ANN-based Automated Scaffold Builder Activity Recognition through Wearable EMG and IMU Sensors

Srikanth Sagar Bangaru[1], Chao Wang[2*], Sri Aditya Busam[3], and Fereydoun Aghazadeh[4]

[1]Ph.D. Candidate, Intelligent Construction Management Lab (ICML), Bert S. Turner Department of Construction Management, Louisiana State University, 237 Electrical Engineering Building, Baton Rouge, LA 70803; PH: (405) 762-3689; email: sbanga3@lsu.edu

[2]Assistant Professor, Intelligent Construction Management Lab (ICML), Bert S. Turner Department of Construction Management, Louisiana State University, 3315D Patrick F. Taylor Hall, Baton Rouge, LA 70803; PH (225) 578-9175; email: chaowang@lsu.edu

[3]Master Student, Department of Mechanical & Industrial Engineering, Louisiana State University, Patrick F. Taylor Hall, Baton Rouge, LA 70803; PH: (225) 288-6287; email: sbusam1@lsu.edu

[4]Professor, Department of Mechanical & Industrial Engineering, Louisiana State University, 3250A Patrick F. Taylor Hall, Baton Rouge, LA 70803; PH (225) 578-5367; email: aghazadeh@lsu.edu

[*]Correspondence: chaowang@lsu.edu; Tel.: +1-225-578-9175

Abstract

Construction worker activity recognition is essential for worker performance and safety assessment. With the development of wearable sensing technologies, many researchers developed kinematic sensor-based worker activity recognition methods with considerable accuracy. However, the limitations of the previous studies remain at the challenge of using smartphones for practical implementation, fewer classified activities, and limited recognized motions and body parts. This study proposes an ANN-based automated construction worker activity recognition method that can recognize complex construction activities. The proposed methodology discusses data acquisition, data fusion, and artificial neural network (ANN) model development. A case study of scaffold builder activities was investigated to validate the proposed methodology's feasibility and evaluate its performance compared to other existing methods. The results show that the proposed model can recognize fifteen scaffold builder activities with an accuracy of 94% with 0.94 weighted precision, recall, and F1 Score.

Keywords: Artificial Neural Network, Wearable sensors, EMG, IMU, Activity recognition, Data fusion, Construction Worker, Scaffold Builder

1. Introduction

The US construction industry is one of the world's largest markets, with annual expenditures of over $1,293 billion and seven million employees [1]. However, the construction industry is facing a massive skilled workforce shortage [2]. More than 80% of the construction companies have reported that they have a hard time finding skilled craft workers. It is estimated that more than one million craft professionals are required by 2023, which includes various crafts such as carpentry, masonry, electricians, pipefitters, ironworkers, and scaffold builders, etc. [3]. In particular, the demand for craft professionals is very high in the US Gulf Coast area, with an increase in petrochemical investments [3]. As a result of workforce shortage, there is a significant impact on the project outcomes and worker performance, such as delayed project completion, an increase in overall project cost, and an increase in workload for the existing skilled workers [4]. One of the primary reasons for the workforce shortage in the construction industry is the premature retirement of the experienced, skilled workforce due to safety and health issues [5]. The Construction Industry Institute (CCI) and the Center for Construction Research and Training (CPWR) have established various training programs and investigated various new technologies to improve construction safety and health [6]. Various researchers have proposed different technology-based solutions to prevent the workers' premature retirement and exposure to safety and health issues (e.g., computer-vision technologies, building information modeling, wearable sensing technologies, and data mining and management) [6-10]. Moreover, these technology-based solutions help monitor and improve workers' performance by providing feedback [11-13].

Among these technology-based solutions, wearable sensing technologies have increased attention in recent years since they provide a wide range of opportunities for researchers and practitioners to develop automated and real-time systems for workers' safety and performance monitoring [13], in which the fundamental requirement for the workers' performance and safety assessment is activity recognition [14-17].

Previous studies proposed various activity recognition systems using kinematic-based methods [13,18-20], vision-based methods [11,12,21-23], and audio-based methods [15] to recognize construction worker activities. Each of these methods has its advantages and disadvantages. Computer-vision-based methods use an image or video data captured using optical cameras to provide information on worker activities. Even though vision-based methods provide semi-real-time information and reliable documentation for the future, these methods are sensitive to environmental factors, affected by obstacles, require large data storage, and high equipment cost [13]. Whereas the audio-based methods use sounds captured using audio sensors to recognize activities, they are not suitable for noisy environments and predict activities with accuracy compared to other methods [13]. Among the three methods, the kinematic-based methods have gained increased attention due to ease of use, low-cost, non-intrusive, reliable, and high accuracy activity recognition models compared to vision-based or audio-based methods [13,16,18]. The kinematic-based methods involve wearable sensors such as an inertial measurement unit (IMU) attached to the workers' body to recognize the activities' kinematic patterns. The previously proposed kinematic-based construction workers' activity recognition systems have used a smartphone or IMU sensor attached to the waist, arm, thigh, chest, and wrist to acquire accelerometer and gyroscope data of the worker performing activities such as bricklaying, carpentry, hammering, sawing, wrenching, hauling, unloading, and drilling [13,16,18,24-32].

Although most of the previous studies have achieved good accuracies, there are some the limitations such as restrictions on using smartphones on construction sites pose challenges for practical implementation, use of multi smartphones or IMU sensors is challenging while performing construction activities due to the dynamic nature of the construction work, and the current models have predicted few activities involving limited motions and body parts. Moreover, the recent state-of-the-art review articles on construction worker activity recognition methods [13] and wearable sensor applications in construction safety and health [6] stated that using sensor fusion and hybrid model could be the solution to obtain more precise and generalized methods to commercialize the workers' activity recognition and other construction safety applications such as fatigue monitoring or workload evaluation. Therefore, to overcome the challenges of the current construction workers' wearable sensor-based activity recognition methods, this study proposes developing automated construction workers' activity recognition using forearm electromyography (EMG) and IMU data. Moreover, this study validates the feasibility through a case study of scaffold builders who are important craft to the industrial and commercial construction projects. In the case study, forearm physiological data collected from EMG sensor and kinematic data collected from IMU sensor were analyzed for recognizing complex scaffold builder activities that involve different body parts (wrist, forearm, upper body, lower body, and whole-body) and various motions (repetitive motion, impulsive motion, and free motion) performed in a short time. To achieve the proposed objective, we first collected forearm EMG and IMU data from six participants using the armband sensor while performing scaffold builder activities from six participants. The dataset consists of 38 variables (accelerometer - 3, gyroscope – 3, and EMG – 32) with approximately 150,000 datapoints. Secondly, the collected data were preprocessed and prepared for Artificial Neural Network (ANN) model building and training. Then, the ANN model was trained and evaluated. Finally, the performance of the proposed model was evaluated on real-time un-labeled data and compared the performance of different sensor combinations and classification algorithm.

The rest of the paper is divided as follows. First, we reviewed the background and related work regarding construction workers' activity recognition using wearable sensors. Next, the proposed automatic construction workers' activity recognition model was introduced and followed by the experiment section, including model validation and performance evaluation of the proposed method. In the end, it concludes with the discussions of the findings, limitations of the study, and future research directions.

2. Literature Review

2.1.Human Activity Recognition, Deep Learning, and Sensor Modality

An activity is defined as a group of actions that include a series of consecutive movements [13]. Human activity recognition (HAR) involves predicting a person's movement based on sensor data and machine learning models [33]. HAR is broadly classified into two types, i.e., sensor-based and vision-based [34]. The vision-based activity recognition system detects human motion using images or videos, whereas sensor-based systems focus on the motion data from smart sensors such as accelerometers, gyroscopes, electromyography, audio sensors, vibration sensors, etc. [35].

The wearable sensor-based activity recognition using traditional pattern recognition (PR) methods mainly involves three steps, i.e., sensor data collection, feature extraction, and model training [36]. Firstly, acquiring the data from sensors such as accelerometers, gyroscope, magnetometers, electromyography sensors, audio sensors, vibration sensors, etc. Secondly, features such as time-domain, frequency-domain, or statistical features are manually extracted from the data based on human experience or domain knowledge. Finally, features are used to train the models to recognize activities [37]. The deep learning models are preferred over traditional pattern recognition (PR) because of the following reasons [35,38]:

- In traditional PR methods, the features are extracted through a heuristic and hand-crafted approach, which relies heavily on human experience and domain knowledge [39].
- From the human experience, only statistical features such as mean, median, amplitude, frequency, minimum, maximum, etc., can be learned by the models. These statistical features alone are not sufficient to recognize complex activities.
- The traditional pattern recognition methods require a large amount of labeled data for training models, whereas deep learning networks can utilize the unlabeled data for model training.
- The traditional pattern recognition models focus on training from static data, whereas in real life, the activity data is streamed real-time, which requires robust and incremental learning.

The deep learning models are used to overcome the limitations of traditional pattern recognition models. Unlike traditional pattern recognition models, the feature extraction and model training are performed simultaneously in the deep learning models by extracting the high-level features in deep layers, which helps recognize complex activities. In the case of extensive unlabeled data, deep generative models can exploit unlabeled data for model training. Moreover, the models trained on extensive labeled data can be transferred to new activities with few or none labels [34,35,38].

The sensor modalities for human activity recognition (HAR) can be classified into body-worn sensors, object sensors, ambient sensors, and hybrid sensors [38]. The body-worn sensors, such as accelerometers, gyroscope, magnetometers, etc., wore on the body are among the most common modalities in HAR. The body-worn sensors are found in wristwatches, smartphones, glasses, bands, and helmets [33,35,38]. The object-worn sensors are attached to the objects to recognize the object's movement to infer human actions. The most common object-worn sensors used for HAR are radio frequency identifiers (RFID) tags and accelerometers. However, object-worn sensors are less popular than body-worn sensors due to their deployment [40].

In contrast, ambient sensors such as sound, radar, temperature, and pressure sensors capture the interaction between humans and the environment. Human activities are inferred based on the changes in the environment. Similar to the object-worn sensor, the ambient sensors are difficult to deploy. Moreover, only certain types of activities can be inferred using ambient sensors. In recent years, hybrid sensors (a combination of body-worn, object-worn, and ambient) is gaining importance due to the rich information of human activities provided by the sensors and improving HAR accuracy. The hybrid sensors can recognize the complex activities of multiple occupants of smart homes [41,42]. Various deep learning models such as a deep neural network (DNN), convolution neural network (CNN), recurrent neural network (RNN), deep belief network (DBN), stacked autoencoder (SAE), and hybrid models are available for HAR [35,38]. All these deep learning models are the classes of ANN which are used based on the data type. For example, CNN and RNN models are used for image/video and sequence data, respectively [43].

2.2. Wearable Sensing Technology Applications in Construction

In recent years, wearable sensors are widely used in the construction industry for different applications, especially in construction safety and health. The different types of sensors widely used for construction applications are kinematic sensors (such as IMU), cardiac activity (such as Electrocardiogram (ECG or EKG), and photoplethysmogram (PPG)), skin response (such as

Electrodermal Activity (EDA) and Skin Temperature (ST)), eye movement (such as eye-tracking), muscle engagement (such as EMG), and brain activity (such as electroencephalogram (EEG)). IMU sensors are widely used as wearable sensors in the construction industry to measure the objects' kinematic movement, including construction workers, equipment, and tools. IMU sensors attached to workers' bodies were used to determine workers' body posture, acceleration, and orientation [44-46], and were also used for preventing musculoskeletal disorder by detecting awkward postures [47-49] and fall protection by identifying a sudden change in body acceleration [50-52]. The measure of cardiac activity using ECG and PPG sensors facilitates in determining the workers' physiological status. The metrics such as heart rate variability (HRV), inter-beat-intervals (IBI), pulse-rate variability (PRV), and heart-rate reserve (HRR) derived from heart rate are essential to determine the physical and mental condition of the workers [53,54]. The EMG sensors capture muscle activity used to assess the muscle load and forces used for ergonomic assessment [55]. The PPG, EDA, ST, and heart rate sensors were extensively used for assessing the workers' physical workload and fatigue [8,56-58]. The use of eye-tracking to measure eye positions and movements relative to the participant's head helps evaluate the construction safety training and hazard recognition abilities [59,60]. The EEG sensors which measure brain activity are used to assess the workers' mental status on the job site and the effectiveness of training programs [61,62]. Even though several have shown the feasibility of using wearable sensors for construction safety and health, there exist some challenges such as noise and artifacts in field measurements, variability in standard to assess personal safety and health risks, the uncertainty of return of investments, and user resistance for adoption [6].

2.3. Construction Activity Recognition

Construction activity recognition helps in safety, productivity, and quality control analysis. Advancements in automated data acquisition systems to quantity progress and track resources to streamline the crew activity analysis have shown promising results compared to conventional methods such as direct observation or survey-based methods, which are time-consuming, tedious, and error-prone. However, automated data collection technologies are still being investigated for their feasibility and reliability in construction domain applications. The automated data acquisition systems can be broadly classified into vision-based and wireless sensor-based systems. The vision-based techniques have been proposed and evaluated by various researchers for activity recognition and process monitoring [63]. On the other hand, wireless sensor-based systems are assessed to collect Spatio-temporal activity data [64]. However, vision-based techniques are often prone to illumination variability and occlusions on the job site, whereas wireless sensor-based methods overcome the challenges of the line of sight (LOS) and occlusions. Moreover, sensor-based methods are a low-cost solution for activity analysis.

Wearable sensor-based activity recognition aims at identifying the physical actions from a set of sensor signal data, which can be achieved by utilizing machine learning techniques. The inertial measurement units (IMUs), which include accelerometer, gyroscope, and magnetometer, are the most commonly used wearables sensors used for construction activity recognition. The overall process of developing an activity recognition system using sensor signal data and machine learning techniques is as follows: raw signal data acquisition and annotation, segmentation of labeled data for feature extraction, training machine learning-based classifier algorithms, and validation of the models. Even though the framework for activity recognition using wearable sensors and ML algorithms remains the same, it is essential to investigate the feasibility of using different wearable sensors for an activity or action recognition in the construction domain to improve accuracy, reliability, and usability. The model accuracy depends on various factors such as type of sensor data (acceleration, gyroscope, EMG, etc.), feature set (time-, frequency-, and discrete representation domain), classifier algorithms (k-nearest neighbor, neural network, support vector machine, and decision tree). Various studies have developed using different ML models and investigated the influence of several factors for construction activity recognition using different wearables sensors. Joshua and Varghese [65] investigated the use of wired accelerometers attached to the waist of the mason to recognize brick laying actions for productivity analysis. The study reported that the multilayer perceptron and neural network classifier algorithm best performance with 80% accuracy using features such as mean, maximum, variance, correlation, and energy. Joshua and Varghese [66] developed the accelerometer-based method for ironwork and carpentry activities classification using a decision tree with 90.07 and 77.74 percent accuracy. Cezar [24] has developed a construction activity recognition model using a dominant hand accelerometer and gyroscope data to recognize hammering, sawing, sweeping, and drilling activities with the highest accuracy of 91% Quadratic Discriminant Analysis (QDA). Khan and Sohail [25] have evaluated 17 classification algorithms and three sensor positions to recognize nine construction activities. The study concluded that the waist position had achieved the highest accuracy of 93.90% for the Random Forest classifier. Moreover, Joshua and Varghese [26] have proposed a framework to select the accelerometer sensor's position to obtain the best classification results. A bricklaying case study proved that the sensor's position has a significant effect on classification accuracy. Yang, et al. [27] have developed automated near miss fall incidents in ironworkers using IMU data from waist and support vector machine (SVM), which obtained an accuracy of 91.1%. In contrast, the near-miss classification model of Lim, et al. [28] obtained an accuracy of 94% by using accelerometer data from a smartphone placed at the hip pocket. Akhavian and Behzadan [29] developed a construction activity recognition and classification system using raw accelerometer and gyroscope data from a smartphone placed on the upper arm while performing sawing, hammering, wrenching, loading, hauling, and unloading. The study evaluated the performance of the classification algorithms such as K-nearest neighbor (KNN), ANN, logistic regression (LR), decision trees (DT), and support vector machine (SVM) using the features such as average, minimum, maximum, interquartile range (IQR), and root means square (RMS). The 10-fold cross-validation of the classifiers reported that the NN algorithm performed better than other classifiers with an average accuracy of 93.63%. Further, the study was extended to determine the activity duration using an ANN model with 90.74% accuracy [20]. Ryu, et al. [67] tested the feasibility of using an accelerometer-embedded wrist-worn for construction workers' action recognition such as spreading motor, laying blocks, adjusting blocks, and removing mortar precision performing bricklaying activity. The study investigated the classification accuracy of KNN, DT, multilayer perceptron, and SVM for different window sizes and features (time- and frequency -domain); the 10-fold cross-validation results reported that SVM with 4s window size showed the highest classification accuracy of 88.1%. Cheng, et al. [68] developed a task-level activity analysis using the data fusion of Spatio-temporal and workers' posture data for productivity analysis. The accelerometer and gyroscope data were used to evaluate the construction workers' workload [19] and ergonomic risk [69] using an SVM classifier accuracy of 95.67% 92.7%, respectively.

Previous studies have proved that the sensor placement on the body significantly affects the activity recognition performance because the sensor signal pattern for the same activity varies depending on the sensor's position [70]. For activity recognition using accelerometers, the sensor's location close to the waist represents the significant body motions [71]. However, waist-oriented acceleration signals do not reflect hand and arm movement, challenging to differentiate actions, including the movements [72]. The studies [73,74] reported using a single accelerometer sensor on the dominant wrist to classify daily living activities with an accuracy of around 95%. Although these studies have proved that using a single accelerometer sensor on the upper body was sufficient for recognizing construction activities, the models' robustness needs to be improved to predict real-time un-labeled data. There remains a gap in the area of construction workers' activity recognition and wearable sensors applications in the construction domain, such as sensor data fusion at various levels, robust and reliable model to recognize multiple complex construction activities, and generalization of activity recognition models to convert to the commercialized application [6,13].

2.4.Point of Departure

In the construction domain, the worker activity recognition models are broadly classified into kinematic-based, vision-based, and audio-based methods. The latter two methods have technical and practical implementation challenges such as high initial cost, influence environmental factors, low accuracy, high computation cost, large storage size, and privacy concerns [13]. Whereas the kinematic-based approaches have gained increased attention for worker activity recognition due to ease of use, low cost, non-intrusive, suitable for any environment and trade, and high accuracy. Most previous studies have used smartphones as a cost-effective data collection system for recognizing workers' motion using acceleration and gyroscope signal data acquired from embedded sensors in the smartphone [49,75-78]. However, the use of smartphones for activity recognition has challenges for practical implementation. To overcome the challenges of smartphone sensors, other studies have proposed using accelerometer and gyroscope data to develop machine learning-based activity recognition models for various applications such as the activity analysis of workers, fall risk detection, ergonomic assessment, and equipment detection

[18,27,41,50,79,80]. The limitations of these studies are they can identify a fewer number of construction activities involving either stationary movements or traveling (e.g., bricklaying and walking) and were limited to the forearm or upper body movements (e.g., hammering, sawing, wrenching, power drilling, and hammering) [74,81-84]. Therefore, it is essential to develop an activity recognition system to recognize complex construction activities involving different body parts (wrist, forearm, upper body, lower body, and whole-body) and various motions (repetitive motion, impulsive motion, free motion, and idle) performed in a short time interval. None of the previous studies have used other than motion data for construction worker activity recognition. Since most construction activities involve muscle activity and dynamic motion in a short interval of time, muscle activity and motion data might improve activity recognition. So, it is essential to investigate the fusion of physiological and kinematic data to improve the worker activity classification performance and recognize activities that do not involve the movement of a human body part. Other technical challenges of previous studies include the necessity of large dataset to develop models, use of multiple sensors, need for domain knowledge for feature extraction, human variability, unable to generalize the model, Moreover, there is a necessity to explore various preprocessing techniques such as data augmentation, and hyperparameter optimization to develop robust and reliable models using an optimal number of sensors.

3. Research Methodology

As shown in Figure 1, the proposed research methodology starts with data acquisition from a wearable armband sensor that can collect EMG and IMU data. The collected raw multi-sensor data is then preprocessed and fused to obtain a dataset with EMG and IMU data features. The fused data is labeled with the actual activity class and further used to build and train an ANN model. The proposed methodology's performance is evaluated through a series of the analysis, such as the performance on unlabeled new data, the performance of different sensor combinations, and comparison of performance with other classification algorithms. Each of these steps is further discussed in the following subsections.

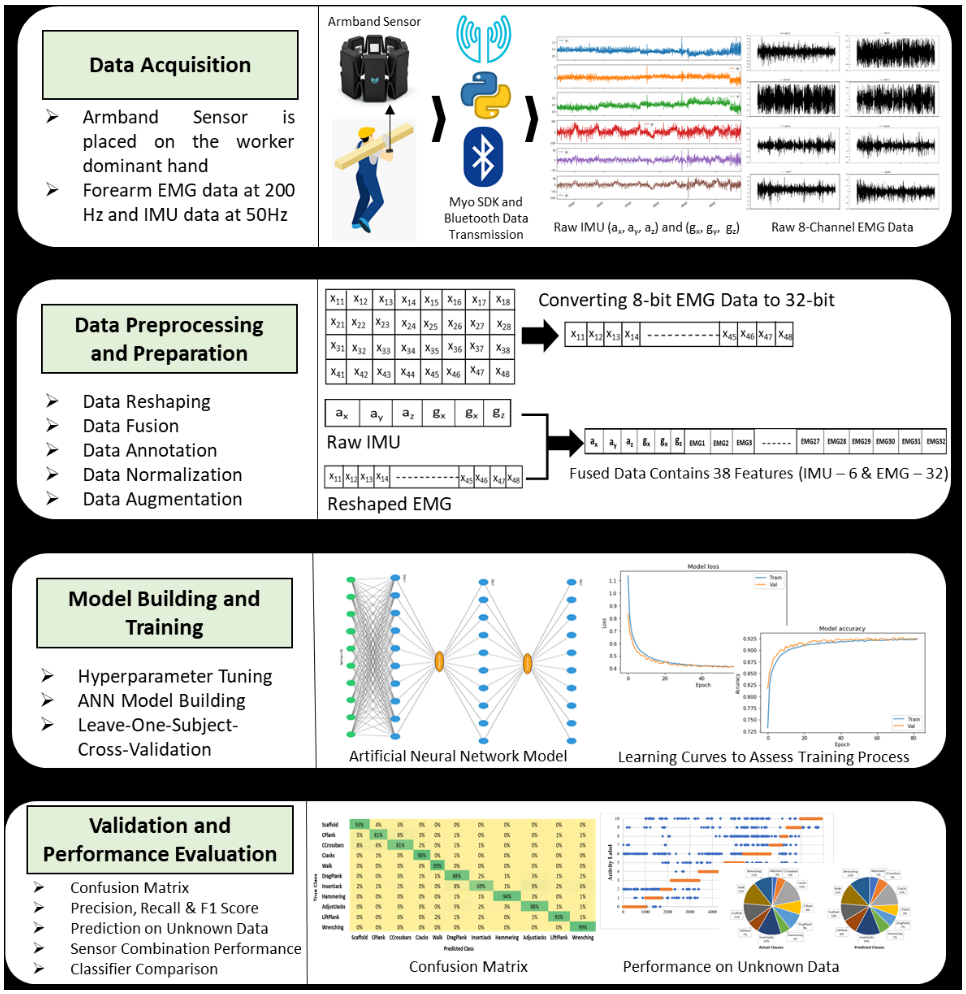

Figure 1. Framework for construction worker activity recognition using forearm-based EMG and IMU armband sensor

3.1. Data Acquisition using Forearm-based Armband Sensor

A forearm-based armband sensor (Myo Armband) developed by Thalmic Labs Inc. was used to collect forearm EMG and IMU data. This armband sensor is a non-intrusive wearable sensor that consists of eight EMG sensors (#1-#8) and one 9-axes IMU sensor (3 for acceleration, 3 for gyroscope, 3 for magnetometer, and embedded within EMG sensor #4). The armband sensor weighs approximately 93grams and needs to be worn at the thickest part of the forearm with EMG sensor #4 in the line of the index finger and LED light towards the lower forearm, as shown in Figure 2. Moreover, Figure 2 shows the electrode locations and IMU axes directions. The data was transmitted in real-time to local or cloud storage via Bluetooth Low Energy (BLE) wireless connection. According to Thalmic Labs Inc., the Myo armband sensor has a built-in rechargeable lithium-ion battery that can last for one full day on a single charge. The armband sensor is also easy to use, comfortable to wear for long periods, do not obstruct ongoing work, and stable Bluetooth connectivity. The Myo armband has achieved an acceptable system usability score (SUS) when tested for usability in other domains such as medical [85] and entertainment [86]. The raw EMG and IMU data can be collected from a program that we developed using the Myo software development kit (SDK) 200 and 50 Hz. The EMG sensors capture the forearm muscle electrical impulses, which are stored as an 8-bit array with values ranging from -128 to 127, which is different from the data collected from conventional EMG sensor values are in a format of volts or millivolts. In comparison, the IMU sensors capture the acceleration, angular velocity, and orientation of the forearm along x, y, and z directions.

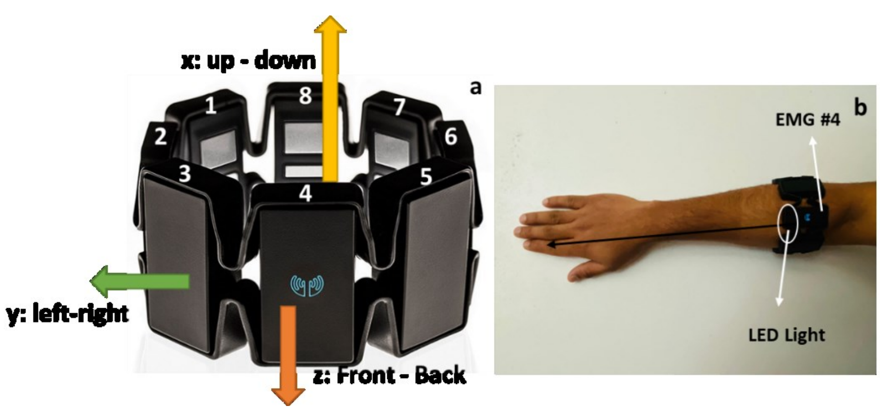

Figure 2. (a) Myo armband electrode location and IMU axes directions; (b) Myo armband placement on the forearm

3.2.Data Preprocessing and Preparation

Sensor data fusion can be performed at different levels, including signal level, feature level, and decision level. The signal level data fusion involves fusing the raw sensor data, the feature level involves fusing features extracted from the sensor data, and the decision level involves fusing the decisions from outputs from sensor data [6]. In this study, the signal level sensor data fusion is considered for processing collected EMG and IMU sensor data. The signal-level fusion of the data eliminates the necessity for feature extraction from the raw data, which requires domain knowledge. Since the EMG and IMU data are collected at different frequencies, the EMG data is first reshaped to match the IMU frequency, which is performed by transposing four rows of 8-bit EMG data to 32-bit data. The reshaped 32-bit EMG data is then fused with the IMU data of each activity using concatenation. The fused data are manually annotated using the class label shown in Table 1. The manual annotation process involves assigning activity ID to each row of the dataset since the training data will be collected for each activity. Since the EMG and IMU data obtained using armband sensors are in different units, the data is normalized using the z-score standardization (feature scaling) technique. The z-score is calculated by subtracting each feature's mean from that feature's values and then dividing the corresponding value by the standard deviation of that feature, as shown in Equation 1. This transforms the data to have a mean value as zero and standard deviation as one. Feature scaling is essential for neural network models to handle data smoothly. Feature scaling is essential for a neural network to handle the data smoothly. If the input data has units in different scales, the features with high range values may get higher derivatives during backpropagation than the features with low range values. Hence, the weights in the connected layers will be updated abnormally, and there will be a bias added to the model. Standardizing makes the model update the weights effectively during forward and backward propagation and avoid model weights and errors. Moreover, it helps in faster convergence of gradient descent to the global minima. After performing standardization, all features have been reduced to the same scale [87].

$$\text{Z-score} = \frac{x_i - \bar{x}}{\sigma} \tag{1}$$

A large amount of synthetic data can be generated using time series data augmentation techniques to improve the ANN model's performance and prevent overfitting the model parameters. Also, data augmentation helps in the model's generalization since it introduces variability in the data without altering the labels. To account for those factors, various augmentation techniques are available such as time-wrapping, pooling, drifting, and reversing. Time-warping has a spatial-temporal characteristic that can generate data with a different warping ratio for different activities and is controlled by a number of speed changes and the maximal ratio of max/min speed. Pooling makes the data reduce the temporal resolution without a change in length. In contrast, the drift changes the data randomly and smoothly and is controlled by parameters like several drift points and maximal drift. Finally, the reverse will help in reversing the timeline in a series of data. Each augmentation technique generates a 1-fold increase in training data, resulting in a 4-fold increase in the number of data points for each user [40,70].

3.3.Model Building, Training, and Evaluation

An ANN-based deep learning model is proposed for construction worker activity recognition. An ANN model can handle complex data by recognizing the hidden patterns in the data and sensing the linear and non-linear relationship between independent and dependent variables by reducing the noise in the data. In this study, the ANN model is built in Keras [88], a high-level neural networks application programming interface (API), written in Python and capable of running on top of TensorFlow. The model building and training module involve three essential steps, i.e., hyperparameter optimization, model building and compiling, and model training. Each of these steps is discussed in this section.

3.3.1. Hyperparameter Optimization

ANN network is designed with significant hyperparameters to achieve desirable activity classification results. To obtain the best classification results, one needs to tune the model with different combinations of hyperparameters, where the manual tuning process is time-consuming and inefficient. To overcome the manual tuning process challenges, various automated hyperparameter optimization techniques were proposed, such as grid search, random search, and Bayesian optimization [89]. Each of these techniques has its advantages and disadvantages. A grid search method selects the grid of parameters and tries every combination to select the best parameters. However, this method is computationally expensive and takes a long time to complete. A random search does not select all the combinations but a random list of parameters and select the best parameters among those combinations. Even though it is computationally efficient, it can probably miss some of the crucial parameters during the evaluation, which is unreliable due to its random selection. In contrast, Bayesian optimization keeps track of the past evaluated results and builds a probabilistic model to map the hyperparameters to the objective function's probability score. They perform better based on a surrogate function, which can help identify the global minima. In this study, the tree-structured Parzen Estimator (TPE) based surrogate model has been used, a sequential model-based optimization (SMBO) approach [89]. TPE is represented as $p(y|x)$, where y is the quality Score, and x represents hyperparameters, as shown in Equation 2.

$$p(y|x) = \frac{p(x|y) * p(y)}{p(x)} \quad (2)$$

$p(x|y)$ is a probability of hyperparameters given the value of an objective function, as shown in Equation 3.

$$p(x|y) = \begin{cases} l(x)\ if\ y < y^* \\ g(x)\ if\ y >= y^* \end{cases} \quad (3)$$

l(x) and *g(x)* are two different distributions of hyperparameters with $l(x)$ used when the value of an objective function is less than the threshold, and $g(x)$ is used when the objective function's value is more significant than a threshold. Y^* is the threshold value. TPE draws a sample of hyperparameters from $l(x)$ and returns the parameters which yielded the highest value with the ratio $l(x)/g(x)$. Overall, the algorithm selects a new set of hyperparameters, evaluates the model, and stores them as history. With every iteration and using the history, $l(x)\ and\ g(x)$ is built by an algorithm to evaluate the objective function's probability model. Since the algorithm suggests better candidate hyperparameters for evaluation, the objective function score increases much faster than random or grid search results in less total evaluations of the objective function. Also, TPE can reduce the running time and get the best scores on test data. Sequential model-based optimization approaches vary like the surrogate, but all depend on the knowledge from previous studies to suggest better hyperparameters for the next evaluation. TPE is an algorithm that uses Bayesian reasoning to create a surrogate model and can use expected improvement to pick the next hyperparameter.

3.3.2. Model Building and Compiling

In a neural network architecture, many crucial parameters need to be considered to develop an efficient model. The most important features are the numbers of hidden layers, neurons in each layer, optimizers, activation functions, learning rate, batch size, epochs, and regularization. The number of layers and neurons in each layer depends on the data where the input and output layer nodes are equal to input features and the number of activity classes, respectively. The optimizers in neural networks change attributes such as weights and learning rate to reduce the losses. Adam optimizer is the most commonly used algorithm, an adaptive learning technique for each weight in the neural network; it uses the estimates of both first and second moments of gradient and evaluates individual learning rates for different parameters. Adam optimizer is considered an improvised version of well-known optimizers such as RMSProp, AdaGrad, and SGD [90]. It uses the functional combination of RMSProp and SGD by using squared gradients and moving average of gradients for effective faster convergence to global minima.

The ReLu activation function is used for input and hidden layers. Compared with other functions like sigmoid and tanh, ReLu can handle large layers and tackle the vanishing gradient issue. For the output layer, the Softmax activation function is applied since it is useful for multi-label classification. Also, Softmax best suits for output layer as it gives the probability values for predicting different classes. The choice of batch size decides the number of samples from the training data propagated through the network. Whereas the epoch decides the number of times, all the training samples are passed forward and backward through a neural network. If the class labels were mutually exclusive, the sparse categorical cross-entropy loss function should be applied to the model. Moreover, it is essential to convert target variables into integers for the ANN model.

Regularization involves concepts such as L_1 and L_2 regularization, dropout, and early stopping. Firstly, L_1 and L_2 are lasso and ridge regressions, which add a penalty to the loss function. The loss function is the ordinary least square technique that measures the sum of the squared errors. They are used for feature selection and removing multicollinearity during model training. Both are involved in the process of reducing the weights or coefficients of neural network function. L1 reduces the weights faster than L2 and finally makes the model more straightforward and reduces overfitting. Each has its advantages and disadvantages, but elastic net regularization has been used to optimize the model in the best possible way, combining L_1 and L_2 regularizations. Secondly, the dropout function reduces the number of neurons required for training in a selected layer for each iteration to prevent overfitting. The dropout ratio increase eventually results in underfitting curves. Finally, early stopping is another regularization method that helps stop the model training when the validation loss is no longer decreasing or increasing after performing a certain number of epochs. Early stopping is considered one of the best solutions to tackle the overfitting problem.

3.3.3. Model Training

During model training, backpropagation involves the multiplication of gradients in every layer. If the gradient values are too small, the models suffer from vanishing gradient problems, but if the gradient values are too high, the model suffers from an exploding gradient problem. Selecting a set of optimized parameters plays a significant role in providing an accurate predictive model. Once the optimum parameters are selected through hyperparameter optimization, the model is diagnosed for the underfitting or overfitting issues using learning curves. The learning curves, such as model loss and accuracy, help understand the model's learning performance over time during training. Moreover, the model curves can be used to diagnose the problems of under and overfitting. Two metrics used to assess the performance of learning are loss (error) and accuracy. For a better learning performance, the model loss (error) should be decreasing, and the model accuracy should be increasing. The training learning curve measured on training data indicates how well the model is learning, whereas the validation learning curve calculated on validation data, which is not part of training data, represents how well the model is generalizing. The learning curves' shape and dynamics help diagnose the model's behavior and identify if the model has under fitted or a good fit or overfitted. The model's underfitting occurs when the model cannot learn the training dataset, whereas overfitting refers to a model that has leaned the training data too well, including random fluctuations and noise in the data. A good fit model exists between underfitting and overfitting models, which can be identified from learning if the loss curve decreases to the point of stability and has a small gap between the training and validation curve. The learning curves are developed using the Keras callback history, which records the loss and accuracy of training and validation dataset for each epoch. The batch size and epochs are set to 100 and 150, respectively. To overcome overfitting or underfitting, the regularization concept has also been implemented during the model training.

3.3.4. Model Evaluation Technique

General evaluation of machine learning models can be done by splitting the collected experiment data into train and test data. However, the disadvantage with this technique is that the model's evaluation is done specifically on this split data where they can have data leakage between the train and test on the same subject, especially in human activity recognition and testing any new or unseen data on the trained model may not be reliable [36]. In order to avoid this and make a generalized model, the cross-validation technique has been used. Cross-validation is a technique that holds out test data from a given data in an experiment, trains the model on the remaining data, and tests it on the formerly reserved test data. This process is repeated for the K number of experiments for the entire training data. Splitting of the data depends on the number of splits we required and is represented by K, where K is the number of folds. Depending on the given input parameter K, the K number of experiments will be performed to evaluate the model performance. Popular cross-validation techniques are K-fold, Stratified K-fold, Repeated K-fold, Leave One Out, Leave One Subject Out, and Nested. The dataset consists of different construction activities performed by different subjects. So, in this study, Leave-One-Subject-Out (LOSO) cross-validation technique has been chosen. LOSO is a K-fold cross-validation technique where the number of folds is chosen before the model evaluation. In LOSO, the number of folds is equal to the number of subjects who performed their activities in our experiment, and LOSO evaluates each subject's accuracy in different folds or experiments. Hence, LOSO performance is robust. The LOSO's overall accuracy is determined by finding the average of all the folds in our experiment [91].

3.4.Performance Evaluation Metrics

Once a good fit model is obtained through the training and validation process, the built ANN model's performance will be evaluated by the testing dataset using classification accuracy, confusion matrix, precision, recall, and F1 Score. The most general and first look evaluation for any deep learning techniques are done by classification accuracy. It is calculated as the number of correctly predicted outcomes to the total number of predictions. Higher classification accuracy is required to achieve the desired activity recognition results. However, the classification accuracy alone is not sufficient to decide the robustness and reliability of classification results. Therefore, other metrics such as precision, recall, and F1 Score of the proposed model are also analyzed. A confusion matrix is a matrix with an equal number of rows and columns. It represents the complete performance of the model considering each class. Each row and column of the matrix corresponds to true and predicted classes. The diagonal cells of the matric represent the percentage of correct prediction for each class, and the off-diagonal elements represent the misclassification percentage with respect to other classes. In order to understand the concept of precision and recall, firstly following terms are defined, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) come into the picture. TP is the number of correct positive predictions done by a positive model. TN refers to the number of negative predictions done by a model that is negative. FP is the number of classes predicted incorrectly where the model thinks predicted classes are positive (true) but, it is not true. FN is the only misclassified metric where the model thinks the predicted activity is not positive (true), but it is true. For the multi-classification model, the values of TP, TN, FP, and FN were calculated using the confusion matrix where TP – value in the diagonal cell, FN – for a class is the sum of values in the corresponding column excluding TP value, and FP – for a class is the sum of values in the corresponding rows excluding TP value. Using the TP, TN, and FN values, the metrics precision and recall were calculated using Equations 2 and 3, respectively. The prediction value indicates how often the prediction is correct, which is defined as the ratio of the number of true positive predictions (TP) to all total number of positive predictions of the model (TP+FP) (Equation 4). In contrast, the recall indicates the correctly predicted rate of a class, which is the ratio of the number of true positive predictions (TP) to a total number of predictions (TP+FP) (Equation 5).

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}}=\frac{\text{Value of the Diagonal Cell of the Class}}{\text{Total Number of Predictions of the Class}} \tag{4}$$

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}}=\frac{\text{Value of the Diagonal Cell of the Class}}{\text{Total Number of Instances of the Class}} \tag{5}$$

If the classes are imbalanced, the most useful and reliable metric to assess the model performance is the F1 Score, a harmonic mean of precision and recall, as shown in Equation 6.

$$\text{F1 Score}=2*\frac{\text{Precision * Recall}}{\text{Precision + Recall}} \tag{6}$$

The above formulas are used to calculate the performance metrics for individual classes, whereas the weighted precision, recall, and F1 Score following Equation 7 are applied to evaluate the overall model performance. The weighted average of a metric is the sum of the metric (precision, recall, and F1 Score) multiplied by the samples of each class (i), then divided by the samples of all the classes.

$$\text{Weighted Metric}= \frac{\sum_{i=1}^{m}(\text{Metric}_i)*(\text{Samples}_i)}{\sum_{i=1}^{m}\text{Samples}_i} \tag{7}$$

Where m is the total number of classes, $Metric_i$ is the value of metric for class i (i = 1, 2, …, m), and $Samples_i$ is the number of samples in each class i (i = 1, 2, …, m).

In addition to the performance evaluation of the proposed model, another analysis is conducted to evaluate the activity prediction accuracy on an entirely new dataset, i.e., an unknown dataset. We also compare the results with other classification algorithms to examine the robustness of the proposed model. The new dataset's prediction includes performing activity recognition using the proposed model on the dataset that is not used either in the training or testing process. Moreover, the data was collected from an individual who performed a whole sequence of activities at his own pace. The proposed model's performance is compared with the most common classification algorithms previously used in other construction activity recognition studies [13].

4. System Feasibility Validation and Performance Evaluation

4.1. Case Study of Scaffold Builder Activities

To validate and evaluate the proposed construction worker activity recognition model's performance, a case study of scaffold builder activities was considered since it involves various body parts and different movements, which allows testing the proposed model on complex construction activities. According to OSHA, a scaffold is defined as an elevated, temporary structure [92]. Based on the construction work, the type of scaffold may vary. Two basic types of scaffolds are supported and suspended scaffolds. The supported scaffolds consist of one or more platforms supported by load-bearing or rigid supports, whereas the suspended scaffold is supported by an overhead structure using non-rigid support such as ropes [92]. The supported scaffolds are extensively used in industrial and commercial construction projects [93]. The scaffold building requires scaffold erection skills, carpentry hand tools, and heavy labor-intensive tasks [94]. By reviewing various scaffolding activities onsite and online, we have recognized fifteen common activities in building a supported scaffold. The activities include carrying a metal 5 ft. x 5ft. scaffold frame (38 Lbs.) sideward, carrying a 6 ft. x 12 in. wooden plank (15 Lbs.), carrying a 7 ft. x 4 ft. scaffold cross brace (10 Lbs.), carrying 24 in. leveling jacks (6.5 Lbs.), walking around, dragging wooden plank along the frame, lifting the plank from elbow to overhead, adjusting and inserting the leveling jacks, hammering, wrenching, climbing stairs with and without tool bag, and walking downstairs with and without tool bag to set up and fix scaffold pin. The fifteen scaffold builder activities and the activity ID used for the ANN model are summarized in Table 1. Some of the key scaffold builder activities are shown in Figure 3. All these activities require extensive manual efforts and involve different body parts (wrist, forearm, upper body, lower body, and whole-body) movements, and various motions (such as repetitive motion, impulsive motion, and free motion). Moreover, it involves manual material handling tasks such as carrying different weights, lifting at different heights, and pushing activities.

Table 1. Scaffold builder activities and activity ID

ID No.	Activity Description	Activity ID
0	Adjusting Leveling Jacks	AdjustJacks
1	Carrying Crossbars	CCrossbars
2	Carrying Leveling Jacks	CJacks
3	Carrying Scaffold Plank	CPlank
4	Carrying Scaffold Frame	CScaffold
5	Dragging Scaffold Plank	DragPlank
6	Hammering	Hammering
7	Inserting Jacks into Scaffold Frame	InsertJack
8	Lifting Scaffold Plank from Elbow to Overhead	LiftPlank
9	Walking	Walk
10	Wrenching	Wrenching
11	Climb	Climb
12	Downstairs	Downstairs
13	Climb with Tool Bag	ClimbW
14	Downstairs with Tool Bag	GDownstairsW

Figure 3. Shows few scaffold builder activities performed in outdoor environment (a) scaffold carrying, (b) plank carrying, (c) crossbars carrying, (d) leveling jacks carrying, (e) adjusting leveling jacks, and (f) insert leveling jacks into the scaffold

4.2.Experiment Setup

4.2.1. Data Collection and Augmentation

To validate the feasibility and evaluate the proposed automated activity recognition model's performance using forearm EMG and IMU data, an experiment was performed, which involved participants performing scaffold builder activities in the outdoor environment. Seven male college students have voluntarily participated in the experiment. The participants' age ranged from 24 to 28 years (mean ± SD: 26.43 ± 1.40 years), weight ranged from 62.60 to 100 kgs (mean ± SD: 80.98 ± 13.38 kg), and height ranged from 1.65 to 1.83 m (mean ± SD: 1.73 ± 0.06 m). All participants are right-handed, healthy, and have no musculoskeletal disorders record. None of the participants have prior scaffold building experience, but all the activities were demonstrated to all the participants before starting the experiment. The armband sensor was placed on the dominant hand of each participant during the experiment. Each activity was clearly explained to the participants and asked to perform the activity for at least 30 seconds, with enough rest provided between the activities. EMG and IMU data collected from the participants' forearms were transmitted to the computer via Bluetooth, and the data were stored and labeled with the activity ID. The six participants' data were used for model building and evaluation, whereas the seventh participant (age = 26 years, weight = 76 kgs., and height = 1.65 m) was asked to perform all the activities in any sequence without any time constraint. The seventh participant's data (referred to as the new unlabeled data previously) was used to test the performance of the proposed ANN model. The whole experiment of the seventh participant was videotaped and later used for evaluating the model performance. The six participants' dataset was manually labeled, and it consists of 38 features (EMG – 32, Acc -3, and Gyro – 3) with 149,491 samples. Therefore, the input layer's size was defined as 38 nodes or neurons to hold the 38 raw data features, whereas the number of nodes in the output layer is equal to a number of activities (i.e., 15). Moreover, each activity's sample count is different since the participants performed each activity for a different duration. The imbalanced classes represent a real scenario because not all construction activities are performed for the same duration.

Once the field data was collected and labeled, data augmentation techniques such as time-wrapping, pooling, drifting, and reversing were applied to each user data, which increases the data by 4-folds. The number of samples per participant before and after the data augmentation are 22,000 and 88,000, respectively. Therefore, the total number of samples for all the participants after the data augmentation is 524,218.

4.2.2. Hyperparameter Optimization and ANN Model

The optimum hyperparameters were determined using the Bayesian TPE algorithm are set to a wide range to test different combinations, as shown in Table X. From the Bayesian TPE results, it can be observed that optimum performance was achieved for the ANN architecture shown in Figure 4. The parameters for the optimized model include four hidden layers, two dropout layers, no. of neurons for hidden layers as [2048, 2048, 512, 256], batch size = 256, epochs = 100, optimizer = Adam, and activation function = ReLu. Since the batch size of 256 is used during the model training, 256 samples from training data will be used and sent to the network in both forward and backward propagation. The number of epochs selected for model training is 100, which means the model will train 100 times for the selected batches. Since the proposed model uses the early stopping function, it stops the model training process once the model performance is stable. To overcome overfitting or underfitting, the regularization concept has also been implemented during the model training.

Table X. Parameters used for Bayesian Tree-structure Parzen Estimator optimization

Parameter	**#**	**Values**
No. of Hidden Layers	7	1 to 7
No. of Neurons	6	64, 128, 256, 512, 1024, 2048
No. of Dropout Layers	5	0.1, 0.2, 0.3, 0.4
Batch Sizes	3	128, 192, 256
Epochs	3	50, 100, 150
Optimizers	3	SGD, Adam, RMSprop
Activation Functions	3	ReLu, Tanh, Sigmoid

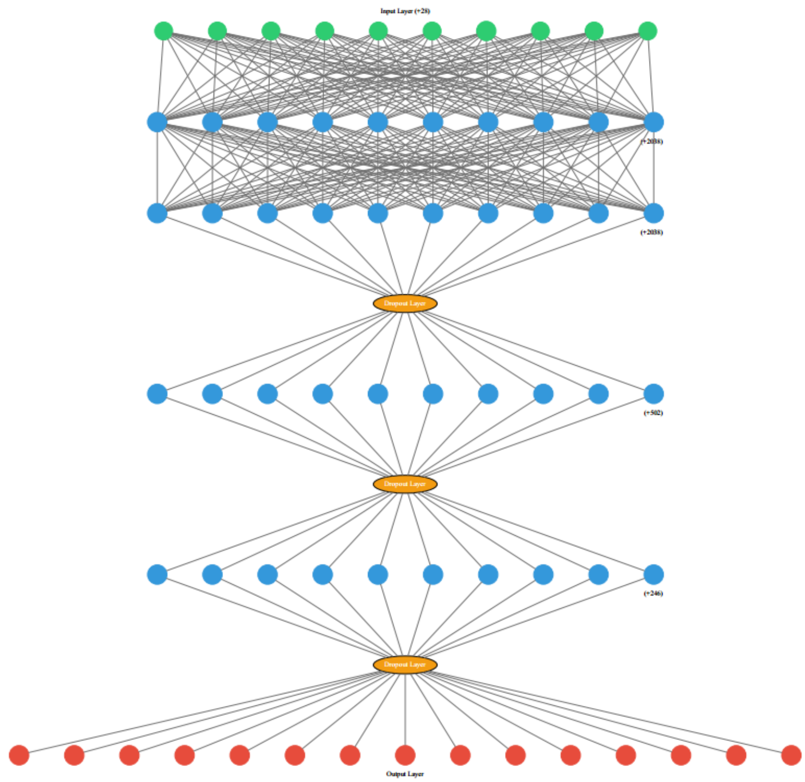

Figure 4. Optimal artificial neural network architecture for scaffold builder activities prediction

4.3.Model Learning Curves

To assess the model performance and achieve the bias-variance trade-off, learning curves have been plotted. Learning curves show the relation between training instances and accuracy. These curves show the plot of training and cross-validation scores from a given model with different training sizes and using these curves can help understand how the training and cross-validation scores are moving as the number of training instances increase. These curves tell whether the model is suffering from bias or variance. Leave One Subject Out (LOSO) cross-validation has been used to generate these learning curves. Figure 5 shows the plots of training and cross-validation scores between accuracy and different training sizes. Ten different training set sizes have been used. The model slightly shows underfitting for the initial training set, but as the training size increases, the curves are so close and stable with low bias and low variance. Moreover, the graph shows the increase in model generalization with an increase in the training set. Average training and cross-validation scores generated from the learning curves are 94.90% and 93.29%.

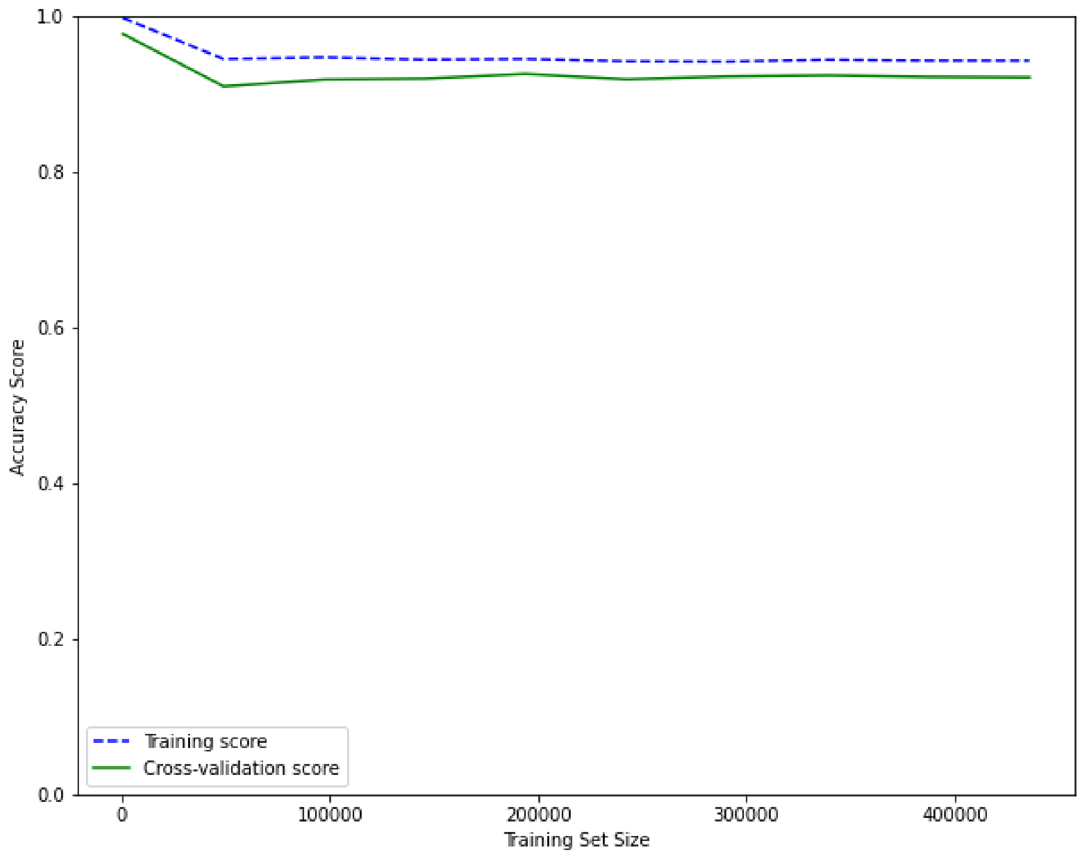

Figure 5. Learning curve of the proposed ANN model using EMG and IMU data

4.4.Performance Evaluation on Testing Data

After training the model, performance evaluation is required to understand the model's overall and class performance. Leave One Subject Out (LOSO) Cross-validation has been performed, and it splits the data into train and test based on the number of subjects. In each experiment or fold, one of the subjects is used as test data, and the remaining subjects are used as training data. As our data has six users, six experiments are performed by LOSO to evaluate the model performance using confusion matrix and classification report. Figure 6 shows the normalized confusion matrix of the proposed ANN model generated after cross-validation on the six subjects where X and Y axes represent the predicted and true classes. The diagonal cells represent the percent of correctly classified instances, whereas the off-diagonal elements represent the percent of misclassified instances for each activity. From Figure 6, it can be observed that the "Downstairs" (0.13) activity was highly misclassified among all other classes, followed by "AdjustJacks" with values (0.03), "Climb" (0.03), and "GDownstairsW" (0.03). The highest misclassification of "Downstairs" was identified with "Walk." Whereas, the highest classification was observed in "Hammering" (0.99), "CJacks" (0.97), "CScaffold" (0.97) and "ClimbW" (0.97) followed by "LiftPlank" (0.96), and " GDownstairsW " (0.96). Table 2 presents the precision, recall, and F1 score values for all the activities. The "Hammering" (0.99), "CJacks" (0.97), and "ClimbW" (0.97) activities shows highest precision. Whereas the least precision value of 0.86 was observed in "Downstairs." The highest recall value of 0.98 was observed in "CJacks" followed by "Wrenching" (0.97), "GDownstairsW" (0.97), and "LiftPlank" (0.97), whereas the lowest recall value was observed in "Downstairs" (0.89). Similarly, the F1 score is highest for "Hammering" (0.97), "CJacks" (0.97), and " LiftPlank " (0.97) and lowest is for "Downstairs" (0.87). The overall prediction accuracy of 93.68% was obtained on the testing dataset with 0.94 weighted average precision, recall, and F1 Score.

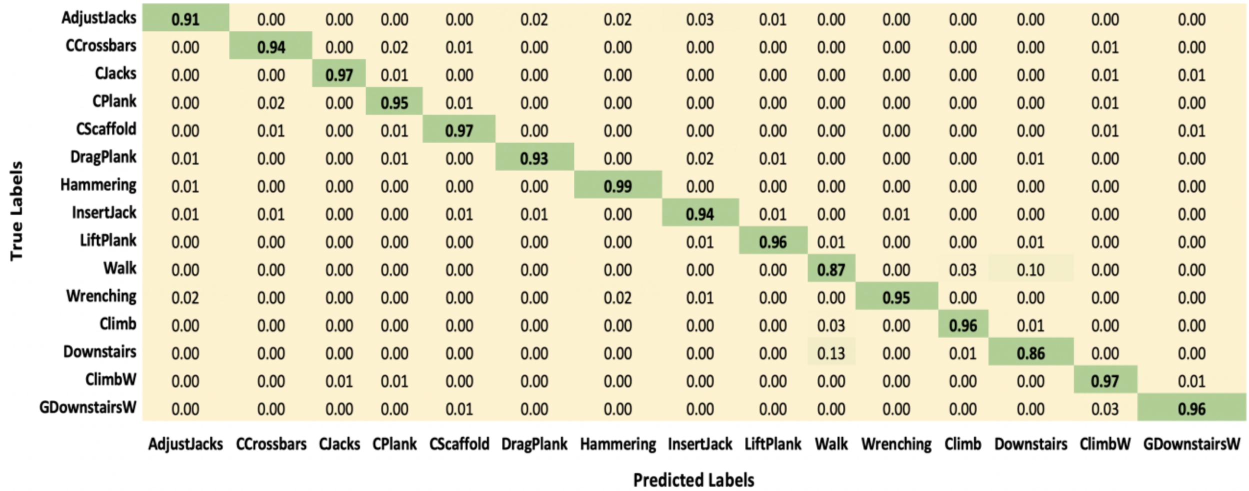

Figure 6. Confusion matrix of the proposed ANN model using EMG and IMU data

Table 2. Class report of the proposed ANN model using EMG and IMU data

	Precision	Recall	F1 Score
AdjustJacks	0.91	0.95	0.93
CCrossbars	0.94	0.96	0.95
CJacks	0.97	0.98	0.97
CPlank	0.95	0.93	0.94
CScaffold	0.96	0.95	0.96
DragPlank	0.93	0.96	0.94
Hammering	0.99	0.95	0.97
InsertJack	0.94	0.94	0.94
LiftPlank	0.96	0.97	0.97
Walk	0.87	0.93	0.9
Wrenching	0.95	0.97	0.96
Climb	0.96	0.96	0.96
Downstairs	0.86	0.89	0.87
ClimbW	0.97	0.93	0.95
GDownstairsW	0.96	0.97	0.96
Accuracy			0.94
Weighted Average	0.94	0.95	0.94

4.5. Real-Time Evaluation

The prediction was performed on the dataset collected from a new individual (seventh participant) to evaluate the model's robustness. The seventh participant's evaluation has been performed using the trained weights of ANN generated from LOSO cross-validation. The data was not used either in training or testing the model. The dataset consists of 15,850 samples and fifteen activities. During the seventh participant experiment session, the video recorded was reviewed for activity sequence and actual class labeling to build the benchmark activities for performance evaluation by matching the time stamp in video and sensor data. Figure 7 shows the confusion matrix of the proposed ANN model on the new dataset. From the confusion matrix, it can be observed the "Downstairs" (0.27) activity was highly misclassified, followed by "GDownstairsW" (0.15), "CScaffold" (0.09), and "CJacks" (0.08). The highest misclassification of "Downstairs" was observed with the "Walk" activity. Whereas the highest classification was observed in "Wrenching" (0.98) followed by "Hammering" (0.94), "LiftPlank" (0.93), and "ClimbW" (0.91). Table 3 shows the precision, recall, and F1 score results of the ANN model on an unknown dataset. The highest precision of 0.96 was observed in the "CJacks," "Wrenching," and "Climb" activities, followed by "Hammering" (0.95), "CPlank" (0.91), and "CScaffold" (0.90). Whereas, the highest recall value of 0.98 was observed in "Wrenching" followed by "Hammering" (0.94) and "LiftPlank" (0.93). The lowest recall value was observed in "Downstairs" (0.64). The F1 score was highest in "Wrenching" (0.97) activity followed by "Hammering" (0.94), "Climb" (0.92), and "CJacks" (0.91). Overall, the prediction accuracy of the proposed ANN model on an unknown dataset is 86.87% with weighted average precision (0.86), recall (0.87), and F1 Score (0.86). Moreover, Figure 8 shows the predicted and actual sequence of activities performed by the seventh participant. The proposed model recognized the activities' sequence with the highest errors in "AdjustJacks" and "Walk" activities. Also, the time ratio difference between actual and predicted classes range between 1% to 2%, as shown in Figure 9.

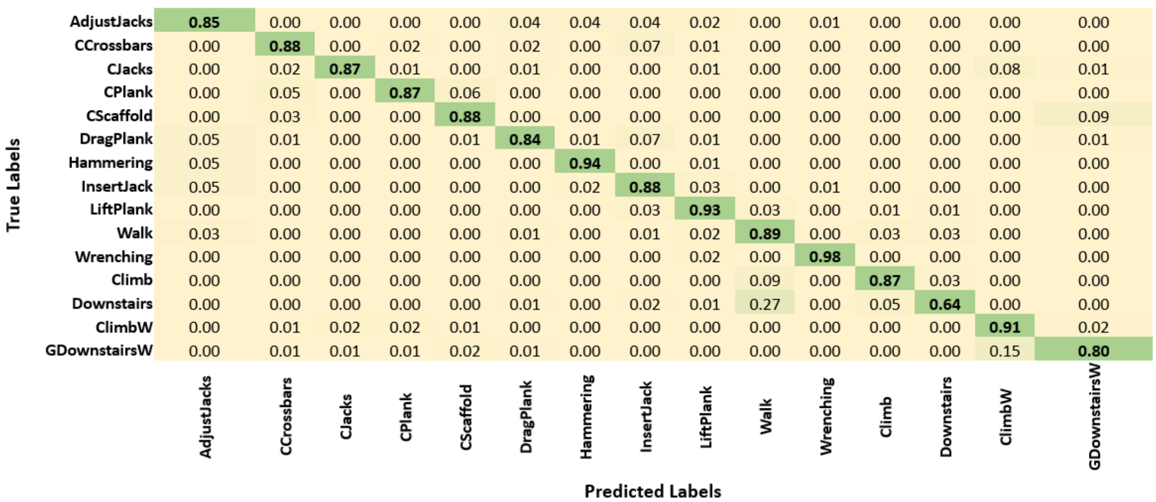

Figure 7. Confusion matrix of the proposed ANN model on the unknown dataset

Table 3. Class report of the proposed ANN model for the unknown dataset

	Precision	Recall	F1 Score	Support
AdjustJacks	0.82	0.85	0.83	1000
CCrossbars	0.78	0.88	0.83	450
CJacks	0.96	0.87	0.91	1150
CPlank	0.91	0.87	0.89	700
CScaffold	0.90	0.88	0.89	900
DragPlank	0.88	0.84	0.86	800
Hammering	0.95	0.94	0.94	1600
InsertJack	0.84	0.88	0.86	1100
LiftPlank	0.80	0.93	0.86	600
Walk	0.63	0.89	0.74	950
Wrenching	0.96	0.98	0.97	900
Climb	0.96	0.87	0.92	2400
Downstairs	0.86	0.64	0.73	900
ClimbW	0.80	0.91	0.85	1200
GDownstairsW	0.88	0.80	0.84	1200
Accuracy	0.87	0.87	0.87	0.87
Macro Average	0.86	0.87	0.86	15850
Weighted Average	0.88	0.87	0.87	15850

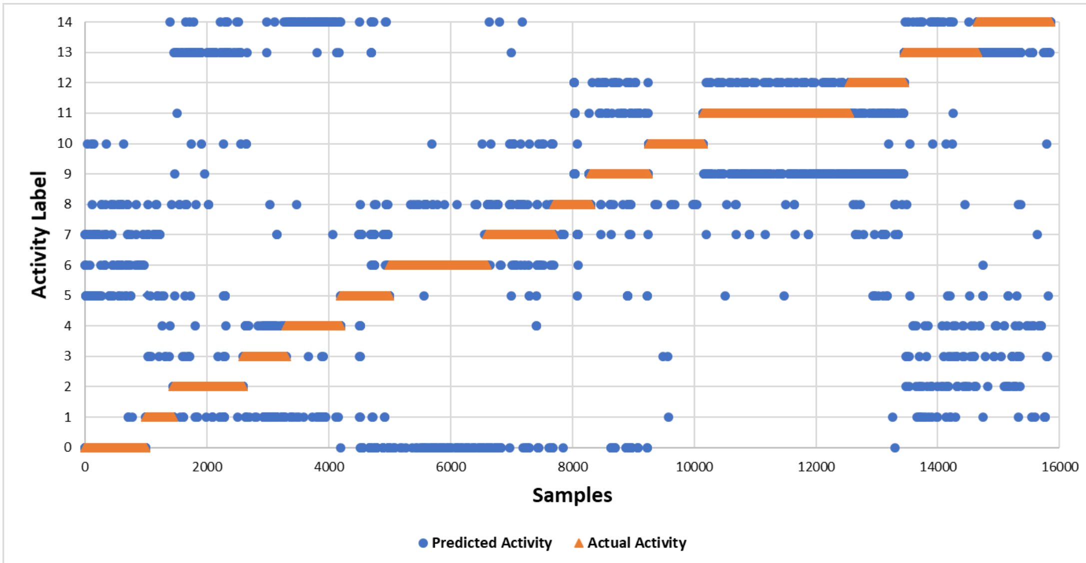

Figure 8. Performance of proposed ANN model on the unknown dataset – A plot showing predicted and actual classes over the entire session

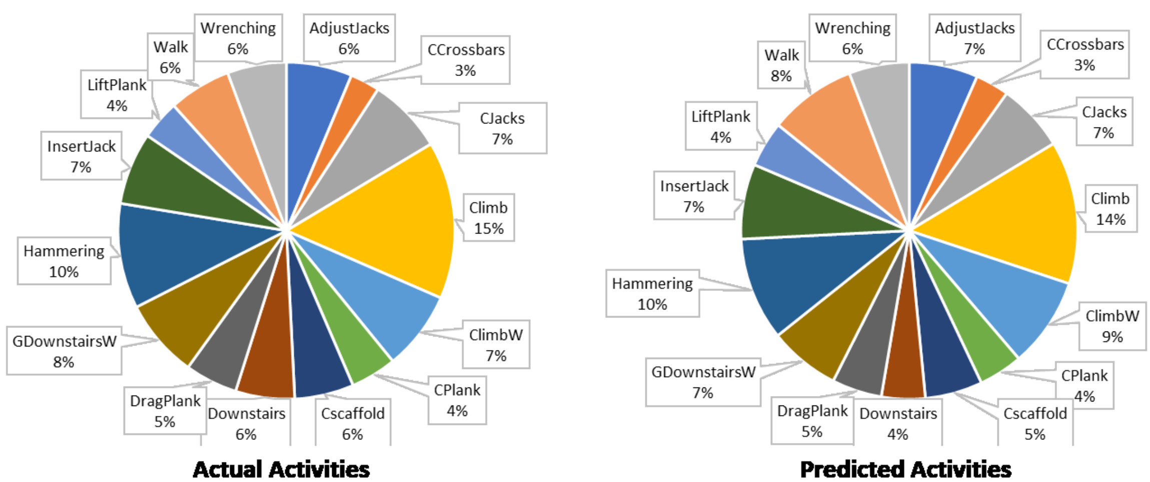

Figure 9. Time ratio of actual and predicted activities

4.6. Comparison of Activity Recognition Performance for Different Sensor Combination

To understand the ANN model's performance for different sensor combinations, individual ANN models were built using various sensor combination data, namely, EMG+IMU, IMU, EMG, Acc, and Gyro. All these models were built using the framework proposed in this study, and all the models were evaluated for performance and diagnosed for overfitting or underfit problems. The overall accuracy, weighted precision, weighted recall, and weighted F1 Score of each sensor combination's best performance model are presented in Figure 10. From Figure 11, it can be observed the EMG+IMU model achieved the highest accuracy of 93.68%, followed by EMG (75.12%), IMU (71.45%), Acc (48.00%), and Gyro (32.30%). From this analysis, it can be concluded that EMG+IMU data helps improve classification accuracy. Also, the sensor combination model performance was analyzed for different classes. As shown in Figure 11, the EMG+IMU model has outperformed other models in all classes. Between EMG and IMU models, EMG outperformed IMU in the majority of the activities. EMG models performed better for "CCrossbars", "CJacks", "CPlank", "CScaffold," "DragPlank", "Hammering", "InsertJack", "Walk", "Climb", "Downstairs", "ClimbW", and "GDownstairsW". Whereas, IMU models have better accuracy than EMG models for "AdjustJacks," "LiftPlank," and "Wrenching." For "Wrenching" activity, both EMG and IMU performed equally. The acceleration and gyroscope models have performed poorly compared to EMG+IMU, IMU, and EMG. Among acceleration and gyroscope, the acceleration models have better accuracy. From Figure 10 and Figure 11, it can be concluded that EMG+IMU features yield higher accuracy for all the classes compared to other sensor combinations.

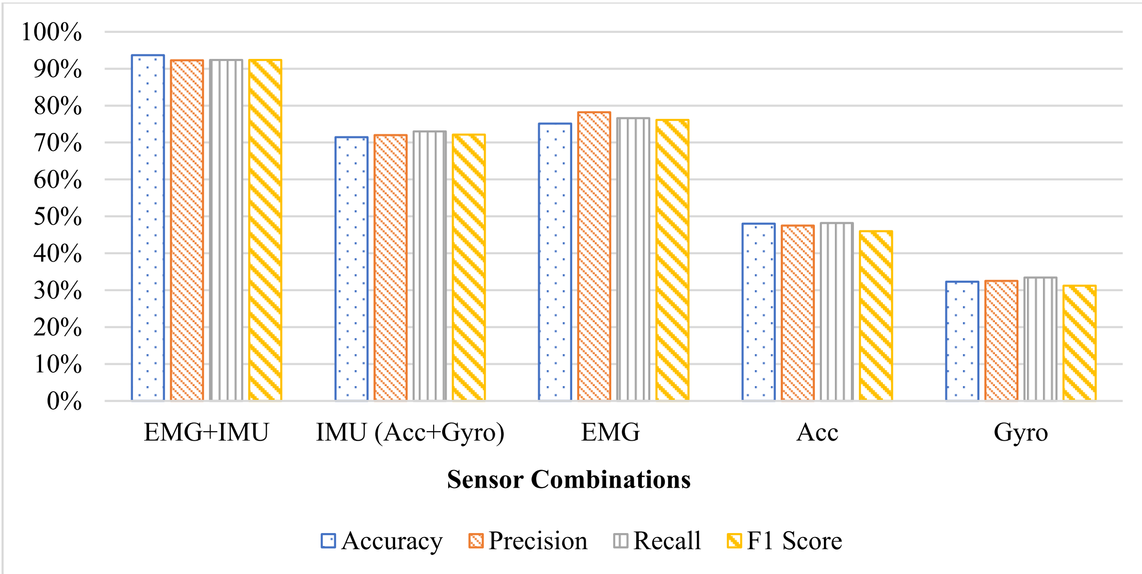

Figure 10. Comparison of classification performance of the ANN model using different sensor combination

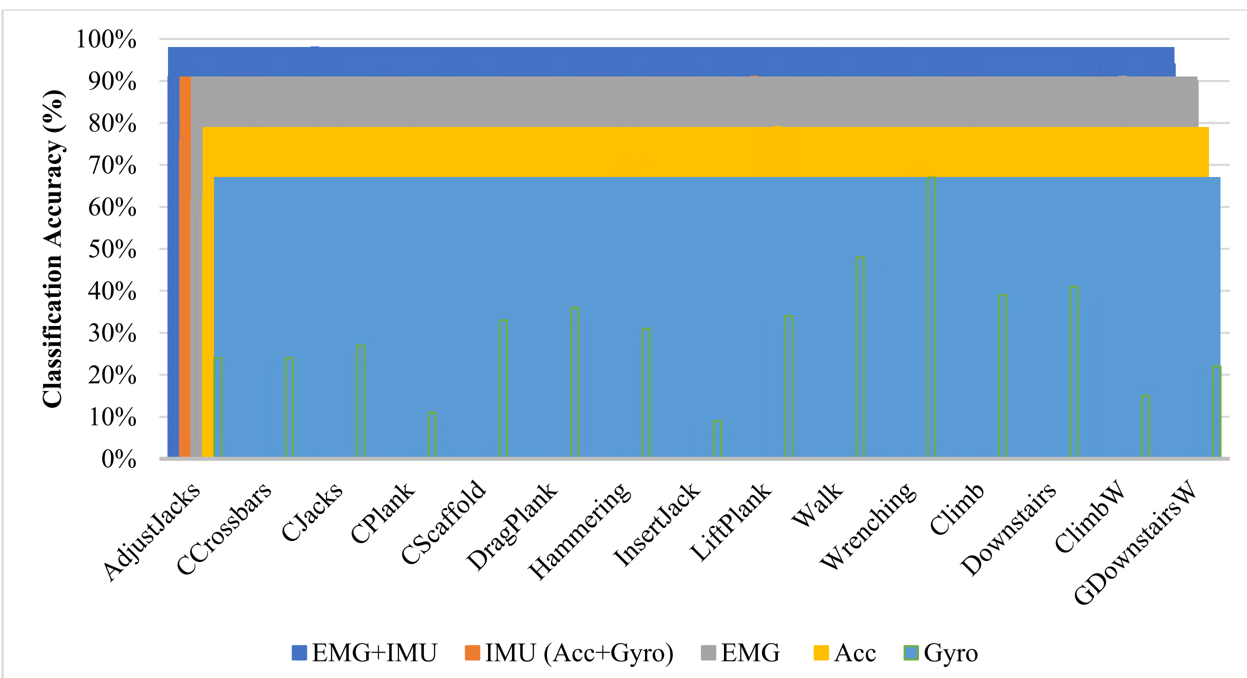

Figure 11. Comparison of classification performance of the ANN model for all activities for different sensor data

4.7. Comparison of Proposed Model with other Classification Models

It is essential to determine how well the proposed ANN model performs compared to other classification algorithms. Therefore, the EMG+IMU dataset was further used to test with other existing classification models such as Random Forest (RF), Decision Trees (DT), Gradient Boosting (XGB), Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), K-nearest Neighbors (KNN), and Multilayer Perceptron (MLP). Leave One Subject Out cross-validation technique was used to evaluate all the classifiers' performance with six-folds. Figure 12 compares the cross-validation accuracy of all the classification algorithms on the EMG+IMU dataset. It can be observed that the highest accuracy was obtained using the proposed ANN (93.68%) model, followed by XGB (85.45%), RF (84.10%), and DT (79.11%). Whereas the least accuracy was obtained in the SVM classifier case with 0.26, 0.25, and 0.24 recall, precision, and F1 Score, respectively.

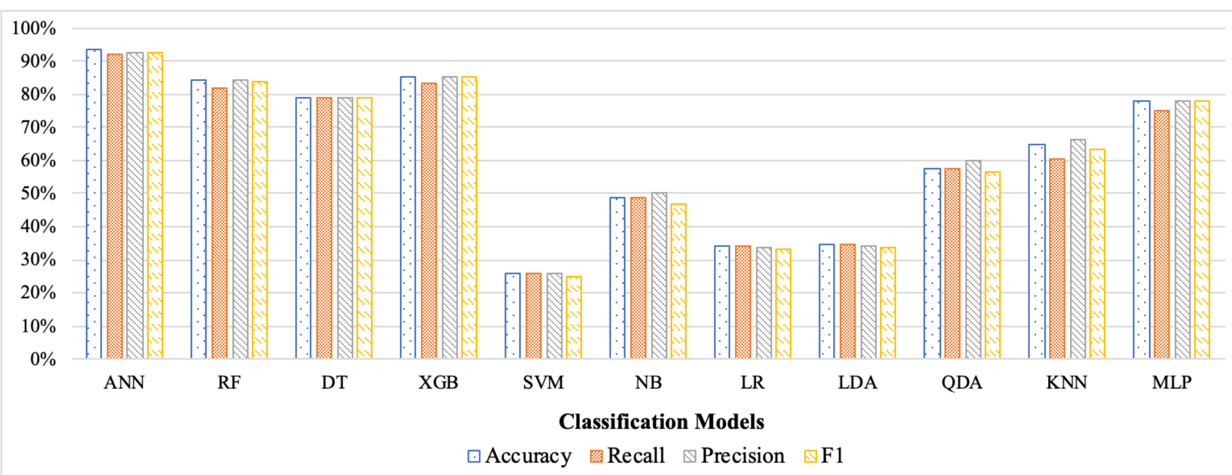

Figure 12. Comparison of activity recognition performance using the ANN model for different sensor data

5. Discussion

In this study, a case study of scaffold builder activities was conducted to evaluate the proposed worker activity recognition framework's performance using forearm EMG and IMU data from the dominant hand. The use of armband sensors on the dominant hand can recognize whole-body activities highly suitable for construction applications. The case study results show that the ANN model developed using EMG and IMU has achieved the highest average classification accuracy of 93.29% for all fifteen activities. Since the construction activities involve either muscle activity and body movement, the use of EMG and IMU helps in recognizing complex construction activities involving different motions and body parts. From the results of the current case study, the activities involving motion such as adjusting jacks, dragging plank, lifting plank, walking, wrenching, and hammering, IMU data model have shown better accuracy compared to EMG model. Whereas in the case of activities involving muscle activity or material handling such as carrying scaffold, carrying plank, carrying jack, and carrying tool bag, the EMG data model has higher accuracy than the IMU data model. These results conclude that the proposed framework can recognize activities that do not involve considerable body movement of human body parts or repetitive activities, which is one of the significant challenges of previous activity recognition models [13]. Besides, the proposed framework can recognize a more significant number of activities compared to previous models. The high precision, recall, and F1 Score of the proposed model on real-time predictions show that the model can be used for real-time worker activity monitoring for safety, productivity, and project controls applications.

One of the common issues with deep learning models is the necessity of large datasets, and this issue is addressed by incorporating a data augmentation technique in the framework. The data augmentation trains better, as seen in learning curves, and helps model generalization and removes human variabilities. As previous studies stated [6], the sensor data fusion at a lower-level (signal level) improved accuracy significantly by eliminating the redundancy and considering dependencies and correlation between different features. The signal-level data fusion improved the accuracy and eliminates the process of extracting features from the raw data, which requires domain knowledge. Besides, the hyperparameter optimization using Bayesian TPE automates network parameter selection, which helps in adopting the proposed framework for any construction activity recognition and prevents human errors. Since the proposed framework is fully automated and independent of activities, it can be extended to any trade or multiple trades by retraining the model with new activity data.

6. Limitations and Future Work

Test Subjects: As this study was initially designed to investigate the testbed before actual production on large-scale workers in a real-world environment, the experiment was performed with limited non-construction workers in a semi-construction environment. Since the proposed framework is independent of human variability and environment, retraining the model with data from construction workers enables producing large-scale and field-ready models.

Future Work: We further expect to understand how the armband sensor position (slid or rotated) on the forearm influences the activity recognition results. Even though the signal-level sensor data fusion yielded high accuracies, we want to investigate how other data fusion levels (feature-level and decision-level) will affect the activity recognition performance. The authors plan to develop one generic model to recognize multiple construction trades' activities using the proposed framework. Future research investigates the performance of recurrent neural networks such as long short-term memory (LSTM) for construction activity recognition.

7. Conclusions

This study proposes an automated construction worker activity recognition method using forearm EMG and IMU data. The proposed framework is fully automated and can be applied for any number of activities and different construction trades by retraining the model with additional training data. Moreover, data augmentation and hyperparameter optimization help achieve high accuracy with limited participant data. The proposed method was validated and evaluated through a case study on scaffold builder activities, including complex construction activities involving different body parts (wrist, forearm, upper body, lower body, and whole-body) and various motions (repetitive motion, impulsive motion, and free motion). The proposed ANN model was able to classify fifteen scaffold builder activities with an overall testing accuracy of 93.29% and real-time prediction accuracy of 87%. The sequences and time ratio plots showed that the model could successfully predict the sequence and time spent on each activity with minimal error. The performance evaluation of the ANN model on different sensor combinations showed that the classification accuracy was highest for EMG+IMU (93.29%), followed by EMG alone (75.1%) alone and IMU (71.4%) alone. The results also show that the EMG data alone performed better than IMU data alone and acceleration data alone for carrying scaffold, carrying plank, carrying crossbars, inserting jacks, and climbing stairs with weight. In contrast, the IMU data alone performed better than EMG data alone for the rest of the activities. Since most construction activities involve motion and muscle activity, EMG and IMU data have increased the accuracy of activity recognition. The proposed model was also compared with the other machine learning-based classification algorithms, and the comparisons show that the proposed model outperformed all the other classifiers.

Compared to previous studies, the main advantages of the proposed worker activity recognition system are inexpensive equipment cost, fully automated framework, low computation cost, ability to recognize complex construction activities, and recognizing more activities. Since the proposed framework is fully automated, scalable, robust, and adaptable, the system be can be commercialized. As the future direction, we will further explore the feasibility of workload assessment, fatigue monitoring, and productivity assessment using the proposed system and methodology.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. #2026575. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] R.d. Best, U.S. Construction Industry - Statistics & Facts, 2020.https://www.statista.com/topics/974/construction/, (January 12 2021)

[2] S. Kim, S. Chang, D. Castro-Lacouture, Dynamic modeling for analyzing impacts of skilled labor shortage on construction project management, Journal of Management in Engineering 36 (1) (2020) 04019035

[3] B.Y. Future, Craft Labor Demand, 2020, (July 5th 2020)

[4] H. Karimi, T.R. Taylor, P.M. Goodrum, C. Srinivasan, Quantitative analysis of the impact of craft worker availability on construction project safety performance, Construction innovation (2016)

[5] O.A. Ayodele, A. Chang-Richards, V. González, Factors affecting workforce turnover in the construction sector: a systematic review, Journal of Construction Engineering and Management 146 (2) (2020) 03119010

[6] C.R. Ahn, S. Lee, C. Sun, H. Jebelli, K. Yang, B. Choi, Wearable sensing technology applications in construction safety and health, Journal of Construction Engineering and Management 145 (11) (2019) 03119007

[7] I. Awolusi, E. Marks, M. Hallowell, Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices, Automation in construction 85 (2018) 96-106

[8] A. Aryal, A. Ghahramani, B. Becerik-Gerber, Monitoring fatigue in construction workers using physiological measurements, Automation in construction 82 (2017) 154-165

[9] J. Häikiö, J. Kallio, S.-M. Mäkelä, J. Keränen, IoT-based safety monitoring from the perspective of construction site workers, International Journal of Occupational and Environmental Safety 4 (1) (2020) 1-14

[10] S. Hwang, H. Jebelli, B. Choi, M. Choi, S. Lee, Measuring workers' emotional state during construction tasks using wearable EEG, Journal of Construction Engineering and Management 144 (7) (2018) 04018050

[11] H. Luo, C. Xiong, W. Fang, P.E. Love, B. Zhang, X. Ouyang, Convolutional neural networks: Computer vision-based workforce activity assessment in construction, Automation in construction 94 (2018) 282-289

[12] D. Roberts, M. Golparvar-Fard, End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level, Automation in construction 105 (2019) 102811

[13] B. Sherafat, C.R. Ahn, R. Akhavian, A.H. Behzadan, M. Golparvar-Fard, H. Kim, Y.-C. Lee, A. Rashidi, E.R. Azar, Automated Methods for Activity Recognition of Construction Workers and Equipment: State-of-the-Art Review, Journal of Construction Engineering and Management 146 (6) (2020) 03120002

[14] P. Chan, A. Kaka, Construction productivity measurement: A comparison of two case studies, 20th Annual ARCOM Conference, Edinburgh, Scotland, 2004

[15] C.-F. Cheng, A. Rashidi, M.A. Davenport, D.V. Anderson, C. Sabillon, Software and hardware requirements for audio-based analysis of construction operations, Reston, VA: ASCE, 2018

[16] L. Joshua, K. Varghese, Accelerometer-based activity recognition in construction, Journal of computing in civil engineering 25 (5) (2011) 370-379

[17] D. Wang, J. Chen, D. Zhao, F. Dai, C. Zheng, X. Wu, Monitoring workers' attention and vigilance in construction activities through a wireless and wearable electroencephalography system, Automation in construction 82 (2017) 122-137

[18] J. Ryu, J. Seo, H. Jebelli, S. Lee, Automated action recognition using an accelerometer-embedded wristband-type activity tracker, Journal of Construction Engineering and Management 145 (1) (2019) 04018114

[19] Z. Yang, Y. Yuan, M. Zhang, X. Zhao, B. Tian, Assessment of construction workers’ labor intensity based on wearable smartphone system, Journal of Construction Engineering and Management 145 (7) (2019) 04019039

[20] R. Akhavian, A.H. Behzadan, Coupling human activity recognition and wearable sensors for data-driven construction simulation, ITcon 23 (2018) 1-15

[21] J. Yang, Z. Shi, Z. Wu, Vision-based action recognition of construction workers using dense trajectories, Advanced Engineering Informatics 30 (3) (2016) 327-336

[22] A. Albert, M.R. Hallowell, B. Kleiner, A. Chen, M. Golparvar-Fard, Enhancing construction hazard recognition with high-fidelity augmented virtuality, Journal of Construction Engineering and Management 140 (7) (2014) 04014024

[23] A. Khosrowpour, I. Fedorov, A. Holynski, J.C. Niebles, M. Golparvar-Fard, Automated worker activity analysis in indoor environments for direct-work rate improvement from long sequences of RGB-D images, Construction Research Congress 2014: Construction in a Global Network, 2014, pp. 729-738

[24] G. Cezar, Activity recognition in construction sites using 3D accelerometer and gyrometer, Accessed Octobar 10 (2012) 2018

[25] S.H. Khan, M. Sohail, Activity monitoring of workers using single wearable inertial sensor, 2013 International Conference on Open Source Systems and Technologies, IEEE, 2013, pp. 60-67

[26] L. Joshua, K. Varghese, Selection of accelerometer location on bricklayers using decision trees, Computer-Aided Civil and Infrastructure Engineering 28 (5) (2013) 372-388

[27] K. Yang, S. Aria, C.R. Ahn, T.L. Stentz, Automated detection of near-miss fall incidents in iron workers using inertial measurement units, Construction Research Congress 2014: Construction in a Global Network, 2014, pp. 935-944

[28] T.-K. Lim, S.-M. Park, H.-C. Lee, D.-E. Lee, Artificial neural network–based slip-trip classifier using smart sensor for construction workplace, Journal of Construction Engineering and Management 142 (2) (2016) 04015065

[29] R. Akhavian, A.H. Behzadan, Smartphone-based construction workers' activity recognition and classification, Automation in construction 71 (2016) 198-209

[30] T. Zhang, Y.-C. Lee, M. Scarpiniti, A. Uncini, A supervised machine learning-based sound identification for construction activity monitoring and performance evaluation, Construction Research Congress 2018, 2018, pp. 358-366

[31] S. Deb, D.W. Carruth, R. Sween, L. Strawderman, T.M. Garrison, Efficacy of virtual reality in pedestrian safety research, Appl Ergon 65 (2017) 449-460.'doi': 10.1016/j.apergo.2017.03.007

[32] J. Bosch, N. Gracias, P. Ridao, D. Ribas, Omnidirectional underwater camera design and calibration, Sensors (Basel) 15 (3) (2015) 6033-6065.'doi': 10.3390/s150306033

[33] P. Vepakomma, D. De, S.K. Das, S. Bhansali, A-Wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities, 2015 IEEE 12th International conference on wearable and implantable body sensor networks (BSN), IEEE, 2015, pp. 1-6

[34] D. Cook, K.D. Feuz, N.C. Krishnan, Transfer Learning for Activity Recognition: A Survey, Knowl Inf Syst 36 (3) (2013) 537-556.'doi': 10.1007/s10115-013-0665-3

[35] J. Wang, Y. Chen, S. Hao, X. Peng, L. Hu, Deep learning for sensor-based activity recognition: A survey, Pattern Recognition Letters 119 (2019) 3-11

[36] O.D. Lara, M.A. Labrador, A survey on human activity recognition using wearable sensors, IEEE communications surveys & tutorials 15 (3) (2012) 1192-1209

[37] E. Munguia Tapia, Using machine learning for real-time activity recognition and estimation of energy expenditure, Massachusetts Institute of Technology, 2008

[38] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, Y. Liu, Deep learning for sensor-based human activity recognition: overview, challenges and opportunities, arXiv preprint arXiv:2001.07416 (2020)

[39] Y. Bengio, Deep learning of representations: Looking forward, International Conference on Statistical Language and Speech Processing, Springer, 2013, pp. 1-37

[40] K.M. Rashid, J. Louis, Times-series data augmentation and deep learning for construction equipment activity recognition, Advanced Engineering Informatics 42 (2019) 100944

[41] H. Lee, C.R. Ahn, N. Choi, T. Kim, H. Lee, The effects of housing environments on the performance of activity-recognition systems using Wi-Fi channel state information: An exploratory study, Sensors 19 (5) (2019) 983

[42] J.A. Álvarez-García, P. Barsocchi, S. Chessa, D. Salvi, Evaluation of localization and activity recognition systems for ambient assisted living: The experience of the 2012 EvAAL competition, Journal of Ambient Intelligence and Smart Environments 5 (1) (2013) 119-132

[43] F.J. Ordonez, D. Roggen, Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition, Sensors (Basel) 16 (1) (2016) 115.'doi': 10.3390/s16010115

[44] R. Khusainov, D. Azzi, I.E. Achumba, S.D. Bersch, Real-time human ambulation, activity, and physiological monitoring: taxonomy of issues, techniques, applications, challenges and limitations, Sensors (Basel) 13 (10) (2013) 12852-12902.'doi': 10.3390/s131012852

[45] S. Kim, M.A. Nussbaum, Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks, Ergonomics 56 (2) (2013) 314-326.'doi': 10.1080/00140139.2012.742932

[46] T. Seel, J. Raisch, T. Schauer, IMU-based joint angle measurement for gait analysis, Sensors (Basel) 14 (4) (2014) 6891-6909.'doi': 10.3390/s140406891

[47] A. Alwasel, E.M. Abdel-Rahman, C.T. Haas, S. Lee, Experience, productivity, and musculoskeletal injury among masonry workers, Journal of Construction Engineering and Management 143 (6) (2017) 05017003

[48] J. Chen, J. Qiu, C. Ahn, Construction worker's awkward posture recognition through supervised motion tensor decomposition, Automation in construction 77 (2017) 67-81

[49] N.D. Nath, R. Akhavian, A.H. Behzadan, Ergonomic analysis of construction worker's body postures using wearable mobile sensors, Applied ergonomics 62 (2017) 107-117

[50] T.-K. Lim, S.-M. Park, H.-C. Lee, D.-E. Lee, Artificial neural network–based slip-trip classifier using smart sensor for construction workplace, Journal of Construction Engineering and Management 142 (2) (2015) 04015065

[51] K. Yang, C.R. Ahn, M.C. Vuran, S.S. Aria, Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit, Automation in construction 68 (2016) 194-202

[52] W. Wu, H. Yang, D.A. Chew, S.-h. Yang, A.G. Gibb, Q. Li, Towards an autonomous real-time tracking system of near-miss accidents on construction sites, Automation in construction 19 (2) (2010) 134-141

[53] S. Hwang, J. Seo, H. Jebelli, S. Lee, Feasibility analysis of heart rate monitoring of construction workers using a photoplethysmography (PPG) sensor embedded in a wristband-type activity tracker, Automation in construction 71 (2016) 372-381

[54] H. Jebelli, B. Choi, H. Kim, S. Lee, Feasibility study of a wristband-type wearable sensor to understand construction workers' physical and mental status, Construction Research Congress, 2018, pp. 367-377

[55] A.D. Nimbarte, F. Aghazadeh, L.H. Ikuma, C.M. Harvey, Neck disorders among construction workers: understanding the physical loads on the cervical spine during static lifting tasks, Industrial health 48 (2) (2010) 145-153

[56] T.S. Abdelhamid, J.G. Everett, Physiological demands during construction work, Journal of Construction Engineering and Management 128 (5) (2002) 427-437

[57] F.L. Chang, Y.M. Sun, K.H. Chuang, D.J. Hsu, Work fatigue and physiological symptoms in different occupations of high-elevation construction workers, Appl Ergon 40 (4) (2009) 591-596.'doi': 10.1016/j.apergo.2008.04.017

[58] Z.S. Maman, M.A.A. Yazdi, L.A. Cavuoto, F.M. Megahed, A data-driven approach to modeling physical fatigue in the workplace using wearable sensors, Applied ergonomics 65 (2017) 515-529

[59] S. Hasanzadeh, B. Esmaeili, M.D. Dodd, Measuring the impacts of safety knowledge on construction workers' attentional allocation and hazard detection using remote eye-tracking technology, Journal of Management in Engineering 33 (5) (2017) 04017024

[60] I. Jeelani, K. Han, A. Albert, Automating and scaling personalized safety training using eye-tracking data, Automation in construction 93 (2018) 63-77

[61] H. Jebelli, S. Lee, Feasibility of Wearable Electromyography (EMG) to Assess Construction Workers' Muscle Fatigue, Advances in Informatics and Computing in Civil and Construction Engineering, Springer, 2019, pp. 181-187.'doi':

[62] H. Jebelli, M.M. Khalili, S. Lee, Mobile EEG-based workers' stress recognition by applying deep neural network, Advances in Informatics and Computing in Civil and Construction Engineering, Springer, 2019, pp. 173-180.'doi':

[63] J. Gong, C.H. Caldas, An intelligent video computing method for automated productivity analysis of cyclic construction operations, Computing in Civil Engineering (2009), 2009, pp. 64-73.'doi':

[64] C. Kim, T. Park, H. Lim, H. Kim, On-site construction management using mobile computing technology, Automation in construction 35 (2013) 415-423

[65] L. Joshua, K. Varghese, Accelerometer-based activity recognition in construction, Journal of computing in civil engineering 25 (5) (2010) 370-379

[66] L. Joshua, K. Varghese, Automated recognition of construction labour activity using accelerometers in field situations, International Journal of Productivity and Performance Management 63 (7) (2014) 841-862

[67] J. Ryu, J. Seo, H. Jebelli, S. Lee, Automated action recognition using an accelerometer-embedded wristband-type activity tracker, Journal of Construction Engineering and Management 145 (1) (2018) 04018114

[68] T. Cheng, J. Teizer, G.C. Migliaccio, U.C. Gatti, Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data, Automation in construction 29 (2013) 24-39

[69] N.D. Nath, T. Chaspari, A.H. Behzadan, Automated ergonomic risk monitoring using body-mounted sensors and machine learning, Advanced Engineering Informatics 38 (2018) 514-526

[70] L. Bao, S.S. Intille, Activity recognition from user-annotated acceleration data, International conference on pervasive computing, Springer, 2004, pp. 1-17

[71] C.C. Yang, Y.L. Hsu, A review of accelerometry-based wearable motion detectors for physical activity monitoring, Sensors (Basel) 10 (8) (2010) 7772-7788.'doi': 10.3390/s100807772

[72] N. Ravi, N. Dandekar, P. Mysore, M.L. Littman, Activity recognition from accelerometer data, Aaai, Vol. 5, 2005, pp. 1541-1546

[73] S. Chernbumroong, A.S. Atkins, H. Yu, Activity classification using a single wrist-worn accelerometer, 2011 5th International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA) Proceedings, IEEE, 2011, pp. 1-6

[74] M. Shoaib, S. Bosch, O. Incel, H. Scholten, P. Havinga, Complex human activity recognition using smartphone and wrist-worn motion sensors, Sensors 16 (4) (2016) 426

[75] R. Akhavian, A.H. Behzadan, Productivity analysis of construction worker activities using smartphone sensors, Proc., 16th Int. Conf. Comput. Civil Building Eng, 2016

[76] A. Bayat, M. Pomplun, D.A. Tran, A study on human activity recognition using accelerometer data from smartphones, Procedia Computer Science 34 (2014) 450-457

[77] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers, ACM SigKDD Explorations Newsletter 12 (2) (2011) 74-82

[78] S.S. Intille, L. Bao, E.M. Tapia, J. Rondoni, Acquiring in situ training data for context-aware ubiquitous computing applications, Proceedings of the SIGCHI conference on Human factors in computing systems, 2004, pp. 1-8

[79] J. Yang, Toward physical activity diary: motion recognition using simple acceleration features with mobile phones, Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics, ACM, 2009, pp. 1-10

[80] M.-K. Tsai, Automatically determining accidental falls in field surveying: A case study of integrating accelerometer determination and image recognition, Safety science 66 (2014) 19-26

[81] K. Yang, H. Jebelli, C. Ahn, M. Vuran, Threshold-based approach to detect near-miss falls of iron workers using inertial measurement units, Computing in Civil Engineering 2015, 2015, pp. 148-155.'doi':

[82] H. Koskimaki, V. Huikari, P. Siirtola, P. Laurinen, J. Roning, Activity recognition using a wrist-worn inertial measurement unit: A case study for industrial assembly lines, 2009 17th Mediterranean Conference on Control and Automation, IEEE, 2009, pp. 401-405

[83] O.C. Ann, L.B. Theng, Human activity recognition: a review, 2014 IEEE international conference on control system, computing and engineering (ICCSCE 2014), IEEE, 2014, pp. 389-393

[84] Y.K. Cho, K. Kim, S. Ma, J. Ueda, A robotic wearable exoskeleton for construction worker's safety and health, ASCE Construction Research Congress, 2018, pp. 19-28

[85] M. Sathiyanarayanan, S. Rajan, MYO Armband for physiotherapy healthcare: A case study using gesture recognition application, 2016 8th International Conference on Communication Systems and Networks (COMSNETS), IEEE, 2016, pp. 1-6

[86] R. Guérit, A. Cierro, J. Vanderdonckt, J.L. Pérez-Medina, Gesture Elicitation and Usability Testing for an Armband Interacting with Netflix and Spotify, International Conference on Information Technology & Systems, Springer, 2019, pp. 625-637

[87] M. Shanker, M.Y. Hu, M.S. Hung, Effect of data standardization on neural network training, Omega 24 (4) (1996) 385-397

[88] F. Chollet, others. 2015, Keras: Deep learning library for theano and tensorflow. URL: https://keras. io/k (2015)

[89] L. Yang, A. Shami, On hyperparameter optimization of machine learning algorithms: Theory and practice, Neurocomputing 415 (2020) 295-316

[90] Z. Zhang, Improved adam optimizer for deep neural networks, 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), IEEE, 2018, pp. 1-2

[91] D. Gholamiangonabadi, N. Kiselov, K. Grolinger, Deep Neural Networks for Human Activity Recognition With Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection, IEEE Access 8 (2020) 133982-133994

[92] OSHA, Scaffolding eTool, 2020.https://www.osha.gov/SLTC/etools/scaffolding/index.html, (Juky 6th 2020)

[93] K.M. Halperin, M. McCann, An evaluation of scaffold safety at construction sites, Journal of safety research 35 (2) (2004) 141-150

[94] J.C. Rubio-Romero, M.C.R. Gámez, J.A. Carrillo-Castrillo, Analysis of the safety conditions of scaffolding on construction sites, Safety science 55 (2013) 160-164