every year. Numerous solutions have been proposed to enhance safety on construction sites, one of which is the use of personal protective equipment (PPE). The use of PPE can significantly reduce workplace risks and fatal accident [1, 2]. Previous studies have demonstrated that safety behaviors, including the consistent use of PPE, are strongly associated with a reduction in work-related injuries. Conversely, unsafe behaviors, such as the failure to wear PPE, are key risk factors in workplace accidents [3].

One of the challenges related to construction worker safety is the ineffective use of PPE. Factors such as discomfort, inadequate training, and poor communication between safety managers and workers can hinder the effective use of PPE. As a result, one of the main responsibilities of safety managers is to ensure that workers are wearing PPE on site. However, due to the dynamic and expansive nature of construction sites, manual supervision of PPE usage by safety managers is often inefficient. Consequently, efforts have been made to develop automated solutions for monitoring PPE usage by workers. Solutions include wearable devices, computer vision, and deep learning models [4-6].

These technologies enable real-time monitoring of PPE usage and can proactively warn workers of potential safety hazards, ensuring that they are properly wearing their safety gear, thereby reducing the likelihood of accidents. However, the adoption of these technologies is not without challenges [7-9]. For instance, the use of cameras on construction sites, in addition to the high costs associated with purchasing and processing equipment, raises concerns about worker privacy.

One technology used to improve worker safety is real-time localization through Ultra-wideband (UWB) sensors. This technology is typically employed in proximity warning systems to locate workers on construction sites and provide alerts when they approach hazardous areas. While these systems generally offer accurate two-dimensional localization of workers, three-dimensional localization, which involves the z-axis, often results in significant errors [10]. Consequently, using the z-axis for tasks that require high precision may not yield satisfactory results. However, if machine learning can be used to detect helmet usage based on the relatively imprecise z-axis data from these systems, it would be possible to extend their functionality without additional equipment.

Therefore, the objective of this research is to develop a model for detecting helmet usage by workers on construction sites using machine learning and UWB technology. By developing this model, it will be possible to monitor both the location of workers and their PPE status using current proximity warning systems without the need for additional equipment. This method not only preserves worker privacy but also reduces the financial and computational costs associated with system development.

2 Literature Review

Given the importance of safety in construction sites, numerous previous studies have proposed various solutions to ensure the appropriate use of PPE by workers. These solutions typically follow two main approaches: sensor-based methods and image processing techniques. Sensor-based studies often use proximity sensors, such as Bluetooth and RFID, to monitor the use of PPE [8, 11-13]. However, due to the limited applicability of these solutions, using such equipment solely to track PPE usage can increase initial costs. Another limitation of these methods is their inability to detect individuals outside of the predefined system. Moreover, the need for separate hardware to address different safety-related applications ultimately reduces the feasibility of implementing these solutions in construction sites.

On the other hand, image processing and deep learning models have gained significant attention in recent years. In this line of research, scholars typically collect images from construction site environments and process them using deep learning models. These models first detect target equipment, such as helmets and vests, and then assess whether workers are wearing the equipment by analyzing the relationship between the worker and the detected gear [4, 6, 9, 14-16]. Researchers have worked to incorporate equipment of various colors into their models and evaluate the performance of different models, such as YOLO and YOLACT-based approaches. However, this method also presents challenges.

One such challenge is the dependence of cameras on direct line of sight. Additionally, using cameras to monitor workers on construction sites raises concerns about worker privacy. Furthermore, many studies that employ this approach often integrate several deep learning models to achieve their goals, which significantly increases computational load. As a result, to achieve near-real-time performance, costly hardware is required to run the models. Considering the advantages and disadvantages of these approaches, there is a need to develop a model that retains the strengths of both methods while addressing their associated challenges.

3 Methodology

This research followed a three-step process to achieve its objective of developing a model for detecting helmet usage in construction sites using UWB sensors. These steps included data collection, machine learning model training, and performance evaluation. In the following sections, the scope of the research will first be outlined, followed by a detailed description of each step in the methodology.

3.1 Research Scope

Construction workers engage in various tasks on site, during which wearing a helmet is mandatory. Since this study aims to explore the use of UWB sensors and machine learning models for helmet usage detection, four possible worker states were considered: (1) standing while wearing a helmet (CL1), (2) standing while holding a helmet (CL2), (3) walking while wearing a helmet (CL3), and (4) walking while holding a helmet (CL4). These states represent typical conditions encountered in construction sites. Accordingly, the necessary data for developing a machine learning model to differentiate between these states were collected, followed by the model's development.

3.2 Data Collection

Data collection was conducted in a controlled laboratory environment. UWB Real-Time Location Systems (RTLS) generally consist of anchors and tags, where the anchors serve as reference points and the tags represent moving objects. In this study, the DWM1001C 6.5 GHz transceiver was used, which can function as both tags and anchors. Four anchors were placed at a height of 1.9 meters, covering a square area of 7 by 7 meters. Additionally, a tag was attached to a helmet, which was used for data collection in this study (Figure 1c).

To gather the data, participants were asked to perform each of the states outlined in the research scope for a duration of two minutes, during which their localization data, including x, y, and z coordinates, were recorded. The UWB modules were set to a frequency of 10 Hz, ensuring that data points were collected at 10 readings per second. Prior to data collection, participants' demographic information, such as height and gender, was recorded. A total of 12 participants took part in the experiment, resulting in the collection of 66,446 data records. Their collective demographic information is presented in Table 1.

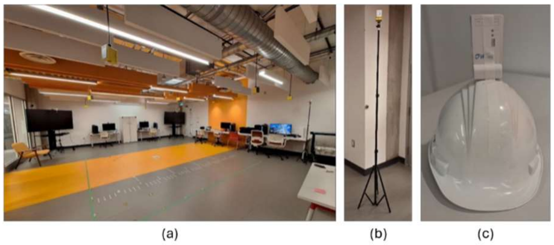

Figure 1. (a) Laboratory setup with UWB anchors. (b) UWB anchor at 1.9 m. (c) Helmet with UWB tag.

Table 1. Participant demographics: gender, count, height stats.

Gender	Count	Avg. Height (m)	Min. Height (m)	Max. Height (m)	Std. of Height
F	3	1.64	1.62	1.69	0.03
M	9	1.80	1.68	1.93	0.07
Total	12	1.76	1.62	1.93	0.09

3.3 Data Preprocessing

3.3.1 Data Cleaning

As previously mentioned, the aim of using machine learning algorithms in this study to detect helmet usage by workers is addressing the significant error in the z-axis measurements provided by UWB sensors. Therefore, outlier removal must be handled with extra care, as removing a substantial portion of the data could limit the model's generalizability. Additionally, given the wide range of data values due to variations in participant heights, removing outliers based on aggregated user data could introduce bias into the model. To avoid this, the Isolation Forest [17] algorithm was applied individually to each participant's data, with a contamination rate of 2% used to identify and remove outliers. The data preprocessing was implemented using Python 3.8.9 with Scikit-learn 1.3.2 and Pandas 2.0.3. After performing outlier detection and cleaning, the total number of data records decreased from 66,446 to 64,380.

3.3.2 Feature Engineering

Feature selection and extraction are critical steps in developing a machine learning model when working with location data collected in a laboratory environment. Since the laboratory space is limited, the data range in all three dimensions is also constrained, which can lead to overfitting. In such cases, the model may associate different states with the participants' specific locations in the lab, rather than identifying meaningful relationships between the data points. For instance, if participants remain in one specific area of the lab while performing a particular state, the model might incorrectly link that state to those specific coordinates. However, this would not hold true in real-world construction sites.

To mitigate overfitting, instead of using the raw x and y coordinates, the changes in these coordinates relative to previous steps were used as features for model training. The specific time intervals (10, 50, 100, and 150 steps) were selected based on temporal considerations. Since the UWB modules were set to a frequency of 10 Hz, each 10 steps correspond to approximately 1 second of data collection. The selected intervals therefore represent time windows of approximately 1, 5, 10, and 15 seconds, respectively. This range was chosen to balance the trade-off between model accuracy and real-time performance capabilities. Shorter intervals (e.g., 10 steps) capture immediate movements but may miss longer patterns, while longer intervals (e.g., 150 steps) capture more comprehensive movement patterns but increase the latency of detection in real-time applications. Table 2 lists the features selected for developing the machine learning model. It should be noted that, in addition to the rationale mentioned above, these features were also chosen based on their significance as determined through feature importance analysis.

3.4 Machine Learning Model Development

3.4.1 Algorithm Selection

The model training process involved training multiple models based on different machine learning algorithms, evaluating the performance of each, selecting the best-performing algorithm, and subsequently optimizing the chosen model and reporting its performance. In the initial selection of algorithms, a range of options was considered, from less complex algorithms such as SVM to more complex ones like deep learning algorithms. The rationale behind this approach was to identify an algorithm with lower computational cost and greater transparency. All models were implemented using Scikit-learn 1.3.2, with the Gradient Boosting Decision Trees (GBDT) implementation utilizing XGBoost 2.1.1. Additional libraries used in the analysis include NumPy 1.24.4 and Seaborn 0.13.2 for data manipulation and visualization. After evaluating the models developed in the first step, the GBDT algorithm was selected for further development and optimization due to its superior performance compared to other algorithms. A list of the evaluated algorithms and their performance is presented in Table 3.

Table 2. Selected features for ML model to mitigate overfitting

Feature	Description
z	Z coordinate values
z_norm	Z coordinate values normalized by the participant's height
dz_10	Change in Z coordinate compared to the value 10 steps earlier
dist_10	Distance change compared to the position 10 steps earlier (using X, Y, and Z)
dz_norm_10	Normalized Z changes compared to 10 steps earlier
dz_50	Change in Z coordinate compared to the value 50 steps earlier
dz_norm_50	Normalized Z change compared to 50 steps earlier
dist_50	Distance change compared to the position 50 steps earlier (using X, Y, and Z)
dz_100	Change in Z coordinate compared to the value 100 steps earlier
dz_norm_100	Normalized Z changes compared to 100 steps earlier
dist_100	Distance change compared to the position 100 steps earlier (using X, Y, and Z)
dz_150	Change in Z coordinate compared to the value 150 steps earlier
dz_norm_150	Normalized Z changes compared to 150 steps earlier
dist_150	Distance change compared to the position 150 steps earlier (using X, Y, and Z)

Table 3. Performance metrics of ML algorithms.

Model	AUC	F1	Prec.	Recall
Random Forest	0.945	0.712	0.721	0.724
SVM	0.763	0.398	0.415	0.395
Decision Tree	0.803	0.671	0.674	0.682
AdaBoost	0.773	0.626	0.629	0.635
Logistic Regression	0.928	0.585	0.674	0.615
kNN	0.843	0.489	0.49	0.491
Gradient Boosting	0.946	0.717	0.731	0.728
Naive Bayes	0.891	0.571	0.596	0.596

3.4.2 Model Training and Evaluation

The XGBoost library was used to implement the GBDT algorithm. One critical aspect to consider was the division of the available data for training, validation, and testing of the model's performance. In this study, instead of merging the data from all participants and splitting it based on a percentage of the total data, the data were divided based on the number of participants, without merging. Specifically, data from 6 out of the 12 participants were used for training the model, 3 participants' data were used for testing, and 3 participants' data were reserved for validation during training.

The model training process involved optimizing the model's parameters to achieve the best possible results. This was done using a Grid Search approach with 5-fold cross-validation to find the optimal hyperparameters. The parameter grid explored ranges for learning rate (0.05-1), max depth (3-9), number of estimators (50-150), subsample ratio (0.5-0.9), column sample by tree (0.5-1), and gamma (0-0.1). The XGBoost model was configured to utilize GPU acceleration via CUDA, and hyperparameter tuning was performed by selecting the parameter combination that achieved the highest accuracy. The training and optimization process was performed on an Nvidia RTX 4070 Ti GPU. Once the optimal parameters were obtained, the model was trained and tested 100 times to assess its performance. In each iteration, the participants were randomly assigned to training, testing, and validation sets. The results of the model evaluation will be discussed in the next section.

4 RESULTS

After repeating the training and evaluation process 100 times, the model's performance for each class was recorded in every iteration. The average performance metrics from these evaluations can be seen in Table 4 and Figure 2. The overall accuracy of the model, considering all classes, was 0.744. The developed model demonstrated satisfactory capability in distinguishing between safe and unsafe conditions. In Classes 2 and 4, where workers are not wearing helmets and are thus considered in an unsafe state, the model performed well. The average accuracy for Class 2 was 69%, with a precision of 74.2%, while for Class 4, the accuracy was 83%, and the precision was 86.7%. These results indicate that the model effectively identifies unsafe conditions.

However, the model's performance was more variable for Classes 1 and 3, where workers are wearing helmets and are considered to be in a safe state. Class 1 achieved an average accuracy of 66.6% and a precision of 66.4%, while Class 3 had a higher accuracy of 79.1%. Overall, the model was able to adequately differentiate between safe and unsafe conditions, but improvements in accurately detecting safe workers could help reduce false negatives and enhance the system's overall accuracy.

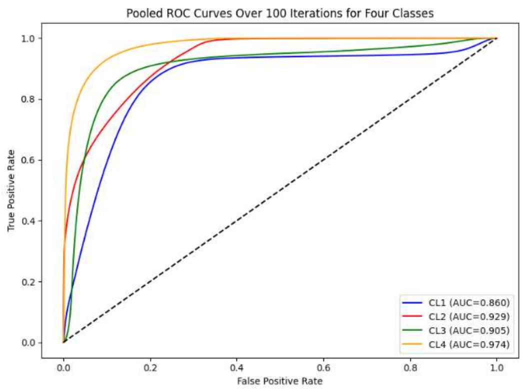

Figure 2. One-vs-Rest ROC curve showing AUC values across four classes.

Table 4. Performance metrics for each class.

Class	Acc.	Precision	Recall	F1	AUC
CL1	0.666	0.664	0.666	0.652	0.860
CL2	0.690	0.742	0.690	0.696	0.929
CL3	0.791	0.750	0.791	0.766	0.905
CL4	0.830	0.867	0.830	0.845	0.974

As the model demonstrated, it not only performed well in distinguishing safe conditions—where workers are wearing helmets—from unsafe ones, but it also showed the ability to differentiate between various worker states, such as standing or walking. Based on this capability, it was decided to assess the model's performance with the addition of another class. A new class was introduced, which involved the worker sitting on a chair while wearing a helmet. The entire training and testing process was repeated to evaluate the model's performance with this additional class. The results of this extended evaluation can be found in Table 5 and Figure 3.

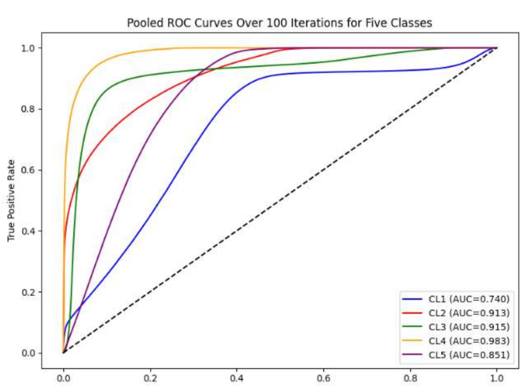

Figure 3. One-vs-Rest ROC curve showing AUC values across five classes.

Table 5. Performance metrics for five classes.

Class	Acc.	Precision	Recall	F1	AUC
CL1	0.318	0.415	0.318	0.341	0.740
CL2	0.673	0.733	0.673	0.684	0.913
CL3	0.817	0.792	0.817	0.801	0.915
CL4	0.852	0.888	0.852	0.868	0.984
CL5	0.601	0.474	0.601	0.521	0.851

5 DISCUSSION

5.1 Advantages and Limitations of the Proposed UWB and Machine Learning Approach

This study aimed to develop a model that not only detects whether workers are wearing helmets but also identifies various worker states using UWB RTLS, which are commonly used for tracking workers and issuing warnings when they approach hazardous areas or to prevent collisions. Machine learning algorithms were employed to achieve this goal.

The solution presented in this study offers both advantages and limitations compared to previous research. One of the key advantages of this approach over other sensor-based methods is that the UWB RTLS are not limited solely to PPE detection. Because this method can be integrated as an additional layer of safety monitoring within proximity warning systems, due to its software-based nature, it addresses the challenge of high initial costs typically associated with sensor-based methods. Moreover, the proposed model demonstrated that this approach can also be used to recognize different worker gestures, which could provide additional information to safety systems, facilitating their further development.

Additionally, the proposed model, being independent of cameras, avoids the disadvantages of image processing-based methods, such as privacy concerns, reliance on direct line of sight, and heavy computational loads. This independence allows for near-real-time performance in the dynamic environments of construction sites. However, it is important to note that this study focused solely on detecting helmet usage. In contrast, image processing methods can potentially detect the use of various other safety equipment, such as vests.

Moreover, the proposed model was tested in a controlled laboratory environment. However, previous studies have demonstrated that the accuracy of UWB RTLS can be impacted by various factors present in real construction sites, such as metal structures, equipment, and environmental conditions. Therefore, it is important to validate the developed model in construction environments in future studies. Future work should also extend this model to include the detection of additional PPE to further enhance its applicability.

5.2 Impact of Participant Height on Z-Axis Localization Error

As demonstrated in the results section, while the models presented in this study are capable of distinguishing between safe conditions—indicated by helmet usage—and unsafe conditions, this capability fulfills the primary research objective of providing a model for differentiating these states using UWB. However, to improve the model’s ability to differentiate between similar states, it is essential to identify the root cause of the existing error. One potential factor contributing to the increase in sensor error in the 3D space is the placement of the tags and anchors on the same plane (at the same height), which can reduce localization accuracy along the Z-axis.

To investigate this issue further, localization errors were evaluated based on the participants’ heights. As shown in Figure 4, as the participants' height increases, the absolute localization error along the Z-axis also increases. The correlation between participant height and absolute error is 0.75, and the statistical significance of this relationship, with a p-value of 0.0052, suggests that as participants’ height approaches the height at which the anchors are placed (1.9 meters), the error significantly increases. This finding is important because individuals with varying heights experience differing levels of error, and using the same system for individuals of different heights would result in varying safety levels. Therefore, it is necessary to adopt more flexible approaches, such as machine learning models, that can account for these variations.

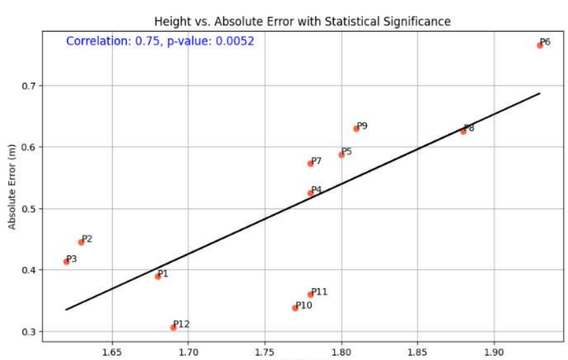

Figure 4. Correlation between participant height and Z-axis error.

6 Conclusion

This study successfully demonstrated the feasibility of using UWB sensors and machine learning algorithms to detect helmet usage and differentiate between various worker states on construction sites. The proposed approach offers significant advantages over existing sensor- and image-based methods by integrating PPE detection into existing proximity warning systems, thereby reducing costs and addressing privacy concerns. The model proved capable of accurately identifying unsafe conditions, though further improvements are necessary to enhance the detection of safe worker states, particularly in distinguishing between specific postures. Additionally, this study highlighted the impact of worker height on localization errors, suggesting that future systems should account for this factor to improve overall performance.

Future research should focus on several directions. First, expanding the model to detect other types of PPE, such as safety vests, would enhance its applicability in construction safety monitoring. Second, field validation in real construction site environments is needed to test the solution's reliability and performance under authentic conditions. Finally, as this study built its solution on top of existing proximity warning systems to decrease adoption costs, future studies can explore this approach to expand current safety systems and increase the practicality of proposed solutions for different applications in construction.

References

[1] Sawacha, E., S. Naoum, and D. Fong, *Factors affecting safety performance on construction sites.* International Journal of Project Management, 1999. **17**(5): p. 309-315.

[2] Li, J., et al., *Evaluating the impact of mental fatigue on construction equipment operators' ability to detect hazards using wearable eye-tracking technology.* Automation in Construction, 2019. **105**: p. 102835.

[3] Wong, T.K.M., S.S. Man, and A.H.S. Chan, *Critical factors for the use or non-use of personal protective equipment amongst construction workers.* Safety Science, 2020. **126**: p. 104663.

[4] Lee, Y.-R., et al., *Deep learning-based framework for monitoring wearing personal protective equipment on construction sites.* Journal of Computational Design and Engineering, 2023. **10**(2): p. 905-917.

[5] Han, K. and X. Zeng, *Deep Learning-Based Workers Safety Helmet Wearing Detection on Construction Sites Using Multi-Scale Features.* IEEE Access, 2022. **10**: p. 718-729.

[6] Ahmed, M.I.B., et al., *Personal Protective Equipment Detection: A Deep-Learning-Based Sustainable Approach.* Sustainability, 2023. **15**(18): p. 13990.
[7] Chen, H., et al., *The Impact of Wearable Devices on the Construction Safety of Building Workers: A Systematic Review.* Sustainability, 2023. **15**(14): p. 11165.
[8] Yang, X., et al., *Automated PPE-Tool pair check system for construction safety using smart IoT.* Journal of Building Engineering, 2020. **32**: p. 101721.
[9] Nath, N.D., A.H. Behzadan, and S.G. Paal, *Deep learning for site safety: Real-time detection of personal protective equipment.* Automation in Construction, 2020. **112**: p. 103085.
[10] Maalek, R. and F. Sadeghpour, *Accuracy assessment of Ultra-Wide Band technology in tracking static resources in indoor construction scenarios.* Automation in Construction, 2013. **30**: p. 170-183.
[11] *Automated PPE Misuse Identification and Assessment for Safety Performance Enhancement*, in *ICCREM 2015*. p. 204-214.
[12] Barro-Torres, S., et al., *Real-time personal protective equipment monitoring system.* Computer Communications, 2012. **36**(1): p. 42-50.
[13] Kelm, A., et al., *Mobile passive Radio Frequency Identification (RFID) portal for automated and rapid control of Personal Protective Equipment (PPE) on construction sites.* Automation in Construction, 2013. **36**: p. 38-52.
[14] Wang, Z., et al., *Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches.* Sensors, 2021. **21**(10): p. 3478.
[15] Lo, J.-H., L.-K. Lin, and C.-C. Hung, *Real-Time Personal Protective Equipment Compliance Detection Based on Deep Learning Algorithm.* Sustainability, 2023. **15**(1): p. 391.
[16] Chen, S. and K. Demachi, *Towards on-site hazards identification of improper use of personal protective equipment using deep learning-based geometric relationships and hierarchical scene graph.* Automation in Construction, 2021. **125**: p. 103619.
[17] Liu, F.T., K.M. Ting, and Z.H. Zhou. *Isolation Forest*. in *2008 Eighth IEEE International Conference on Data Mining*. 2008.