www.itcon.org - Journal of Information Technology in Construction - ISSN 1874-4753

INTEGRATING DOMAIN KNOWLEDGE WITH DEEP LEARNING MODEL FOR AUTOMATED WORKER ACTIVITY CLASSIFICATION IN MOBILE WORK ZONES

SUBMITTED: June 2023
REVISED: February 2024
PUBLISHED: April 2024
EDITOR: Esther Obonyo
DOI: 10.36680/j.itcon.2024.013

Chi Tian, Ph.D. Candidate
School of Construction Management Technology, Purdue University, West Lafayette, IN, 47907, United States
tian154@purdue.edu

Yunfeng Chen, Associate Professor
School of Construction Management Technology, Purdue University, West Lafayette, IN, 47907, United States
chen428@purdue.edu

Jiansong Zhang, Associate Professor
School of Construction Management Technology, Purdue University, West Lafayette, IN, 47907, United States
zhan3062@purdue.edu

Yiheng Feng, Assistant Professor
Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, 47907, United States
feng333@purdue.edu (*corresponding author)

SUMMARY: *Accurate classification of workers' activity is critical to ensure the safety and productivity of construction projects. Previous studies in this area are mostly focused on building construction environments. Worker activity identification and classification in mobile work zone operations is more challenging, due to more dynamic operating environments (e.g., more movements, weather, and light conditions) than building construction activities. In this study, we propose a deep learning (DL) based classification model to classify workers' activities in mobile work zones. Sensor locations are optimized for various mobile work zone operations, which helps to collect the training data more effectively and save cost. Furthermore, different from existing models, we innovatively integrate transportation and construction domain knowledge to improve classification accuracy. Three mobile work zone operations (trash pickup, crack sealing, and pothole patching) are investigated in this study. Results show that although using all sensors has the highest performance, utilizing two sensors at optimized locations achieves similar accuracy. After integrating the domain knowledge, the accuracy of the DL model is improved. The DL model trained using two sensors integrated with domain knowledge outperforms the DL model trained using three sensors without integrating domain knowledge.*

KEYWORDS: *Activity Classification, Mobile Work Zone, Wearable Sensors, Sensor Location Optimization, Domain Knowledge, Deep Learning.*

REFERENCE: *Chi Tian, Yunfeng Chen, Jiansong Zhang, Yiheng Feng (2024). Integrating Domain Knowledge with Deep Learning Model for Automated Worker Activity Classification in mobile work zone. Journal of Information Technology in Construction (ITcon), Vol. 29, pg. 264-280, DOI: 10.36680/j.itcon.2024.013*

1. INTRODUCTION

The construction industry is a critical component of the U.S. economy in terms of its contribution to Gross Domestic Product (GDP) and the number of job creation. According to data from the U.S. Bureau of Economic Analysis, the construction industry contributed \$832.8 billion to U.S. GDP in 2022, which accounts for 4.5 percent of the total GDP (U.S. Bureau of Economic Analysis, 2023). As of February 2023, 8 million workers are employed by the construction industry in the U.S. (U.S. Bureau of Labor Statistics, 2023). However, the construction industry faces two major challenges, which are high levels of hazards and low levels of productivity. Fatal injuries in construction sites account for 22.4% of all workers fatal injuries in the U.S. in 2019 (U.S. Bureau of Labor Statistics, 2019). Furthermore, the construction industry has lower productivity compared to other industries (Fulford & Standing, 2014). To prevent potential hazards and collect data for workers' productivity analysis, monitoring workers' activities is an essential step (Kim & Cho, 2021). Several emerging technologies such as wearable sensors (al Jassmi et al., 2021) and computer vision methods (Li & Li, 2022) have been applied to enable automated activity monitoring. These technologies are mostly found to be effective in the building construction sector (Bangaru et al., 2021; Ryu et al., 2019; Zhao & Obonyo, 2020). However, few studies have explored the automated construction activity classification in transportation construction operations, which has a more dynamic environment, for example, more movements than building construction sites.

In transportation construction operations, work zones, especially mobile work zone operations such as crack sealing, pothole patching, and trash pickup are critical to keeping the roadways in good conditions. Though numerous studies have been performed on building construction workers' activity classification, few studies have focused on workers' activity classification in transportation work zones (Kim et al., 2021; Tian, Chen, et al., 2022). Similar to building construction, accurate and automated worker activity classification is a necessary step to ensure that mobile work zones are being operated safely and productively. Therefore, the identification and monitoring of workers' activities in mobile work zones becomes a critical task, which is the focus of this study.

Data collection for worker activity classification can be roughly divided into two categories, kinematic-based methods (Mekruksavanich et al., 2022; Zhao & Obonyo, 2020) and vision-based methods (Luo et al., 2018; Torabi et al., 2022; Tian et al., 2024). The kinematic-based methods are considered as an effective technique to classify workers' activity in construction zones since it is not affected by occlusion and noisy environments. Inertial measurement unit (IMU) sensors are commonly used as wearable sensors to record workers' kinematic information during their work. The acceleration data can be collected by the IMU sensors from different locations of workers' body parts. The collected acceleration can be used as the input to develop an activity classification model (Akhavian & Behzadan, 2016). However, the accuracy of the activity classification depends on the locations and number of wearable sensors on workers' body parts. Kim & Cho (2020) found wearing 21 IMU sensors on workers achieved the highest classification accuracy. However, excessive number of sensors usually indicates a higher budget and more intrusiveness to the workers' normal operations. Therefore, optimizing the locations and reducing the number of wearable sensors can greatly reduce costs and improve data collection efficiency. On the other hand, vision-based methods are non-intrusive and do not require sensor installation on workers' bodies. However, the accuracy of vision-based methods is impacted by the occlusion, and background clutter (Bux et al., 2017; Khosrowpour et al., 2014). Due to the "moving" nature of mobile work zones, the implementation of vision-based methods for activity classification is more difficult and less accurate. As a result, in this study, we implement the kinematic-based method.

Deep-learning (DL) models, for example, artificial neural network (ANN) has been employed by previous studies to classify workers' activity based on the collected kinematic data (Bangaru et al., 2021; Yu et al., 2023). The ANN model's accuracy purely depends on the training data quality and optimization of the ANN structure (Bangaru et al., 2021). However, workers' activities and sequences are highly dependent on domain knowledge such as activity procedures, and human motion constraints, which are not considered in existing ANN models.

This study aims to address the four research gaps identified above, namely lack of research in worker activity classification in mobile work zones, lack of studies in optimizing sensor locations for different operations, and lack of interpretation of classification results generated from DL models. The objectives of this study are 1) developing a DL based framework to classify workers' activities in mobile work zones; 2) optimizing the sensor locations for different mobile work zone operations; 3) exploring the corner cases when the DL model fails; and 4) refining the classification accuracy by integrating transportation and construction domain knowledge. The results of this paper show the proposed framework classifies mobile work zone activities with a high accuracy.

The optimization of sensor location is a necessary step for each activity to save the budget and reduce intrusiveness. In addition, rules derived from the domain knowledge can improve the DL model classification accuracy especially when fewer sensors are used. To summarize, this study aims to answer the following four research questions.

- Question 1: Can the proposed DL-based framework be applied to workers' activity classification in mobile work zones?
- Question 2: Do sensors need to be placed at different locations on workers to collect data for different mobile work zone operations?
- Question 3: Under what situations that the DL-based framework cannot classify the activities accurately?
- Question 4: Can we improve the DL-based framework performance by integrating domain knowledge from transportation and construction?

The rest of the paper is organized as follows. First, a brief literature review on workers' activity classification using kinematic-based methods is provided in Section 2. Second, the proposed DL model-based framework is proposed to classify workers' activity in mobile work zones in Section 3. Third, the framework is validated by the data collected from three real-world mobile work zone operations. The results are discussed in Section 4 and the four research questions are answered in Sections 4.1, 4.2, 4.3, and 4.4 respectively. Lastly, the conclusion is provided.

2. LITERATURE REVIEW

Wearable sensors are commonly used sensors to collect workers' kinematic data. The wearable sensors are affixed to workers' bodies which are not impacted by occlusions. The occlusions are common problems for vision-based sensors. This literature review section first reviews how wearable sensors are applied in kinematic-based activity classification and different DL-based models applied in this topic. Then, the research gaps are identified based on the literature review results.

As discussed in the introduction section, the kinematic-based methods are ideal methods for construction workers' activity classification. The wearable sensors, such as wristbands and IMU are attached to workers' bodies directly for the data collection. The wristbands are worn on workers' wrists to collect acceleration from their wrists when they are performing various activities. One study used three-axis acceleration data from the wristbands to classify masonry workers' activity into four sub-activities (Ryu et al., 2019). The results reached 88.1% accuracy for the classification. In addition, some types of wristbands can also collect workers' physiological data, which are fused with kinematic data for activity classification. The implementation of physiological data for activity classification achieved 88% accuracy for pre-fabrication stone construction (al Jassmi et al., 2021). The classification accuracy can be further improved by fusing acceleration data and physiological data collected from one wristband, including heart rate, electrodermal activity (EDA), and skin temperature as shown in Tian, Chen, et al. (2022). Compared with wristbands, IMU sensors can be installed at multiple body locations. Previous studies selected various body locations for wearing IMU sensors and indicated that the locations of IMU sensors have significant impacts on classification accuracy (Kim & Cho, 2020). For example, the workers were asked to wear IMU sensors on their wrists and legs to classify nine construction activities, which achieved over 84% accuracy (Sonhood et al., 2021). In another study, IMU sensors equipped with an accelerometer, gyroscope, and magnetometer were used to collect scaffold builders' arm movement for the activity classification (Bangaru et al., 2021). For permanent patching activity in highway work zones, one study found that wearing IMU sensors on the hip and neck is sufficient to conduct the classification (Kim et al., 2021). Therefore, sensor location optimization is a necessary step for different work zone operations to improve classification accuracy and reduce data collection costs.

Machine learning models capture the hidden patterns from the training data set and can be used to make predictions on unseen test data (Shah et al., 2020). Therefore, current studies in workers' activity classification mostly utilized machine learning models, such as support vector machine (SVM) (Ryu et al., 2019), and K-Nearest Neighbors (KNN) algorithm (Sanhudo et al., 2021; Tian, Chen, et al., 2022). Recent studies also applied DL models for activity classification using different neural network structures such as a fully connected neural network (Bangaru et al., 2021) and Long Short-Term Memory (LSTM) neural networks (Kim et al., 2021; Mekruksavanich et al., 2022) for workers' activity classification. The neural network-based methods achieved 94% accuracy in workers' activity classification which is higher than using other machine learning methods such as SVM, random forest, etc. (Bangaru et al., 2021). The accuracy of machine learning based methods usually depends on the data quality and the DL model's structure and parameters. Integrating domain knowledge with the DL model can further improve the DL model's accuracy in prediction. For example, the domain knowledge (i.e., time-work relationship and equipment-work probability) is integrated with a tunneling workflow identification DL model (Wu et al., 2020). The prediction accuracy is improved from 71.6% by purely using DL model to 81.3% after combining the domain knowledge. Inspired by such studies, this paper aims to improve DL model's accuracy by integrating transportation and construction knowledge workers' activity classification.

3. METHODOLOGY

The proposed framework for automated mobile work zone activity classification using DL-based model is depicted in Figure 1. The framework starts with the acceleration data collection from various body locations of workers by using wearable sensors. Three mobile work zone operations are chosen in this study, namely trash pickup, crack sealing, and pothole patching. These operations are selected because they are all indispensable mobile work zone operations to ensure the highway system functionality (Tian, Xiao, et al., 2022). Furthermore, these three mobile work zone operations are among the top operations that consume the most man-hours in the Indiana Department of Transportation (INDOT). The raw data are pre-processed and serve as input to the DL models for activity classification. Next, the sensor locations are optimized through feature importance analysis. Lastly, a domain knowledge-based refinement process is conducted to further improve the classification results generated by the DL model. Each step is described in detail below.

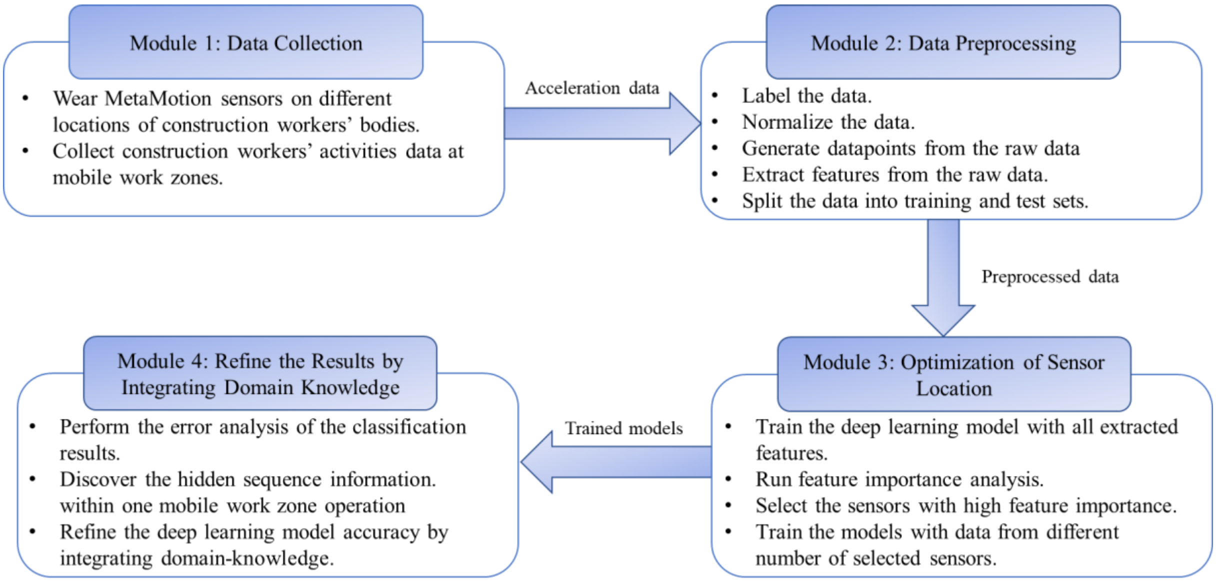


Figure 1. Framework for Construction Workers' Activity Classification using the Acceleration Data.

3.1 Data Collection

The MetaMotion sensor, developed by MBIENTLAB was utilized to collect workers' acceleration data in this study. Workers were asked to wear four MetaMotion sensors (shown in Figure 2 (a)) on their lower arm, ankle, upper spine, and lower spine during the mobile work zone operations as indicated in Figure 2 (b). The initial four sensor locations were selected based on the suggestions from previous studies that the sensor shall be put at a certain distance to capture the movement of different body parts (Kim & Cho, 2020). In addition, the sensors on the upper spine and lower spine were found effective for the infrastructure work zone activity classification (Kim & Cho, 2021). Each sensor collected three-axis acceleration data, which are referred to as X, Y, and Z axes. The MetaMotion sensors were equipped with 512 MB flash memory to store the collected data, which is sufficient for one-day operation. Due to the low weight of MetaMotion sensors, it collects acceleration data in a non-intrusive manner (Zhao et al., 2021), which are applied by previous studies to capture people's movement information (Pribadi & Shinoda, 2022). In addition, the MetaMotion sensors can be connected to mobile phones via Bluetooth. The MetaBase application configures the data collection frequency, starts the sensors for data collection, and exports the raw data in .csv format. In total, 279 seconds, 1,459 seconds, and 1,703 seconds of trash pickup, crack sealing and pothole patching are collected. The frequency to collect the acceleration is configured as 100 Hz. After each day's data collection, the data is exported in .csv format. Videos are also recorded during mobile work zone operations for ground truth label generation.

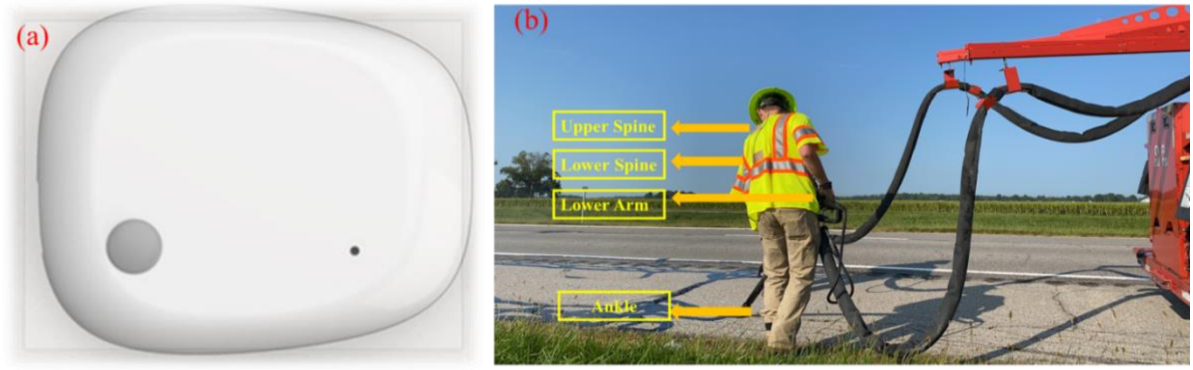


Figure. 2. (a) Data Collection Device - MetaMotion Sensor; (b) Sensor Installation Locations.

3.2 Data Preprocessing

The data preprocessing includes four steps: data labelling, data normalization, feature extraction, and training/test data splitting. The data is labelled manually by reviewing the recorded videos for each operation, which contains several activities. The mean, standard deviation (std.), maximum, and minimum values of acceleration data in each axis of each sensor are considered as the features to train the ANN model. These features were used in previous studies for the construction workers' activity classification (Ryu, 2019). Each sensor collects three-axis acceleration, which results in 48 features (four features × four sensors × three axes of each sensor) in each data point. Then, the data is normalized to ensure that the fully connected artificial neural network (ANN) model can be trained effectively (Tan et al., 2020). Each feature is normalized by subtracting the mean of the feature and then dividing it by its standard deviation. Next, data points are generated using 0.2 second window size and 0.1 second sliding window from the normalized data set. In the end, 2,784, 14,588, and 17,018 data points are extracted for trash pickup, crack sealing, and pothole patching operations respectively. Finally, the data points are divided into training data, validation data, and testing data using a stratified random split with a ratio of 55%, 15%, and 30% respectively.

3.3 Sensor Location Optimization

The ANN model is composed of an input layer, multiple hidden layers, and an output layer, with several neurons in each layer connected by different weights. The output of the ANN is calculated using equation (1).

$$h_{\mathrm{W},b}(X) = \phi(X\mathrm{W} + b) \tag{1}$$

where X denotes the input matrix, W denotes the weight matrix for all connections, b denotes the bias vector for all connections, and the activation function is denoted by ϕ. Equation (2) updates the connection weight between the i^{th} and j^{th} nodes at each training step,

$$w_{i,j}{}^{(next\ step)} = w_{i,j} + \eta\,(y_j - \hat{y}_j)x_i \tag{2}$$

Where $w_{i,j}$ is the connection weight between the i^{th} and j^{th} neurons, η is the learning rate, y_j is the output of the j^{th} neuron, and $\hat{y}$ is the target output of the j^{th} neuron, xi is the input of the i^{th} neuron. The model's accuracy is calculated using equation (3).

$$Accuracy = \frac{Number\ of\ correct\ predications}{Total\ number\ of\ predications} \tag{3}$$

Tuning hyperparameters of the ANN models, such as the number of layers, number of nodes in each layer, and learning rate can significantly improve the accuracy of the ANN models (Agatonovic-Kustrin & Beresford, 2000). Bayesian optimization is applied to tune the hyperparameters of each ANN model (Victoria & Maragatham, 2021). Keras Tuners, a Python hyperparameter tuning package is used to tune the parameters. The objective function was to find the maximum validation accuracy. The hyperparameter tuning process terminates after 50 epochs of search.

The hyperparameters are tuned iteratively by proposing a new combination of hyperparameters based on the best hypermeters found so far (Klein et al., 2017). Table 1 displays the hyperparameters and their ranges in this study. The range of the number of hidden layers is between 2 and 7. The number of nodes can take the values such as 64, 128, 256, 512, and 1024. The learning rate is sampled between 0.01 and 0.0001 using logarithmic sampling. Additionally, the activation functions are tuned between Relu and Tanh. In addition, since the dataset is imbalanced, different weights are added for each class during the training. The weights are added as the inversely proportional to class frequencies in the dataset. During the training, more penalization is added for the classes with higher weights, which addresses the issues of imbalanced datasets.

Table 1. Hyperparameters and their Ranges used for Bayesian Optimization of the ANN Models.

Hyperparameters	Values
Number of hidden layers	2 to 7
Number of nodes	64, 128, 256, 512, 1024
Learning rate	0.01 to 0.0001
Activation functions	Relu, Tanh

Next, to optimize the sensor locations for each mobile work zone operation, the feature importance is calculated for the 48 features extracted from each sensor. The feature importance measures the contribution of the features in determining the output label in machine-learning models (Megantara & Ahmad, 2020). The higher value of feature importance indicates the feature has a higher impact on the model to determine the classification results. The recursive feature elimination (RFE) method is applied as an iterative process to select the most important features. The RFE trains the model by using all features at the beginning (Yan & Zhang, 2015). Then, one feature with the least importance is removed at each iteration until the desired number of features is selected. The feature importance is calculated after each iteration. The feature importance for each sensor is calculated by summing up the feature importance scores of all features collected from that sensor. The RFE calculates the feature importance using various methods, such as random forest, support vector machine, and decision trees (Sharma & Yadav, 2021). In this study, the decision trees are used as the estimator, and the feature importance is calculated based on how much the feature can reduce the Gini impurity. A similar RFE feature selection method is applied by previous studies to improve the accuracy of deep learning models (Ustebay, et al. 2018). The feature importance is calculated before training the ANN model to remove unimportant features in the training process.

3.4 Classification Results Refinement

This section discusses how to refine the DL model classification results by integrating the domain knowledge. First, the error analysis is performed, and two error patterns are found. The first pattern is that a large proportion of classification errors occur during the translation from one activity to the other activity, due to rapid changes in the acceleration data patterns. To address this issue, sequential information of mobile work zone operation, which defines the order of activities for one operation, is applied to eliminate the error. The second pattern is that the prediction results fluctuate between similar activities (e.g., walking and walking with a shovel). The fluctuation errors are removed by considering the activities in the previous and next time steps. For example, if the activities in previous time step (t-1) and next time step (t+1) are the same while the activity in time step (t) is classified differently, the activity in time step (t) is corrected to the same label as time step (t-1) and time step (t+1). A concrete example is discussed in the next section.

4. RESULTS AND DISCUSSIONS

To validate the proposed classification framework, three mobile work zone operations including trash pickup, crack sealing, and pothole patching are tested. For this study, two workers participated in trash pickup, three workers participated in crack sealing, and two workers participated in pothole patching. All participants are INDOT workers and experienced in these tasks. The data collection was conducted in September and October 2022 during on-road mobile work zone operations. The mobile work zone operations contain one leader truck and one follower truck, which carry INDOT workers, tools, and materials required for the specific operation. The follower truck is attached to a truck-mounted attenuator (TMA) to protect workers' safety. The leader and follower trucks layout for crack sealing operations is shown in Figure 3 (a). Each truck is installed with front-facing and rear-facing cameras to record videos for ground truth labelling during the mobile work zone operations as indicated in Figure 3 (b). When the trucks travel on the roadway, the workers are vigilant to search for trash, cracks, and potholes and perform the tasks accordingly.

Figure 3. (a) Mobile Work Zone Layout; (b) Sensor Installation on the Trucks.

Figures 4-6 show the activities identified for each mobile work zone operation. The trash pickup operation consists of three activities, which are “walking”, “picking up trash”, and “loading trash”. The crack sealing operation includes four activities, which are “crack cleaning”, “crack sealing”, “standing”, and “walking”. The pothole patching operation has six activities, which are “walking”, “shoveling”, “walking with shoveling”, “placing mix”, “patching”, and “standing”.

Figure 4. Activities of Trash Pickup Operation: (a) Walking, (b) Picking up Trash, and (c) Loading Trash.

Figure 5. Activities of Crack Sealing Operation: (a) Crack Cleaning, (b) Crack Sealing, (c) Standing, and (d) Walking.

Figure 6. Activities of Pothole Patching Operation: (a) Walking, (b) Shoveling, (c) Walking with Shoveling, (d) Placing Mix, (e) Patching, and (f) Standing.

4.1 Training ANN Models to Classify Activities for Each Mobile Work Zone Operation

Three ANN models are trained and the hyperparameters are tuned for each mobile work zone operation. All 48 features from four sensors are used to train the basic ANN models. The hyperparameters tuning results for each ANN model are shown in Table 2. The results show that the ANN model reached 87.64% overall accuracy for trash pickup operation, 98.06% accuracy for crack sealing operation, and 93.73% for porthole patching operation respectively. Figs. 7-9 illustrate the confusion matrix for each operation. The horizontal axis represents the predicted label, and the vertical axis represents the true label. Each cell in the unnormalized confusion matrix represents the number of data points. For example, 184 cases of “loading trash” is classified as “loading trash”, 11 cases of “loading trash” is classified as “picking up trash”, and 10 cases of “loading trash” is classified as “walking”. The cell in the normalized confusion matrix represents the percentage of each class. For example, 184 / (184+11+10) = 89.76% percent of “loading trash” cases are classified as “loading trash”. Similarly, 5.37% percent of “loading trash” cases are classified as “picking up trash”, and 4.88% percent of “loading trash” cases are classified as “walking”. The results show the accuracy for loading trash, picking up trash and walking is 83.90%, 76.62%, and 89.50% respectively for the trash pickup operation. The accuracy for crack cleaning, standing, walking, and crack cleaning is 97.71%, 99.33%, 88.62%, and 96.27% respectively. For the pothole patching operation, the accuracy for the six activities classification is 81.20% for patching, 95.16% for placing mix, 96.07% for shoveling, 95.31% for standing, 93.04% for walking, and 92.26% for walking with shovels. In addition, the three models’ performances are used as the baseline to evaluate the model’s performance with sensor location optimization.

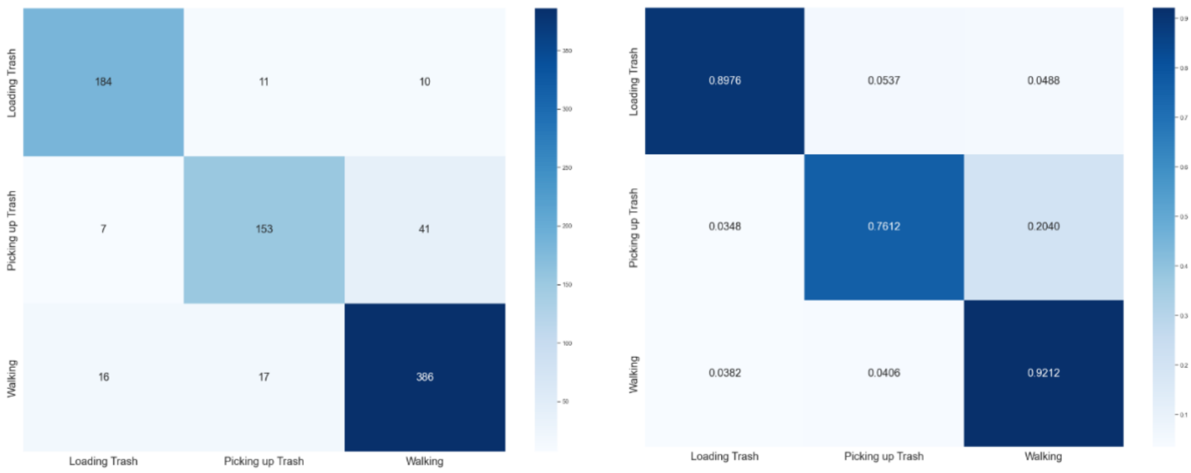
Table 2. Hyperparameters Tuning Results and Model Accuracy for the ANN Model for Each Mobile Work Zone Operation

Hyperparameters	ANN Model for Trash Pickup	ANN Model for Crack Sealing	ANN Model for Pothole Patching
Number of hidden layers	7	4	3
Number of nodes for each layer	1,024,1,024,1,024,1,024, 1024,1,024,1,024	1,024,1,024,1,024,1,024	1,024,1,024,1,024,
Learning rate	0.0001	0.0001	0.0001
Activation functions	Relu	Relu	Relu
Accuracy	87.64%	98.06%	93.73%

4.2 Sensor Location Optimization

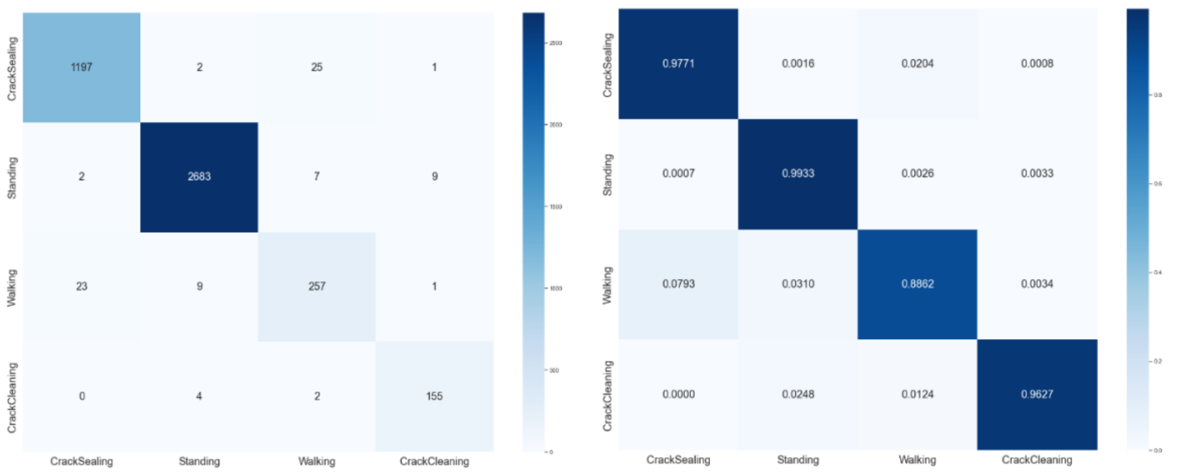
The objective of sensor location optimization aims to find the importance of different IMU sensors in various mobile work zone operations. The feature importance for all the features contained in one sensor is calculated as the indicator using the RFE method. The feature importance by sensor location for each operation is shown in Table 3. The higher value of feature importance means the sensor is more important in determining the classification results. Then, three additional ANN models are trained for each mobile work zone operation using top 1, 2 and 3 sensors of feature importance respectively. The results from the three additional models and the original model are compared. The findings are summarized and discussed below.

First, the sensor’s feature importance ranking varies in different operations. For the trash pickup operation, the lower spine sensor has the highest feature importance, then followed by the ankle, lower arm, and upper spine. However, for crack sealing, the sensor on the upper spine has the highest feature importance which is 0.6654 and then followed by 0.2111 for the lower arm. The ankle and lower spine only accounted for 0.0743 and 0.0493 feature importance respectively. For the pothole patching, the sensor on the upper spine has the highest feature importance, which is 0.4776, then followed by 0.2433 for the lower arm, 0.1542 for the ankle, and 0.1249 for the lower spine. The results also show adding sensors in the location where most movements happen can improve the classification results most significantly. For example, “walking” accounts for 50.79% of the total number of activities in trash pickup, while “walking” only accounts for 6.63% and 29.85% for crack sealing and pothole patching respectively. This is consistent with the observations from the fieldwork since the workers need to walk around to pick up the trash distributed in different locations, while the workers spend most of their time “standing” when performing the crack sealing operation. Therefore, the sensor on the ankle has a higher importance in trash pickup operation classification than crack sealing and pothole patching operations. In addition, activities of crack



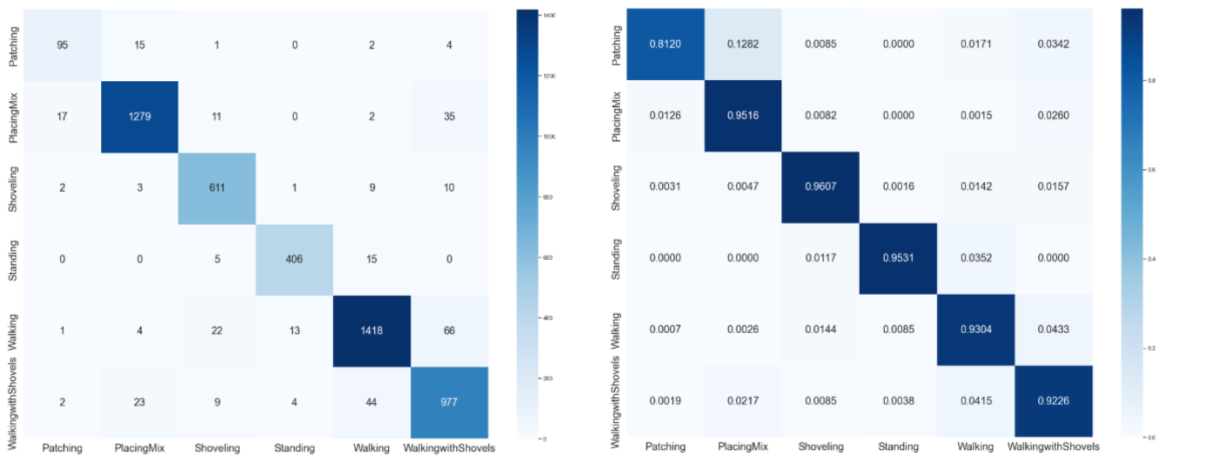
(a) *(b)*

Figure 7. Confusion Matrix for Trash Pickup Operation Classification: (a) Unnormalized and (b) Normalized



(a) *(b)*

Figure 8. Confusion Matrix for Crack Sealing Operation Classification: (a) Unnormalized and (b) Normalized



(a) *(b)*

Figure 9. Confusion Matrix for Pothole Patching Operation Classification: (a) Unnormalized and (b) Normalized

sealing and pothole patching operations involve more upper body movements such as "crack sealing", "crack cleaning", "shoveling", "placing mix", and "patching". Therefore, the lower arm sensor is more important in crack sealing and pothole patching operations.

Table 3. Feature Importance by Sensor Location for Each Mobile Work Zone Operation.

Feature Importance Rank	Trash Pickup		Crack Sealing		Pothole Patching	
	Sensor Location	Feature Importance Value	Sensor Location	Feature Importance Value	Sensor Location	Feature Importance Value
1	Lower spine	0.3280	Upper spine	0.6654	Upper spine	0.4776
2	Ankle	0.3083	Lower Arm	0.2111	Lower Arm	0.2433
3	Lower arm	0.1864	Ankle	0.0743	Ankle	0.1542
4	Upper spine	0.1773	Lower spine	0.0493	Lower spine	0.1249

Second, training the ANN models with the data from sensors with the top two feature importance can achieve similar accuracy as models trained by all four sensors. The results from Tables 4 to 6 indicate the ANN models achieve the highest accuracy for all three operations when trained by the data collected by all four sensors. However, the accuracy difference between the models trained by the four sensors and the sensors with top two feature importance is not significant. For the trash pickup operation, the accuracy of ANN model trained by data collected from the lower spine and ankle sensors is 83.64%, which is an 11.88% improvement over only using one sensor with the highest feature importance. However, adding the third sensor (i.e., the sensor on lower arm) only improves the model accuracy further by 1.33%. The conclusions from crack sealing and pothole patching operations are consistent with the findings from the trash pickup operation. Therefore, if the sensors are carefully optimized, using two sensors is sufficient to achieve satisfactory results. Utilizing fewer sensors reduces the budget, particularly when taking into consideration the large workforce employed for infrastructure maintenance work. For example, INDOT alone employs over 1,000 employees as the maintenance team (INDOT Division of Maintenance, 2022). In addition, using fewer sensors reduces the intrusiveness of the workers' normal operations.

Table 4. ANN Models' Performance for Trash Pickup Operation.

Model #	Sensors	Accuracy			
		Overall	Loading Tash	Picking up trash	Walking
Baseline Model	All four sensors	87.64%	89.76%	76.12%	92.12%
Model 1	Lower spine	71.76%	63.90%	61.69%	80.43%
Model 2	Lower spine and ankle	83.64%	83.90%	68.16%	90.93%
Model 3	Lower spine, ankle, and lower arm	84.97%	83.90%	76.62%	89.50%

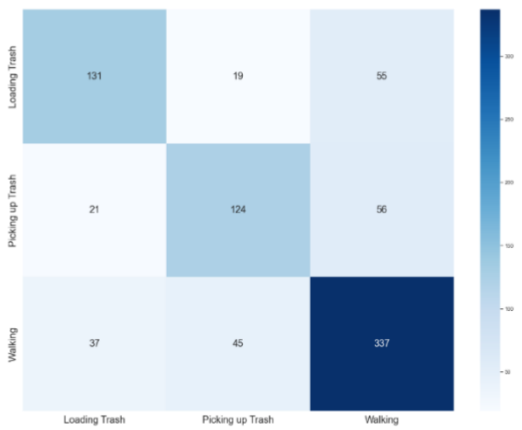
Table 5. ANN Models' Performance for Crack Sealing Operation.

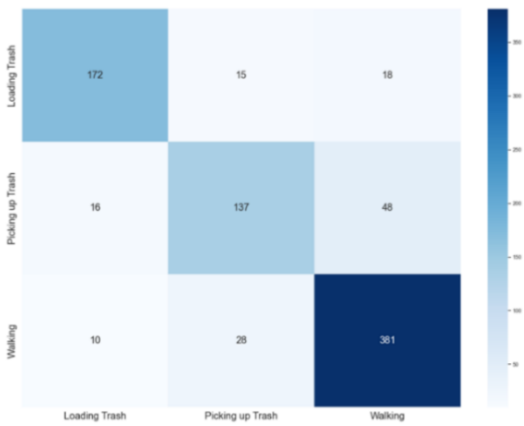
Model #	Sensors	Accuracy				
		Overall	Crack cleaning	Standing	Walking	Crack sealing
Baseline Model	All four sensors	98.06%	97.71%	99.33%	88.62%	96.27%
Model 1	Upper spine	90.36%	91.51%	93.15%	64.14%	81.99%
Model 2	Upper spine and lower arm	97.41%	99.18%	99.15%	77.93%	90.06%
Model 3	Upper spine, lower arm, and ankle	97.76%	98.78%	98.82%	85.52%	94.41%

Table 6. ANN Models' Performance for Pothole Patching Operation.

Model #	Sensors	Accuracy						
		Overall	Patching	Placing mix	Shoveling	Standing	Walking	Walking with Shovels
Baseline Model	All four sensors	93.73%	81.20%	95.16%	96.07%	95.31%	93.04%	92.26%
Model 1	Upper spine	79.49%	75.21%	85.42%	80.66%	90.85%	81.04%	64.97%
Model 2	Upper spine and lower arm	90.80%	71.79%	94.27%	91.04%	94.37%	90.42%	87.44%
Model 3	Upper spine, lower arm, and ankle	91.85%	69.23%	92.78%	93.08%	94.37%	91.80%	91.50%

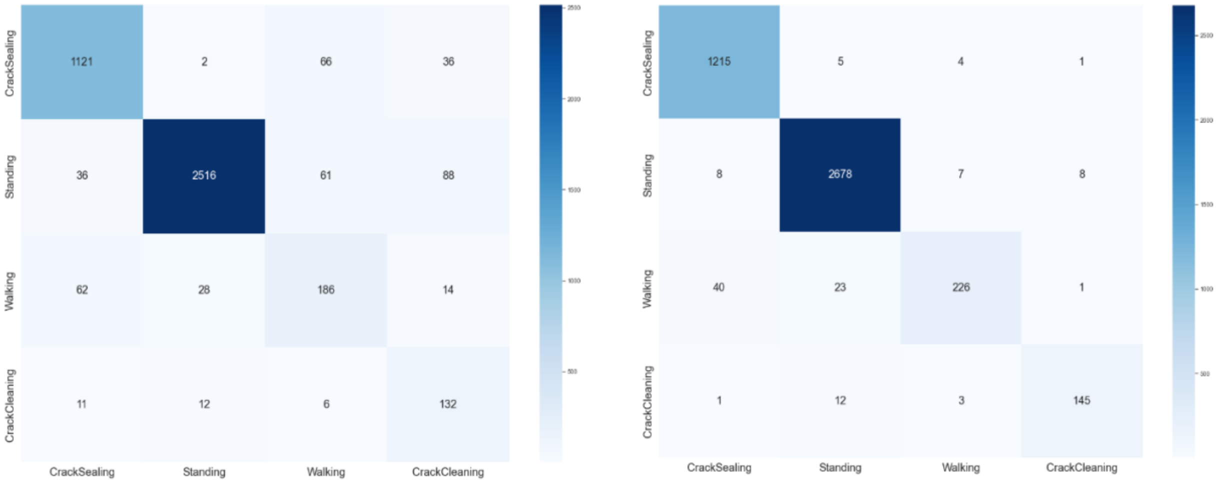
Third, the study finds that analyzing the movement patterns of an operation can guide the location of sensor installation, in which adding a sensor at a specific location can significantly enhance the accuracy of activity classification. For example, Figure 10 shows the confusion matrix for the trash pickup operation trained by (a) data collected by the lower spine sensor and (b) the data collected by both lower spine and ankle sensors. Figure 10 (a) indicates the ANN model classified "loading trash" as "walking" 55 times and "walking" as "loading trash" 37 times. By adding the data collected by the ankle sensor, more walking-related acceleration information is fed into the ANN model. The errors where "loading trash" was classified as "walking" decrease from 55 to 18 times. The errors where "walking" was classified as "loading trash" reduce from 37 to 10 times. For crack sealing, adding the second lower arm sensor in addition to the first upper spine sensor decreases the errors such as in classifying "walking" as "crack sealing" from 62 to 40 times, and in classifying "crack sealing" as "walking" from 66 to 4 times as shown in Figure 11. The reason is workers are still walking slowly in the "crack sealing" activity to move from one location to another location to continue filling the asphalt into the cracks. The slow walking characteristics of the "crack sealing" activity led to the similarity between data for classifying the "crack sealing" and "walking" activities when data is only collected from the ankle sensor. However, the workers' hands need to hold the crack sealing hose during the "crack sealing" activity, while their hands can move freely during the "walking" activity. Therefore, adding data collected in the lower arm sensor could help the ANN model to distinguish walking and crack sealing better. A similar analysis can be performed for the pothole patching operation. The six activities in pothole patching mainly involve arm movement (i.e., patching, shoveling, walking with shovels, placing mix) and spine movement (i.e., patching, placing mix). As a result, the combination of upper





(a) *(b)*

Figure 10. Confusion Matrix for Trash Pickup Operation Classification: (a) Trained by Data Collected by the Lower Spine, and (b) trained by Data Collected by the Lower Spine and Ankle.



(a) *(b)*

Figure 11. Confusion Matrix for Crack Sealing Operation Classification: (a) Trained by Data Collected by the Upper Spine, and (b) Trained by Data Collected by the Upper Spine and Lower Arm.

spine and lower arm sensors achieves the best results (90.80%) as shown in Table 6. However, if two sensors are randomly selected (e.g., ankle and lower spine), using the same training and hyperparameter tuning method, the model only achieved 84.06% accuracy.

The findings can be utilized to design the sensor layouts solely by observing the mobile work zone operations. This is beneficial especially there exists a large number of operations that need to be classified (e.g., the INDOT work performance standards document outlines over 80 types of work zone operations (INDOT Division of maintenance, 2022). It is very time-consuming to build and test DL models for each operation.

4.3 Error Analysis

The error analysis was performed to find the patterns of incorrect classification cases and propose potential solutions. Two major patterns are found as described below.

First, most incorrect classifications happen when workers are transitioning from one activity to another activity, which makes the ANN model make the incorrect classification due to variations of acceleration data. For example, in the trash pickup operation, the error in classifying "walking" as "picking up trash" mostly happens when the workers just start to walk. Figure 12 shows two of these typical error cases. Similarly, most errors in classifying loading trash" as "walking" happen when the workers just start to load trash. They are walking close to the truck before "loading trash". Figure 13 shows some examples of this error.

Figure 12. Error Cases – "walking" Classified as "picking up trash".

Second, in the pothole patching operation, when exploring error cases between "walking" and "walking with shovels", it is found that most errors happen for only one time step. For example, the model classifies the activity at t time step as "walking", while classifying the activity at t-1 and t+1 time steps as "walking with shovels" and vice versa. In the real pothole patching operation, these cases may never happen. It is impossible at one time step, the workers are walking with shovels, but the previous and following time step they threw the shovels away. Similarly, the worker does not only carry the shovel for a single time step.

Figure 13. Error Cases – "loading trash" Classified as "walking".

Enlightened by these observations, we discover that domain knowledge (such as engineering procedures and human behaviors) can be integrated into the DL model to further improve the classification results in real mobile work zone operations. Next, we demonstrate this idea using the pothole patching operation as an example since it has the most complicated procedure and the greatest number of activities.

4.4 Model Improvement with Domain Knowledge

Based on observations from the error analysis, two sets of rules are developed to refine the DL model classification results for the pothole patching operation. The first set of rules forbids the DL model from making predictions for impossible transition activities. Based on our observation from field pothole patching, the workers need to perform the activities in the order of "walking", "shoveling", "walking with shovels", "placing mix", and "patching" for each cycle of pothole patching operation. The "standing" is only needed when workers need to wait for vehicles to proceed first before they can reach the pothole locations. Therefore, based on the operational cycle, the following four rules are applied to inform the DL model from making predictions of impossible transition activities.

1. It is impossible that the "placing mix" activity follows right after the "patching" activity. The reason is the workers need to "place the mix" first and then "patch" the pavement.
2. It is impossible that "patching" or "placing mix" activity follows right after the "shoveling" activity. The reason is that workers has to "walk with shovels" to get close to the pothole before "placing the mix", and then "patching".
3. It is impossible that "patching" or "placing mix" activity follows the "walking" activity. The reason is workers still need to carry the shovels to perform the placing mix and patching. The "walking" activity happens when workers need to walk from the location where the vehicle stops to the hot box containing the asphalt.
4. It is impossible that the "patching" activity follows the "walking with shovels" activity. The reason is that "placing mix" must be performed before the "patching".

The first set of four rules is applied to correct the DL model's original prediction results. If consecutive prediction results in two time steps (e.g., t-1 and t) fall into the abovementioned rules, the probability of prediction at time t is manually set as 0 (i.e., impossible prediction). The prediction result is selected as the one that has the highest probability among the remaining classes. For example, if at time of t-1, the model outputs "patching", the probability of classifying the activity as "placing mix" at time t is set as 0. The DL model outputs the class with the highest probability among the remaining four classes which are "patching", "walking with shovels", "walking", and "standing".

The second set of rules aims to inform the DL model that it is impossible a second activity to be classified between the same activity for a very short time (e.g., one time step), which violates the human physical movement limits, such as the "walking" and "walking with shovel" example illustrated above. The second set of rules is applied by taking the previous (t-1) and following (t+1) time step prediction results into consideration. For example, if the activity is classified as “walking” at time t, but predictions are “walking with shovel” at both t-1 and t+1. The prediction result is changed to “walking with shovel”.

The two sets of rules are added to the four ANN models for pothole patching classification, which are baseline model trained by the data from all four sensors, model 1 trained by the data only from the upper spine sensor, model 2 trained by the data from upper spine and lower arm sensors, model 3 trained by the data from upper spine, lower arm, and ankle sensors. Table 7 summarizes the accuracy of each model before and after adding the rules. The results indicate that, for the baseline model, the accuracy is improved from 93.73% to 93.91% by adding the first set of rules, and the accuracy is further improved to 94.91% after adding the second set of rules. The improvement is not significant since the model trained by four sensors already achieves high accuracy in distinguishing “walking” and “walking with shovels”. However, for Model 1, since only one sensor is installed and the least number of features are used for model training, adding rules can significantly improve the classification results. For example, significant improvement is observed in classifying “walking” and “walking with shovels”. After applying both sets of rules, the errors classifying “walking” as “walking with shovels” reduce by 65 cases, and errors classifying “walking with shovels” as “walking” reduce by 42 cases (Figure 14). Interestingly, for model 2, adding two sets of rules improved the model accuracy from 90.80% to 92.13% which is even higher than adding the third sensor, which is 91.85%. Comparing the accuracy improvement from different models, we find that 1) the rules developed by domain knowledge and human behaviors have more significant impacts when the number of sensors is fewer; and 2) integrating domain knowledge-based rules may outperform adding more sensors in certain situations.

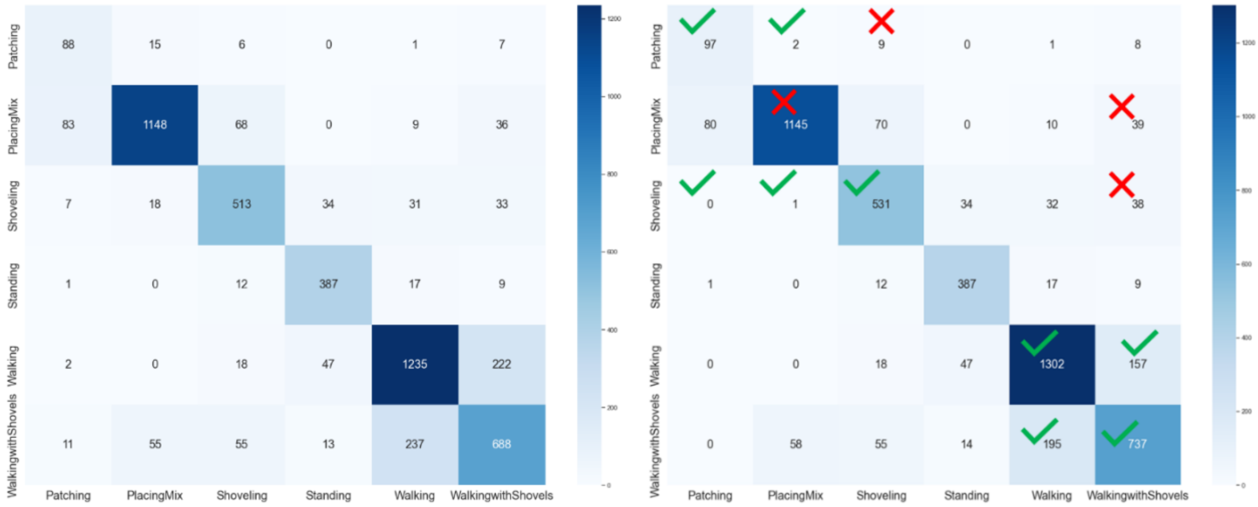
Table 7. Accuracy for Each Model after Adding Rules for Pothole Patching Operation.

Model #	Sensors	Overall Accuracy		
		Without Rules	Adding the First Set of Rules	Adding both Sets of Rules
Baseline Model	All four sensors	93.73%	93.91%	94.91%
Model 1	Upper spine	79.49%	80.10%	82.24%
Model 2	Upper spine and lower arm	90.80%	91.25%	92.13%
Model 3	Upper spine, lower arm, and ankle	91.85%	92.44%	93.32%

Figure 14 further shows the confusion matrix for model 1, and after adding two sets of rules. The model 1 has the most significant improvement after adding two sets of rules. The cells with significant improvements (i.e., increased correct predictions or decreased wrong predictions) are marked with a green cross mark, while the cells with decreasing performance are marked with a red cross mark.

After adding both sets of rules, the number of errors classifying “patching” as “placing mix” decreases from 15 to 2, and the number of correct classifying “patching” improves from 88 to 97. The reason is the model has been informed that “placing mix” after “patching” was impossible. In addition, number of errors such as classifying “shoveling” as “patching” or “placing mix” decrease to 0 and 1 after adding the rules. The errors classifying “walking” as “patching” and classifying “walking with shovels” as “patching” do not happen after adding the rules. As a result, the prediction accuracy for “patching”, “shoveling”, “walking” and “walking with shovels” all increases. In addition, the number of errors in classifying “walking” or “walking with shovels” as each other decreases. The number of errors classifying “walking” as “walking with shovels” decreases from 222 to 157, and the number of errors classifying “walking with shovels” as “walking” decreased from 237 to 195. The reason is the model is informed to stop instantaneous transitions between “walking” and “walking with shovels”.

However, the accuracy for “placing mix” is decreased, and the number of errors classifying “placing mix” as “walking with shovels” increases. Also, the number of errors classifying “shoveling” as “walking with shovels” increases. In the real pothole patching, both pairs are consequential activities since the activity “placing mix” follows “walking with shovels”, and the activity “walking with shovels” follows “shoveling”. As indicated before, the errors mainly happen during the transition between two activities. Since the DL model only stops the impossible transition activities as discussed in rule set 1, the errors for real transition activities are not controlled by the added rules.



(a) (b)

FIgure 14. Confusion Matrix for (a) Pothole Patching Trained by Model 1, (b) after Adding Both Sets of Rules

5. CONCLUSIONS

Though many studies were conducted on construction workers' activity classification, few studies focused on workers' activity classification in mobile work zones. In addition, the sensor location optimization was not well understood based on the findings from previous studies. Various mobile work zone operations require different body parts' movements, which leads to the need to optimize the sensor's location to capture the most distinguished movement information. Furthermore, to the best of our knowledge, this is the first study that integrated domain knowledge and human behaviors to improve learning-based worker activity classification models. The findings from this study can provide insights for the research community in 1) transportation construction site activity identification and classification; 2) data collection efficiency and cost reduction; and 3) integration of white-box models (i.e., domain knowledge) and black-box models (i.e., DL models) in construction workers' activity classification.

6. ACKNOWLEDGMENTS

The authors would like to thank the Joint Transportation Research Program (Project SPR-4630) administered by the Indiana Department of Transportation and Purdue University for financial support. The authors also would like to thank the support received from INDOT crews during the data collection process. It should be noted that the paper reflects the opinions of the authors only and does not reflect the official views of the sponsoring organizations.

REFERENCES

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. In Journal of Pharmaceutical and Biomedical Analysis (Vol. 22). www.elsevier.com/locate/jpba

Akhavian, R., & Behzadan, A. H. (2016). Smartphone-based construction workers' activity recognition and classification. Automation in Construction, 71(Part 2), 198–209. https://doi.org/10.1016/j.autcon.2016.08.015

al Jassmi, H., al Ahmad, M., & Ahmed, S. (2021). Automatic recognition of labor activity: a machine learning approach to capture activity physiological patterns using wearable sensors. Construction Innovation, 21(4), 555–575. https://doi.org/10.1108/CI-02-2020-0018

Bangaru, S. S., Wang, C., Busam, S. A., & Aghazadeh, F. (2021). ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. Automation in Construction, 126. https://doi.org/10.1016/j.autcon.2021.103653

Bux, A., Angelov, P., & Habib, Z. (2017). Vision based human activity recognition: A review. Advances in Intelligent Systems and Computing, 513, 341–371. https://doi.org/10.1007/978-3-319-46562-3_23

Fulford, R., & Standing, C. (2014). Construction industry productivity and the potential for collaborative practice. International Journal of Project Management, 32(2), 315–326. https://doi.org/10.1016/j.ijproman.2013.05.007

INDOT Division of Maintenance. (2022). INDOT work performance standards. Retrieved on April 30th, 2023. https://www.in.gov/indot/div/pubs/INDOT-Work-Performance-Standards.pdf

Joshua, L., & Varghese, K. (2014). Automated recognition of construction labour activity using accelerometers in field situations. International Journal of Productivity and Performance Management, 63(7), 841–862. https://doi.org/10.1108/IJPPM-05-2013-0099

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422-440.

Khosrowpour, A., Niebles, J. C., & Golparvar-Fard, M. (2014). Vision-based workface assessment using depth images for activity analysis of interior construction operations. Automation in Construction, 48, 74–87. https://doi.org/10.1016/j.autcon.2014.08.003

Kim, K., & Cho, Y. K. (2020). Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition. Automation in Construction, 113. https://doi.org/10.1016/j.autcon.2020.103126

Kim, K., & Cho, Y. K. (2021). Automatic Recognition of Workers' Motions in Highway Construction by Using Motion Sensors and Long Short-Term Memory Networks. Journal of Construction Engineering and Management, 147(3), 04020184. https://doi.org/10.1061/(asce)co.1943-7862.0002001

Klein, A., Falkner, S., Bartels, S., Hennig, P., & Hutter, F. (2017). Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets.

Li, Z., & Li, D. (2022). Action recognition of construction workers under occlusion. Journal of Building Engineering, 45. https://doi.org/10.1016/j.jobe.2021.103352

Luo, H., Xiong, C., Fang, W., Love, P. E. D., Zhang, B., & Ouyang, X. (2018). Convolutional neural networks: Computer vision-based workforce activity assessment in construction. Automation in Construction, 94, 282–289. https://doi.org/10.1016/j.autcon.2018.06.007

Megantara, A. A., & Ahmad, T. (2020). Feature Importance Ranking for Increasing Performance of Intrusion Detection System. 2020 3rd International Conference on Computer and Informatics Engineering, IC2IE 2020, 37–42. https://doi.org/10.1109/IC2IE50715.2020.9274570

Mekruksavanich, S., Jantawong, P., Hnoohom, N., & Jitpattanakul, A. (2022). Wearable-based Activity Recognition of Construction Workers using LSTM Neural Networks. ITC-CSCC 2022 - 37th International Technical Conference on Circuits/Systems, Computers and Communications, 807–810. https://doi.org/10.1109/ITC-CSCC55581.2022.9894868

Pribadi, T. W., & Shinoda, T. (2022). Analysis of Wrist Hand Motion for Monitoring of Basic Welder Training using Wearable Sensors. IOP Conference Series: Earth and Environmental Science, 972(1). https://doi.org/10.1088/1755-1315/972/1/012010

Ryu, J., Seo, J., Jebelli, H., & Lee, S. (2019). Automated Action Recognition Using an Accelerometer-Embedded Wristband-Type Activity Tracker. Journal of Construction Engineering and Management, 145(1). https://doi.org/10.1061/(asce)co.1943-7862.0001579

Sanhudo, L., Calvetti, D., Martins, J. P., Ramos, N. M. M., Mêda, P., Gonçalves, M. C., & Sousa, H. (2021). Activity classification using accelerometers and machine learning for complex construction worker activities. Journal of Building Engineering, 35. https://doi.org/10.1016/j.jobe.2020.102001

Seo, J. O., & Lee, S. H. (2021). Automated postural ergonomic risk assessment using vision-based posture classification. Automation in Construction, 128. https://doi.org/10.1016/j.autcon.2021.103725

Shah, D., Patel, S., & Bharti, S. K. (2020). Heart Disease Prediction using Machine Learning Techniques. SN Computer Science, 1(6), 345. https://doi.org/10.1007/s42979-020-00365-y

Sharma, N. V., & Yadav, N. S. (2021). An optimal intrusion detection system using recursive feature elimination and ensemble of classifiers. Microprocessors and Microsystems, 85, 104293.

Tan, M., Zhang, S., & Wu, L. (2020). Mutual kNN based spectral clustering. Neural Computing and Applications, 32(11), 6435–6442. https://doi.org/10.1007/s00521-018-3836-z

Tian, C., Chen, Y., Feng, Y., & Zhang, J. (2022). Worker Activity Classification using Multimodal Data Fusion from Wearable Sensors. 19th International Conference on Computing in Civil & Building Engineering.

Tian, C., Chen, Y., Feng, Y., & Zhang, J. (2024). Fine-Tuning Vision Transformer (ViT) to Classify Highway Construction Workers' Activities. Construction Research Congress 2024. (Accepted)

Tian, C., Xiao, J., Chen, Y., Feng, Y., & Zhang, J. (2022). Implementation, Benefits, and Challenges of Autonomous Truck-Mounted Attenuator. International Conference on Transportation and Development.

Torabi, G., Hammad, A., & Bouguila, N. (2022). Two-Dimensional and Three-Dimensional CNN-Based Simultaneous Detection and Activity Classification of Construction Workers. Journal of Computing in Civil Engineering, 36(4). https://doi.org/10.1061/(asce)cp.1943-5487.0001024

U.S. Bureau of Economic Analysis (2023), Gross Output by Industry. Retrieved April 3, 2023. https://www.bea.gov/

U.S. Bureau of Labor Statistics. (2023). Industries at a Glance. Retrieved April 3, 2023. https://www.bls.gov/iag/home.htm

U.S. Bureau of Labor Statistics (2019) Fatal occupational injuries by industry and event or exposure, all United States. Retrieved April 3, 2023. https://www.bls.gov/iif/oshwc/cfoi/cftb0331.htm

Ustebay, S., Turgut, Z., & Aydin, M. A. (2018). Intrusion detection system with recursive feature elimination by using random forest and deep learning classifier. In 2018 international congress on big data, deep learning and fighting cyber terrorism (IBIGDELFT) (pp. 71-76). IEEE.

Victoria, A. H., & Maragatham, G. (2021). Automatic tuning of hyperparameters using Bayesian optimization. Evolving Systems, 12(1), 217–223. https://doi.org/10.1007/s12530-020-09345-2

Wu, R., Fujita, Y., & Soga, K. (2020). Integrating domain knowledge with deep learning models: An interpretable AI system for automatic work progress identification of NATM tunnels. Tunnelling and Underground Space Technology, 105, 103558.

Yan, K., & Zhang, D. (2015). Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sensors and Actuators, B: Chemical, 212, 353–363. https://doi.org/10.1016/j.snb.2015.02.025

Yu, S., Yang, J., Huang, T. H., Zhu, J., Visco, C. J., Hameed, F., Stein, J., Zhou, X., & Su, H. (2023). Artificial Neural Network-Based Activities Classification, Gait Phase Estimation, and Prediction. Annals of Biomedical Engineering. https://doi.org/10.1007/s10439-023-03151-y

Zhao, J., & Obonyo, E. (2020). Convolutional long short-term memory model for recognizing construction workers' postures from wearable inertial measurement units. Advanced Engineering Informatics, 46. https://doi.org/10.1016/j.aei.2020.101177

Zhao, J., Obonyo, E., & Bilén, S. G. (2021). Wearable inertial measurement unit sensing system for musculoskeletal disorders prevention in construction. Sensors (Switzerland), 21(4), 1–28. https://doi.org/10.3390/s21041324