

Louisiana State University

## LSU Scholarly Repository

---

LSU Doctoral Dissertations

Graduate School

---

May 2021

# Fatigue Monitoring Through Wearable Sensors for Construction Workers

Srikanth Sagar Bangaru

*Louisiana State University and Agricultural and Mechanical College*

Follow this and additional works at: [https://repository.lsu.edu/gradschool\\_dissertations](https://repository.lsu.edu/gradschool_dissertations)



Part of the Biomechanical Engineering Commons, Construction Engineering and Management Commons, Ergonomics Commons, and the Robotics Commons

---

### Recommended Citation

Bangaru, Srikanth Sagar, "Fatigue Monitoring Through Wearable Sensors for Construction Workers" (2021). *LSU Doctoral Dissertations*. 5544.

[https://repository.lsu.edu/gradschool\\_dissertations/5544](https://repository.lsu.edu/gradschool_dissertations/5544)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Scholarly Repository. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Scholarly Repository. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

# **FATIGUE MONITORING THROUGH WEARABLE SENSORS FOR CONSTRUCTION WORKERS**

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

Engineering Science

by

Srikanth Sagar Bangaru  
B.S., National Institute of Technology Warangal, 2015  
M.S., Oklahoma State University, 2017  
M.S., Louisiana State University, 2020  
August 2021

© 2021 Srikanth Sagar Bangaru

This dissertation is dedicated to:

**My Dear Parents, Vimala Bangaru and Hanmanthu Bangaru**  
for their endless love, sacrifices, prayers, support, and encouragement.

**Wonderful Akila Reddy Palamoor,**  
for her constant source of support and encouragement.

**My Advisors, Teachers, Friends, Family, and Fellow Members,**  
without whom I would never have been able to finish this dissertation.

## **ACKNOWLEDGEMENTS**

First and foremost, I would like to extend my sincere gratitude to my advisor, Dr. Chao Wang, who gave me the opportunity to join his research group Intelligent Construction Management Lab (ICML) and work under his guidance. I could not thank him enough for his motivation, constant support, guidance, and encouragement that he has given me in the research. Furthermore, I would like to thank him for being a remarkable mentor throughout my Ph.D. career. Ultimately, I would like to mention that the work included in this thesis and the work that I conduct beyond graduate school would not be possible without his support.

I sincerely thank my co-advisor Dr. Fereydoun Aghazadeh, for his support, motivation, and encouragement throughout my Ph.D. study at Louisiana State University. Professor Aghazadeh provided me with his insights and support for the research study and encouraged me to grow as a multifaceted individual.

I want to thank Dr. Yimin Zhu and Dr. Hsiao-Chun Wu for their valuable inputs, suggestions, support to complete my Ph.D. study.

I want to thank my friend Sri Aditya Busam for spending many brainstorming sessions to solve complex problems related to the study.

I want to thank my friends Shashank Muley, Vamsi K V Pusapati, Tarun Ayiluri, Girish Rantala, Rahul Paramkusam, Kaushik Sunder, Nandaguru Durai, Vikas Reddy, Srikanth Reddy Palamoor, and David Imuetinyan for their constant support and for helping me in collecting the data for the Ph.D. study.

I want to thank my ICML lab friend Entai Xu, Xu Zhou, and Nick Hardy for their motivation, encouragement, and brainstorming sessions.

I want to take this opportunity to thank my friends Nikhil Ranga, Madhu Gunda, Chandra Deepak Reddy, Sindhura Reddy, and others for their love, support, and encouragement. Finally, I thank my parents Mr. & Mrs. Hanmanthu Bangaru, and my family members, for their unconditional love and constant support throughout my life.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iv
LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
ABSTRACT .....	xiii
CHAPTER 1. INTRODUCTION .....	1
1.1. Background and Problem Statement .....	1
1.2. Research Objectives and Approach.....	5
1.3. Organization of the Dissertation .....	11
CHAPTER 2. LITERATURE REVIEW .....	14
2.1. Definition and Causes of Fatigue .....	14
2.2. Current Approaches for Fatigue Measurement .....	15
2.3. Point of Departure .....	20
CHAPTER 3. SYSTEM DEVELOPMENT, DATA QUALITY, AND RELIABILITY ASSESSMENT .....	21
3.1. System Development.....	21
3.2. Data Quality and Reliability Assessment.....	22
3.3. Results .....	39
3.4. Discussion .....	54
3.5. Conclusions .....	58
CHAPTER 4. AUTOMATED CONSTRUCTION ACTIVITY RECOGNITION .....	61
4.1. Introduction .....	61
4.2. Literature Review .....	64
4.3. Research Methodology .....	73
4.4. Model Building, Training, and Evaluation.....	76
4.5. System Feasibility Validation and Performance Evaluation .....	86
4.6. Discussion .....	98
4.7. Limitations and Future Work .....	99
4.8. Conclusions .....	100
CHAPTER 5. MAXIMUM AEROBIC CAPACITY FOR CONSTRUCTION ACTIVITIES ..	102
5.1. Introduction .....	102
5.2. Materials and Methods .....	105
5.3. Results .....	109
5.4. Discussion .....	111
5.5. Conclusions .....	112

<b>CHAPTER 6. OXYGEN UPTAKE PREDICTION DURING CONSTRUCTION ACTIVITIES.....</b>	<b>114</b>
6.1. Introduction .....	114
6.2. Materials and Methods .....	117
6.3. Results .....	127
6.4. Discussion .....	131
6.5. Conclusion.....	135
<b>CHAPTER 7. FATIGUE MONITORING SYSTEM FEASIBILITY VALIDATION AND PERFORMANCE EVALUATION.....</b>	<b>137</b>
7.1. Introduction .....	137
7.2. Aerobic Fatigue Threshold – Feasibility Validation and Performance Evaluation ....	137
7.3. Fatigue Monitoring System – Feasibility Validation and Performance Evaluation....	144
7.4. Discussion .....	152
7.5. Conclusions .....	155
<b>CHAPTER 8. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH.....</b>	<b>156</b>
8.1. Summary of Research .....	156
8.2. Recommendations for Future Research .....	161
<b>APPENDIX A. IRB APPROVAL FORMS.....</b>	<b>164</b>
<b>APPENDIX B. SAMPLE RAW ARMBAND SENSOR DATA .....</b>	<b>166</b>
<b>APPENDIX C. FATIGUE LEVEL CLASSIFIER PERFORMANCE .....</b>	<b>167</b>
<b>APPENDIX D. COPYRIGHT PERMISSION (CHAPTER 3) .....</b>	<b>168</b>
<b>APPENDIX E. COPYRIGHT PERMISSION (CHAPTER 4).....</b>	<b>169</b>
<b>LIST OF REFERENCES .....</b>	<b>170</b>
<b>VITA .....</b>	<b>190</b>

## LIST OF TABLES

3.1. Comparison of EMG and IMU data quality of Myo armband and conventional sensors. ....	41
3.2. Comparison of EMG and IMU data quality between indoor and outdoor for at-rest activity. ....	41
3.3. Comparison of EMG and IMU data quality between two armband sensors for at-rest activity. ....	41
3.4. Effect of confounding factors on data quality for at-rest activity. ....	41
3.5. Accelerometer trial-to-trial reliability .....	47
3.6. Gyroscope trial-to-trial reliability assessment .....	48
3.7. EMG trial-to-trial reliability assessment.....	48
3.8. Accelerometer, gyroscope, and EMG day-to-day reliability assessment. ....	49
3.9. ML-based classifier performance on Day-1 dataset. ....	49
3.10. ML-based classifier performance on Day-2 dataset. ....	50
3.11. Confusion matrix and class report of random forest classifier on Day-1 dataset. ....	50
3.12. Confusion matrix and class report of random forest classifier on Day-2 dataset. ....	51
3.13. Classifier performance on lifting different weights. ....	52
3.14. Confusion matrix and class report of random forest classifier on different lifting weights.	53
3.15. Overall classification accuracy for different sensor combinations for controlled and uncontrolled datasets. ....	54
4.1. Scaffold builder activities and activity ID .....	87
4.2. Features extracted from raw EMG and IMU data .....	89
4.3. Parameters used for Bayesian Tree-structure Parzen Estimator optimization.....	90
4.4. Class report of the proposed BiLSTM model using EMG and IMU data .....	92
4.5. Class report of the proposed BiLSTM model for the unseen dataset .....	94

5.1. VO <sub>2</sub> max estimation equations for construction activities using submaximal exercise test	109
5.2. Heart rate and oxygen consumption response to different activity intensities .....	111
6.1. Scaffold building activities .....	118
6.2. Feature extraction and selection from raw EMG and IMU data.....	123
6.3. LOSO cross-validation metrics of proposed BiLSTM model on test and unseen data .....	127
6.4. Measured and estimated oxygen uptake to build one scaffold unit using the proposed model .....	130
6.5. LOSO CV metrics for different recurrent neural networks and sensor combinations.....	131
6.6. Comparison of activities and sensor of the current and previous studies on oxygen uptake prediction .....	133
6.7. Comparison of model performance of the current and previous studies on oxygen uptake prediction .....	134
7.1. MAC values for different construction activities.....	138
7.2. Scaffold building activities description and activity ID. ....	140
7.3. Fatigue rating scale along with verbal anchors and level of fatigue. ....	140
7.4. Average AFT for each fatigue level.....	142
7.5. Classification accuracies for different feature combinations.....	143
7.6. Classification F1 Score for different feature combinations. ....	143
7.7. Activity sequence and duration of unseen dataset .....	146
7.8. Fatigue level assessment using predicted AFT features for every 1-min, 2-min, and each activity .....	150
7.9. Chi-Squared test between actual and predicted fatigue levels for different window sizes..	152
C. 1. Classification recall for different feature combinations .....	175
C. 2. Classification precision for different feature combinations .....	175

## LIST OF FIGURES

2.1.Causes of physical fatigue .....	15
2.2. Physical fatigue measurement methods .....	15
3.1. Framework of the proposed fatigue monitoring system .....	21
3.2. (a) Myo armband electrode location and (b) Myo armband placement on the forearm. ....	29
3.3. Position of conventional (a) inertial measurement unit IMU and (b) electromyography (EMG) along with armband sensor.....	32
3.4. Three armband sensor positions to test the effect of sensor position on data quality: (a) rotated, (b) standard, and (c) slid down. ....	32
3.5. Evolution of yaw angle of armband sensor while (a) stationary on the body and (b) lying on the floor.....	42
3.6. Acceleration magnitude while lifting using (a) conventional sensor and (b) armband sensor... ...	43
3.7. EMG RMS plots for lifting activity using (a) conventional EMG sensor and (b) armband sensor.....	43
3.8. Comparison of (a) acceleration and (b) gyroscope magnitude for three (rotated, standard, and slid) sensor positions. ....	44
3.9. Comparison of RMS values of EMG-3, 4, 5, and 6 channels for three (rotated, standard, and slid) sensor positions.....	45
3.10. Correlation of features for three (10 lbs, 25 lbs, and 50 lbs) lifting weights. ....	53
4.1. Framework for construction worker activity recognition using forearm-based EMG and IMU armband sensor .....	74
4.2. (a) Myo armband electrode location and IMU axes directions; (b) Myo armband placement on the forearm .....	75
4.3. LSTM cell structure .....	78
4.4. Shows a few scaffold builder activities performed in a warehouse environment (a) installing guardrail, (b) going up vertical ladder, (c) hammering, (d) carrying baseboard, (e) installing baseboard on a different level, (f) installing crossbars, and (g) carrying scaffold frame .....	88

4.5. Optimal BiLSTM architecture for scaffold builder activities prediction.....	90
4.6. Confusion matrix of the proposed BiLSTM model using EMG and IMU data .....	92
4.7. Confusion matrix of the proposed BiLSTM model on the unseen dataset .....	94
4.8. Performance of proposed BiLSTM model on the unseen dataset – A plot showing predicted and actual classes over the entire session .....	95
4.9. Time ratio of actual and predicted activities.....	95
4.10. Comparison of classification performance of the BiLSTM model using different sensor combination on the unseen dataset .....	96
4.11. Comparison of classification performance of the BiLSTM model for all activities for different sensor data on the unseen dataset.....	97
4.12. Comparison of activity recognition performance using the ANN model for different sensor data .....	98
5.1. Average MAC value for different construction activities.....	110
6.1. Shows participants' performing different scaffold building activities. (a) hammering, (b) carrying scaffold frame, (c) carrying baseboard, (d) carrying guardrail, (e) going up the vertical ladder, (f) installing baseboard on a different level, (g) installing crossbars, and (i) adjusting leveling jacks .....	119
6.2. Participant wearing forearm Myo armband sensor and metabolic analyzer.....	120
6.3. Proposed approach to develop BiLSTM-based oxygen prediction model using forearm EMG and IMU data .....	121
6.4. The structure of LSTM cell.....	124
6.5. Architecture of proposed BiLSTM-based oxygen prediction using EMG and IMU.....	125
6.6. Oxygen uptake prediction on test Subject#1 using proposed BiLSTM model.....	128
6.7. Learning curve of the proposed BiLSTM model using EMG and IMU data .....	128
6.8. Correlation analysis for measured and predicted VO <sub>2</sub> on the unseen dataset .....	129
7.1.Data analysis protocol to evaluate the feasibility and performance of using aerobic fatigue threshold for fatigue monitoring .....	138

7.2. Average AFT for each activity .....	142
7.3. Confusion matrix for decision tree classifier using AFT features .....	144
7.4. Data analysis protocol to evaluate the feasibility and performance of the proposed fatigue monitoring system .....	145
7.5. Actual and predicted AFT for every 1-second on the unseen dataset.....	147
7.6. Actual and predicted AFT for every 1-minute on the unseen dataset.....	148
7.7. Actual and predicted AFT for every 2-minute on the unseen dataset.....	148
7.8. Actual and predicted AFT for each activity on the unseen dataset.....	148
7.9. Correlation analysis for actual and predicted AFT values on the unseen dataset for (a) every 1-second, (b) average of 1-min, (c) average of 2-min, and (d) average for each activity .....	149
7.10. Confusion matrix of fatigue level assessment using predicted AFT features (a) for 1-min, (b) for 2-min, and (c) for each activity .....	150
7.11. Predicted and actual fatigue level for every 1-min on the unseen dataset .....	151
7.12. Predicted and actual fatigue level for every 2-min on the unseen dataset .....	152
7.13. Predicted and actual fatigue level for each activity on the unseen dataset .....	152
B. 1. Sample raw EMG data from armband sensor .....	174
B. 2. Sample raw IMU data from armband sensor .....	174

## **ABSTRACT**

About 40% of the US construction workforce experiences high-level fatigue, which leads to poor judgment, increased risk of injuries, a decrease in productivity, and a lower quality of work. Excessive fatigue from working in unpleasant working conditions, long working hours, or heavy workloads can aggravate fatigue's adverse effects, leading to work-related musculoskeletal disorders (WMSDs) and productivity loss. Therefore, it is essential to monitor fatigue to reduce the adverse effects and preventing long-term health problems. However, since fatigue demonstrates itself in several complex processes, there is no single standard measurement method for fatigue detection. This research aims to develop a system for continuous workers' fatigue monitoring by predicting aerobic fatigue threshold (AFT) according to forearm muscle activity and motion data. The forearm muscle activity and motion data were acquired using a low-cost, non-invasive, wearable sensor. The proposed fatigue monitoring system consists of multiple measurable frameworks with five objectives: (1) assess the data quality and reliability of forearm motion and muscle activity data, (2) develop and validate the construction workers' activity recognition framework, (3) estimate construction activity-specific maximum aerobic capacity, (4) develop and validate continuous oxygen uptake prediction framework, and (5) develop fatigue level classifier using AFT features and validate the proposed fatigue monitoring system. The proposed system was evaluated on the participants performing fourteen scaffold building activities. The results show that the AFT features have achieved a higher accuracy of 92.31% in assessing the workers' fatigue level compared to heart rate (51.28%) and percentage heart rate reserve (50.43%) features. Moreover, the overall performance of the proposed fatigue monitoring system on unseen data using average 2-min AFT features was 76.74%. The study validates the feasibility of using forearm muscle activity and motion data to monitor the workers' fatigue level

continuously. The performance of the proposed system shows some promising potentials that it can be applied on the construction field to help assess worker's physiological status, evaluate the physical workload of the activity, quantify the direct impacts of the fatigue level on the accidents, and enhance the workers' safety, health, and productivity through early detection of risk.

# **CHAPTER 1. INTRODUCTION**

## **1.1. Background and Problem Statement**

The construction industry is one of the leading industries in the world, which spends \$10 trillion on construction-related goods and services every year (Wang 2019). However, the construction industry is facing a massive workforce shortage of skilled craft workers (Kim et al. 2019). More than 8 out of 10 construction firms report having a hard time finding qualified workers. One of the significant causes of workforce shortage is the premature retirement of skilled craft workers due to safety and health issues. Due to a lack of proper safety training and monitoring systems, the construction workforce is exposed to various fatal and non-fatal injuries such as Work-related Musculoskeletal Disorders (WMSDs). According to the report released by the International Labor Organization (ILO) in 2015, it was estimated that there are at least 60,000 construction-related fatalities all over the world each year (Juan 2015). Similarly, in the United States, 971 (18.9%) out of 5147 fatal injuries have occurred in the construction industry in 2017, based on occupational injury reports released by the Bureau of Labor Statistics (BLS 2017). In addition, the estimated non-fatal injury rate in the construction industry is 5.3 cases per 10,000 full-time workers in 2018, according to BLS (BLS 2018). According to the 2018 Liberty Mutual Workplace Safety index, all businesses spent more than one billion dollars per week for non-fatal severe workplace injuries. The high rate of non-fatal injuries in the construction industry is mainly due to WMSDs. WMSDs are among the most prevalent occupational health problems among construction workers due to high labor-intensive construction tasks.

Construction activities involving lifting heavy loads, repetitive tasks, working in the same posture for an extended period, whole-body vibration, and static work in awkward posture results in work-related injuries and illnesses (Wang et al. 2017). WMSDs are injuries or disorders that

affect the muscles, joints, ligaments, tendons, and the nervous system. Even though the safety performance improved significantly between 1973 and 2004 in the United States because of the implementation of highly effective injury and illness prevention program, no significant improvement was observed in safety incidents over the past decade, demonstrating that the construction industry has reached saturation concerning traditional injury prevention approaches and new safety innovations are needed (Esmaeili and Hallowell 2011; Tixier et al. 2016).

The construction industry often involves high labor-intensive and repetitive tasks, which results in worker physical fatigue. About 40% of the US construction workforce experiences high-level fatigue, which leads to poor judgment, increased risk of injuries, a decrease in productivity, and a lower quality of work (Abdelhamid and Everett 2002; Cheng et al. 2012). Further, excessive fatigue due to working in unpleasant working conditions, long working hours, and heavy workloads can aggravate the adverse effects of fatigue and leads to WMSDs and productivity loss. Moreover, fatigue has been shown to result in impairing the physical and cognitive functions (Zhang et al. 2015) and identified as a possible risk factor for slip-induced falls, which is one of the "fatal four" causes of fatalities in the construction industry, according to Occupational Safety and Health Administration (OSHA 2018).

A significant number of craft workers (20% to 40%) routinely exceed generally accepted physiological thresholds for manual work shifts (Abdelhamid and Everett 2002). Physical fatigue and impaired mental capacity lead to a high risk of accidents in any environmental conditions (Aryal et al. 2017) and affects worker's safety performance (Fang et al. 2015). Since physical fatigue is a predominant risk factor for injuries and illnesses in the construction industry, it is essential to monitor fatigue to reduce the adverse effects and prevent long-term health problems. However, since fatigue demonstrates itself in several complex processes, there is no single

standard measurement method for fatigue detection. For example, if a certain physiological function is altered, it is only to reflect the body's adaptive behavior instead of the level of fatigue (Gatti et al. 2014; Hwang and Lee 2017). Moreover, overall physical fatigue is a result of interaction between local (muscular fatigue) and central factors (such as metabolic, cardiovascular, and thermoregulatory) (Romain et al. 2006). Therefore, fatigue quantification typically involves a combination of kinematic and kinetic measurements, which is often supplemented or substituted by physiological (body temperature, heart rate, or muscle activity) and subjective measures (perceived exertion or discomfort) measures. Physical fatigue is always associated with a high workload, which is measured as physical demand. The evaluation of measured workloads involves two phases, namely, assessment and evaluation phases. The assessment phases involve the measurement of physiological response to work as a measure of physical demand. The evaluation phase involves determining whether the physical demand (workload) of a task is excessive, and workers performing the task may suffer from physical fatigue. The physical demand evaluation techniques include classification of work severity based on recommendations for oxygen uptake, energy expenditure, and heart rate and evaluation of physical fatigue based on aerobic fatigue threshold, absolute energy expenditure, and heart rate limits (Abdelhamid and Everett 2002).

Even though there is no gold standard for fatigue measurement, several subjective and objective techniques are adapted for occupational use. The subjective evaluation of fatigue involves workers' feedback to the questionnaire, and several construction studies used various fixed sets of questions and feedback scales related to fatigue (Arellano et al. 2015; Borghini et al. 2014; Chan et al. 2012; Debnath et al. 2015; Dittner et al. 2004; Fang et al. 2015; Lu et al. 2017; Mitropoulos and Memarian 2012). However, subjective assessments rely on workers' internal perceptions, previous experience and interrupt the ongoing work. Furthermore, most of these

measurement techniques are cumbersome and impractical on construction sites, emphasizing the need for a continuous fatigue monitoring system with minimal obstruction to construction tasks (Aryal et al. 2017; Jebelli et al. 2019).

With advancements in wearable sensing technology, a few researchers have developed objective techniques using physiological sensors to assess the workers' overall physical fatigue by monitoring physiological responses of the worker to physical demand (Aryal et al. 2017; Gatti et al. 2014; Hwang and Lee 2017; Jebelli et al. 2019). In recent studies, Jebelli et al. (2019) have recognized physical demand during on-site work by training machine learning model on workers' photoplethysmogram (PPG), electrodermal activity (EDA), and skin temperature (ST) with an energy expenditure of the task, which was determined using Energy Expenditure Prediction Program (EEPP). Aryal et al. (2017) used skin temperature and heart rate for fatigue detection based on workers' ratings of perceived exertion. Hwang and Lee (2017) used heart rate reserve (%HRR) as a metric to distinguish different levels of physical demand. Maman et al. (2017) have estimated the RPE fatigue level of an individual performing assembly, manual material handling, and supply pick-up tasks using four inertial measurement units (IMU) attached to the human body. Even though these studies have established the potential of physiological responses to determine the worker physical demand for a long duration, they still have limitations such as unable to identify the physical demand of individuals with different characteristics (such as work experience, work conditions, age, and health status), previous studies are limited to classifying individual physical demand based on work severity, not capable of continuous workers' fatigue level monitoring for multiple tasks performed in short intervals, the measurements such as heartrate, skin temperature, and electrodermal activity are highly influenced by external factors which may

not yield reliable results on construction sites. Moreover, these studies do not help in quantifying the direct impacts of fatigue on construction safety performance or accidents.

To overcome these challenges or limitations, the present study proposes an automated continuous workers' fatigue monitoring system by measuring aerobic fatigue threshold (AFT) using forearm muscle activity and kinematic data for an activity. The aerobic fatigue threshold has been set for the proposed system because AFT is activity-dependent which is appropriate for construction. Unlike heart rate, electrodermal activity, and skin temperature, which are highly influenced by external factors, the forearm muscle activity and kinematic data are activity specific. Since the proposed system is workers' activity-centric, the system is highly suitable for construction workers' fatigue monitoring as they involve in various labor-intensive tasks throughout the day.

## **1.2. Research Objectives and Approach**

The overall objective of this doctoral research is to promote a safe construction workplace and healthy and productive workers by developing quantifiable frameworks for continuous physical fatigue monitoring by capturing muscle activity and kinematic data of construction workers using a non-invasive, low-cost wearable armband sensor. The armband sensor measures the muscle activity using electromyography (EMG) sensors and kinematic data using the inertial measurement unit (IMU). The overall framework involves construction workers' activity recognition and continuous oxygen prediction using forearm EMG and IMU data, which enables to monitor of Aerobic Fatigue Threshold (average oxygen consumption [ $\text{VO}_2$ ]/maximum oxygen consumption [ $\text{VO}_{2\text{max}}$ ])) and determine the fatigue level of a worker. Continuous fatigue monitoring aids in quantifying the direct impacts of fatigue on construction safety performance or accidents. Moreover, this system can perform workload evaluation of a construction activity that involves

the classification of work severity. In order to accomplish the overall objective of this study, the framework is divided into multiple stand-alone studies with the following research objectives and research questions, and the hypothesis examined to answer them, and the significance of associated research tasks as outlined below:

- **Objective-1:** Investigating the muscle activity (EMG) and kinematic (IMU) signals provide an understanding of workers' physiological responses to workload. Furthermore, the signals facilitate activity analysis and workers' behavior towards the work. An affordable, non-invasive, lightweight, and wireless wearable armband sensor is available off-the-shelf to collect workers' forearm muscle activity and kinematic data. Many researchers used these signal data for different applications in various domains. To the best of the authors' knowledge, none of the studies in the construction domain have explored the use of wearable EMG and IMU data for construction applications. In order to use a wearable sensor for any construction application, it is essential to investigate the data quality and reliability. This is because the muscle activity and motion sensor signals from the forearm may inevitably be contaminated due to noise signals and artifacts that originate at the skin-electrode interface or due to external sources. It is required to reduce the noise and artifact contamination and preserve the required information from the signals. Moreover, the sensor should provide consistent and reliable signals for activity throughout the data collection process. Therefore, the first objective of this research is to assess the data quality and reliability of forearm EMG and IMU data for construction activity recognition.
  - **Questions:** What is the noise level in armband signal data? What is the rotational drift in the IMU sensor? What is the quality of EMG and IMU data for at-rest and in-motion activities? Does the armband provide reliable data and classification results? Does the

- sensor position affect EMG and IMU data? What is the accuracy of activity classification using EMG and IMU data? Are EMG and IMU data sufficient to recognize lifting weights?
- **Hypothesis:** The armband sensor provides reliable EMG and IMU data and activity classification results. Armband sensor provides consistent and reliable forearm muscle activity and kinematic data for construction applications.
  - **Significance:** The answer to these questions establishes the reliability and applicability of using muscle activity and motion data from the armband sensor for construction applications.
- **Objective-2:** Since the proposed fatigue monitoring framework is activity-dependent, it is essential to develop a reliable and highly accurate activity recognition model using forearm EMG and IMU data. Furthermore, activity recognition helps evaluate the direct impacts of fatigue on safety, productivity, and quality of work. The forearm muscle activity and motion data from armband sensor facilities recognize construction tasks using machine learning algorithms. Previous studies in the construction domain have proposed various machine learning-based activity recognition models using accelerometer and gyroscope data. However, limitations exist, such as the inability to classify a greater number of complex activities involving various body parts and motions, challenges for the practical implementation of smartphones on construction sites. None of these studies have explored the use of a combination of muscle activity and kinematic data for construction activity recognition, which has the potential to recognize complex activities with high accuracy. Moreover, it is essential to investigate the feasibility of using the forearm armband sensor to classify complex construction activities which involve different body parts (wrist,

forearm, upper body, lower body, and whole body) and various motions (repetitive motion, impulsive motion, and free motion). This issue leads to the second objective of this research, develop automated construction workers' activity recognition using forearm electromyography (EMG) and inertial measurement units (IMU) sensor data fusion.

- **Questions:** What classification conditions can better differentiate whole-body and upper-body construction activities using forearm data? What is the accuracy of the construction activity recognition framework? Do a combination of muscle activity and kinematic data improve activity recognition accuracy? What combination of data and features will lead to the highest recognition accuracy? In what combination (s) and to what level of detail (s) can a construction activity be recognized using forearm EMG and IMU data and machine learning?
- **Hypothesis:** The use of forearm EMG and IMU data can recognize complex construction activities involving the upper body and lower body movements at higher accuracy compared to other sensor combinations.
- **Significance:** An accurate construction activity recognition framework contributes to automating the process of measuring aerobic fatigue threshold and continuous fatigue monitoring. Moreover, the activity recognition model helps in workers' performance and safety assessment.
- **Objective-3:** Construction workers' physical fatigue is a critical factor resulting in workers' poor safety performance, reduced productivity, and increased accidents. To prevent their adverse effects on work, continuous monitoring of Aerobic Fatigue Threshold (AFT) by measuring physiological responses of the body. The AFT represents the level of oxygen uptake below which a worker can work for a prolonged duration without excessive fatigue. The AFT

is measured using physiological responses and is expressed as a percentage of average oxygen uptake to the maximum aerobic capacity (MAC or  $\text{VO}_{2\text{max}}$ ). Moreover, according to the National Institute of Occupational Safety and Health (NIOSH), the AFT should not exceed 33% of activity-specific MAC, where MAC is the maximum oxygen uptake that a worker can consume during a physical task or individual's work capacity. The MAC and physiological responses (such as oxygen consumption, heart rate, blood pressure, and lactic acid accumulation) vary based on the worker, task, and environmental characteristics. Therefore, the third objective is to develop and validate construction activity-specific maximum aerobic capacity (MAC) estimation models using submaximal exercise methods.

- **Questions:** What is the experimental procedure for measuring construction activity-specific maximum aerobic capacity? What is the accuracy of the MAC estimation models?
- **Hypothesis:** Submaximal exercise variables can be used to develop construction-specific maximum aerobic capacity estimation models.
- **Significance:** The answer to these questions establishes the experimental method and procedure to develop MAC estimation models for construction-specific activities.
- **Objective-4:** Continuous monitoring of physiological responses of the worker to workload will enable continuous fatigue monitoring. In this regard, several studies have validated the use of physiological responses such as heart rate, heart rate variability, body temperature, electrodermal response, and electrodermal activity. However, these studies have several drawbacks, such as the workers' physiological responses are associated with subjective fatigue measurement, and these studies focus on work severity classification rather than an

evaluation of workers' physical fatigue. The continuous monitoring of the aerobic fatigue threshold (AFT) helps in determining the workers' physical fatigue. The monitoring of AFT is achieved by continuous measuring of oxygen uptake. Some of the studies in the healthcare domain have validated the use of multiple wearable sensors such as accelerometers, respiratory bands, and heart rate to predict oxygen uptake for daily living activities. None of the studies have investigated the use of wearable sensors for oxygen uptake prediction during construction activities. Moreover, the use of multiple sensors in the construction field is not practically feasible. Therefore, it is essential to investigate the use of forearm EMG and IMU sensor for oxygen uptake prediction during construction activities. To address this issue, the fourth objective of this research is to develop continuous  $\text{VO}_2$  (volume of oxygen uptake) prediction models using forearm muscle activity and kinematic signal data.

- **Questions:** Can forearm muscle activity and kinematic signal data be used for the volume of oxygen uptake measurement? Which machine learning algorithms are suitable for  $\text{VO}_2$  predictions? Does a combination of sensor data or individual data produce accurate prediction models? Does data fusion of armband and heart rate yield higher prediction accuracy?
- **Hypothesis:** The forearm muscle activity and kinematic data from the armband contain latent features for oxygen uptake prediction during construction activities.
- **Significance:** An accurate prediction of the oxygen uptake using muscle activity and kinematic data from the armband sensor contributes to the continuous measurement of the aerobic fatigue threshold for fatigue level assessment.

- **Objective-5:** Automated system for continuous fatigue monitoring by measuring aerobic fatigue threshold by integrating all the frameworks mentioned earlier (i.e., activity recognition framework, activity-specific MAC estimation, and continuous  $\text{VO}_2$  measurement framework). However, it is required to test the feasibility and performance of the automated continuous fatigue monitoring system. In order to achieve that, the fifth objective of this research is to assess the feasibility and performance of the automated physical fatigue monitoring system compared to other physiological and conventional fatigue measurement systems.
  - **Questions:** How accurate is the proposed fatigue monitoring system compared to other fatigue measurement techniques? Are there any implementation challenges?
  - **Hypothesis:** Fatigue monitoring using the proposed continuous fatigue monitoring system by measuring the aerobic fatigue threshold is more accurate and reliable compared to fatigue measurement using heart rate physiological measurement.
  - **Significance:** Accurate and reliable continuous fatigue monitoring system helps the construction industry monitor their workers' fatigue levels to prevent accidents and improve productivity.

### **1.3. Organization of the Dissertation**

The overall goal of this research study is to develop an automated workers' fatigue monitoring system using muscle activity and kinematic data from wearable sensors. In order to develop the proposed system, the study has multiple stand-alone studies with specific objectives, as mentioned above. This dissertation is the result of compiling these multiple studies to achieve the proposed continuous workers' fatigue monitoring system. In terms of content organization,

chapters 3 - 7 elaborate on the individual studies to reach the objectives mentioned above. Overall, this dissertation is divided into eight chapters, and the outline is as follows:

- **Chapter 1: Introduction** - This chapter includes the background and problem statement, outlines the focus of this dissertation, specifies research objectives and approaches to achieve stated objectives.
- **Chapter 2: Literature Review** - This chapter includes the review of relevant literature on the definition of physical fatigue, available fatigue measurement methods or techniques, necessity of fatigue monitoring, available fatigue monitoring methods, wearable technologies for fatigue monitoring, fatigue monitoring studies in construction, metrics used for fatigue monitoring, aerobic fatigue threshold measurement methods using traditional and wearable sensors, classification and recognition machine learning algorithms.
- **Chapter 3: System Development, Data Quality, and Reliability Assessment** - This chapter introduces the overall framework of the proposed automated fatigue monitoring system. Moreover, this chapter includes the study on forearm EMG and IMU data quality and reliability for construction applications.
- **Chapter 4: Automated Construction Workers' Activity Recognition** – Since the whole framework is activity-centric, it is essential to develop a reliable and accurate activity recognition model. This chapter presents an automated construction activity recognition method using scaffold builder activities as a case study.
- **Chapter 5: Maximum Aerobic Capacity for Construction Activities** – Since there do not exist the MAC values for construction activities, it presents the experimental procedure

to develop construction activity-specific MAC models using submaximal exercise methods.

- **Chapter 6: Oxygen Uptake Prediction during Construction Activities** – This chapter presents the framework for automated continuous oxygen uptake prediction using forearm EMG and IMU data. The proposed framework was evaluated for feasibility and performance using scaffold building activities.
- **Chapter 7: Fatigue Monitoring System Feasibility Validation and Performance Evaluation** – This chapter evaluates the feasibility and performance of using AFT for continuous fatigue monitoring and performance of the proposed system in assessing the workers' fatigue level.
- **Chapter 8: Conclusions and Recommendations for Future Research** - This chapter summarizes findings, practical implications, and directions for future research.

## CHAPTER 2. LITERATURE REVIEW

### 2.1. Definition and Causes of Fatigue

Fatigue refers to a declination of a persons' ability to maintain a normal level of performance and impair mental alertness (Edwards 1981). In general, fatigue is defined as a state of feeling tired, sleepy, or weary that results from loss of sleep, the extended period of anxiety, exposure to an adverse environment, and prolonged physical and mental work. The unidimensional characterization of fatigue usually describes it as mental fatigue and physical fatigue (Grandjean 1979). Mental fatigue results in a decrease in cognitive and behavioral performance, whereas physical fatigue leads to a decline in the capacity to perform physical activity (Michielsen et al. 2004). Fatigue is a complex phenomenon caused by various factors in the workplace and out of the workplace. Outside the workplace, the lack of restorative sleep is the most common cause of fatigue. Whereas in the workplace, fatigue is caused due to excess physical and mental workload. The workload refers to the amount of work assigned to a worker, categorized into a physical load, a mental load, and an environmental load (Horrey et al. 2011). The fatigue might be due to an individual factor or a combination of interrelated factors. Figure 2.1 shows the work-related causes of fatigue (Sadeghniiat-Haghighi and Yazdi 2015). Moreover, physical fatigue is identified as localized muscular fatigue and overall physical fatigue. Compared to localized muscular fatigue, overall physical fatigue is challenging to quantify as it is caused by the interactions between local (muscular) and central (metabolic, thermoregulatory, cardiovascular, etc.) factors.

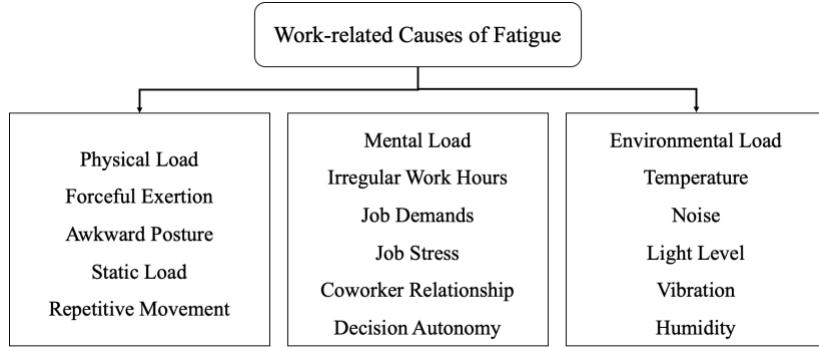


Figure 2.1.Causes of physical fatigue.

## 2.2.Current Approaches for Fatigue Measurement

Since the human body demonstrates physical fatigue in several ways, there exist numerous ways to measure fatigue (Yung and Wells 2016). However, these methods are limited to their application since they are developed for specific contexts and purposes (Zhang et al. 2015). The fatigue measurement methods adopted for occupational use can be broadly classified into subjective and objective assessment techniques (Figure 2.2).

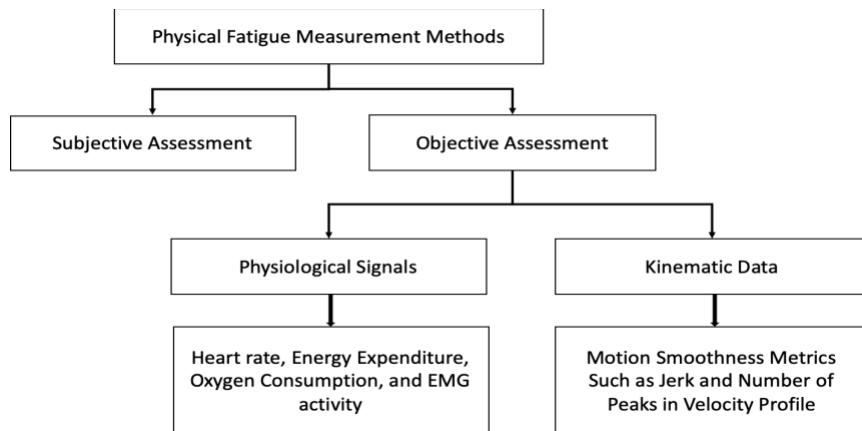


Figure 2.2. Physical fatigue measurement methods.

Early attempts in measuring fatigue involve subjective assessment using a fixed questionnaire related to physical and mental fatigue (Chalder et al. 1993; Lee et al. 1991). Several studies in construction used different questionnaires and subjective feedback scales to quantify fatigue involved in construction activities (Chan et al. 2012; Fang et al. 2015; Mitropoulos and

Memarian 2012; Yi et al. 2016; Zhang et al. 2015). Fang et al. (2015) have developed an experimental method to understand the effect of fatigue on construction workers' safety performance where authors used the Fatigue Assessment Scale for Construction Workers (FASCW) developed by Zhang et al. (2015) to determine fatigue level. The experimental study has concluded that above fatigue level 20, there was a linear relationship between workers' fatigue and error rate (a measure of safety performance). Mitropoulos and Memarian (2012) used the NASA Task load index (TLX) rating scale to determine the task demands in masonry work, where NASA TLX measures mental load, physical load, temporal load, and performance of the worker in a particular task. Measuring the TLX index facilitated to determine the various factors such as task features, supervisor practices, and work conditions that yield high task demands (Mitropoulos and Memarian 2012). Chan et al. (2012) used the Physiological Strain Index (PSI) to determine the recovery time after fatigue state, which was identified by the ratings of perceived exertion (RPE) using the Borg CR10 Scale. Yi et al. (2016) developed an early-warning system to monitor workers' heat-strain levels when working in a hot and humid environment using subjective index perception rating of perceived exertion (RPE) and artificial neural network (ANN). The ANN-based prediction model in the early-warning system uses wet bulb globe temperature (WBGT), age, BMI, job nature, work duration, alcohol drinking habit, smoking habit as input features to predict RPE to monitor workers' heat-strain level. However, subjective fatigue assessment has two significant limitations for the field. First, the feedback assessment is strongly biased due to the workers' internal perception, experience, ethics, and socioeconomic backgrounds (Zhang et al. 2015). Second, the subjective feedback collection on construction sites by stopping the worker while performing a task is cumbersome and not practical.

The objective measurement of overall physical fatigue involves quantifying workers' physiological processes and kinematic data. The physiological processes involve heart rate, energy consumption, oxygen consumption, and EMG activity, whereas the kinematic data includes the body motion data collected using motion capture systems such as kinetic camera and inertial measurement unit sensors (IMUs). Optoelectrical measurement systems are considered the gold standard for body motion analysis within a research setting. However, due to the high cost, large installation spaces, and extensive post-processing of optoelectrical measurement systems, IMU sensors are widely used for full-body motion data collection. The IMUs are non-intrusive wearable sensors integrated with accelerometers, gyroscopes, and magnetometers to measure the body segments' acceleration, orientation, and velocity. IMUs are used for overall physical fatigue detection by monitoring the reduction of motor control (Zhang et al. 2018). Motor performance and control are assessed using the motion smoothness metrics such as the ratio between the maximum and the mean velocity during the movement, number of peaks in velocity profile, and jerk derived from kinematic data (Bosecker et al. 2010; Hogan and Sternad 2009). Jerk is the first derivative of acceleration, which is used to determine motor control and motion smoothness. Van Dieën et al. (1996) investigated that jerk at various joints such as the ankle, hip, knee, and lumbosacral joint is increased due to the fatigue during repetitive lifting of a barbell. Maman et al. (2017) developed logistic and MLR-based physical fatigue detection models using the features which included wrist and hip jerk during simulated manufacturing tasks. The features from the sensor data are extracted using the Least Absolute Shrinkage and Selection Operator (LASSO). The study reported that the accelerometer located at the hip and wrist are strong predictors of physical fatigue than heart rate features. Zhang et al. (2018) investigated the feasibility of using jerk as the metric to detect physical fatigue in repetitive bricklaying activity. The results indicate

that the jerk values obtained from the upper arms and pelvis are significant compared to the values from hands and forearms. However, the motion smoothness metrics such as the jerk values are task-dependent and highly influenced by the worker's repeated shocks, impacts, and skill level.

Harnessing the workers' physiological processes such as oxygen consumption, heart rate, skin temperature, muscle engagement, and blood pressure to determine physical workload or fatigue level (Abdelhamid and Everett 2002; Gatti et al. 2014; Hwang et al. 2016; Hwang et al. 2016a; Jebelli et al. 2018). Measuring the physiological workload can be used to assess the level of physical fatigue. The physiological workload can be determined by measuring oxygen uptake while performing work. With advancements in wearable sensing technologies and machine learning, the oxygen uptake or  $\text{VO}_2$  can be estimated using sensor data such as heart rate and IMU. The measurement of physiological workload involves two phases, namely the assessment phase and evaluation phase. The assessment phase involves a measure of physiological response to an activity as a measure of physical demand. The evaluation phase involves determining whether the physical demand is excessive and the worker performing the task is susceptible to fatigue. The workload evaluation techniques include work severity classification based on published guidelines and physical fatigue detection based on oxygen consumption (Abdelhamid and Everett 2002). According to NIOSH recommendation, the average oxygen uptake during an 8-hour workday should not exceed 33% of activity-specific maximum aerobic capacity (Health 1981; Ilmarinen 1992; Snook and Irvine 1968). Abdelhamid and Everett (2002) reported that 20-40% of craft workers exceed physiological thresholds on a daily basis by measuring the workers' oxygen uptake and heart rate.

Most of the studies in construction focused on measuring workload and work severity classification based on physiological responses (Abdelhamid and Everett 2002; Gatti et al. 2014;

Mital et al. 1994; Wong et al. 2014). Abdelhamid and Everett (2002) used oxygen uptake and heart rate to determine the physical demands required for different construction activities. Wong et al. (2014) proved that the energy required to perform bar fixing tasks was more than bar bending tasks in a hot and humid environment. Chan et al. (2012) determined the optimal recovery time for rebar workers after working to exhaustion in a hot and humid environment using blood pressure, heart rate, and subjective rating fatigue. However, heart rate alone is insufficient for monitoring fatigue. This is because heart rate is influenced by various physiological and behavioral factors such as cigarette smoking, mentally stressful situation, alcohol consumption, and energy drinks intake (Bosquet et al. 2008). To address this issue, Aryal et al. (2017) used heart rate in combination with human body thermoregulatory changes to monitor fatigue in construction workers. Hwang and Lee (2017) showed the potential of using a wristwatch-based heart rate sensor to determine the levels of physical demands by measuring heart rate variability metrics. Maman et al. (2017) used jerk as a metric to derive from IMU sensors placed at the ankle, wrist, hip, and torso to detect physical fatigue. The study presents logistic regression models trained on the rating of perceived exertion. Jebelli et al. (2019) used PPG, EDA, and ST physiological signals of the worker in association with energy expenditure to determine the workers' physical demands. The energy expenditure of the task was determined using the energy-expenditure prediction program (EEPP). However, several limitations exist in the current objective measurement systems for fatigue monitoring such as the individual variability were not considered in the models, not applicable in case where several tasks were performed in a short interval of time, and most of studies considered subjective measurement of fatigue.

### **2.3.Point of Departure**

The construction industry often involves high-intensity physical demand tasks, which results in workers' physical fatigue. The worker experiencing high-level fatigue leads to poor judgment, increased risk of injuries, decreased productivity, and a lower quality of work. Therefore, it is essential to monitor workers' physical fatigue to prevent adverse effects continuously. Previous studies have developed various subjective and objective fatigue measurement methods. The subjective measurements are not practically applicable to the construction industry because they interrupt the ongoing work and result in an individual's perception and experience. Since physical fatigue is associated with heavy workloads, the physiological response of a worker to workload can be used to determine the fatigue level. Various objective measurement systems were developed by measuring physiological responses such as heart rate, oxygen uptake, and muscle activity. The recent studies in the construction domain have proposed the use of PPG, EDA, ST, and %HRR to develop machine learning-based fatigue detection models (Aryal et al. 2017; Hwang and Lee 2017; Jebelli et al. 2018). Even though these studies have established the potential of physiological response to determining worker physical demand for a long duration, there are still limitations in the current studies. Some of the limitations of current methods are the inability to identify physical fatigue of individuals with different characteristics (such as work experience, work conditions, age, and health status), unable to determine physical fatigue when activities are performed in a short interval of time, use of subjective measurements for baseline fatigue measurement, and inability to determine direct impacts of fatigue on construction activities. Moreover, the previous studies focused on work severity classification rather than determining the workers' real-time physical fatigue.

## CHAPTER 3. SYSTEM DEVELOPMENT, DATA QUALITY, AND RELIABILITY ASSESSMENT<sup>†</sup>

### 3.1. System Development

The framework for the proposed automated continuous fatigue monitoring system is shown in Figure 3.1. The framework involves six steps: data acquisition, activity recognition, oxygen uptake prediction, maximum aerobic capacity (MAC) estimation, continuous aerobic fatigue threshold (AFT) monitoring, and fatigue assessment. The workflow of the proposed framework involves forearm EMG and IMU data collection from the worker, and the collected data will be used for activity recognition and oxygen uptake prediction. Once the activity is recognized, the respective activity-specific MAC value is estimated from the MAC models. Using the oxygen uptake and MAC values, continuous measurement of the aerobic fatigue threshold is performed, and the AFT features are used for fatigue level assessment. Each of these steps has a specific objective and integrating all these steps facilitates developing a system for monitoring workers' physical fatigue assessment.

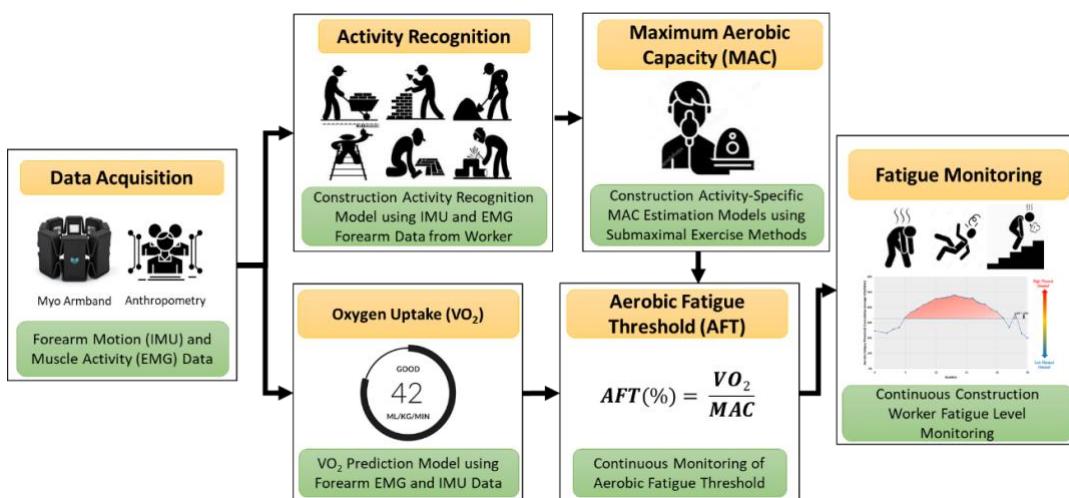


Figure 3.1. Framework of the proposed fatigue monitoring system.

<sup>†</sup> This chapter is adapted from Bangaru, S. S., Wang, C., & Aghazadeh, F. (2020). Data Quality and Reliability Assessment of Wearable EMG and IMU Sensor for Construction Activity Recognition. *Sensors*, 20(18), 5264. (See Appendix D for permission).

### **3.2.Data Quality and Reliability Assessment**

#### **3.2.1. Introduction**

The construction industry is one of the leading industries in the world, which spends \$10 trillion on construction-related goods and services every year (Wang 2019). However, the construction industry is facing a massive workforce shortage of skilled craft workers (Kim et al. 2020). More than 8 out of 10 construction firms report having a hard time finding qualified workers. One of the significant causes of workforce shortage is the premature retirement of skilled craft workers due to safety and health issues. Due to a lack of proper safety training and monitoring systems, the construction workforce is exposed to various fatal and non-fatal injuries such as work-related musculoskeletal disorders (WMSDs). To overcome these challenges, various researchers have proposed wearable sensor-based systems in the area of construction safety and health (Ahn et al. 2019; Aryal et al. 2017; Awolusi et al. 2018; Häikiö et al. 2020; Hwang et al. 2018; Valero et al. 2017). Various applications in the area of safety and health involve preventing musculoskeletal disorders, fall prevention, mental and physical workload assessment, and fatigue monitoring (Ahn et al. 2019; Aryal et al. 2017; Awolusi et al. 2018; Häikiö et al. 2020; Hwang et al. 2018; Valero et al. 2017). These applications can be categorized as a classification problem since they involve identifying different postures, classifying different physical and mental workloads, or detecting different motions or gestures using the sensor data. Moreover, classifying workers' activities helps monitor and manage the productivity, safety, and quality of work (Zhang et al. 2018).

In the construction domain, the wearable sensor-based activity recognition models have gained increased attention due to low cost, ease of use, high accuracy, and non-intrusiveness. Most previous studies used accelerometers and gyroscopes embedded in smartphones to recognize

construction workers' activity (Akhavian and Behzadan 2016; Bayat et al. 2014; Kwapisz et al. 2011; Nath et al. 2017). A study by Cezar (2012) used an accelerometer and gyroscope embedded in the smartphone placed on the dominant hand to recognize hammering, sawing, sweeping, and drilling activities with an accuracy of 91% using a quadratic discriminant analysis (QDA) algorithm. Lim et al. (2016) and Akhavian and Behzadan (2018) have developed artificial neural network (ANN) based models for identifying falls and manual material handling activities with an accuracy of 94% and 90.74% using the smartphone placed in the hip pocket and upper arm respectively. The ironwork activities recognition models developed by (Yang et al. 2019) and (Zhang et al. 2018) using support vector machine (SVM) and decision trees (DT) were able to recognize activities with 94.83% and 92.98% accuracy. Even though these smartphone sensors-based models have achieved considerable accuracy, there are practical implementation challenges. In order to overcome the smartphone challenges, inertial measurement unit (IMU), sensor-based activity recognition models have been proposed for various construction applications such as work sampling (Joshua and Varghese 2014; Khan and Sohail 2013; Ryu et al. 2019) and fall detection (Yang et al. 2014). The wearable IMU sensor-based models used various machine learning algorithms such as DT, random forest, and SVM to recognize ironwork (Joshua and Varghese 2014), fall detection (Khan and Sohail 2013), and bricklaying (Ryu et al. 2019) with an accuracy of 90.4%, 93.90%, and 88.1% respectively. However, the current activity classification methods are limited to fewer activities involving either upper body or lower body, use of multiple sensors, and do not consider activities with multiple intensities. Moreover, none of these studies have discussed the reliability of sensor data and classification results. Therefore, there is a necessity for low-cost, easy to use, and non-obstructive sensors that can provide reliable data for complex construction activity classification.

Despite the fact that sensors provide rich and detailed information, not all sensors can be used for construction applications due to the dynamic nature of construction work (Awolusi et al. 2018). It was recommended that multisensory data fusion, which was applied in other domains, provides an opportunity for enhancing the accuracy of activity classification (Ahn et al. 2019; Chen et al. 2020). The sensor for construction applications should be simple and easy to wear, unobtrusive, affordable, and wireless. Moreover, the sensor should provide reliable data and involve minimal or no preprocessing for noise removal. Therefore, it is essential to identify a suitable and reliable sensor for construction activity classification, which helps develop construction workers' safety and health monitoring systems to prevent work-related injuries such as WMSDs, which is one of the significant reasons for the workforce shortage.

The armband sensor is an affordable, non-invasive, lightweight, and wireless wearable armband sensor that is available off-the-shelf to collect workers' forearm electromyography (EMG) and inertial measurement unit (IMU) data (Benalcázar et al. 2017). Many researchers used these signal data for different applications in various domains. To the best of the authors' knowledge, none of the studies in the construction domain have explored the use of armbands and a combination of EMG and IMU data for construction applications. Furthermore, investigating muscle activity and kinematic signals provides an understanding of workers' physiological responses to workload. Furthermore, the signals facilitate activity analysis and workers' behavior towards the work. In order to choose a wearable sensor for any construction applications specifically for activity classification, it is essential to investigate the data quality and reliability because the muscle activity and motion sensor signals from the forearm may inevitably be contaminated due to noise signals and artifacts that originate at the skin-electrode interface or due to external sources. A reduction in the noise and artifact contamination is required and the

preservation of the required information from the signals. Moreover, the sensor should provide consistent and reliable signals for activity throughout the data collection process. Therefore, the objective of this study is to assess the data quality and reliability of forearm EMG and IMU data for construction activity classification by following the guidelines, recommendations, and methods for data quality and reliability assessment proposed by previous studies on the sensor (Düking et al. 2018; Jensenius et al. 2012; Nymoen et al. 2015; Nymoen et al. 2012; Skogstad et al. 2011; Vigliensoni and Wanderley 2012).

In order to achieve the proposed objective, the study is divided into seven experiments. The first three experiments involve evaluating the data quality, understanding the effect of armband position on data quality, and forearm EMG and IMU data reliability. Later, four experiments involve building and evaluating activity classification models, assessing the reliability of classification results, understanding the effect of lifting weights on classification results, and evaluating the classification performance of different sensor combinations. The results of these experiments answer various questions such as noise level in armband signal data, drift in the IMU sensor data, quality of EMG and IMU data for at-rest and in-motion activities, the effect of armband position on signal quality, the accuracy of construction activity classification using EMG and IMU, reliability of sensor data and classification results, effect of lifting weights on classification accuracy, and classification performance of different sensor combinations. It was hypothesized that the armband sensor provides reliable EMG and IMU data and activity classification results. The answers to the above questions establish the reliability and applicability of forearm EMG and IMU data for construction activity classification.

### **3.2.2. Materials and Methods**

#### **3.2.2.1. Participants**

Eight healthy college male students voluntarily participated in all the experiments. The participants' ages ranged from 24 to 28 years (mean  $\pm$  SD:  $26.13 \pm 1.55$  years), height ranged from 1.65 to 1.83 m ( $1.74 \pm 0.06$  m), and weight ranged from 62.60 to 100 kg ( $81.35 \pm 12.44$  kg). All the participants were right-handed, healthy, and had no musculoskeletal disorders at the time of the experiments. All the procedures involving human participants were approved by the Louisiana State University Institutional Review Board (IRB #: IRBAM-20-0112). The purpose of the research was demonstrated to all the participants before starting the experiment, and their signatures were obtained on the informed consent forms (See Appendix A). The sample size required to assess the reliability of the sensor using the intraclass correlation (ICC) was determined using the tables from Bujang and Baharum (2017). An ICC score greater than or equal to 0.75 indicates excellent reliability (Šerbetar 2015; Zhang et al. 2014). At least seven participants are required to achieve a minimum of 0.75 ICC scores with two assessments per subject at a 0.05 significance level and a power of 0.80 (Bujang and Baharum 2017).

#### **3.2.2.2. Measurements and Instrumentation**

A forearm-based wearable armband sensor (Myo armband) developed by Thalmic Labs Inc. was used to collect the EMG and IMU data. Myo armband sensor is a non-intrusive wearable sensor that consists of eight dry surface EMG sensors and a 9-axes IMU sensor (3-axes gyroscope, 3-axes accelerometer, and 3-axes magnetometer). The sensor weighs approximately 93 g (Benalcázar et al. 2017). The data from the sensor is transmitted to the computer or cloud storage via Bluetooth Low Energy (BLE) wireless connection. The raw EMG and IMU data can be assessed through the Myo software development kit (SDK). The Myo SDK was used to acquire

real-time forearm EMG and IMU data at a frequency of 200 Hz and 50 Hz, respectively. The device goes into an idle state if there is no activity for more than 30 s. The configuration of Myo armband electrodes is shown in Figure 3.1(a), where the electrode with the LED light and Myo logo is channel-4, followed by channel-3 in a clockwise direction and channel-5 in a counter-clockwise direction. Moreover, Figure 3.1(a) shows the direction of x, y, and z of the IMU sensor. The armband was worn on the thickest part of the forearm, as shown in Figure 3.1(b) with the channel-4 in the line of the index finger, and the blue marker was in the lower forearm for the experiments unless otherwise stated (Arief et al. 2015). After wearing the armband sensor, the participant calibrates their motion by performing predefined gestures such as finger spread, wave-in, wave-out, and relaxed state gestures by connecting with Thalmic Labs' Myo Connect manager (Chae et al. 2018).

The eight EMG sensors capture the electrical impulses generated by the forearm muscles, which are returned as an 8-bit array, in other words, each EMG sensor outputs an integer value between -128 and 127, representing muscle activation levels. The armband sensor captures the muscle activity of various forearm muscles such as the brachioradialis, flexor digitorum superficialis, medial epicondyle of humerus, palmaris longus, flexor carpi ulnaris, flexor carpi radialis, and pronator teres (Visconti et al. 2018). Whereas the IMU unit captures the motion of the forearm by measuring acceleration, angular velocity, and orientation along the x, y, and z axes. It was ensured that the armband was always synced with the application and calibrated throughout the experiments.

High-precision conventional wearable EMG and IMU sensors such as FREEEMG (BTS Bioengineering Corp., Quincy, MA) and YEI 3-Space IMU sensor (Yost Engineering In., Portsmouth, OH), respectively, were used to compare the armband sensor data quality. The conventional sensor measures the acceleration and gyroscope in units of g and radians/s,

respectively. In comparison, the conventional EMG sensor measures muscle activity in millivolts (mV). Besides, the conventional IMU sensor was calibrated using a gradient descent calibration procedure, and no preprocessing was performed on any sensor data before data quality calculations.

To assess the reliability of the armband sensor data, features such as absolute acceleration, absolute angular velocity, and mean absolute value of EMGsum (sum of EMG values) were calculated from raw data (De Pasquale and Somà 2010). These sensor features are widely used in activity/gesture/motion recognition applications (Arief et al. 2015; De Pasquale and Somà 2010; Frank et al. 2019; He et al. 2017; Koskimäki et al. 2017; Kubota et al. 2019; Mendez et al. 2017; Tao et al. 2018). The acceleration along x, y, and z axes was used to compute the absolute acceleration or magnitude of the acceleration vector (Acc) at any given timestamp (t) using Equation (3.1) (Janidarmian et al. 2017; Murcia and Triana 2019; Zhang et al. 2014). Similarly, the angular velocity along the three axes provided by the gyroscope sensor was used to calculate the absolute gyroscope angular velocity or magnitude of gyroscope vector (Gyro) at any given timestamp (t) using Equation (3.2) (Murcia and Triana 2019; Stančin and Tomažič 2011; Xu et al. 2016; Zhang et al. 2014). For simplicity, the angular velocity along the axes was represented as Gyro in Equation (3.2). Using the eight EMG values, a new feature EMGsum was calculated by summing up all the eight EMG values at any timestamp (t) (Kim et al. 2017; Koskimaki and Siirtola 2016). Further, the mean absolute value (MAV) of EMGsum was evaluated using Equation (3.3), which was later used for reliability assessment (Arief et al. 2015; Koskimäki et al. 2017; Mendez et al. 2017). For each trial, an average of acceleration magnitude, an average of gyroscope magnitude, and MAV of EMGsum were computed to assess the trial-to-trial (intra-day) reliability

of the sensor. Whereas in the case of day-to-day (inter-day) reliability test, the mean values of three trials of each day were used for ICC analysis.

$$Acc(t) = \sqrt{Acc(t)_x^2 + Acc(t)_y^2 + Acc(t)_z^2} \quad (3.1)$$

$$Gyro(t) = \sqrt{Gyro(t)_x^2 + Gyro(t)_y^2 + Gyro(t)_z^2} \quad (3.2)$$

$$MAV = \frac{1}{N} \sum_{k=1}^N |EMGsum_k| \quad (3.3)$$

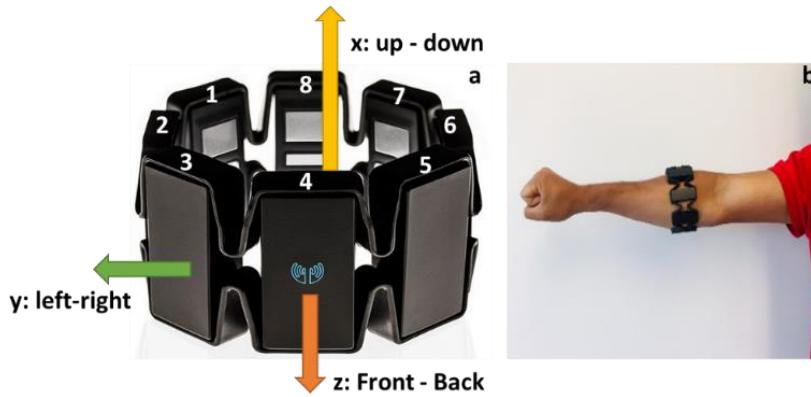


Figure 3.2. (a) Myo armband electrode location and (b) Myo armband placement on the forearm.

### 3.2.3. General Procedures of the Study

This study consists of seven experiments, including (a) evaluating the forearm EMG and IMU data quality for “at-rest” and “in- motion” activities (Experiment I); (b) investigating the effect of armband sensor position on EMG and IMU data (Experiment II); (c) assessing the reliability of forearm EMG and IMU data obtained while performing construction activities (Experiment III); (d) classification model building, performance evaluation, and classifier comparison (Experiment IV); (e) investigating the reliability of results obtained from classification models using EMG and IMU data while performing construction activities on different days (Experiment V); (f) investigating the effect of lifting weight on forearm EMG and IMU data and

activity classification results (Experiment VI); and (g) comparison of activity classification performance for different sensor combinations. The activities performed by the participants are standardized across all the experiments. The “at-rest” activities include the armband lying stationary on the floor or placed on the arm of a person sitting still with an arm resting on a desk. In contrast, the “in- motion” activities include screwing at elbow height at a frequency of 1 turn/6 s, wrenching while kneeling at a frequency of 1 turn/6 s, lifting a 25 lbs sandbag from elbow to shoulder height at a frequency of 1 lift/6 s, and carrying a 25 lbs sandbag on the shoulder with the dominant hand at the bottom of the sandbag for 30 s. Activities were designed in such a way that they represent a wide range of construction activities involving forearm (lifting), wrist (screwing and wrenching), and whole-body (carrying). Moreover, these activities represent controlled natural motions such as repeated motion (lifting), impulsive motion (screwing or wrenching), and free motion (carrying). All the activities were performed for 30 s (i.e., each trial of activity was 30 s). Each participant performed three trials for an activity on a testing day. There were two testing periods (i.e., Day-1 and Day-2) where participants performed all five activities (i.e., stationary on the body, screwing, wrenching, lifting, and carrying) on both days. Therefore, each participant performed a total of 15 activities (3 trials × 5 activities) in one day. There was no gap between the testing periods. The activities were randomized for all the participants for both days. Before starting the experiment, all the participants were given enough time to familiarize themselves with the tools to eliminate systematic bias due to learning effects (Hopkins 2000). The participants were asked to warm up their bodies before the session, and enough rest was provided between the trials to prevent injuries and fatigue (Bangaru et al. 2019). Once the armband was worn on the body and synced with the computer, a two-minute settling time was considered before the start of the experiment to prevent the rotational drift. In order to test the reliability using the test-retest

approach, all the activities were performed in an indoor environment under control conditions unless stated otherwise. The eight participants' EMG, accelerometer, and gyroscope data were recorded and stored for all five activities for both days. The data were processed and analyzed accordingly based on the experiment requirements. The seven experiments mentioned above are further explained in the following sections and broadly divided into three categories: data quality assessment, data reliability assessment, and activity classification performance evaluation.

### **3.2.3.1.Data Quality Assessment**

- Experiment I – Evaluating the Forearm EMG and IMU Data Quality for “At-Rest” and “In-Motion” Activities

The wearable sensor data is highly susceptible to various confounding factors that affect the quality of data. In this experiment, the data quality of EMG, acceleration, and gyroscope measurements were assessed by evaluating the signal-to-noise ratio (SNR) and compared to a conventional sensor. Furthermore, the influence of confounding factors (communication devices, another sensor, power tools, and smartwatches) and environments (indoor and outdoor) on the data quality were studied in this experiment. First, the data quality was determined for the armband sensor and compared with the conventional sensors for at-rest and in-motion activities. In order to compare the data quality of the armband sensor, the conventional sensors were placed along with the armband sensor while performing activities, as shown in Figure 3.2. Each in-motion activity was performed three times by all eight participants. The average SNR value was used for the comparison. The influence of various confounding factors and environmental conditions on the armband sensor data quality was assessed when the armband sensor was lying on the floor by computing SNR values for three trials. Inter-device data quality was assessed using two armbands lying on the floor at the same time to check if the data is consistent across different devices under

the same conditions. All the at-rest activities were conducted three times, and the average value was considered to represent the influence of confounding factors, environment, and inter-device variability on the data quality.

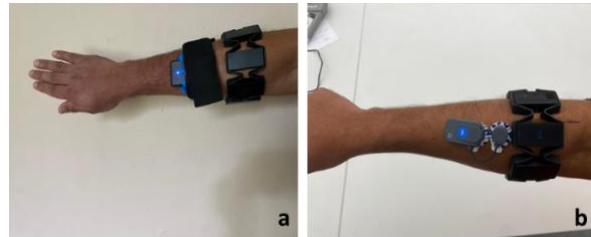


Figure 3.3. Position of conventional (a) inertial measurement unit IMU and (b) electromyography (EMG) along with armband sensor.

- Experiment II - Investigating the Effect of Armband Sensor Position on EMG and IMU Data

In order to explore the effect of sensor position on the EMG and IMU data, a lifting activity was performed for three different sensor positions, as shown in Figure 3.4. The standard position refers to wearing an armband with sensor-4 in the direction of the index finger. Whereas the rotated and slid positions refer to rotating the armband in an anticlockwise direction (sensor-5 in the direction of the index finger) and sliding the armband downwards with respect to the standard position, respectively. A qualitative analysis was performed on the root mean square value of EMG and the absolute magnitude of IMU data collected while performing lifting activity with three sensor positions.



Figure 3.4. Three armband sensor positions to test the effect of sensor position on data quality: (a) rotated, (b) standard, and (c) slid down.

The sensor data quality was assessed by evaluating the noise level in the data using the signal-to-noise ratio (SNR). The SNR value of a signal is the ratio of the power of the signal to the power of the noise (Muppalla et al. 2017). Alternatively, it is defined as the ratio of the mean of the measurements ( $\mu$ ) to the standard deviation of the measurements ( $\sigma$ ) as shown in Equation (3.4). Where mean and standard deviation (SD) of measurements represent the power of signal and power of noise in the measurements. The signal power of acceleration and gyroscope measurements were determined as mean values of absolute magnitude. In contrast, the mean value of the EMG measurements was calculated as mean-absolute-value (MAV) (Nymoen et al. 2015; St-Amant et al. 1998).

$$\text{SNR} = \frac{\mu}{\sigma} \quad (3.4)$$

### **3.2.3.2.Data Reliability Assessment**

- Experiment III - Assessing the Reliability of Forearm EMG and IMU Data Obtained while Performing Construction Activities

The EMG and IMU data collected from eight participants while performing construction activities on two different days were assessed for reliability. In this experiment, the raw EMG and IMU data collected from eight participants were processed to calculate the mean absolute value (MAV) of EMG, absolute acceleration (Acc), and absolute gyroscope (Gyro) for each trial of both the days (Day-1 and Day-2). The MAV, Acc, and Gyro values of each trial were used to assess trial-to-trial reliability for both days. Further, the MAV, Acc, and Gyro values of all three trials were averaged for an activity for a participant for each day to evaluate reliability between days. The relative reliability was assessed using the intraclass correlation coefficient, and absolute reliability was evaluated using standard error of measurement (SEM) and smallest detectable difference (SDD).

All statistical analyses were performed using IBM SPSS statistical package version 25. The trial-to-trial and day-to-day reliability were assessed on the accelerometer, gyroscope, and electromyography measurements obtained while performing five construction activities on two different testing periods (i.e., Day-1 and Day-2). Moreover, the assessment of the trial-to-trial and day-to-day reliability measures intradevice reliability. The reliability was assessed between the trials and between the days using test-retest reliability, which consists of relative and absolute reliability (Bruton et al. 2000; Zhang et al. 2014). The relative reliability refers to the magnitude of the correlation of repeated measurements, which was evaluated using the intraclass correlation coefficient (ICC) (Šerbetar 2015; Zhang et al. 2014). The relative reliability was expressed using ICC form (3, k), which includes a two-way mixed effect model, mean of k measurement type, and a definition of a relationship as absolute agreement (Koo and Li 2016; McGraw and Wong 1996). Moreover, the ICC form (3, k) considers both systematic and random errors and uses the mean value of the repeated measurements as evaluation scores (Zhang et al. 2014). Based on the ICC score, the strength of relative reliability can be interpreted as excellent (if ICC score is higher than 0.75), good (if ICC score is between 0.59 and 0.75), fair (if ICC score is between 0.48 and 0.58), and poor (if ICC score is less than 0.40) (Hsu et al. 2016; Koo and Li 2016; McGraw and Wong 1996; Šerbetar 2015; Xi et al. 2019; Zhang et al. 2014).

Whereas absolute reliability refers to variability in the repeated measurements of an individual (Šerbetar 2015; Zhang et al. 2014). The absolute reliability was evaluated by estimating the standard measurement error (SEM). SEM estimates how the repeated measures of an individual on the same device tend to distribute around true value (Zhang et al. 2014). SEM is estimated as defined in Equation (3.5), where SD is the standard deviation of the measurements of a test and retest of all participants, and ICC is the average trial-to-trial or day-to-day test-retest relative

reliability (Guillén-Rogel et al. 2019; Hsu et al. 2016; Regterschot et al. 2014; Šerbetar 2015; Zhang et al. 2014). The SEM% was used to compare the absolute test-retest reliabilities of different scenarios, which was evaluated using Equation (3.6), where the SEM score is represented as a percentage of SEM divided by the mean of test and retest measurements. The SEM% value below 10% indicates excellent absolute test-retest reliability. Moreover, the smallest detectable difference (SDD) was calculated from SEM at a 95% confidence interval using Equation (3.7), which is the smallest change in the measurement that is required to be considered as a real change in the measurement but not due to error (Hsu et al. 2016; Šerbetar 2015; Zhang et al. 2014). Similar to SEM%, the SDD score is expressed as a percentage of the mean of measurements (SDD%), which is computed using Equation (3.8) (Hsu et al. 2016; Šerbetar 2015; Zhang et al. 2014). Before performing the parametric reliability testing, a nonparametric Kolmogorov–Smirnov test was performed to verify the normality of the data.

$$\text{SEM} = \text{SD} \sqrt{1-\text{ICC}} \quad (3.5)$$

$$\text{SEM\%} = \frac{\text{SD} \sqrt{1-\text{ICC}}}{\text{Mean}} \times 100 \quad (3.6)$$

$$\text{SDD} = 1.96 \times \sqrt{2} \times \text{SEM} \quad (3.7)$$

$$\text{SDD\%} = \frac{1.96 \times \sqrt{2} \times \text{SEM}}{\text{Mean}} \times 100 \quad (3.8)$$

### **3.2.3.3. Activity Classification, Performance Evaluation, and Classification Reliability**

- Experiment IV - Classification Model Building, Performance Evaluation, and Classifier Comparison

The data obtained from the armband sensor worn by the eight participants performing five activities on two different days (Day-1 and Day-2) was used to build machine learning (ML) based classifiers for respective days. A typical machine learning methodology, which includes data

preparation, model building, model training, hyperparameter tuning, and model evaluation, was used to develop ML classifiers for activity classification, which was implemented using the PyCatet classification module in Google Colab. First, the dataset was prepared using the raw acceleration ( $a_x$ ,  $a_y$ ,  $a_z$ ), gyroscope ( $g_x$ ,  $g_y$ ,  $g_z$ ), and EMG (8-channel) features for both days. The 8-channel EMG data were downsampled by converting 8-bit to 32-bit to match the frequency of accelerometer and gyroscope data. Therefore, the final dataset for each day consists of 38 (3-acceleration, 3-gyroscope, and 32-EMG) input features. Further, the data were manually labeled for five different activities (i.e., stationary on the body, screwing, wrenching, lifting, and carrying). Once the datasets were prepared, the labeled data was used to build the machine learning (ML) based classifier models using the default classifier settings. Besides, the hyperparameters of the model were tuned by optimizing the model accuracy to obtain a finely tuned model. The ten most common ML-based classifier models such as random forest, J48 decision trees, support vector machine (SVM), naïve Bayes, k-nearest neighbors (KNN), logistic, multi-layer perceptron (MLP), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and gradient boosting (Xgboost) were built using each day dataset. Additionally, 10-fold cross-validation was performed to evaluate the performance of the classifiers. In the cross-validation technique, the dataset is randomly shuffled and divided into ten groups. Each unique group is considered a holdout or test dataset, and the remaining nine groups are used for model training. Once the model has been fitted on the training dataset, the model is evaluated on the test set. The evaluation score is retained, and the model is discarded. This process is repeated for each unique group. The performance of the trained ML classifier was evaluated using metrics such as accuracy, recall, precision, F1 score, kappa, and confusion matrix. The performance of different classifiers was compared to determine the best performing classifier for each dataset.

- Experiment V - Investigating the Reliability of Results Obtained from Classification Models Using EMG and IMU Data While Performing Construction Activities on Different Days

The reliability of results obtained from the classification models using Day-1 and Day-2 datasets was investigated. The best classifier obtained in Experiment III was further used to run ten iterations on each dataset. The accuracies of the classifier on the Day-1 dataset were compared to accuracies of the same classifier on the Day-2 dataset using paired t-test at a 0.05 significance level.

- Experiment VI - Investigating the Effect of Lifting Weight on Forearm EMG and IMU Data and Activity Classification

Detecting different weights is useful for many construction applications. This experiment investigates if the weight affects forearm EMG and IMU data and activity classification. For this experiment, an activity of lifting three different weights (10 lbs, 25 lbs, 50 lbs) with three trials from four participants was considered. The raw data with 38 features (acceleration-3, gyroscope-3, and EMG-32) was manually labeled for three activities (Lift10, Lift25, and Lift50). The ML-based classification models such as random forest, J48 decision trees, support vector machine (SVM), naïve Bayes, k-nearest neighbors (KNN), logistic, multi-layer perceptron (MLP), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and gradient boosting (Xgboost) were built using the raw data and evaluated using 10-fold cross-validation technique. The best classifier results were analyzed for three different classes to check if the sensor data could classify different weights.

- Experiment VII - Comparison of Activity Classification Performance for Different Sensor Combinations

This experiment focuses on comparing the performance of various ML-based classifier models built using different sensor feature combinations such as EMG+IMU, IMU alone, and EMG alone. For this analysis, two datasets were considered, namely, controlled and uncontrolled activity datasets. The controlled activity dataset was prepared by combining the Day-1 and Day-2 data of five controlled activities, namely screwing, wrenching, lifting, and carrying. Whereas the uncontrolled dataset was prepared by collecting forearm armband data from the participants while performing nine construction activities at varied intensities and pace such as walking at random speed (walk), carrying (10 lbs, 25 lbs, and 50 lbs), lifting (10 lbs, 25 lbs, and 50 lbs), and screwing (at elbow height, kneeling, and overhead). Both the datasets consist of 38 features (3-acceleration, 3-gyroscope, and 32-EMG). Once the datasets were prepared, the ML-based classifiers were built with different sensor feature combinations. As explained in Experiment IV, the ten most used ML-based classifier models were built for three sensor data combinations for both datasets. The finely tuned ML-based classifiers were evaluated using 10-fold cross-validation, and the accuracy of the classifiers was combined across all the sensor feature combinations for both datasets.

The classification involves identifying a set of classes using the input features. The performance of a classification algorithm is evaluated using metrics such as accuracy, recall, precision, F1 score, and kappa. In order to define these metrics, one needs to understand the terms true positives (TP), true negative (TN), false positive (FP), and false-negative (FN). The classification accuracy is the ratio of correct predictions ( $TP + TN$ ) to the total number of predictions ( $TP + TN + FP + FN$ ). Precision measures the number of correct positive predictions, which is the ratio of true positives (TP) to total positive predictions ( $TP+FP$ ). In contrast, recall measures the number of correct positive predictions from all the positive predictions, which are the ratio of true positives to true positives (TP) and false negatives (FN). F1 score is the weighted

average of precision and recall, as shown in Equation (3.9) (Bangaru et al. 2019). Cohen's kappa value measures the agreement between the predicted and actual labels. Apart from these metrics, the performance of the classifier on individual classes was assessed by using a confusion matrix.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.9)$$

### 3.3. Results

#### 3.3.1. Forearm EMG and IMU Data Quality for “At-Rest” and “In-Motion” Activities

First, the EMG and IMU data quality of the armband sensors was compared with conventional EMG (FREEEMG) and IMU (Yost) using standard deviation and signal-to-noise ratio. Table 3.1 shows the standard deviation (noise level) and SNR (signal quality) for the accelerometer, gyroscope, and EMG for both the conventional and armband sensors for at-rest and in-motion activities. The at-rest activities include stationary on the body for the EMG and stationary on the floor for IMU. For at-rest and in-motion activities, the armband sensor's noise levels in acceleration and gyroscope data are comparable to a conventional sensor. The SNR values are higher in the case of armband acceleration data compared to the conventional sensor for both at-rest and in-motion activities. In contrast, the SNR values of gyroscope and EMG armband data are comparable to conventional sensors (Table 3.1). However, the signal quality measured as SNR is better in armband data compared to conventional sensors for both EMG and IMU (Table 3.1). Second, the noise level and data quality were compared between the indoor and outdoor environments. The results show that the noise level slightly increased in case of gyroscope ( $SD_{\text{Indoor}} = 0.121$ , and  $SD_{\text{Outdoor}} = 0.138$ ) and EMG ( $SD_{\text{Indoor}} = 3.006$ , and  $SD_{\text{Outdoor}} = 2.974$ ) data for outdoor environment (Table 3.2). However, the signal quality is comparably the same for both environments (Table 3.2). Third, two different armband sensors under same conditions have similar noise level and data quality for acceleration ( $SD_1 = 0.002$ ,  $SNR_1 = 514.120$ ;  $SD_2 = 0.002$ ,

$\text{SNR}_2 = 515.192$ ), gyroscope ( $\text{SD}_1 = 0.121$ ,  $\text{SNR}_1 = 1.325$ ;  $\text{SD}_2 = 0.138$ ,  $\text{SNR}_2 = 1.469$ ) and EMG ( $\text{SD}_1 = 3.006$ ,  $\text{SNR}_1 = 0.865$ ;  $\text{SD}_2 = 2.947$ ,  $\text{SNR}_2 = 0.881$ ) data (Table 3.3). The acceleration, gyroscope, and EMG data of stationary on the body were assessed for potential confounding factors, as shown in Table 3.4. The results show that the noise level in the acceleration is almost similar for all the factors; however, it is slightly affected in the presence of a communication device (Table 3.4). The noise level in gyroscope and EMG data have slightly increased in the presence of other sensors and power tools, respectively. However, the data quality of gyroscope and EMG data is similar in the presence and absence of confounding factors (Table 3.4). Finally, the rotational drift was determined by observing the evolution of the yaw angle for the data collected during the stationary on the body and Myo lying on the floor. Figure 3.5 shows the evolution of a yaw angle for 80 s of a stationary experiment. The results indicate that there was 0.13 deg/s drift initially, and it reached a steady orientation when the armband was stationary on the body (Figure 3.5a). Whereas in the case of the armband lying on the floor, the yaw angle drifts at a rate of 0.17 deg/s before it reached a steady orientation, as shown in Figure 3.5b. Besides, it can be observed that the rotational drift was reduced considerably when worn on the body compared to the armband lying on the floor. Furthermore, the drift was higher in the initial frames and reached a steady orientation in a few seconds. Therefore, a settling time of two minutes was considered to prevent rotational drift.

Table 3.1. Comparison of EMG and IMU data quality of Myo armband and conventional sensors.

	Accelerometer (Units of g)				Gyroscope (rad/s)				EMG			
	Myo		Conv.		Myo		Conv.		Myo		Conv.	
	SD	SNR	SD	SNR	SD	SNR	SD	SNR	SD	SNR	SD	SNR
At-rest Activities	0.00	514.12	0.00	340.64	0.00	1.32	0.04	0.32	3.01	0.87	0.00	0.83
Screwing	0.02	60.42	0.03	35.30	0.31	0.52	0.62	0.43	4.39	0.78	0.02	0.67
Wrenching	0.03	37.10	0.04	25.88	0.39	0.66	0.40	0.71	5.64	0.70	0.02	0.67
Lifting	0.13	8.02	0.15	6.96	0.95	0.90	0.97	0.89	19.50	0.49	0.07	0.41
Carrying	0.06	16.07	0.07	15.35	0.45	1.22	0.50	1.16	11.25	0.70	0.01	0.66

Table 3.2. Comparison of EMG and IMU data quality between indoor and outdoor for at-rest activity.

	Indoor		Outdoor	
	Std. Dev.	SNR	Std. Dev.	SNR
Accelerometer (Units of g)	0.002	514.120	0.002	495.712
Gyroscope (rad/s)	0.121	1.325	0.223	1.093
EMG	3.006	0.865	3.389	0.846

Table 3.3. Comparison of EMG and IMU data quality between two armband sensors for at-rest activity.

	Myo-1		Myo-2	
	Std. Dev.	SNR	Std. Dev.	SNR
Accelerometer (Units of g)	0.002	514.120	0.002	515.192
Gyroscope (rad/s)	0.121	1.325	0.138	1.469
EMG	3.006	0.865	2.974	0.881

Table 3.4. Effect of confounding factors on data quality for at-rest activity.

	Communication Device		Other Sensor		Power Tool		Smart Watch	
	Std. Dev.	SNR	Std. Dev.	SNR	Std. Dev.	SNR	Std. Dev.	SNR
Accelerometer (Units of g)	0.002	443.860	0.002	470.081	0.002	493.564	0.002	479.49
Gyroscope (rad/s)	0.100	1.588	0.166	1.157	0.133	1.147	0.111	1.25
EMG	3.270	0.851	3.100	0.855	4.610	0.830	3.030	0.85

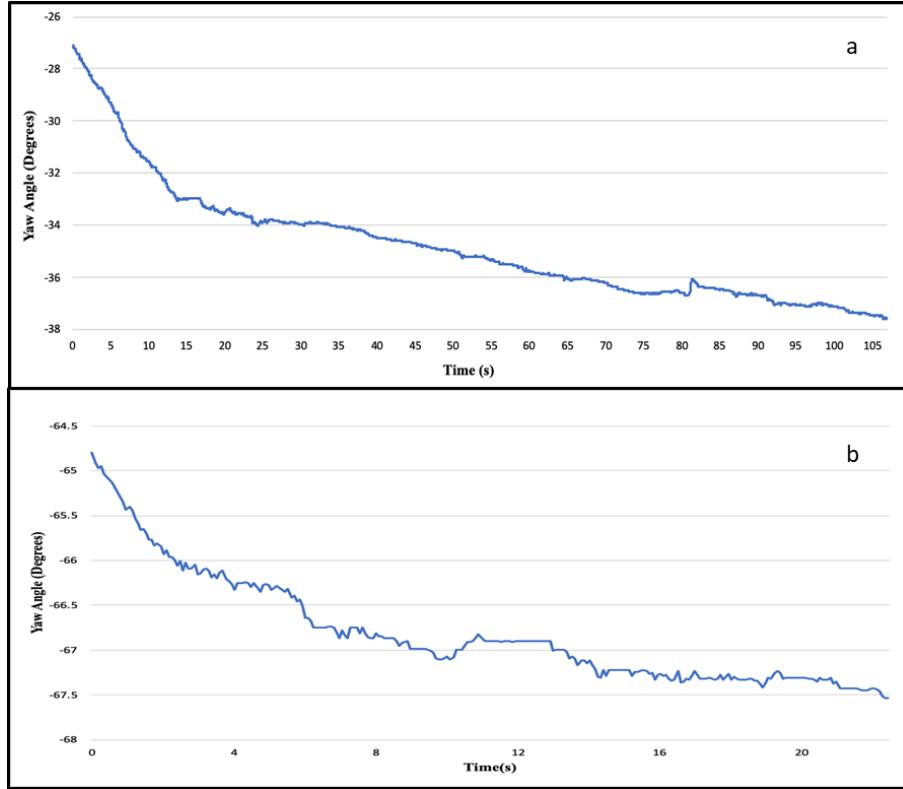


Figure 3.5. Evolution of yaw angle of armband sensor while (a) stationary on the body and (b) lying on the floor.

Further, a qualitative comparison was performed by inspecting the in-motion activity data from the armband and conventional sensors. The acceleration and EMG data of lifting activity of armband and conventional sensor wore at the same time was plotted in Figures 3.6 and 3.7, respectively. In Figure 3.6, the acceleration magnitude was compared for both the sensors, and it is evident that the acceleration data pattern is similar to the conventional IMU sensor. In Figure 3.7, the root mean square (RMS) of EMG channel-4 was compared with conventional EMG RMS, which shows that they follow a similar trend. Moreover, the Myo armband can capture more detailed information compared to a single FREEEMG sensor.

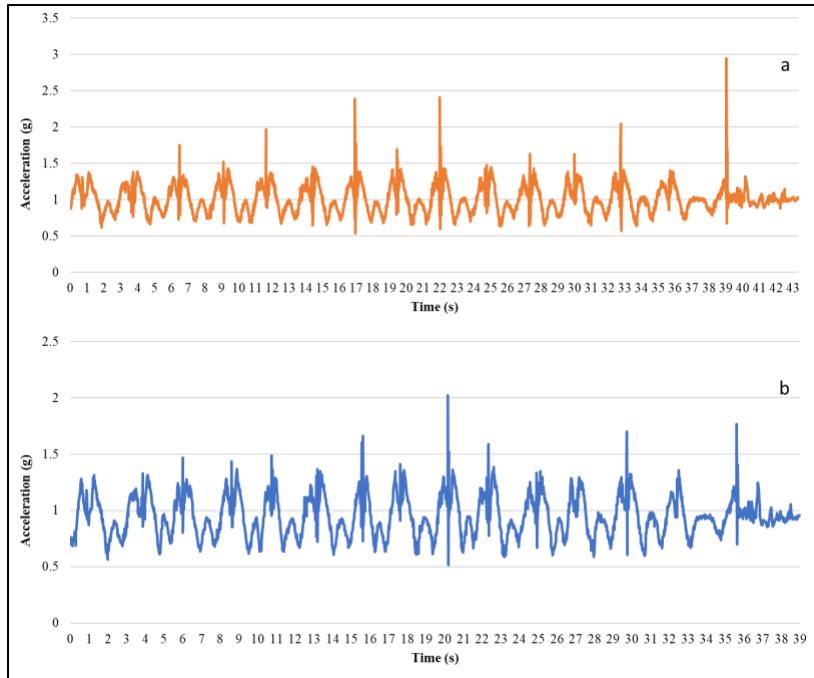


Figure 3.6. Acceleration magnitude while lifting using (a) conventional sensor and (b) armband sensor.

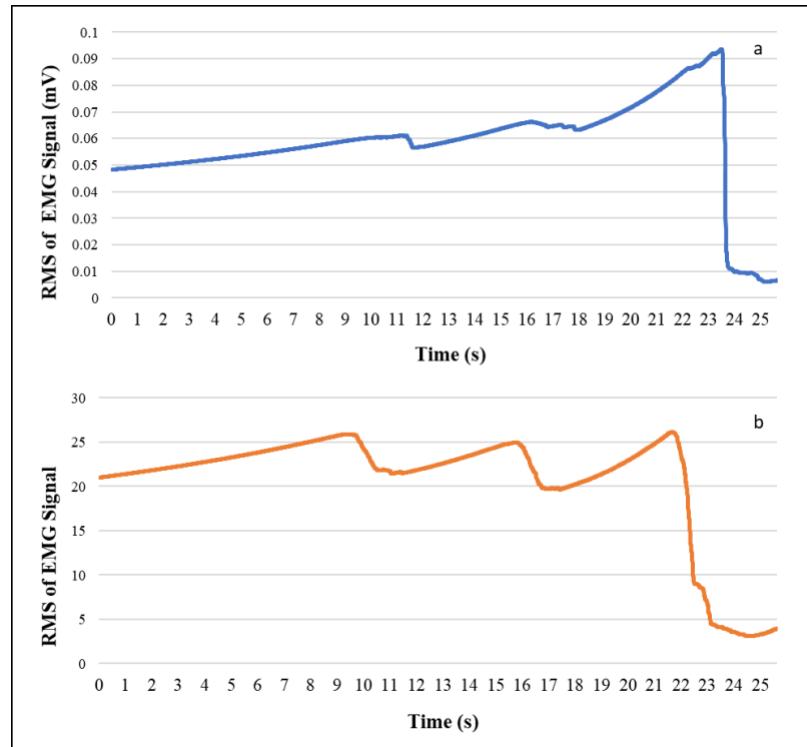


Figure 3.7. EMG RMS plots for lifting activity using (a) conventional EMG sensor and (b) armband sensor.

### 3.3.2. Effect of Sensor Position on Forearm EMG and IMU Data Quality

The effect of three sensor positions, such as “rotated,” “standard,” and “slid down” are compared for lifting activity. Figure 3.8 (a, b) shows the acceleration and gyroscope magnitude for three positions, the range of magnitude and median is the same for all the three positions, and this shows that the IMU data of the forearm is almost the same irrespective of the armband position, whereas the RMS plots of EMG vary for different sensor positions, as shown in Figure 3.9. When the armband is rotated by one sensor, channel-5 takes the position of channel-4, and channel-6 takes the position of channel-5. Similarly, the RMS plot of EMG-5 and EMG-6 in rotated positions are similar to the RMS plot of EMG-4 and EMG-5 in standard positions, respectively. Whereas the RMS plots in slid down position have a lower magnitude range compared to the standard position due to lesser contact with the muscle in slid position.

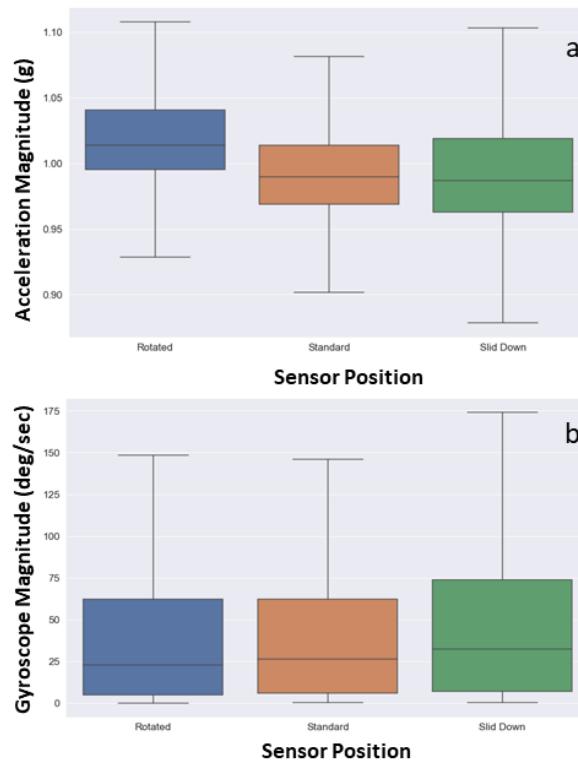


Figure 3.8. Comparison of (a) acceleration and (b) gyroscope magnitude for three (rotated, standard, and slid) sensor positions.

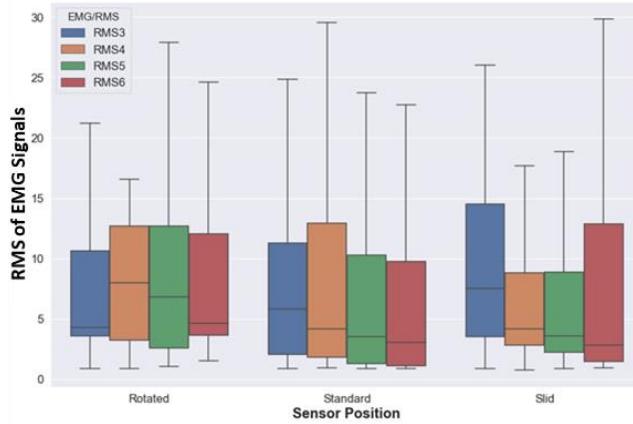


Figure 3.9. Comparison of RMS values of EMG-3, 4, 5, and 6 channels for three (rotated, standard, and slid) sensor positions.

### 3.3.3. Reliability of Forearm EMG and IMU Data of Construction Activities

The forearm acceleration, gyroscope, and EMG data from eight participants while performing construction activities such as screwing, wrenching, lifting, carrying, and at-rest were assessed for trial-to-trial and day-to-day reliability using the ICC test. Tables 3.5–3.7 summarize the test-retest reliability evaluation of accelerometer, gyroscope, and EMG measurements. For each activity, the mean and standard deviation of the measurements is the average of three trials (test mean (SD)) for each day. The average ICC value of three trials at a 95% confidence interval (CI), SEM%, and SDD% for all five activities for both days are shown in Tables 3.5–3.7. For acceleration measurements of both days, the average ICC values range from 0.844 to 0.995 for all five activities (Table 3.5). Similarly, for gyroscope and EMG, the values range from 0.839 to 0.987 and 0.864 to 0.988, respectively (Tables 3.6 and 3.7). The results from Tables 3.5–3.7 indicate excellent relative reliability between trials of acceleration, gyroscope, and EMG measurements for all five activities for both days. Moreover, SEM% for all the activities for acceleration, gyroscope, and EMG measurements is below 10%, which indicates excellent absolute reliability between trials for both days. The SDD% for all activities for both days ranges from 0.098% to 0.669% for acceleration, 5.953% to 32.225% for gyroscope, and 6.709% to 29.130% for EMG.

The day-to-day reliability assessment for accelerometer measurements shows that the ICC value is greater than 0.75, and SEM% is below 10% for all the activities, which indicates an excellent relative and absolute reliability for all activities (Table 3.8). For gyroscope, the ICC values are greater than 0.75 except for lifting activity (ICC = 0.724), indicating excellent relative reliability of gyroscope data except for lifting. Whereas for the absolute reliability, SEM% values are below 10% except for stationary on the body (SEM% = 11.36%) and screwing (SEM% = 16.322%) activity. For EMG measurements, the ICC values are greater than 0.75 for all the activities indicating excellent relative reliability. Whereas the SEM% is slightly above 10% except for lifting activity (SEM% = 7.75%). The SDD% values range from between 14.48% to 31.48% and 24.49% to 39.89% for gyroscope and EMG measurements. The higher SDD% values of gyroscope and EMG suggest that caution should be taken when using gyroscope and EMG measurements for activity recognition because the change in the measurements might be due to error. Therefore, later experiments investigate if the data quality and reliability of the armband data are sufficient to yield accurate and reliable activity classification results.

### **3.3.4. Validating the Classifier Performance on Day-1 and Day-2 Dataset**

Tables 3.9 and 3.10 present the classification performance results of the classifiers built using Day-1 and Day-2 datasets. The performance of both classifiers was evaluated using overall accuracy, recall, precision, F1 score, and kappa, as shown in Tables 3.9 and 3.10. The best classification performance was obtained for the random forest for Day-1 (accuracy—96.48%) and Day-2 datasets (accuracy—96.48%). Further, the random forest classifier was used to assess performance between the classes using the confusion matrix and class report, as shown in Tables 11 and 12. The recall values above 90% for both the classifiers show that a specific activity can be predicted with fewer false-positive values. The F1 score demonstrated high overall performance

for stationary on the body, carrying, lifting, screwing, and the lowest for wrenching (93.2% and 94.9%) for both the classifiers (Tables 3.11 and 3.12). Finally, the association between the actual activities and the predicted classes was measured with Cohen's kappa coefficient, and the values indicate strong agreement with the reality in both Day-1 ( $95.6\% \pm 0.003$ ) and Day-2 ( $96.73\% \pm 0.0019$ ) classifiers (Tables 3.11 and 3.12)

Table 3.5. Accelerometer trial-to-trial reliability

	Day-1				Day-2			
	Test Mean (SD)	ICC	SEM%	SDD%	Test Mean (SD)	ICC	SEM %	SDD %
Stationary on the Body	0.97 (0.004)	0.961 (0.943 - 0.990)	0.10%	0.28%	0.97 (0.0050)	0.995 (0.993 - 0.998)	0.03%	0.09%
Screwing	0.98 (0.009)	0.965 (0.952 - 0.979)	0.19%	0.53%	0.98 (0.0093)	0.991 (0.988 - 0.995)	0.09%	0.24%
Wrenching	0.99 (0.012)	0.978 (0.964 - 0.985)	0.19%	0.54%	0.99 (0.0123)	0.980 (0.969 - 0.995)	0.18%	0.49%
Lifting	1.00 (0.010)	0.959 (0.952 - 0.963)	0.21%	0.58%	0.99 (0.0067)	0.923 (0.892 - 0.962)	0.20%	0.52%
Carrying	1.01 (0.007)	0.888 (0.868 - 0.900)	0.24%	0.67%	1.01 (0.0055)	0.844 (0.779 - 0.931)	0.26%	0.59%

Table 3.6. Gyroscope trial-to-trial reliability assessment

	Day-1				Day-2			
	Test Mean (SD)	ICC	SEM%	SDD%	Test Mean (SD)	ICC	SEM%	SDD%
Stationary on the Body	0.575 (0.155)	0.921 (0.840 - 0.966)	7.981%	22.1%	0.581 (0.168)	0.839 (0.757 - 0.924)	11.626%	32.2%
Screwing	9.448 (3.850)	0.987 (0.986 - 0.987)	4.706%	13.0%	7.316 (1.900)	0.893 (0.873 - 0.930)	8.507%	23.5%
Wrenching	14.243 (3.291)	0.967 (0.954 - 0.988)	4.176%	11.6%	14.239 (2.538)	0.885 (0.854 - 0.935)	6.036%	16.7%
Lifting	46.043 (5.164)	0.963 (0.948 - 0.979)	2.148%	5.9%	45.945 (3.897)	0.824 (0.759 - 0.873)	3.562%	9.9%
Carrying	30.804 (3.960)	0.899 (0.864 - 0.921)	4.092%	11.3%	24.838 (5.609)	0.919 (0.880 - 0.939)	6.440%	17.8%

Table 3.7. EMG trial-to-trial reliability assessment.

	Day-1				Day-2			
	Test Mean (SD)	ICC	SEM %	SDD %	Test Mean (SD)	ICC	SEM %	SDD %
Stationary on the Body	10.183 (4.678)	0.981 (0.963– 0.995)	6.3%	17.5%	7.80 (1.163)	0.864 (0.817– 0.917)	5.5%	15.2%
Screwing	23.622 (8.206)	0.946 (0.914– 0.992)	8.0%	22.3%	27.32 (7.602)	0.983 (0.973– 0.990)	3.6%	10.0%
Wrenching	31.719 (11.066)	0.988 (0.985– 0.989)	3.8%	10.7%	30.84 (11.966)	0.983 (0.977– 0.988)	5.1%	14.2%
Lifting	40.497 (4.963)	0.961 (0.946– 0.976)	2.4%	6.7%	40.65 (8.490)	0.948 (0.918– 0.979)	4.7%	13.2%
Carrying	42.326 (16.037)	0.949 (0.937– 0.962)	8.6%	23.7%	35.77 (10.282)	0.866 (0.806– 0.914)	10.5%	29.1%

Table 3.8. Accelerometer, gyroscope, and EMG day-to-day reliability assessment.

	Accelerometer			Gyroscope			EMG		
	Day-1 vs. Day-2			Day-1 vs. Day-2			Day-1 vs. Day-2		
	ICC	SEM%	SDD%	ICC	SEM%	SDD%	ICC	SEM%	SDD%
Stationary on the Body	0.82	0.24%	0.57%	0.80	11.36%	31.48%	0.92	12.57%	34.83%
Screwing	0.86	0.40%	0.97%	0.80	16.32%	45.24%	0.85	12.14%	33.65%
Wrenching	0.86	0.52%	1.26%	0.84	8.22%	22.79%	0.79	16.21%	44.92%
Lifting	0.86	0.35%	0.86%	0.72	5.22%	14.48%	0.82	7.75%	21.49%
Carrying	0.88	0.24%	0.59%	0.78	6.94%	19.25%	0.85	14.39%	39.89%

Table 3.9. ML-based classifier performance on Day-1 dataset.

Classifier	Accuracy	Recall	Precision	F1 Score	Kappa
Random Forest	96.48% (0.0024)	96.39% (0.0025)	96.49% (0.0024)	96.48% (0.0024)	95.60% (0.0030)
J48	94.30% (0.0017)	94.17% (0.0017)	94.78% (0.0016)	94.38% (0.0016)	96.33% (0.0022)
SVM	58.85% (0.0084)	57.59% (0.0086)	58.01% (0.0172)	54.31% (0.0165)	48.44% (0.0105)
Naïve Bayes	70.45% (0.0022)	69.87% (0.0021)	70.64% (0.0023)	70.06% (0.0023)	63.06% (0.0027)
KNN	79.43% (0.0024)	78.853% (0.0025)	79.55% (0.0025)	78.95% (0.0026)	74.25% (0.0030)
Logistic	61.64% (0.0044)	60.57% (0.0045)	64.28% (0.0052)	60.51% (0.0046)	51.96% (0.0055)
MLP	94.81% (0.0029)	94.67% (0.0029)	94.80% (0.0028)	94.80% (0.0028)	93.51% (0.0036)
LDA	59.11% (0.0037)	57.92% (0.0037)	63.11% (0.0081)	56.21% (0.0036)	48.76% (0.0046)
QDA	75.23% (0.0028)	74.65% (0.0029)	75.45% (0.0032)	74.86% (0.0031)	69.03% (0.0035)
Xgboost	93.42% (0.0029)	93.22% (0.0030)	93.50% (0.0028)	93.36% (0.0030)	91.77% (0.0037)

Table 3.10. ML-based classifier performance on Day-2 dataset.

Classifier	Accuracy	Recall	Precision	F1 Score	Kappa
Random Forest	97.43% (0.0015)	97.35% (0.0015)	97.44% (0.0015)	97.43% (0.0015)	96.73% (0.0019)
J48	96.01% (0.0014)	95.99% (0.0014)	96.24% (0.0013)	96.05% (0.0014)	95.02% (0.0017)
SVM	66.50% (0.0069)	65.64% (0.0072)	66.03% (0.0061)	63.69% (0.0090)	58.03% (0.0087)
Naïve Bayes	73.27% (0.0036)	72.80% (0.0037)	73.81% (0.0034)	72.93% (0.0034)	66.61% (0.0046)
KNN	90.98% (0.0015)	91.00% (0.0012)	90.90% (0.0012)	90.90% (0.0012)	88.72% (0.0015)
Logistic	68.31% (0.0046)	67.82% (0.0046)	69.55% (0.0044)	67.97% (0.0046)	60.34% (0.0057)
MLP	97.13% (0.0019)	97.04% (0.0019)	97.13% (0.0018)	97.13% (0.0018)	96.41% (0.0023)
LDA	68.43% (0.0054)	67.88% (0.0055)	71.43% (0.0054)	67.70% (0.0056)	60.45% (0.0068)
QDA	76.32% (0.0035)	75.90% (0.0036)	77.00% (0.0033)	75.79% (0.0033)	70.42% (0.0043)
Xgboost	96.51% (0.0012)	96.39% (0.0012)	96.50% (0.0012)	96.50% (0.0012)	95.64% (0.0015)

Table 3.11. Confusion matrix and class report of random forest classifier on Day-1 dataset.

True Class	Predicted Class					
		Stationary on the Body	Screwing	Wrenching	Lifting	Carrying
Stationary on the Body	8953	27	2	0	0	
Screwing	13	8244	345	38	0	
Wrenching	10	409	8035	209	0	
Lifting	1	129	187	8131	97	
Carrying	0	1	5	64	9598	
Overall Accuracy						96.48% (0.0024)
Precision		99.70%	93.60%	93.60%	96.30%	99.00%
Recall		99.70%	95.40%	92.80%	95.10%	99.30%
F1 Score		99.70%	94.50%	93.20%	95.70%	99.10%

Table 3.12. Confusion matrix and class report of random forest classifier on Day-2 dataset.

	Predicted Class					
		Stationary on the Body	Screwing	Wrenching	Lifting	Carrying
<b>True Class</b>	Stationary on the Body	11,857	25	0	0	0
	Screwing	14	10,941	210	75	0
	Wrenching	0	368	9949	220	0
	Lifting	0	75	247	11,184	59
	Carrying	0	21	32	47	11,653
Overall Accuracy						96.33% (0.0022)
Precision		99.90%	95.70%	95.30%	97.00%	100%
Recall		99.80%	97.30%	94.40%	96.70%	99%
F1 Score		99.80%	96.50%	94.90%	96.90%	99%

### 3.3.5. Reliability of Classification Results

The classification results obtained using the classifiers of Day-1 and Day-2 were further analyzed for reliability using paired t-test on overall accuracy. A paired t-test ( $p = 0.63$ ) at a significance level of 0.05 shows no significant difference between the Day-1 and Day-2 classifier accuracy. The difference between the overall accuracy of the Day-1 and Day-2 random forest classifier is 0.15%.

### 3.3.6. Effect of Lifting Weight on Classification Results

The ten most common classification algorithms' performances were analyzed on the lifting different weights dataset. Table 3.13 shows the accuracy, recall, precision, F1 score, and kappa values of all the classifiers. The random forest classifier showed the best performance in classifying three different weights with an overall accuracy of 83.89%, recall value of 84.06%, and kappa value of 75.82%. The results indicate that using the forearm EMG and IMU data, the random forest classifier can classify all three weights at 83.89% accuracy. Further, the confusion matrix and class report show that the high overall performance for Lift10 (F1 score = 91%) activity

followed by Lift25 (F1 score = 80%) and Lift50 (F1 score = 77%) (Table 3.14). The results confirm that the forearm EMG and IMU data can not only classify lifting activity but is also able to detect weight. In addition, the correlation of raw features shows that the gyroscope and EMG features are highly correlated compared to accelerometer features (Figure 3.10). Therefore, it can be concluded that the gyroscope and EMG features provide an opportunity to classify different weights of lifting activity.

Table 3.13. Classifier performance on lifting different weights.

Classifier	Accuracy	Recall	Precision	F1 Score	Kappa
Random Forest	83.89% (0.0051)	83.71% (0.0053)	84.06% (0.0053)	83.93% (0.0052)	75.82% (0.0077)
J48	72.71% (0.0061)	72.73% (0.0061)	76.14% (0.0067)	73.05% (0.0060)	59.91% (0.0091)
SVM	54.31% (0.0079)	53.43% (0.0082)	53.26% (0.0163)	48.94% (0.0134)	31.00% (0.0123)
Naïve Bayes	43.21% (0.0080)	42.17% (0.0082)	43.86% (0.0169)	36.94% (0.0109)	13.61% (0.0123)
KNN	65.07% (0.0040)	64.47% (0.0003)	64.98% (0.0014)	63.91% (0.0001)	47.45% (0.0006)
Logistic	56.70% (0.0054)	55.96% (0.0055)	54.78% (0.0072)	53.57% (0.0062)	34.70% (0.0082)
MLP	71.838% (0.0043)	82.90% (0.0015)	83.00% (0.0054)	82.90% (0.0003)	31.14% (0.0093)
LDA	55.07% (0.0054)	54.35% (0.0056)	53.08% (0.0071)	52.15% (0.0060)	32.31% (0.0082)
QDA	45.69% (0.0069)	44.68% (0.0070)	46.11% (0.0104)	40.70% (0.0104)	17.50% (0.0106)
Xgboost	75.84% (0.0070)	75.58% (0.0072)	75.91% (0.0074)	75.81% (0.0072)	63.73% (0.0106)

Table 3.14. Confusion matrix and class report of random forest classifier on different lifting weights.

True Class		Predicted Class					
		Lift10	Lift25	Lift50	Lift10	Lift25	Lift50
ay	1	-0.31	-0.19	-0.1	-0.022	-0.047	0.004 -0.00083 0.00047 0.0013 -0.00056 -0.00025 0.004 0.0018
az	-0.31	1	-0.2	0.037	-0.097	-0.097	-0.006 0.00033 0.00059 0.0027 0.0018 0.0064 -0.0016 -0.00056
gz	-0.19	-0.2	1	-0.055	-0.018	0.078	0.0065 -0.0026 -0.0036 -0.0065 -0.0021 -0.0034 -0.00013 0.0015
gx	-0.1	0.037	-0.055	1	0.55	0.27	-0.0025 -0.0036 -0.002 -0.0074 -0.0063 -0.0011 -0.0017 0.00069
gy	-0.022	-0.097	-0.018	0.55	1	0.49	0.0046 -0.0031 -0.0013 -0.0035 -0.00088 -0.0014 0.0025 0.0024
ax	-0.047	-0.097	0.078	0.27	0.49	1	0.0074 -0.0032 -0.0025 -0.0045 -0.0024 -7.8e-05 0.0018 0.0021
emg1	-0.004	-0.006	0.0065	-0.0025	0.0046	0.0074	1 0.26 0.084 0.069 0.083 0.087 0.15 0.39
emg2	-0.00083	0.00033	-0.0026	-0.0036	-0.0031	-0.0032	0.26 1 0.38 0.12 0.1 0.087 0.092 0.14
emg3	-0.00047	0.00059	-0.0036	-0.002	-0.0013	-0.0025	0.084 0.38 1 0.36 0.21 0.14 0.087 0.077
emg4	-0.0013	0.0027	-0.0065	-0.0074	-0.0035	-0.0045	0.069 0.12 0.36 1 0.61 0.31 0.13 0.087
emg5	-0.00056	0.0018	-0.0021	-0.0063	-0.00088	-0.0024	0.083 0.1 0.21 0.61 1 0.44 0.18 0.099
emg6	-0.00025	0.0064	-0.0034	-0.0011	-0.0014	-7.8e-05	0.087 0.087 0.14 0.31 0.44 1 0.41 0.17
emg7	-0.004	-0.0016	-0.00013	-0.0017	0.0025	0.0018	0.15 0.092 0.087 0.13 0.18 0.41 1 0.4
emg8	-0.0018	-0.00056	0.0015	0.00069	0.0024	0.0021	0.39 0.14 0.077 0.087 0.099 0.17 0.4 1



Figure 3.10. Correlation of features for three (10 lbs, 25 lbs, and 50 lbs) lifting weights.

### 3.3.7. Comparison of Activity Classification Performance for Different Sensor Combinations

The comparison of overall accuracies for different sensor combinations is shown in Table 3.15. For the controlled dataset, the EMG + IMU and IMU alone are better compared to EMG. The classification accuracy is higher for EMG + IMU in the case of random forest, SVM, naïve

Bayes, and MLP, whereas the classification accuracy is higher for IMU alone in the case of KNN, logistic, LDA, QDA, and Xgboost. However, except for KNN and MLP, the accuracy is not significantly different for EMG + IMU and IMU alone. For the uncontrolled activity dataset, the accuracy is significantly higher for EMG + IMU compared to EMG and IMU alone, except in the case of KNN. For the KNN classifier, the IMU alone has higher accuracy compared to EMG + IMU. However, the highest classification accuracy (98.13%) for nine activities with various intensities was obtained for the EMG + IMU feature combination. The combination of EMG and IMU features yields higher accuracy compared to individual sensor data for complex activities.

Table 3.15. Overall classification accuracy for different sensor combinations for controlled and uncontrolled datasets.

Classifier	Controlled Activity Dataset			Uncontrolled Activity Dataset		
	EMG + IMU	IMU	EMG	EMG + IMU	IMU	EMG
Random Forest	96.21%	94.65%	44.97%	98.13%	84.80%	47.60%
J48	94.94%	95.33%	48.54%	96.55%	78.55%	30.83%
SVM	73.23%	73.33%	21.21%	96.55%	48.39%	14.19%
Naïve Bayes	71.40%	69.39%	45.95%	82.52%	54.79%	23.05%
KNN	86.16%	96.95%	45.58%	71.03%	84.62%	29.83%
Logistic	64.65%	64.69%	18.86%	88.76%	45.63%	14.11%
MLP	90.87%	81.99%	52.27%	90.82%	62.50%	37.51%
LDA	62.78%	62.87%	18.87%	88.26%	26.57%	14.19%
QDA	74.79%	75.73%	46.44%	62.33%	52.66%	29.24%
Xgboost	41.53%	41.53%	27.51%	85.52%	21.72%	15.70%

### 3.4.Discussion

In this study, the data quality of low-cost forearm-based wearable sensors was explored by comparing the standard deviation and signal-to-noise ratio of the armband sensor and the conventional sensor for at-rest and in-motion activities. The noise levels in the armband acceleration data ( $SD = 0.002$ ) when lying on the floor are comparable to the high precision conventional IMU sensor ( $SD = 0.003$ ), which is in agreement with the previous study ( $SD =$

0.0019) (Nymoen et al. 2015). Similarly, the noise levels in the acceleration and gyroscope data for in-motion activities are comparable to conventional sensors. Besides, the signal quality of armband sensor data is higher compared to the conventional sensor, which shows that the armband sensor is less sensitive compared to high precision and high-frequency sensors. Moreover, the data quality test in the presence of confounding factors also proves that the armband data is not affected much by the confounding factors, environment, and inter-device variability. Drift is one of the most common issues of IMU when used to estimate position and orientation (Nymoen et al. 2012). The rotational drift of the armband sensor was assessed by observing the evolution of the yaw angle for at-rest activities. The yaw angle drifts at a rate of 0.17 deg/s before it reaches the steady orientation, which agrees with a previous study (Hall et al. 2019). This experiment proves that the drift reduced when the Myo was worn on the body compared to lying on the floor. Moreover, the rotational drift was highest in the initial frames and reached a steady state in a few seconds. Similar to the other studies (Nymoen et al. 2015; Phinyomark et al. 2018), the in-motion (i.e., lifting) activity data of the armband and the conventional sensor were visually compared since the quantitative comparison of both sensor signal data would not be appropriate. For the comparison of EMG and accelerometer signals, RMS and absolute magnitude plots were considered, as shown in Figures 5 and 6. The result shows that the armband and conventional sensor both pick the same peaks and follow a similar trend for lifting activity. The qualitative assessment of armband sensor position on EMG and IMU data quality shows that accelerometer and gyroscope data is almost similar for three (rotated, standard, and slid down) sensor positions. A previous study (Kefer et al. 2017) reported similar results where the classification accuracy using accelerometer data at different sensor positions made no significant difference. However, the EMG data for three

armband positions are significantly different, which conforms with the fact that the IMU sensor captures the motion of the forearm, whereas the EMG signal depends on the muscle contact.

The study assessed the relative and absolute reliability of forearm EMG and IMU data of construction activities. The test-retest evaluation of accelerometer data indicated an excellent trial-to-trial ( $ICC = 0.844$  to  $0.995$  and  $SEM\% = 0.087\%$  to  $0.258\%$ ) and day-to-day ( $ICC = 0.824$  to  $0.881$  and  $SEM\% = 0.245\%$  to  $0.526\%$ ) relative and absolute reliability for all the activities as shown in Table 5. Whereas for the gyroscope data, an excellent relative reliability was observed for trial-to-trial ( $ICC = 0.824$  to  $0.987$ ) and day-to-day ( $ICC = 0.801$  to  $0.844$ ) except for lifting where  $ICC = 0.724$  (Tables 6 and 7). The absolute reliability of gyroscope data for day-to-day was slightly greater than 10% ranging from  $5.224\%$  to  $16.322\%$ . The EMG data has shown excellent relative ( $ICC = 0.864$  to  $0.988$ ) and absolute ( $SEM\% = 2.420\%$  to  $10.509\%$ ) reliability between trials but the absolute reliability between the days ( $SEM\% = 7.75\%$  to  $16.21\%$ ) is slightly greater than 10% (Table 8). Overall, the results show that armband sensor data (acceleration, gyroscope, and EMG) exhibited excellent relative reliability between trials and days, which indicates a strong correlation of the repeated measurements. Furthermore, the armband sensor data exhibited excellent absolute reliability between the trials and moderate absolute reliability between days, indicating a slight increase in  $SEM\%$  and  $SDD\%$ . As shown in Equations (6) and (8),  $SEM\%$  and  $SDD\%$  are directly correlated to the ratio of SD and mean of the measurements. The higher  $SEM\%$  and  $SDD\%$  between days are due to the larger SD to mean ratio. Further investigation was performed to determine if the armband data obtained at this level of reliability is sufficient to yield accurate and reliable activity classification results.

The ML-based classification results using both days' datasets show that the forearm EMG, acceleration, and gyroscope features are capable of classifying activities involving different body

parts such as wrist, forearm, and whole-body and various motions such as repetitive motion, repeated impulsive motion, and free motion with high accuracy ( $\text{Day-1}_{\text{accuracy}} = 96.48\% \pm 0.0024$  and  $\text{Day-2}_{\text{accuracy}} = 96.33\% \pm 0.0022$ ). Furthermore, the overall classification accuracy of 98.13% achieved for nine uncontrolled activity datasets shows that the model is capable of recognizing activity with different intensities, which is one of the limitations of current construction activity recognition models (Akhavian and Behzadan 2018; Nath et al. 2017; Ryu et al. 2019). The accuracy of proposed activity recognition models using EMG and IMU forearm data ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 98.13\%$ ) is higher than previously published construction activity recognition models such as carpentry activities (91%) (Cezar 2012), fall identification (94%) (Lim et al. 2016), manual material handling activities (90.74%) (Akhavian and Behzadan 2016), ironworker activities (94.83%, 92.98%) (Yang et al. 2019; Zhang et al. 2018), and bricklaying activities (88.1%) (Ryu et al. 2019).

Moreover, the reliability assessment of classification results using Day-1 and Day-2 classifiers showed that there exists excellent reliability of classification results using the forearm EMG and IMU features. Later, the forearm EMG and IMU data were used to classify different weights of lifting activity, which is useful for various construction applications. The results show that the overall classification accuracy of three classes (Lift10, Lift25, and Lift50) is 83.89% (0.0051), which is higher than the accuracy obtained by Ho et al. (2017) (77.1%) in classifying barbell weights from 20 to 70 lbs using forearm EMG features. Moreover, for three lifting weights, the gyroscope and EMG features are highly correlated, which contributed to higher classification accuracy. The comparison of classification performance for different sensor combinations on controlled ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 96.21\%$ ,  $\text{Accuracy}_{\text{IMU}} = 94.65\%$ ,  $\text{Accuracy}_{\text{EMG}} = 44.97\%$ ) and uncontrolled ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 98.21\%$ ,  $\text{Accuracy}_{\text{IMU}} = 84.80\%$ ,  $\text{Accuracy}_{\text{EMG}} = 47.60\%$ )

dataset showed that the highest accuracy is obtained in case of EMG + IMU which is in agreement with the previous studies on forearm gym activities ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 71.6\%$ ,  $\text{Accuracy}_{\text{IMU}} = 67.8\%$ ,  $\text{Accuracy}_{\text{EMG}} = 20.7\%$ ) (Koskimäki et al. 2017), forearm manufacturing activities ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 87.4\%$ ,  $\text{Accuracy}_{\text{IMU}} = 85.0\%$ ,  $\text{Accuracy}_{\text{EMG}} = 50.7\%$ ) (Tao et al. 2018), and gym exercises ( $\text{Accuracy}_{\text{EMG} + \text{IMU}} = 84.2\%$ ,  $\text{Accuracy}_{\text{IMU}} = 77.7\%$ ,  $\text{Accuracy}_{\text{EMG}} = 85.2\%$ ) (Koskimaki and Siirtola 2016). Further, the increase in classification accuracy due to combined features shows that the gyroscope and EMG features obtained at higher SEM% and SDD% are suitable for activity classification. The fusion of forearm muscle activity (EMG) and kinematic (IMU) data has resulted in the highest classification accuracy for a greater number of complex activities with different intensities. The advantage of using an armband sensor is that both forearm muscle activity and motion data are obtained from the single device and avoids the use of multiple sensors that obstructs construction work.

Some of the limitations of the study are that the data quality of the sensor data was assessed only on at-rest activities. All the in-motion activities were performed in residential settings by participants with little construction experience. All the participants in this study were right-handed and male. In addition to acceleration, gyroscope, and EMG data, the armband sensor provides orientation quaternion and Euler angles of the forearm. However, the orientation angles were not assessed for reliability in this study. Moreover, one can consider performing a validity assessment for forearm EMG and IMU data of armband sensors.

### **3.5. Conclusions**

This study assessed the data quality and reliability of forearm EMG and IMU data from a low-cost wearable sensor for activity classification. In order to achieve the objective, the study was divided into seven experiments. The data was inferred from the first experiment that the

armband sensor data is comparable to conventional EMG and IMU data. Moreover, there was a minimal effect of environment, confounding factors (communication device, power tools, other sensors, and smartwatches), and inter-device variability. Second, a qualitative comparison was performed to understand the effect of armband position on forearm EMG and IMU data, and it was concluded that the armband position does not affect IMU data, but EMG data was affected due to the sensor position. Third, the trial-to-trial and day-to-day reliability of acceleration, gyroscope, and EMG data were assessed for five construction activities. The results conclude that the forearm EMG and IMU data for all five activities have excellent relative and absolute reliability between the trials and between the days, except for EMG data between the days has SEM% slightly higher than 10%. Next, both days, the EMG and IMU data were used to build and evaluate building ML-based activity classification models. The most common classification models were compared for the performance on the Day-1 and Day-2 datasets. The random forest classification algorithm showed the best performance on both datasets. The reliability test on the classification results of both the classifiers confirmed that the classification results are high reliability with minimal change inaccuracies for both days. The effect of lifting weight on classification performance was assessed, which concluded that the forearm EMG and IMU data could classify three different weights. Further, it was observed that a strong correlation in gyroscope and EMG features exists compared to accelerometer data for three classes. Finally, the comparison of classification performance for different sensor combinations showed that the forearm muscle activity and motion data fusion yield higher classification accuracy for construction activities with various intensities. The armband data is highly reliable, and the scientific evaluation of the armband sensor builds trustworthiness among researchers, policymakers, stakeholders, and customers to use the sensor for various applications. The data quality and reliability assessment of armband sensors shows that

the quality of muscle and motion-sensing data is sufficient for various construction applications related to construction skill training, safety training, and monitoring. Moreover, the classification results of the study conclude that the forearm-based EMG and IMU data can be used to generate reliable construction activity recognition models.

## **CHAPTER 4. AUTOMATED CONSTRUCTION ACTIVITY RECOGNITION<sup>‡</sup>**

### **4.1. 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 (Raynor 2020). However, the construction industry is facing a massive skilled workforce shortage (Kim et al. 2020). 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. (Future 2020). In particular, the demand for craft professionals is very high in the US Gulf Coast area, with an increase in petrochemical investments (Future 2020). 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 (Karimi et al. 2016). 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 (Ayodele et al. 2020). 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 (Ahn et al. 2019). 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,

---

<sup>‡</sup> This chapter is adapted from Bangaru, S. S., Wang, C., Busam, S. A., & Aghazadeh, F. (2021). ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. *Automation in Construction*, 126, 103653 (See Appendix E for permission).

wearable sensing technologies, and data mining and management) (Ahn et al. 2019; Aryal et al. 2017; Awolusi et al. 2018; Häikiö et al. 2020; Hwang et al. 2018). Moreover, these technology-based solutions help monitor and improve workers' performance by providing feedback (Luo et al. 2018; Roberts and Golparvar-Fard 2019; Sherafat et al. 2020). 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 an automated system for continuous workers' safety and performance monitoring (Sherafat et al. 2020), in which the fundamental requirement for the workers' performance and safety assessment is activity recognition (Chan and Kaka 2004; Cheng et al. 2018; Joshua and Varghese 2011; Wang et al. 2017).

Previous studies proposed various activity recognition systems using kinematic-based methods (Akhavian and Behzadan 2018; Ryu et al. 2019; Sherafat et al. 2020; Yang et al. 2019), vision-based methods (Albert et al. 2014; Khosrowpour et al. 2014; Luo et al. 2018; Roberts and Golparvar-Fard 2019; Yang et al. 2016), and audio-based methods (Cheng et al. 2018) 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 (Sherafat et al. 2020). 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 (Sherafat et al. 2020). 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 (Joshua and Varghese 2011; Ryu et al. 2019; Sherafat et al. 2020). 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 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 (Akhavian and Behzadan 2016; Bosch et al. 2015; Cezar 2012; Deb et al. 2017; Joshua and Varghese 2011; Joshua and Varghese 2013; Khan and Sohail 2013; Lim et al. 2016; Ryu et al. 2019; Sherafat et al. 2020; Yang et al. 2014; Zhang et al. 2018). Although most previous studies have achieved good accuracy, some limitations, such as restrictions on using smartphones on construction sites, pose challenges for practical implementation, the use of multi smartphones or IMU sensors is challenging while performing construction activities due to the dynamic nature of the construction work. 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 (Sherafat et al. 2020) and wearable sensor applications in construction safety and health (Ahn et al. 2019) 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 ten participants using the armband sensor while performing scaffold builder activities. Second, the collected data were preprocessed and prepared for bidirectional long short-term memory (BiLSTM) model building and training. Then, the BiLSTM model was trained and evaluated. Finally, the performance of the proposed model was evaluated on unseen data and compared the performance of different sensor combinations and classification algorithms.

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. It concludes with the discussions of the findings, limitations of the study, and future research directions.

## **4.2.Literature Review**

### **4.2.1. Human Activity Recognition, Deep Learning, and Sensor Modality**

An activity is defined as a group of actions that include a series of consecutive movements (Sherafat et al. 2020). Human activity recognition (HAR) involves predicting a person's movement based on sensor data and machine learning models (Vepakomma et al. 2015). HAR is broadly classified into two types, i.e., sensor-based and vision-based (Cook et al. 2013). 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. (Wang et al. 2019).

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 (Lara and Labrador 2012). First, acquiring the data from sensors such as accelerometers, gyroscope, magnetometers, electromyography sensors, audio sensors, vibration sensors, etc. Second, 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 (Munguia Tapia 2008). The deep learning models are preferred over traditional pattern recognition (PR) because of the following reasons (Chen et al. 2020; Wang et al. 2019):

- In traditional PR methods, the features are extracted through a heuristic and hand-crafted approach, which relies heavily on human experience and domain knowledge (Bengio 2013).
- Only statistical features such as mean, median, amplitude, frequency, minimum, maximum, etc., can be learned by the models from the human experience. 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 no labels (Chen et al. 2020; Cook et al. 2013; Wang et al. 2019).

The sensor modalities for human activity recognition (HAR) can be classified into body-worn sensors, object sensors, ambient sensors, and hybrid sensors (Chen et al. 2020). 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 (Chen et al. 2020; Vepakomma et al. 2015; Wang et al. 2019). 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 (Rashid and Louis 2019).

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 (Álvarez-García et al. 2013; Lee et al. 2019). 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 (Chen et al. 2020; Wang et al. 2019). 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 (Ordonez and Roggen 2016).

#### **4.2.2. 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 (Khusainov et al. 2013; Kim and Nussbaum 2013; Seel et al. 2014), and were also used for preventing musculoskeletal disorder by detecting awkward postures (Alwasel et al. 2017; Chen et al. 2017; Nath et al. 2017) and fall protection by identifying a sudden change in body acceleration (Lim et al. 2015; Wu et al. 2010; Yang et al. 2016). 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 (Hwang et al. 2016; Jebelli et al. 2018). The EMG sensors capture muscle activity used to assess the muscle load and forces used for ergonomic assessment (Nimbarte et al. 2010). The PPG, EDA, ST, and heart rate sensors were extensively used for assessing the workers' physical workload and fatigue (Abdelhamid and Everett 2002; Aryal et al. 2017; Chang et al. 2009; Maman et al. 2017). 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 (Hasanzadeh et al. 2017; Jeelani et al. 2018). 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 (Jebelli et al. 2019; Jebelli and Lee 2019). Even though several showed 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 (Ahn et al. 2019).

#### **4.2.3. 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 achieved 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 (Gong and Caldas 2009). On the other hand, wireless sensor-based systems are assessed to collect Spatio-temporal activity data (Kim et al. 2013). 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 (2010) 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 (2014) developed the accelerometer-based method for ironwork and carpentry activities classification using a decision tree with 90.07 and 77.74 percent accuracy. Cezar (2012) 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 (2013) 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 (2013) 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. (2014) 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. (2016) obtained an accuracy of 94% by using accelerometer data from a smartphone placed at the hip pocket. Akhavian and Behzadan (2016) 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 (Akhavian and Behzadan 2018). Ryu et al. (2018) 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 report that SVM with 4s window size showed the highest classification accuracy of 88.1%. Cheng et al. (2013) 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 (Yang et al. 2019) and ergonomic risk (Nath et al. 2018) 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 (Bao and Intille 2004). For activity recognition using accelerometers, the sensor's location close to the waist represents the significant body motions (Yang and Hsu 2010). However, waist-oriented acceleration signals do not reflect hand and arm movement, challenging to differentiate actions, including the movements (Ravi et al. 2005). The studies (Chernbumroong et al. 2011; Shoaib et al. 2016) 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 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 (Ahn et al. 2019; Sherafat et al. 2020).

#### **4.2.4. 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 (Sherafat et al. 2020). 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 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 (Akhavian and Behzadan 2016; Bayat et al. 2014; Intille et al. 2004; Kwapisz et al. 2011; Nath et al. 2017). 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 (Lee et al. 2019; Lim et al. 2015; Ryu et al. 2019; Tsai 2014; Yang 2009; Yang et al. 2014). The limitations of these studies are that 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) (Ann and Theng 2014; Cho et al. 2018; Koskimaki et al. 2009; Shoaib et al. 2016; Yang et al. 2015). 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 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 a large dataset to develop models, multiple sensors, the need for domain knowledge for feature extraction, human variability, and the inability to generalize the model. Moreover, there is a necessity to explore various preprocessing techniques such as feature engineering and hyperparameter optimization to develop robust and reliable models using an optimal number of sensors.

### **4.3. 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 a BiLSTM model. The proposed methodology's performance is evaluated through a series of analyses, such as the performance on unseen data, the performance of different sensor combinations, and the

comparison of performance with other classification algorithms. Each of these steps is further discussed in the following subsections.

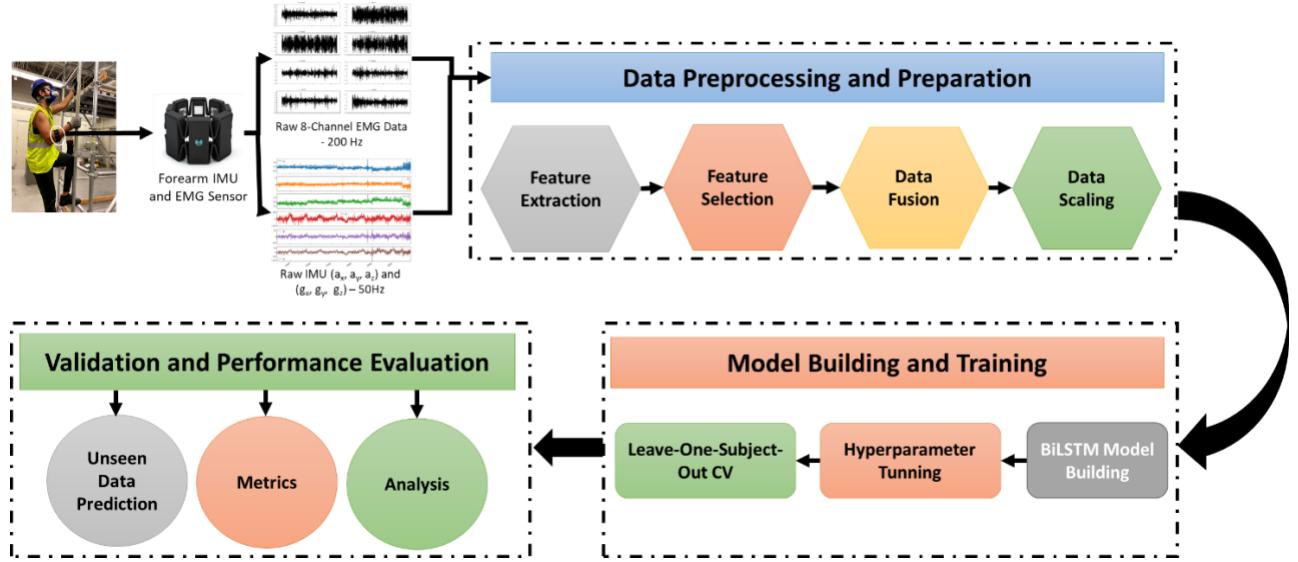


Figure 4.1. Framework for construction worker activity recognition using forearm-based EMG and IMU armband sensor

#### 4.3.1. 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 whole day on a single charge. The armband sensor is also easy to use, comfortable to wear for long periods, does 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 (Sathiyanarayanan and Rajan 2016) and entertainment (Guérat et al. 2019). 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 (See Figure B.1). In comparison, the IMU sensors capture the acceleration, angular velocity, and orientation of the forearm along x, y, and z directions (See Figure B.2).

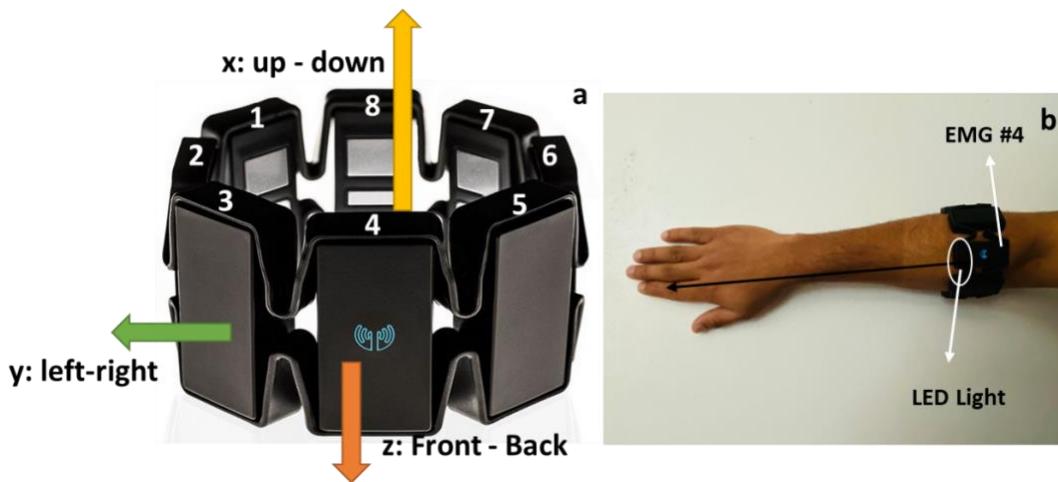


Figure 4.2. (a) Myo armband electrode location and IMU axes directions; (b) Myo armband placement on the forearm

#### 4.3.2. 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 (Ahn et al. 2019). In this study, the feature level sensor data fusion is considered for processing collected EMG and IMU sensor data. The feature

engineering process helps transform the raw signal data to represent the problem to predictive models better to improve accuracy on the unseen dataset. Since the EMG and IMU data are collected at different frequencies, the statistical features were extracted from EMG and IMU data for a 1s window. Later, both the features were fused at 1s frequency by matching the time stamps. 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 4.1. 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 avoids 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 (Shanker et al. 1996).

$$\text{Z-score} = \frac{x_i - \bar{x}}{\sigma} \quad (4.1)$$

#### **4.4. Model Building, Training, and Evaluation**

This study proposes a BiLSTM-based deep learning model for construction worker activity recognition. BiLSTM is a recurrent neural network that is a variation of long short-term memory

(LSTM). An LSTM consists of three gates (i.e., input, output, and forget gate) and two states (i.e., cell and hidden state), as shown in Figure 3. The LSTM networks are capable of learning long-term and temporal dependencies using gates. In LSTM architecture, the information from the previous hidden state and the current input will be sent to three gates, and the outputs from these are passed to the cell state, which carries the required information. Information in this state gets added or removed through gates. The gates are the neural networks and are responsible for deciding which information to remember or remove during training. The forget gate keeps the required information from the previous hidden state and the current input and passes the information through the sigmoid activation function, which gives the values between 0 and 1. The value closer to zero is to forget, and the value close to one is to remember. The input gate has two activations, sigmoid, and tanh, where information from the previous hidden state and the current input will be passed to a sigmoid function and keep the values required to forget gate. Furthermore, the hidden state and current input information are also passed to the tanh function to regulate the network. These two outputs will get multiplied and keep the required information. Information from the cell state, forget vector, will get added to the input vector to form a new cell state. In the output gate, the information from the previous hidden state and the current input is passed to a sigmoid activation function, and the current new cell state will be passed to the tanh activation function. These two outputs get multiplied and give the required information for the hidden state to carry to the next time step. A BiLSTM is generated from similar LSTM architecture, which can run inputs in both ways (i.e., past to future and future to past). A BiLSTM consists of two recurrent components, i.e., forward and backward components, where a forward component is similar to working in unidirectional LSTM, which computes cell state and hidden state. In contrast, the

backward component takes the input in reverse order. This makes the model robust to capture the required dependencies compared to LSTM.

A BiLSTM model can handle complex data by recognizing the data's hidden patterns and sensing the linear and non-linear relationship between independent and dependent variables by reducing the data's noise. A BiLSTM is capable of preserving information from the past and future at any given time. In this study, the BiLSTM model is built in Keras (Chollet 2015), 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.

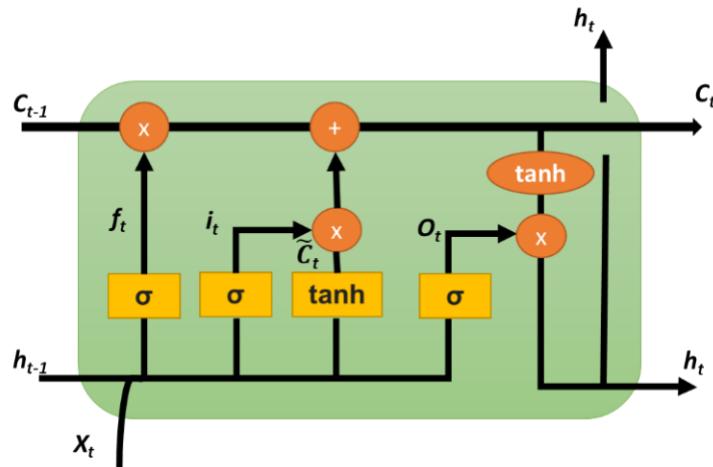


Figure 4.3. LSTM cell structure

#### 4.4.1. Hyperparameter Optimization

The BiLSTM 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 (Yang and Shami 2020). 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 selects 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 (Yang and Shami 2020). TPE is represented as  $p(y/x)$ , where  $y$  is the quality Score, and  $x$  represents hyperparameters, as shown in Equation 4.2.

$$p(y/x) = \frac{p(x|y)* p(y)}{p(x)} \quad (4.2)$$

$p(x/y)$  is a probability of hyperparameters given the value of an objective function, as shown in Equation 4.3.

$$p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases} \quad (4.3)$$

$l(x)$  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 fewer 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 previous studies' knowledge to suggest better hyperparameters for the subsequent evaluation. TPE is an algorithm that uses Bayesian reasoning to create a surrogate model and can use expected improvement to pick the next hyperparameter.

#### **4.4.2. 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 number of BiLSTM layers, the number 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 (Zhang 2018). It uses the functional combination of RMSProp

and SGD by using squared gradients and moving the average 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 helpful 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<sub>1</sub> and L<sub>2</sub> regularization, dropout, and early stopping. First, L<sub>1</sub> and L<sub>2</sub> 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. L<sub>1</sub> reduces the weights faster than L<sub>2</sub> 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<sub>1</sub> and L<sub>2</sub> regularizations. Second, 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.

#### **4.4.3. 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 underfitted 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 training and validation dataset's

loss and accuracy 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.

#### **4.4.4. 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 (Lara and Labrador 2012). 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 (Gholamiangonabadi et al. 2020).

#### **4.4.5. 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 matrix's diagonal cells 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, first and foremost 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 4.2 and 4.3, 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 4.5).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{Value of the Diagonal Cell of the Class}}{\text{Total Number of Predictions of the Class}} \quad (4.4)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{Value of the Diagonal Cell of the Class}}{\text{Total Number of Instances of the Class}} \quad (4.5)$$

If the classes are imbalanced, the most valuable and reliable metric to assess the model performance is the F1 Score, a harmonic mean of precision and recall, as shown in Equation 4.6.

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4.6)$$

The above formulas are used to calculate the performance metrics for individual classes, whereas the weighted precision, recall, and F1 Score following Equation 4.7 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 classes' samples.

$$\text{Weighted Metric} = \frac{\sum_{i=1}^m (\text{Metric}_i) * (\text{Samples}_i)}{\sum_{i=1}^m \text{Samples}_i} \quad (4.7)$$

Where m is the total number of classes, Metric<sub>i</sub> is the value of metric for class *i* (*i* = 1, 2, ..., m), and Samples<sub>i</sub> is the number of samples in each class *i* (*i* = 1, 2, ..., m).

In addition to the proposed model's performance evaluation, another analysis is conducted to evaluate the activity prediction accuracy on an entirely new dataset, i.e., an unseen 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 (Sherafat et al. 2020).

#### **4.5. System Feasibility Validation and Performance Evaluation**

##### **4.5.1. Case Study of Scaffold Builder Activities**

To validate and evaluate the proposed construction worker activity recognition model 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 (OSHA 2020). 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 (OSHA 2020). The supported scaffolds are extensively used in industrial and commercial construction projects (Halperin and McCann 2004). The scaffold building requires scaffold erection skills, carpentry

hand tools, and heavy labor-intensive tasks (Rubio-Romero et al. 2013). By reviewing various scaffolding activities onsite and online, we have recognized fourteen common activities to build a supported scaffold. The activities include free-style walking, carrying, or positioning a metal 5 ft. x 5ft. scaffold frame (38 Lbs.), carrying 24 in. leveling jacks (6.5 Lbs.), inserting or adjusting leveling jacks, carrying 7 ft. x 4 ft. scaffold crossbars (10 Lbs.), installing crossbars, hammering, wrenching, carrying or dragging a baseboard (33 Lbs.), installing baseboard on different level, carrying a 7 ft. x 12 in. guardrail (5 Lbs.), installing guarding, dragging guardrail, going up/down the vertical ladder. The fourteen scaffold builder activities and the activity ID used for the ANN model are summarized in Table 4.1. Some of the scaffold builder activities are shown in Figure 4.4. 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 4.1. Scaffold builder activities and activity ID

ID No.	Activity Description	Activity ID
0	Walking	WALK
1	Carrying / Positioning Scaffold Frame	CPSF
2	Carrying Leveling Jacks	CLJ
3	Inserting / Adjusting Leveling Jacks	IALJ
4	Carrying Crossbars	CC
5	Installing Crossbars	IC
6	Hammering	HAM
7	Wrenching	WRE
8	Carrying / Dragging Baseboard	CDB
9	Installing Baseboard on Different Level	IBDL
10	Carrying Guardrail	CG
11	Dragging Guardrail	DG
12	Installing Guardrail	IG
13	Going Up / Down Vertical Ladder	GUDVL



Figure 4.4. Shows a few scaffold builder activities performed in a warehouse environment (a) installing guardrail, (b) going up vertical ladder, (c) hammering, (d) carrying baseboard, (e) installing baseboard on a different level, (f) installing crossbars, and (g) carrying scaffold frame

#### 4.5.2. Experiment Setup

##### 4.5.2.1. Data Collection and Feature Engineering

To validate the feasibility and evaluate the proposed automated activity recognition model performance using forearm EMG and IMU data, an experiment was performed, which involved participants performing scaffold builder activities in the outdoor environment. Ten male college students have participated in the experiment. The participants' age ranged from 24 to 28 years (mean  $\pm$  SD:  $27 \pm 1.70$  years), weight ranged from 62.60 to 93 kgs (mean  $\pm$  SD:  $76.70 \pm 8.25$  kg), and height ranged from 168 to 178 m (mean  $\pm$  SD:  $171.7 \pm 4.13$  cm). 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 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 five minutes, with enough rest provided between the activities. The 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 ten participants' data were used for model building and evaluation, whereas the eleventh participant (age = 28 years, weight = 74 kgs., and height = 165 cm) was asked to perform all the activities in any sequence without any time constraint. The eleventh participant's data (unseen dataset) was used to test the trained BiLSTM model's performance. The whole experiment of the eleventh participant was videotaped and later used for evaluating the model performance. The ten participants' dataset was manually labeled, and feature engineering was performed on the raw EMG and IMU data. In total, 289 statistical features were extracted from raw acceleration, gyroscope, and EMG data. Among 289 features, the top 100 features were selected using SelectKBest with ANOVA F-value function. After feature engineering, the raw dataset is transformed into 48,515 samples and 100 features. The 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. Further, the transformed data [48515, 100] was reshaped to [timesteps, features, outputs] = [100, 100, 14]. This study observed that 100 timesteps were optimum to achieve the best accuracy without overfitting.

Table 4.2. Features extracted from raw EMG and IMU data

Dataset	Raw Features	Statistical Features
Acceleration	ax, ay, az, ACC	
Gyroscope	gx, gy, gz, GYRO	
EMG	EMG1, EMG2, EMG3, EMG4, EMG4, EMG5, EMG6, EMG7, EMG8, EMGsum	sum, avg, min, max, median, stdev, cv, var, percentiles (5, 10, 25, 75, 90, 95), skew, kurtosis

#### 4.5.2.2.Hyperparameter Optimization and BiLSTM 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 4.3. From the Bayesian TPE results, it can be observed that optimum performance was achieved for the BiLSTM architecture shown in Figure 4.5. The optimized model parameters include two BiLSTM layers with 100 units, two dropout (0.2) layers, and a dense layer (50 units) with ReLu activation function followed by Softmax, batch size = 64, epochs = 100, and optimizer = Adam. Since the batch size of 64 is used during the model training, 64 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. The BiLSTM cells were very effective in handling imbalanced classes.

Table 4.3. Parameters used for Bayesian Tree-structure Parzen Estimator optimization

Parameter	#	Values
No. of BiLSTM Layers	4	1 to 4
No. of BiLSTM Cells	3	50, 100, 150
No. of Dense Layers	3	50, 100, 150
No. of Dropout Layers	6	0.1, 0.2, 0.3, 0.4
Batch Sizes	3	32, 64, 128
Epochs	3	50, 100, 150
Optimizers	3	SGD, Adam, RMSprop
Activation Functions	3	ReLU, Tanh, Sigmoid

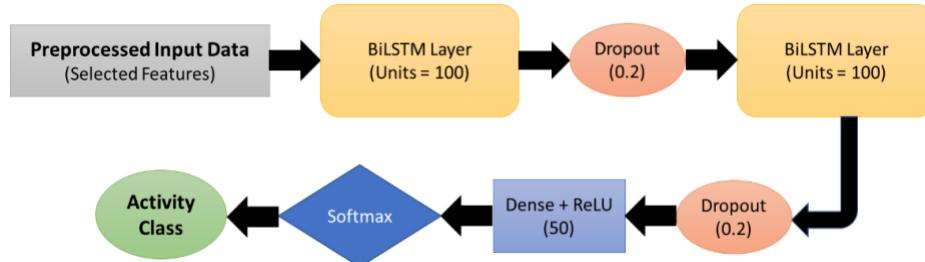


Figure 4.5. Optimal BiLSTM architecture for scaffold builder activities prediction

#### **4.5.3. 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 ten participants, ten 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 BiLSTM model generated after cross-validation on the ten 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 all the activities have achieved 0.99 to 1.00 classification accuracy. Table 3 presents the precision, recall, and F1 score values for all the activities. The results show that the average precision, recall, and F1 score values for all the activities range from 0.99 to 1.00. The overall prediction accuracy of 99.77% was obtained on the testing dataset with 0.99 weighted average precision, recall, and F1 Score.

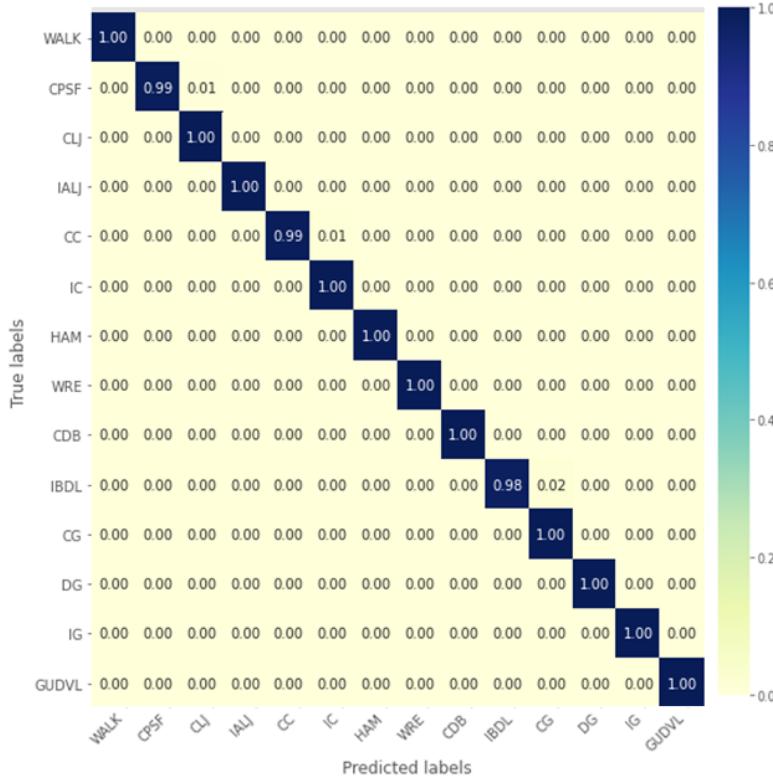


Figure 4.6. Confusion matrix of the proposed BiLSTM model using EMG and IMU data

Table 4.4. Class report of the proposed BiLSTM model using EMG and IMU data

Activity Description	Activity ID	Precision	Recall	F1-Score
Walking	WALK	1.00	1.00	1.00
Carrying / Positioning Scaffold Frame	CPSF	1.00	0.99	1.00
Carrying Leveling Jacks	CLJ	0.98	1.00	0.99
Inserting / Adjusting Leveling Jacks	IALJ	1.00	1.00	1.00
Carrying Crossbars	CC	1.00	0.99	0.99
Installing Crossbars	IC	0.99	1.00	0.99
Hammering	HAM	1.00	1.00	1.00
Wrenching	WRE	1.00	1.00	1.00
Carrying / Dragging Baseboard	CDB	1.00	1.00	1.00
Installing Baseboard on Different Level	IBDL	1.00	1.00	1.00
Carrying Guardrail	CG	0.99	1.00	0.99
Dragging Guardrail	DG	1.00	1.00	1.00
Installing Guardrail	IG	1.00	1.00	1.00
Going Up / Down Vertical Ladder	GUDVL	1.00	1.00	1.00

#### **4.5.4. Evaluation of Unseen Data**

The predictions were performed on the unseen dataset (eleventh participant data) to evaluate the model's robustness. The eleventh participant's evaluation has been performed using the trained weights of the BiLSTM model generated from LOSO cross-validation. The data was not used either in training or testing the model. The dataset consists of 5,188 samples and fourteen activities. During the eleventh 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 4.7 shows the confusion matrix of the proposed BiLSTM model on the unseen dataset. From the confusion matrix, it can be observed that the "IBDL" (0.52) and "CG" (0.53) activities were highly misclassified with "GUDVL" and "CG" activities, respectively. The highest classification was observed in "WALK" (1.00), "CPSF" (0.98), "CLJ" (0.94), "IALJ" (0.99), "CC" (0.99), "HAM" (0.98), "CDB" (0.98), and "DG" (0.96). Table 4.4 shows the precision, recall, and F1 score results of the BiLSTM model on an unseen dataset. The highest precision (1.00), recall (1.00), and F1 (0.98) scores were observed in "IBDL," "WALK," and "WALK." Whereas the lowest precision, recall, and F1 Score were observed in "GUVDL," "IBDL," and "GUDVL." Overall, the prediction accuracy of the proposed BiLSTM model on an unseen dataset is 87.18% with weighted average precision (0.90), recall (0.87), and F1 Score (0.87). Moreover, Figure 4.8 shows the predicted and actual sequence of activities performed by the eleventh participant. The proposed model recognized the activities' sequence as close to the actual activities with few errors in "CC," "CPSF," and "IC." The average time ratio difference between actual and predicted classes ranges between 1.34%, as shown in Figure 4.9. Moreover, the hypothesis test using paired t-test between the actual and predicted time ratio shows no significant difference (p-value = 0.5).

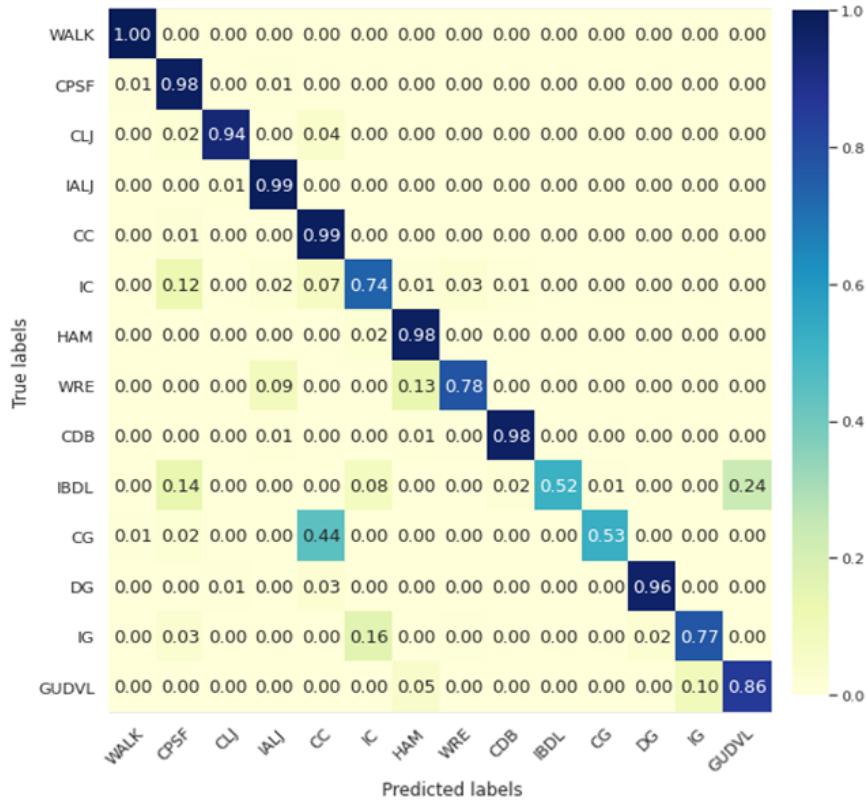


Figure 4.7. Confusion matrix of the proposed BiLSTM model on the unseen dataset

Table 4.5. Class report of the proposed BiLSTM model for the unseen dataset

Activity Description	Activity ID	Precision	Recall	F1-Score
Walking	WALK	0.97	1.00	0.98
Carrying / Positioning Scaffold Frame	CPSF	0.83	0.98	0.90
Carrying Leveling Jacks	CLJ	0.97	0.94	0.95
Inserting / Adjusting Leveling Jacks	IALJ	0.93	0.99	0.96
Carrying Crossbars	CC	0.63	0.99	0.77
Installing Crossbars	IC	0.63	0.99	0.77
Hammering	HAM	0.85	0.98	0.91
Wrenching	WRE	0.94	0.78	0.85
Carrying / Dragging Baseboard	CDB	0.99	0.98	0.98
Installing Baseboard on Different Level	IBDL	1.00	0.52	0.68
Carrying Guardrail	CG	0.98	0.53	0.68
Dragging Guardrail	DG	0.95	0.96	0.96
Installing Guardrail	IG	0.98	0.77	0.86
Going Up / Down Vertical Ladder	GUDVL	0.50	0.86	0.63

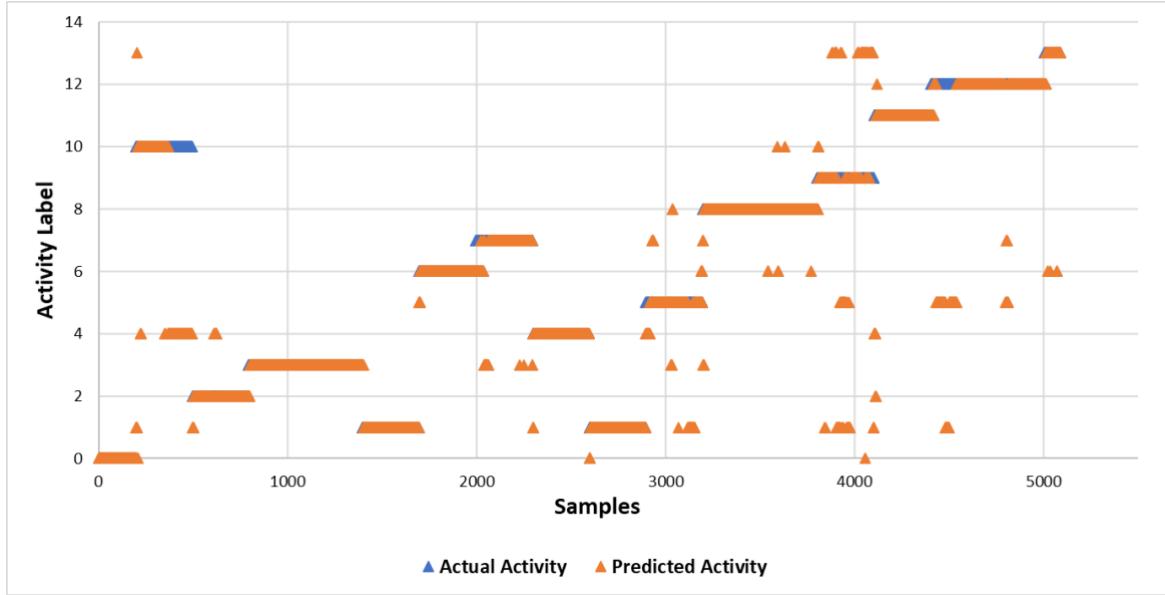


Figure 4.8. Performance of proposed BiLSTM model on the unseen dataset – A plot showing predicted and actual classes over the entire session

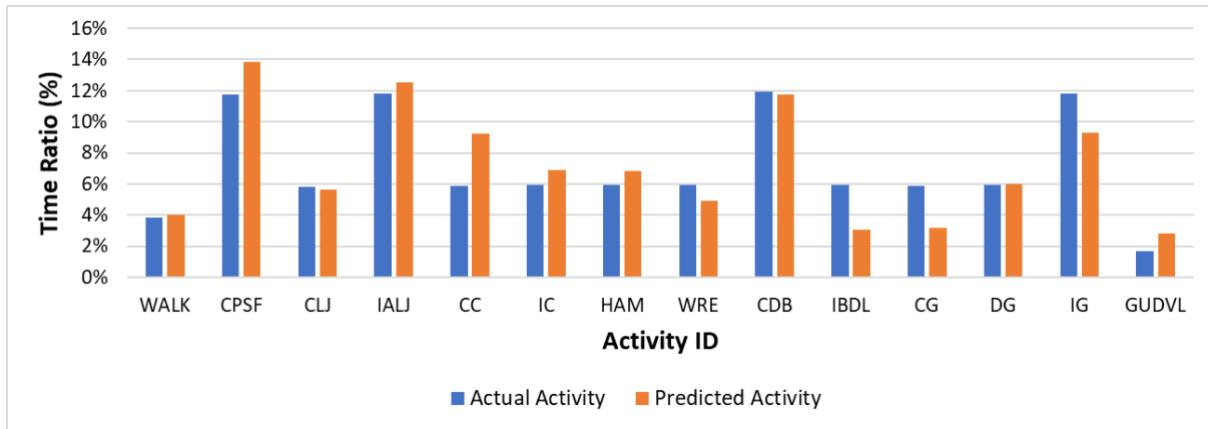


Figure 4.9. Time ratio of actual and predicted activities

#### 4.5.5. Comparison of Activity Recognition Performance for Different Sensor Combination

To understand the BiLSTM model performance for different sensor combinations, individual BiLSTM 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 on an unseen dataset are presented in Figure 10. From Figure 10, it can be observed that the EMG+IMU model achieved the highest accuracy of 87.18%, followed by IMU (85.85%), Acc (70.01%), Gyro (65.69%), and EMG (48.99%). 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 most of the classes except "CPSF," "CLJ," "HAM," and "WRE." Between EMG and IMU models, IMU outperformed EMG in the majority of the activities except for "IC" and "CG." This shows that for some of the activities involving complex motion, muscle activity improves activity classification performance. 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.

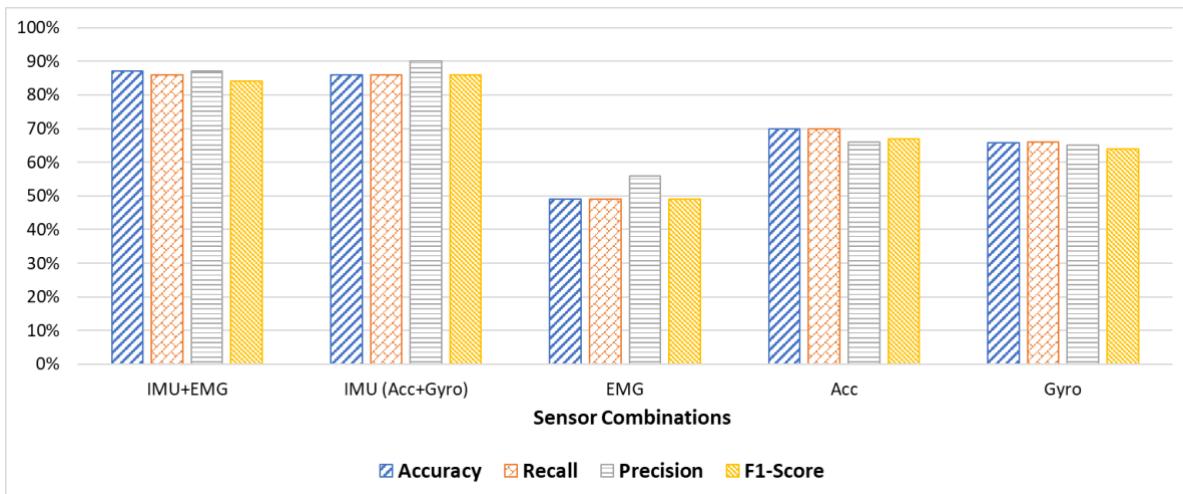


Figure 4.10. Comparison of classification performance of the BiLSTM model using different sensor combination on the unseen dataset

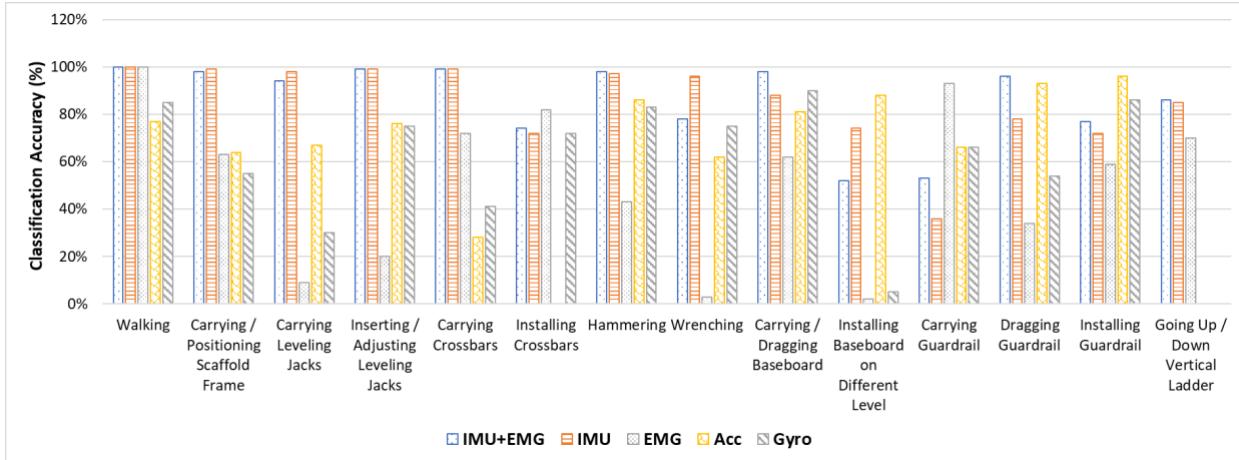


Figure 4.11. Comparison of classification performance of the BiLSTM model for all activities for different sensor data on the unseen dataset

#### 4.5.6. Comparison of Proposed Model with other Classification Models

It is essential to determine how well the proposed BiLSTM model performs compared to other classification algorithms. Therefore, the EMG+IMU dataset was further used to test with other existing classification models such as Light Gradient Boosting Machine (Lightgbm), Extra Trees Classifier (ET), Random Forest Classifier (RF), Gradient Boosting Classifier (GBC), Linear Discriminant Analysis (LDA), Decision Tree Classifier (DT), Ridge Classifier (Ridge), Logistic Regression (LR), K-nearest Neighbors (KNN), Quadratic Discriminant Analysis (QDA), Ada Boost Classifier (ADA), Naïve Bayes (NB), and Support Vector Machine (SVM). 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 BiLSTM (99.76%) model, followed by Lightgbm (82.47%), ET (79.76%), and RF (79.06%). In contrast, the least accuracy was obtained in the SVM classifier with 0.32, 0.42, and 0.30 recall, precision, and F1 Score, respectively.

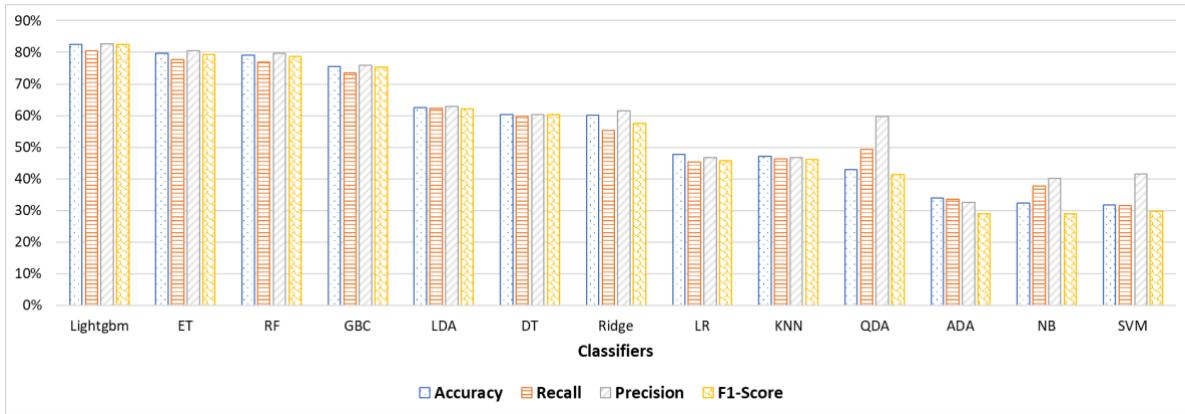


Figure 4.12. Comparison of activity recognition performance using the ANN model for different sensor data

#### 4.6.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 99.77% for all fourteen activities. Since the construction activities involve either muscle activity and body movement, EMG and IMU's use helps recognize complex construction activities involving different motions and body parts. From the current case study results, the activities involving motion such as adjusting jacks, dragging guardrail, lifting baseboard, walking, wrenching, and hammering, IMU data model achieved better accuracy compared to EMG model. Whereas in the case of activities involving muscle activity or material handling, such as installing crossbars and carrying guardrails, 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 (Sherafat et al. 2020). 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 unseen predictions show that the model can be used for continuous worker activity monitoring for safety, productivity, and project controls applications.

The use of recurrent neural networks such as BiLSTM has improved the imbalance classes' performance compared to other classifiers. As previous studies stated (Ahn et al. 2019), the feature extraction and sensor data fusion improved accuracy significantly by eliminating the redundancy and considering dependencies and correlation between different features. The hyperparameter optimization using Bayesian TPE automates network parameter selection, which helps adopt 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.

#### **4.7.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. The authors plan to develop one generic model to recognize multiple construction trades' activities using the proposed framework. Future research investigates recurrent neural networks' performance for various levels of sensor data fusion for construction activity recognition.

## 4.8.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, the use of BiLSTM and hyperparameter optimization helped to 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 BiLSTM model was able to classify fifteen scaffold builder activities with an overall testing accuracy of 99.77% and unseen prediction accuracy of 87.18%. 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 BiLSTM model on an unseen dataset for different sensor combinations showed that the classification accuracy was highest for EMG+IMU (87.18%), followed by IMU alone (85.85%) alone and EMG (48.99%) alone. The results show that the EMG data alone performed better than EMG+IMU and IMU alone for carrying guardrails and installing crossbars. 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 activity recognition accuracy. 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 proposed worker activity recognition system's main advantages 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 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.

## **CHAPTER 5. MAXIMUM AEROBIC CAPACITY FOR CONSTRUCTION ACTIVITIES**

### **5.1. Introduction**

The construction industry faces a considerable workforce shortage due to non-fatal injuries such as Work-related Musculoskeletal Disorders (WMSDs) (Bangaru et al. 2020). The incident rate of non-fatal injuries or illnesses in the construction industry is 2.8 cases per 10,000 full-time workers in 2019 (BLS 2019). According to the 2019 Liberty Mutual workplace safety index, the construction industries have spent \$189.81 million per week on serious non-fatal injuries. Among non-fatal injuries in the construction industry, WMSDs are more prevalent due to high labor-intensity tasks, working in the same posture for more extended periods, static work, repetitive tasks, and whole-body vibrations (Wang et al. 2017). One of the significant reasons for non-fatal injuries or illness is a mismatch between task demands and worker capabilities. The gap between task demands and worker capabilities leads to human errors, accidents, low productivity, and poor work quality. Therefore, it is essential to identify the gap and eliminate the mismatch to prevent overexertion and fatigue (Armstrong et al. 2001). Understanding the physical workload capacity helps design the construction activities that can be performed safely without overexertion or excess fatigue (Nath et al. 2017). The approaches of measuring physical activity's physical workload can be broadly classified into epidemiological, biomechanical, psychophysical, and physiological approaches (Chaffin et al. 2006). The physiological approach uses maximum aerobic capacity (MAC) or physiological fatigue limit (PFL) to determine the physical work capacity (Iridiastadi and Aghazadeh 2005). The MAC or Maximum oxygen uptake ( $\text{VO}_{2\text{max}}$ ) is essential to determine the physiological work capacity, which is defined as the maximum amount of oxygen that a person can consume during a maximal or exhaustive physical task (Åstrand et al. 2003). MAC is usually expressed in liters of oxygen consumed per minute or liters of oxygen consumed per kilogram

body weight per minute. Moreover, MAC is used as an indicator of an individual's endurance and determines the ability to perform moderate-to-high intensity tasks (Abut et al. 2016). In comparison, the PFL or continuous-work capacity is defined as the percentage of MAC which is the maximum allowable oxygen uptake that an individual can consume to perform prolonged physical activity. Below this limit, one can perform physical activity for a long time without getting tired. According to National Institute for Occupational Safety and Health (NIOSH), the average oxygen consumption during an 8-hour day is recommended to be no more than 33% of the exercise-specific MAC (Sharp et al. 1988). In practice, most construction jobs are designed using the MAC values of standard tasks such as treadmills, cycle ergometer, and hand cranks (Ayoub 1989; Stanton et al. 2013). Consequently, standard MAC values for the work capacity evaluation lead to injuries and excessive fatigue (Iridiastadi et al. 1997; Sharp et al. 1988).

MAC can be measured using two methods, namely direct (maximal) and indirect (submaximal) methods (Åstrand et al. 2003). In the direct measurement method, the subject is stressed progressively to the maximum exercise, and the corresponding oxygen consumption at that level is considered the MAC. The direct measurement of MAC using the maximal exercise test is highly accurate (Ladyga and Faff 2005). However, the maximal methods result in exhaustion and require trained staff with costly equipment. Moreover, maximal methods have a slight risk of heart attack, serious injury, or even death (Acevedo 2012), and they are not applicable for construction applications since it requires sophisticated equipment and medical supervision. Whereas, in the indirect (submaximal) methods, the subject must perform at least three different workloads. The heart rate and oxygen consumption are recorded at three workloads and plotted on an X-Y graph. The oxygen consumption at the maximum heart rate is considered as the MAC. The maximum heart rate is determined as (220-age) (Søgaard et al. 1996). The MAC resulting from

the submaximal method deviates at most 15% of the direct method. The submaximal methods are considered safer to administer and practically applicable (Åstrand et al. 2003). The submaximal techniques are suitable for construction applications, but there are no experiment protocols or guidelines for construction activities. Due to the practical limitations of measurement methods, various regression models using machine learning and statistics are developed for MAC predictions.

MAC predictive models are developed using one of the three tests: exercise tests, non-exercise tests, and hybrid tests. The exercise tests consist of maximal and submaximal tests. Maximal tests require the subject to reach exhaustion in terms of the heart rate, which has two substantial advantages over MAC direct measurement. First, the maximal tests are affordable and capable of predicting the MAC at acceptable accuracy. Second, unlike direct measurement, the maximal tests do not require high medical supervision and specialized laboratory (Abut et al. 2016). Due to some of the practical limitations of maximal tests, such as time and medical supervision, submaximal exercise tests have increased attention. The submaximal exercise tests are easy to administer, cheaper, and safer. The only disadvantage of submaximal tests over maximal tests is prediction accuracy (Hunn et al. 2002). Whereas non-exercise tests do not require subjects to perform any exercises, instead use questionnaire data with additional variables such as age, gender, height, and weight to predict  $\text{VO}_{2\text{max}}$ . Non-exercise tests are the most affordable and easy to apply to a large population. The major disadvantage of the non-exercise tests is that they are dependent on the individual self-report (McComb et al. 2006).

Therefore, various studies have developed activity-specific MAC by designing the protocol as close to the actual work situation using the submaximal exercise mode. The previous studies have identified that the MAC value for repetitive lifting and combined manual material handling

(CMMH) tasks is significantly lower than the MAC value of standard treadmill and running exercise modes (Christie and Scott 2005; Iridiastadi and Aghazadeh 2005; Sharp et al. 1988). Moreover, the previous studies have suggested that when using the physiological approach through MAC value, the exercise protocol should be as close to the job requirements when evaluating the physical workload capacity or determining the workers' physiological response to the task to prevent overexertion or excess fatigue (Ayoub 1992; Ayoub 1989; Christie and Scott 2005; Craig et al. 1998; Iridiastadi and Aghazadeh 2005; Iridiastadi et al. 1997; Sharp et al. 1988). Even though previous studies have assessed the work capacity using individual activities such as repetitive lifting, no standard exercise protocol was developed for construction-specific MAC. Moreover, it is unclear that the individual activity MAC values can assess the MAC values of combined activities since most of the construction activities in a real work scenario involve multiple activities performed together. Therefore, it is essential to determine the construction activity-specific MAC values and investigate if those can assess combined activities MAC values. Therefore, the study's objective is to compare the MAC values of individual and combined construction activities. In this study, it is hypothesized that the MAC value of individual activities is higher compared to the combined construction activities. To achieve the proposed objective, we have developed activity-specific MAC estimation models using the submaximal exercise protocol close to the actual construction work situation and performed statistical analysis to verify the hypothesis.

## **5.2.Materials and Methods**

### **5.2.1. Participants**

Ten male active university students between the age group of 24 and 29 years participated in the experiment. The participants' average age, height, and weight were  $27.3 \pm 1.70$  years,  $171.90 \pm 3.87$  cm, and  $76.77 \pm 8.50$  kg. All the participants were in good physical condition with no

musculoskeletal or heart problems. All the participants were untrained construction workers. Written informed consent was obtained from all the participants after explaining the study's objective, experimental procedure, and any consequences resulting from the experiment (See Appendix A). The Institutional Review Board at Louisiana State University (ID: IRBAM-20-0539) approved the research study consistent with the Declaration of Helsinki.

### **5.2.2. Measurements and Instrumentation**

In this study, the oxygen consumption ( $\text{VO}_2$ ) was measured using a wearable metabolic analyzer, the VO<sub>2</sub> Master Analyzer (VO<sub>2</sub> Master Health Sensors Inc., Vernon, British Columbia, CA) (See Figure 5.1). The participant's heart rate was measured using the chest strapped Wahoo Tickr heart rate monitor (Wahoo Fitness, Atlanta, GA, USA). The  $\text{VO}_2$  and HR data were recorded continuously at a frequency of 1 Hz in liters/min and bpm.



Figure 5. 1. Participant wearing wearable metabolic analyzer.

### **5.2.3. Construction Activity Description**

Construction activities are dynamic and involve complex body motion with a heavy workload. In this study, authors have considered construction activities such as walking, carrying, lifting, hammering, wrenching, and dragging related to scaffold building which is one of the labor-intensive tasks. Moreover, scaffold building activities involve different body parts (wrist, forearm,

upper body, lower body, and whole-body) movements, various motions (free motion, impulsive motion, and repetitive motion), and workload intensities (low, moderate, and heavy) (Bangaru et al. 2021). The detailed description of each activity is as follows:

- *Walking Activity*: The walking activity involves walking on a treadmill at a defined speed at zero inclination, which corresponds to a construction site's walking activity.
- *Carrying Activity*: The carrying activity involves carrying weight in a wooden box at elbow height at their own pace. This activity is similar to carrying baseboard, guardrail, crossbars, or scaffold frames.
- *Lifting Activity*: The lifting activity involves lifting 10 lbs. weight in a wooden box (10lbs.) from knuckle height to shoulder height. An adjustable wooden platform is used for lifting activity where the base level is at knuckle height (75.5 cm from the floor) and the shoulder level is at 145 cm from the floor. This activity corresponds to lifting various objects such as baseboard, guardrail, or scaffold frame.
- *Combined Construction Activity*: The combined construction activities involve carrying, dragging, hammering, lifting, and wrenching. The hammering activity involves hammering nails to the baseboard at a defined frequency. In contrast, the wrenching activity involves unscrewing using a heavy-duty pipe wrench for a defined time duration. Finally, the drag activity involves dragging a scaffold platform (7 ft. x 19 in.) at elbow height. The platform weighs 33.19 Lbs. Each of these activities is performed for one minute each.

#### **5.2.4. Experimental Protocol**

Each participant performed all four activities at three different intensities for five minutes each. Before the start of the experiments, the participants were asked to warm up to prevent any injuries. First, the participant performed a walking activity on a treadmill with zero inclination at

three different speeds, i.e., 2.5, 3.5, and 4.5 mph. Second, the participant performed a carrying activity using three different weights, i.e., 20, 25, and 30 lbs. (including box weight). Later, the participant performed lifting activity at 4, 10, and 15 lifts/min. Finally, the participant performed combined construction activities that involve carrying, dragging, hammering, lifting, and wrenching. The combined activities were performed at three different intensities using 20, 25, and 30 lbs. All four activities were performed continuously with one-minute rest between the activities. The heart rate and oxygen consumption were recorded for all the sessions. The order of the activities was randomized and performed in the laboratory at 72°F. The participants' effort was considered maximal if the physical signs suggestive of exhaustion and at least two of the four criteria were achieved: (i) maximal heart rate value ( $HR_{max}$ ) is greater than 85% of age-predicted maximum (220-age), (ii) the respiratory exchange ratio (RER) is more significant than 1.15, (iii) leveling of oxygen uptake despite an increase in workload, and (iv) the rate of perceived exertion (RPE) is 19 on the Borg scale (De Brabandere et al. 2018). The maximum oxygen consumption for a particular intensity of an activity is the highest 15 seconds average value over five minutes, whereas the heart rate is the highest single value recorded during the session (Akay et al. 2011).

The experimental protocol is summarized in Figure 2.

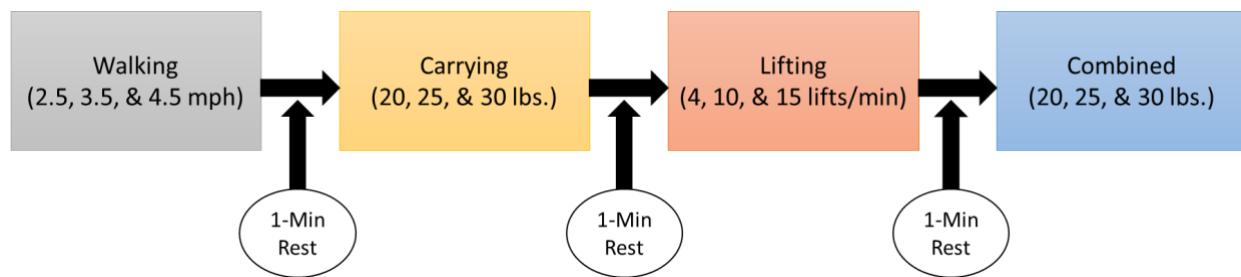


Figure 5. 2. Experimental protocol for MAC estimation using submaximal exercise test.

### **5.2.5. Data Analysis**

The recorded heart rate and oxygen consumption were further used for statistical analysis using SPSS 23.0 for Windows. Regression analysis was performed to develop MAC estimation models where oxygen uptake is the dependent variable and heart rate is the independent variable. Analysis of Variance (ANOVA) was used to test the significant difference in MAC value for different construction activities.

## **5.3. Results**

### **5.3.1. Maximum Aerobic Capacity of Construction Activities**

The regression models were developed using the maximal oxygen uptake and heart rate at three different intensities. Table 5.1 shows the MAC estimation models for each construction activity. The MAC value can be estimated using the predicted maximum heart rate of an individual (220– age) (Vandersmissen et al. 2014). From Table 5.1, it can be observed that there exists a strong positive correlation between oxygen uptake and heart rate for all four activities since R-value ranges from 0.92 to 0.99. Similarly, the high R<sup>2</sup> value in the case of walking (0.99), carrying (0.98), and lifting (0.99) shows a good fit. The combined construction activities, since the R<sup>2</sup> value is 0.85, indicate a fairly good fit. The standard error value is lower for lifting and carrying, followed by combined construction activities and walking.

Table 5.1. VO<sub>2max</sub> estimation equations for construction activities using submaximal exercise test

Activity	MAC Equation	R	R <sup>2</sup>	SEE
Walking	$VO_{2\text{max}} = 0.0192 * (HR_{\text{max}}) - 0.7571$	0.9934	0.9869	5.1358
Carrying	$VO_{2\text{max}} = 0.0219 * (HR_{\text{max}}) - 1.6663$	0.9814	0.9631	2.0761
Lifting	$VO_{2\text{max}} = 0.0225 * (HR_{\text{max}}) - 1.5209$	0.9997	0.9995	0.5400
Combined Construction Activities	$VO_{2\text{max}} = 0.0209 * (HR_{\text{max}}) - 1.7241$	0.9270	0.8593	3.3268

### 5.3.2. Comparison of MAC Values of Different Construction Activities

An average MAC value of each construction activity was estimated using the MAC estimation models and the maximum heart rate of an individual. From Figure 5.3, it can be observed that the average MAC value for ten participants is higher for walking ( $2.95 \pm 0.03 \text{ l/m}$ ) and followed by lifting ( $2.82 \pm 0.04 \text{ l/m}$ ), carrying ( $2.55 \pm 0.04 \text{ l/m}$ ), and combined activities ( $2.30 \pm 0.04 \text{ l/m}$ ). Further ANOVA analysis was performed between the MAC values of four activities to test the significant difference at a significance level of 0.05. The ANOVA results ( $p\text{-value} = 4.12891\text{E-}31$ ) show that the average MAC values of four construction activities are significantly different from each other ( $\alpha = 0.05$ ). The results conclude that the MAC value of combined construction activities is lower than the individual construction activities. The percentage difference in MAC value between the combined construction activities and other activities are 28% (walking), 11% (carrying), and 22% (lifting).

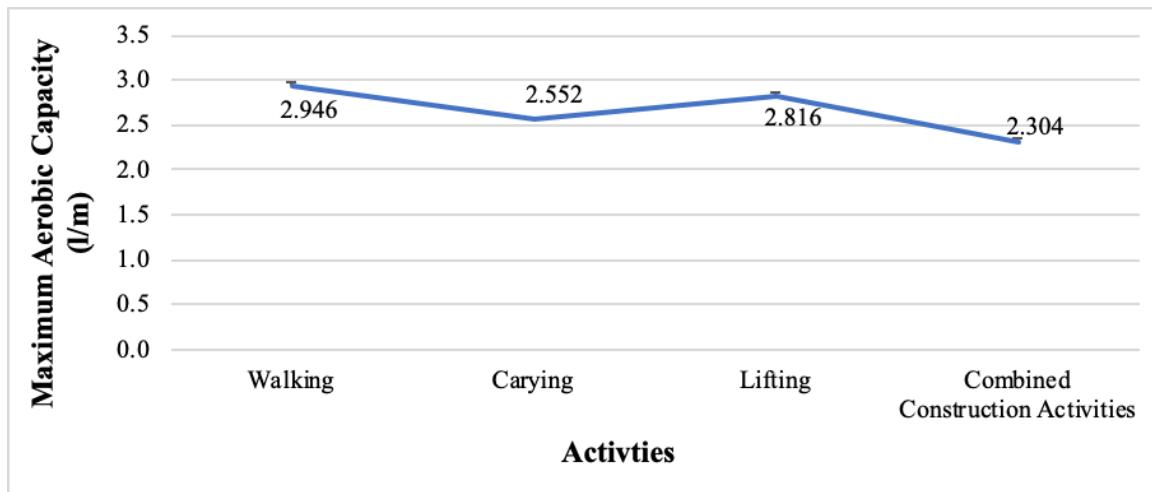


Figure 5.1. Average MAC value for different construction activities

Further, analyzing the heart rate and oxygen consumption for four activities at different intensities shows that for similar heart rate ranges, the oxygen consumption is higher for individual activities than the combined activities (Table 2). This is because the combined activities involve

both static and dynamic work, making the heart work harder. Besides, for the same increase in the heart rate value, the oxygen consumption is higher for individual activities than the combined activities. This shows that the physiological response to individual activities is more efficient than the combined construction activities.

Table 5.2. Heart rate and oxygen consumption response to different activity intensities

Activity	Intensity-1		Intensity-2		Intensity-3	
	HR (bpm)	VO2 (l/m)	HR	VO2 (l/m)	HR	VO2 (l/m)
Walking	106.13	1.23	123.13	1.69	167.50	2.44
Carrying	124.43	1.04	132.14	1.26	139.71	1.37
Lifting	120.83	1.19	134.00	1.50	154.50	1.95
Combined Construction Activity	129.33	0.99	139.45	1.13	140.81	1.27

#### 5.4.Discussion

The main objective of this study was to determine the construction activity-specific maximum aerobic capacity value. The study hypothesized that the MAC value of individual construction activities is higher than the combined construction activities. The results conclude that the MAC value of individual activities such as walking, carrying, and lifting is higher than the MAC value of combined construction activities. Moreover, the ANOVA analysis has shown that the difference between individual and combined activity MAC is significantly different. These results agree with the previous studies that stated that the MAC value of manual material handling tasks is lower than the standard treadmill, running, and bicycle tasks (Christie and Scott 2005; Iridiastadi et al. 1997; Sharp et al. 1988). It is evident from the results that the lower body protocol using treadmill tests indicates a higher value than lower body or combined activity protocol, which would force the workers to exceed the acceptable limits, resulting in injuries (Craig et al. 1998). One of the main reasons for lower MAC value in combined construction activities is significant

static work. This study has proved that combined construction activities' MAC value cannot be assessed from the individual activity MAC values and vice versa.

Some previous studies have attempted to assess the MAC value of manual material handling tasks, namely repetitive lifting. Ayoub (1989) has reported that the MAC value for lifting six lifts/minute activity was between 2.94 l/m to 2.96 l/m. The average lifting MAC value observed in this study is 2.82 l/m is comparable with the previous studies, which stated that lifting at higher frequency results in a higher MAC value (Ayoub 1992; Sharp et al. 1988). This shows that the proposed exercise protocol can develop MAC values for different construction jobs by making necessary modifications. Further, the study identified that the lower oxygen uptake in the combined activities is due to static work, which resulted in increased heart rate and lowers oxygen uptake.

One of the significant limitations of this study is the smaller number of participants and only one construction activity combination. The study assumed that the relationship between oxygen uptake and heart rate is linear. Further studies on this issue can validate this assumption. Future studies on construction activity-specific MAC estimation can consider different construction activity combinations, various heavy workload activities with different task characteristics, and various tasks and individual characteristics in the MAC estimation models such as age, gender, BMI, speed, load, and fatigue level.

## **5.5.Conclusions**

The study concluded that the construction activity-specific MAC value is different from the standard treadmill MAC value. Moreover, the MAC value of combined construction activities is lower than the MAC value of individual activities such as walking (28%), carrying (11%), and lifting (22%). The study determined the MAC value of combined construction activities similar to

real work situations where a worker performs multiple activities. The MAC values and procedure proposed in this study can be used for construction activity design and evaluation. This study's findings recommend that the worker be tested using construction-specific MAC protocol to get an accurate measure of workers' physiological response to the heavy workload jobs instead of using a standard treadmill or bicycle protocol to prevent injuries and overexertion.

## **CHAPTER 6. OXYGEN UPTAKE PREDICTION DURING CONSTRUCTION ACTIVITIES**

### **6.1. Introduction**

The construction industry often involves high labor-intensive and repetitive tasks, which results in worker physical fatigue. About 40% of the U.S. construction workforce experiences high-level fatigue, which leads to poor judgment, increased risk of injuries, a decrease in productivity, and a lower quality of work (Abdelhamid and Everett 2002; Cheng et al. 2012). Further, excessive fatigue due to working in unpleasant working conditions, long working hours, and heavy workloads can aggravate the adverse effects of fatigue and leads to Work-related Musculoskeletal Disorders (WMSDs) and productivity loss (Aryal et al. 2017; Jebelli et al. 2019). Moreover, fatigue has been shown to result in impairing the physical and cognitive functions (Zhang et al. 2015) and identified as a possible risk factor for slip-induced falls, which is one of the "fatal four" causes of fatalities in the construction industry, according to Occupational Safety and Health Administration (OSHA 2018). According to the 2019 Liberty Mutual workplace safety index, the construction industries have spent \$189.81 million per week on serious non-fatal injuries. In order to prevent or reduce the physical fatigue or physiological demands associated with construction activities, it is essential to monitor the physiological effort at which workers perform the construction activity and evaluate if the effort exceeds physiological standards (Abdelhamid and Everett 2002; Gatti et al. 2014). The physical demand or physiological workload of construction work is determined by measuring the average oxygen uptake (or energy expenditure) and heart rate while performing the construction activity (Abdelhamid and Everett 2002; Gatti et al. 2014). To assess the physical demands of construction activities, early efforts were made by Abdelhamid and Everett (2002) to measure the physiological demands such as heart rate and oxygen uptake using the KB1-C metabolic system, which was cumbersome and

uncomfortable. In recent years, with advancements in wearable sensors and machine learning, various authors have developed objective to assess the workers' physical demand using physiological sensors such as heart rate, skin temperature, photoplethysmogram (PPG), and electrodermal activity (EDA) (Aryal et al. 2017; Gatti et al. 2014; Hwang and Lee 2017; Jebelli et al. 2019). The machine learning algorithms help identify non-trivial and complex patterns from biological signal data captured using wearable sensors (Witten et al. 2016). Previous studies used percentage change in heart rate (Gatti et al. 2014), rating of perceived exertion (Aryal et al. 2017), and energy expenditure (Jebelli et al. 2019) to classify the construction work performed by the worker into physical demand level (low – high). Even though heart rate, skin temperature, PPG, and EDA are proved to monitor individual physical workload, but it is not sufficient to continuously monitor the physical demand of workers with different individual characteristics (age, gender, and work experience), task characteristics (complex activities in a short interval of time with varying workloads), and do not help in quantifying the direct impacts of physical workload of activity on construction safety performance or accidents. To overcome the above limitations, there is a necessity to develop a system that can measure physiological demands using signals dependent on the task and individual characteristics. Therefore, this study proposes to use forearm-based inertial measurement units (IMU) and electromyography (EMG) sensors to predict oxygen uptake during construction activities. The continuous monitoring of oxygen uptake ( $\text{VO}_2$ ) helps to evaluate if the physical demand of the activity exceeds the physiological standards, determines the worker's physiological status, and early detection of accidents. Moreover, the motion and muscle activity data obtained from the sensor while performing the activity are influenced by the activity characteristics, individual characteristics, and work conditions, which helps overcome the limitations of the previous studies. The frequency and quality of data obtained

from the armband sensor were proven to be sufficient to recognize the complex construction activities in a short interval of time (Bangaru et al. 2020; Bangaru et al. 2021).

Recent studies have developed oxygen uptake prediction models using wearable sensors and machine learning for light to moderate intensity activities such as treadmill walking (Borror et al. 2019; Shandhi et al. 2020), daily living activities (Altini et al. 2015; Beltrame et al. 2017; Lu et al. 2018; Shandhi et al. 2020), cycling exercise (Zignoli et al. 2020), and outdoor walking (Shandhi et al. 2020). These studies used physiological and motion sensors such as electrocardiogram, seismocardiogram, atmospheric pressure, heart rate, respiratory band, Garmin vector power meter, and accelerometer. Except for Shandhi et al. (2020), all the other studies used multiple sensors on different body parts to capture the input data for  $\text{VO}_2$  prediction models. The use of multiple wearable sensors on construction workers while performing the work is cumbersome and uncomfortable. The oxygen uptake prediction models were developed using machine learning algorithms such as Logistic regression (Altini et al. 2015), XGBoost (Shandhi et al. 2020), Random Forest (Beltrame et al. 2017), Multilayer Perceptron Neural Network (Lu et al. 2018), Artificial Neural Network (Borror et al. 2019), and Recurrent Neural Network (Barut et al. 2020; Zignoli et al. 2020) and features such as absolute acceleration, heart rate, cadence, breathing frequency, and power output. Recurrent neural networks such as Long Short-Term Memory (LSTM) have been effective for sequence data and regression (Barut et al. 2020; Ordóñez and Roggen 2016). Even though these studies have achieved acceptable model performance in predicting oxygen uptake, some of the limitations of these studies are only light to moderate activities were evaluated, use of multiple sensors for data acquisition, and inability to predict  $\text{VO}_2$  for high intensity and complex activities performed in short intervals. Therefore, the objective of

this study is to develop oxygen uptake prediction models for complex construction activities with optimum number of sensors.

In this study, we hypothesized that the combination of forearm muscle activity and motion features predict oxygen uptake more accurately compared to individual features. The study proposes bidirectional long-short term memory (BiLSTM) based oxygen uptake prediction framework using the forearm EMG and IMU data. A scaffold building activity was performed to evaluate the performance of the proposed prediction model. The performance of the proposed method was compared with the existing methods. The results show that the proposed method has improved results for complex high-intensity construction activities. The study also compares the performance of the different sensor features (i.e., IMU alone, EMG alone, and EMG+IMU) and other Recurrent Neural Network (RNN) models. Finally, the study evaluates the average oxygen uptake required to build one scaffolding unit using the proposed model.

## **6.2.Materials and Methods**

### **6.2.1. Data Collection**

#### **6.2.1.2.Participants**

Ten active male university students participated in this study ( $27 \pm 1.70$  years,  $171.7 \pm 4.13$  cm,  $76.70 \pm 8.25$  kg). The activity level of the participants was moderate to vigorous. All the participants were right-handed, non-smokers, and had no lower-back injuries or musculoskeletal disorders. Before starting the experiment, the study's objective was demonstrated to the participants, and written informed consent was obtained from all the participants (See Appendix A). The experiment protocol consistent with the Declaration of Helsinki was approved by the Institutional Review Board (IRB) at Louisiana State University (ID: IRBAM-20-0539).

### **6.2.1.3. Construction Activity Description**

Construction activities involve heavy labor-intense tasks and complex motions. In this study, the authors have considered scaffold building activities considered one of the heavy workload activities (Bangaru et al. 2021). One of the significant reasons to choose scaffold building activities involves moderate to heavy workload tasks involving different body part movements (wrist, upper body, forearm, lower-body, and whole-body) and various motions (free motion, repetitive motion, and impulsive motion). The study has identified fourteen scaffold building tasks, as shown in Table 6.1. The tasks involve carrying/installing different objects such as scaffold frame (38 Lbs.), crossbars (10 Lbs.), leveling jacks (6.5 Lbs.), baseboard (33 Lbs.), and wooden guardrail (5 Lbs.). Also, it involves going up/down vertical ladder. Figure 1 shows some of the scaffold-building tasks performed by the participant.

Table 6.1. Scaffold building activities

SL. No.	Activities
1	Walking
2	Carrying / Positioning Scaffold Frame
3	Carrying Leveling Jacks
4	Inserting / Adjusting Leveling Jacks
5	Carrying Crossbars
6	Installing Crossbars
7	Hammering
8	Wrenching
9	Carrying / Dragging baseboard
10	Installing Baseboard on Different Level
11	Carrying Guardrail
12	Dragging Guardrail
13	Installing Guardrail
14	Going Up / Down Vertical Ladder



Figure 6.1. Shows participants' performing different scaffold building activities. (a) hammering, (b) carrying scaffold frame, (c) carrying baseboard, (d) carrying guardrail, (e) going up the vertical ladder, (f) installing baseboard on a different level, (g) installing crossbars, and (h) adjusting leveling jacks.

### 6.2.2. Measurements and Instrumentation

The study proposes to use a forearm-based wearable armband sensor (Myo Armband) developed by Thalmic Lab Inc. to collect forearm inertial measurement unit (IMU) and electromyography (EMG) data. The armband consists of eight dry EMG surface electrodes and a 9-axes IMU sensor (3-axes accelerometer, 3-axes gyroscope, and 3-axes magnetometer). The IMU sensor is embedded in the EMG channel 4. The EMG electrodes capture the forearm muscle activity and return an 8-bit array of integer values ranging between -128 and 127, acquired at 200 Hz frequency. In contrast, the IMU data captures the forearm's motion by measuring acceleration, angular velocity, and orientation in x, y, and z-direction at 50 Hz frequency. The real-time raw EMG and IMU data were transmitted from the armband sensor to local computer storage via Bluetooth Low Energy (BLE) wireless connection. To capture the gold standard breath-by-breath oxygen uptake while performing construction activities, a portable metabolic analyzer, the VO<sub>2</sub>

Master Analyzer (VO2 Master Health Sensor Inc., Vernon, British Columbia, CA), as shown in Figure 6.2. The metabolic analyzer records the oxygen uptake ( $\text{VO}_2$ ) at a frequency of 1Hz. The armband sensor and metabolic analyzer were calibrated according to the manufacturer's guidelines.



Figure 6.2. Participant wearing forearm Myo armband sensor and metabolic analyzer.

### 6.2.3. Experiment Protocol

Before starting the experiment, participants were asked to warm up their bodies to prevent any injuries. Once the participant was ready, the armband sensor and metabolic analyzer were attached to the participant and calibrated using the manufacturer guidelines. All the participants performed the fourteen scaffold-building activities for five minutes each. The oxygen uptake, IMU, and EMG data were recorded continuously for each activity. The activities were performed in a warehouse environment at an average temperature of 72°F. The ten participants' data were further used for regression model building, training, and evaluation. Besides, all the participants performed a whole sequence of activities to build one scaffolding unit to simulate the actual work situation. While performing the whole sequence activities, the actual oxygen uptake was measured using the  $\text{VO}_2$ .

analyzer. The data (unseen dataset) collected while performing the whole sequence was used to determine the average oxygen consumption required to build one scaffold unit.

#### 6.2.4. BiLSTM-based Oxygen Uptake Prediction

##### 6.2.4.1. Overview of the Proposed Approach

The proposed framework to develop an oxygen uptake prediction model using forearm-based EMG and IMU data is shown in Figure 6.3. The raw EMG and IMU data obtained from the armband sensor were preprocessed and synchronized with the oxygen uptake recorded using a VO<sub>2</sub> analyzer. The preprocessed data were used to train the BiLSTM-based regression model and evaluated using the Leave-One-Subject-Out (LOSO) cross-validation. Further, the trained model was used to predict the oxygen uptake on unseen data to estimate the VO<sub>2</sub> required to build one scaffold unit.

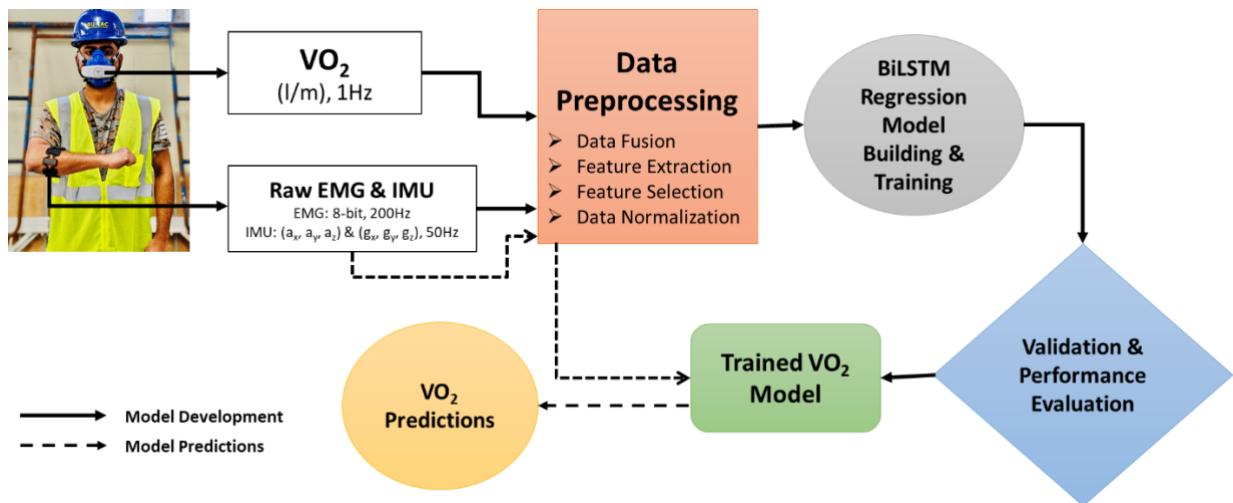


Figure 6.3. Proposed approach to develop BiLSTM-based oxygen prediction model using forearm EMG and IMU data

##### 6.2.4.2. Data Preprocessing

Since the IMU, EMG, and VO<sub>2</sub> data were at different frequencies, 17 statistical features were extracted from the raw acceleration, gyroscope, and EMG data for every one second without any overlap. In addition to these features, a lag feature by shifting the resultant acceleration (ACC),

gyroscope (GYRO), and EMGsum variable and rolling mean for window size 3 were also extracted. Using the lag feature helps the recurrent neural network models to see the sufficient past values relevant for future prediction to make predictions and improves model performance (Bouktif et al. 2018). Later, the extracted features were synchronized with VO<sub>2</sub> data at 1Hz frequency. In total, 289 statistical features and six lag features were extracted from 17 raw data features, as shown in Table 2. However, all the extracted features may not add value to the model performance. Additionally, they might cause overfitting problems. Therefore, feature selection techniques such as Pearson's correlation and mutual information were applied to the 295 features. Pearson's correlation measures the linear correlation between two features. Simultaneously, mutual information measures the amount of information obtained from one feature given another (Bagherzadeh et al. 2021; Sun et al. 2020). This study selected the features if Pearson's correlation and mutual information are more significant than 0.1. Therefore, 69 out of 295 features were selected to build the proposed model. The final step of data preprocessing involves data scaling since all the features are in varied scales. For regression and numeric input variables, the normalization technique is used to scale the data between 0 and 1 (Ding et al. 2020) using Equation 6.1. After the feature extraction, the ten participants dataset consists of 48,515 samples with 69 features.

$$xi_{norm} = \frac{xi - min}{max - min} \quad (6.1)$$

Table 6.2. Feature extraction and selection from raw EMG and IMU data

Features Extracted from Raw Data		
Dataset	Raw Features	Statistical Features
Acceleration	ax, ay, az, ACC	
Gyroscope	gx, gy, gz, GYRO	
EMG	EMG1, EMG2, EMG3, EMG4, EMG4, EMG5, EMG6, EMG7, EMG8, EMGsum	sum, avg, min, max, median, stdev, cv, var, percentiles (5, 10, 25, 75, 90, 95), skew, kurtosis
ACC, GYRO, EMGsum	Lag Feature, Rolling Mean	Shift (1), rolling (window=3)
Features Selected for Proposed Model		
Acc_mean, EMGsum_lag1, Gyro_mean, az_per50, az_median, az_avg, az_per25, az_sum, az_per75, az_per10, az_per90, az_per95, az_per5, az_max, ax_max, gz_stdev, ax_per95, gz_min, gz_max, ax_stdev, gz_per95, ax_per90, ay_stdev, gz_per5, az_min, gz_per90, gz_var, gx_stdev, gz_per10, gx_max, ax_per75, gx_min, gy_stdev, res_gyro_max, gy_max, ay_var, res_acc_skew, res_gyro_per95, gy_min, res_gyro_avg, res_gyro_sum, ay_min, ax_sum, ax_var, gx_var, res_gyro_median, res_gyro_per50, res_gyro_stdev, res_gyro_per90, gx_per5, gy_per95, gz_per75, az_skew, EMG7_per25, ax_skew, res_acc_max, res_gyro_per75, gy_per5, gx_per95, ax_avg, res_acc_per95, ax_median, ax_per50, EMG8_per25, EMG8_per75, gy_per10, EMGsum_var, res_gyro_per25, gy_per90		

#### 6.2.4.3.Bidirectional Long-Short Term Memory (BiLSTM) Model Building and Training

Recurrent Neural Networks are (RNN) are one type of neural network suitable for sequence data. In RNNs, the recurrent layers store the information from the previous step and combine it with the future timestep input. Once all the time steps are evaluated, and output layer generates output using the activation function. Output error generated is backpropagated to the network for updating the weights during training and continues till the error is minimized (Medsker and Jain 2001). The most commonly used RNN models for time-series problems are long short-term memory (LSTM), bidirectional long short-term memory (BiLSTM), and gated recurrent unit (GRU) (Choi et al. 2017). An LSTM cell consists of three gates: input, forget, and output gate, as shown in Figure 4. The input from the previous hidden state and the current input will be sent to

these three gates, and the outputs from these are passed to the cell state, which carries the required information. The gates are the neural networks responsible for retaining or removing the information during the training (Kim and Cho 2020). A BiLSTM based recurrent neural network is a variation of a long-short term memory (LSTM) model that consists of a backward and forward LSTM layer to learn information from the past layer (Schuster and Paliwal 1997). It is duplicating the first LSTM layer so that the two layers are trained side-by-side on all the available input data in the past and future timesteps. The combination of backward and forward LSTM layers helps understand the long-term dependencies between the time steps of the sequence data to improve the model performance (Porta et al. 2021).

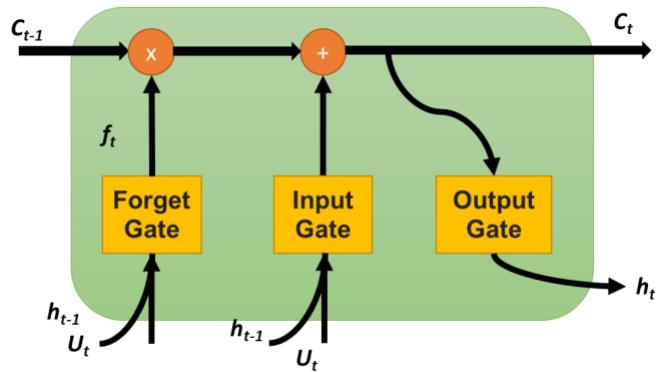


Figure 6.4. The structure of LSTM cell

In this study, the proposed BiLSTM models were implemented in Keras (Chollet 2015), a high-level neural networks API written in Python and capable of running on top of TensorFlow. The proposed BiLSTM model's architecture consists of two-stacked BiLSTM layers, dropout layers, dense layers, and an output layer, as shown in Figure 6.5. The data obtained from preprocessing step was used as input data for the model. The input data consists of 48,515 samples with 69 features for 10 participants was reshaped to three-dimensional input for LSTM models, which is in the format of [samples, timesteps, features]. The timesteps used in this study were 100. The reshaped input sequence data is fed into two-stacked BiLSTM layers with 1024 neurons. The

BiLSTM layers' output was passed through a Dropout (0.3) layer to randomly drop 30% units from the network to prevent overfitting of the model. Later the data was passed through a series of fully connected dense layers (1024) and a dropout layer (0.3) before reaching the output layer. Since the regression model's output is a numerical value, no activation function is applied in the final output layer. The hyperparameters such as the number of layers, neurons, optimizers, batch size, and the number of epochs were chosen using the random search hyperparameter optimization technique (Greff et al. 2016). An Adam optimizer with a learning rate of 0.001 and mean square error (MSE) loss function was used to compile the model. An early stopping method with 100 epochs was used to fit the model. The early stopping method helps prevent the model overfit by stopping the training process if there is no improvement in the metrics.

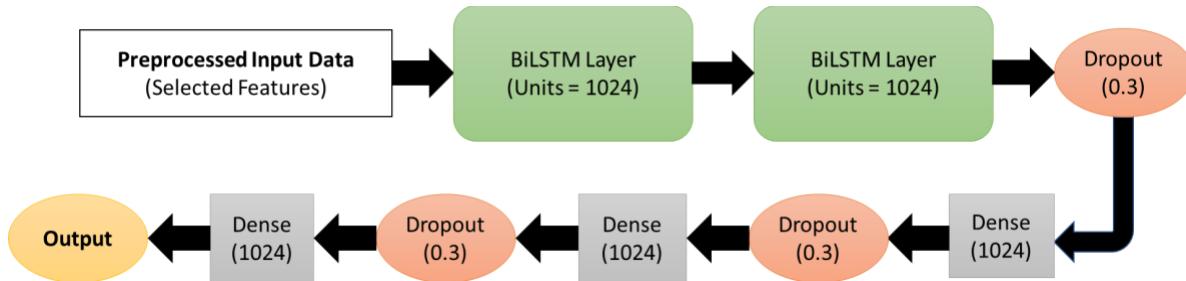


Figure 6.5. Architecture of proposed BiLSTM-based oxygen prediction using EMG and IMU

#### 6.2.4.4.Cross-Validation and Model Evaluation

To evaluate the proposed model's performance, the leave-one-subject-out cross-validation technique was implemented to prevent the overlap between the training and testing dataset, affecting prediction accuracy (Koskimaki and Siirtola 2016). Moreover, LOSO helps in getting realistic and generalized model performance on the unseen dataset. In the LOSO method, the dataset of  $N$  subjects is divided into  $N$  folds or iterations. For each fold, data from  $(N-1)$  subjects are used for model training, and the left-out subject data is used for testing. This is repeated with all the  $N$  subjects' data, and the model performance is the average of the results for all subjects.

Various metrics such as mean absolute deviation (MAE), mean square error (MSE), root mean square error (RMSE), coefficient of correlation (R), coefficient of determination ( $R^2$ ), and mean absolute percentage error (MAPE) was used to assess the performance of the regression models. MSE is calculated as the squared difference between actual ( $y$ ) and predicted ( $\hat{y}$ ) output as shown in Equation 6.2. MAE measures the magnitude of residuals, which is the absolute difference between actual and predicted outputs, as shown in Equation 6.3. Whereas MSE is the most prominent error. Unlike MAE, MSE is highly sensitive to outliers. The metrics RMSE is the standard deviation of the residuals, which helps understand the spread of predicted outputs around expected outputs. RMSE is the square root of MSE, which is calculated using Equation 6.4. The MSE, MAE, and RMSE are expressed in  $\text{ml. kg}^{-1} \cdot \text{min}^{-1}$ . This study also reports MAPE, r, and  $R^2$ , calculated as shown in Equations 6.5, 6.6, and 6.7, respectively. In addition to evaluating the test dataset, the trained model was implemented on the unseen dataset to determine the error between measured and predicted oxygen uptake.

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2 \quad (6.2)$$

$$MAE = \frac{1}{n} \sum |y - \hat{y}| \quad (6.3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (\hat{y} - y)^2} \quad (6.4)$$

$$MAPE = \frac{100\%}{n} \sum \frac{|y - \hat{y}|}{y} \quad (6.5)$$

$$r = \frac{\sum_{i=1}^n (\hat{y} - y)^2}{\sum_{i=1}^n (y - \bar{y})^2} \quad (6.6)$$

$$R^2 = \left( \frac{\sum_{i=1}^n (\hat{y} - y)^2}{\sum_{i=1}^n (y - \bar{y})^2} \right)^2 \quad (6.7)$$

## 6.3. Results

### 6.3.1. Proposed Model Performance

This section presents the LOSO cross-validation results of the proposed BiLSTM based regression model using EMG and IMU features. The proposed model was built using 10 participants' data and 69 selected EMG and IMU features. The average metrics of 10-fold LOSO cross-validation metrics shows are shown in Table 6.3. The average  $R^2$ , RMSE, and MAE values of the proposed model for all the participants are 0.80,  $1.257 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ , and  $1.581 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ , respectively. Moreover, Table 6.3 presents the LOSO cross-validation results on test Subject#1, where  $R^2$ , RMSE, and MAE are 0.75,  $1.685 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ , and  $1.264 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ . The measured and predicted oxygen consumption for each second on Subject#1 test data is shown in Figure 6.6. From Figure 6.6, it can be observed that the model has predicted the pattern perfectly, but some of the extreme values were missed. Even though the desired results are achieved, it is essential to assess the training performance using learning curves to understand if the model is suffering from variance or bias (Lipton et al. 2015). Figure 6.7 shows that the proposed model loss function (i.e., MSE) decreased with the number of epochs. Moreover, the training and validation loss curves are close to each other, showing that the model has a good fit with low bias and variance.

Table 6.3. LOSO cross-validation metrics of proposed BiLSTM model on test and unseen data

Model	R	$R^2$	MAE	MSE	RMSE	MAPE
All Participants Data	0.895	0.800	0.757	1.581	1.257	10%
Subject#1 as Test Data	0.866	0.750	1.264	2.838	1.685	16%
Unseen Data of All Participants	0.833	0.695	1.236	2.525	1.589	13%

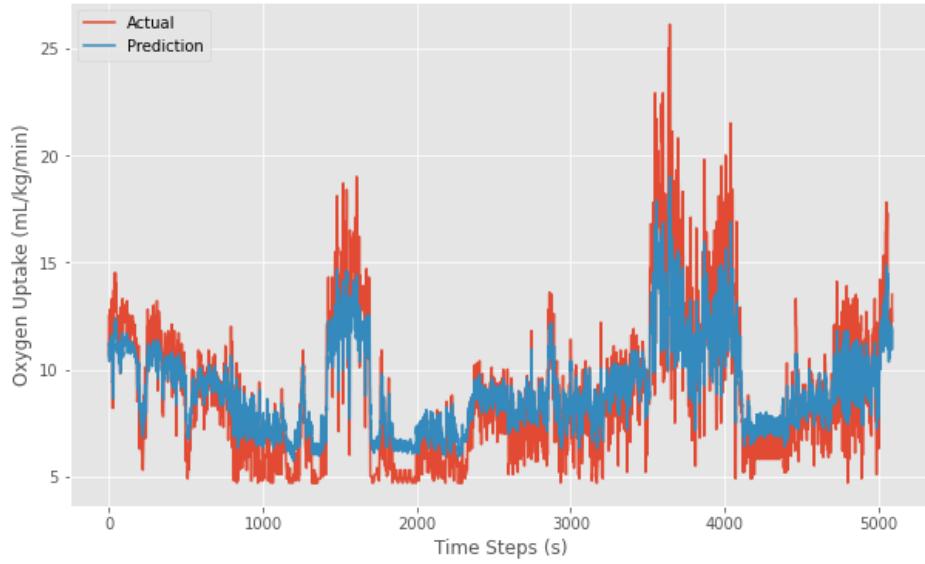


Figure 6.6. Oxygen uptake prediction on test Subject#1 using proposed BiLSTM model

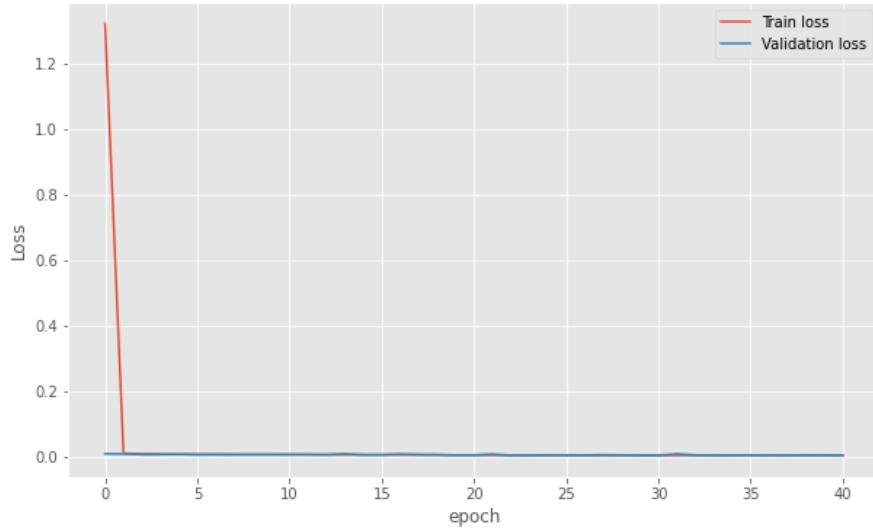


Figure 6.7. Learning curve of the proposed BiLSTM model using EMG and IMU data

A further trained model was used to predict the oxygen consumption on the unseen data (i.e., participants performed an entire sequence of activities to build one unit scaffold frame). The unseen dataset consists of 4,116 number samples from all the participants. The average  $R^2$ , RMSE, and MAE on unseen data of all participants are 0.75,  $1.685 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ , and  $1.264 \text{ ml. kg}^{-1} \cdot \text{min}^{-1}$ .

<sup>1</sup>. Figure 6.8 shows the linear correlation analysis for measured and predicted  $\text{VO}_2$ . It should be noted that these results are from second-by-second oxygen uptake predictions.

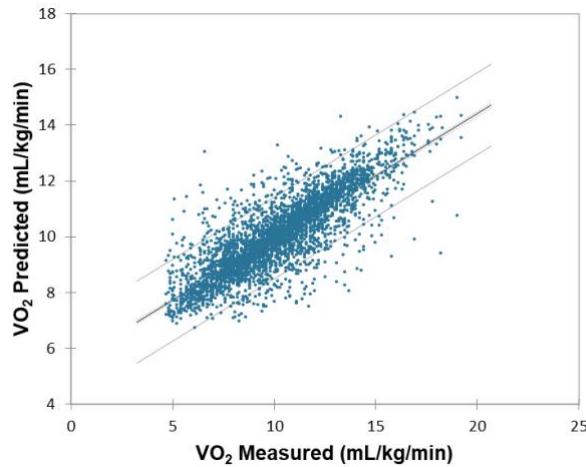


Figure 6.8. Correlation analysis for measured and predicted  $\text{VO}_2$  on the unseen dataset

### 6.3.2. Average Oxygen Consumption to Build One Scaffolding Unit

The prediction of the proposed BiLSTM model on unseen data of the participants was used to determine the average oxygen consumption required to build a scaffold unit involving fourteen activities shown in Table 6.1. The average time taken for the participants to build one unit of scaffold is 6.67 minutes (i.e., approximately 400 samples in the unseen dataset for each subject). The average measured and estimated oxygen uptake to build a scaffold unit is  $9.18 \pm 1.97$  and  $9.22 \pm 1.30$ , as shown in Table 6.4. ANOVA analysis between measured and estimated oxygen uptake shows no significant difference ( $p = 0.9641$ ). It is observed that the error is minimized for the average oxygen consumption over the duration of the build compared to second-by-second predictions. The oxygen consumption estimated using the proposed model can be used to determine the activity's physical workload. However, it should be noted that the physical workload or oxygen consumption increases with an increase in the work duration and the number of units.

Table 6.4. Measured and estimated oxygen uptake to build one scaffold unit using the proposed model

Participant		Measured VO <sub>2</sub>		Estimated VO <sub>2</sub>		Difference
	Weight (Lbs.)	mL/kg/min	L/min	mL/kg/min	L/min	
Participant - 1	75	9.15	0.69	9.33	0.71	-0.18
Participant - 2	74.25	7.67	0.81	8.27	0.78	-0.60
Participant - 3	73	9.26	0.68	9.19	0.70	0.07
Participant - 4	85	8.24	0.96	8.63	0.91	-0.39
Participant - 5	77.70	5.95	0.88	7.09	0.83	-1.13
Participant - 6	75	9.35	0.84	9.33	0.80	0.02
Participant - 7	63	12.80	0.62	11.67	0.62	1.13
Participant - 8	93	8.53	0.82	8.70	0.86	-0.17
Participant - 9	81	8.90	0.89	9.02	0.85	-0.12
Participant - 10	70	11.98	0.67	10.94	0.68	1.04
Average	76.70	9.18	0.79	9.22	0.77	-0.03
SD	8.25	1.97	0.11	1.30	0.09	0.68

#### 6.4. Comparison with Other RNN Models and Different Sensor Combinations

Further, this study has compared other RNN models' performance and different sensor combinations using the proposed framework (i.e., feature extraction, feature selection, model training, hyperparameter tuning, and LOSO CV evaluation). Three commonly used RNN models (i.e., LSTM, BiLSTM, and GRU) were built for three different sensor feature combinations (i.e., EMG+IMU, IMU alone, and EMG alone). The LOSO cross-validation results for all the models are shown in Table 6.5. The BiLSTM model has the lowest error and highest correlation for each sensor combination than LSTM and GRU. Similarly, the EMG+IMU sensor combination has performed better compared to EMG and IMU alone. This shows that the fusion of EMG and IMU features has improved the performance of the model.

Table 6.5. LOSO CV metrics for different recurrent neural networks and sensor combinations

Sensor Combination	Model	R	R-Square	MAE	MSE	RMSE	MAPE
EMG+IMU	LSTM	0.871	0.760	1.005	1.905	1.380	11%
	BiLSTM	<b>0.895</b>	<b>0.800</b>	<b>0.757</b>	<b>1.581</b>	<b>1.257</b>	<b>10%</b>
	GRU	0.817	0.667	1.299	2.639	1.624	18%
IMU	LSTM	0.776	0.603	1.509	3.143	1.773	21%
	BiLSTM	0.887	0.787	0.936	1.687	1.299	11%
	GRU	0.469	0.220	2.001	6.173	2.485	22%
EMG	LSTM	0.797	0.636	1.252	2.870	1.694	16%
	BiLSTM	0.816	0.667	1.082	2.627	1.621	13%
	GRU	0.793	0.629	1.254	2.922	1.709	14%

#### 6.4.Discussion

In agreement with the initial hypothesis, the combination of motion (IMU) and muscle activity (EMG) have achieved the highest performance compared to EMG and IMU alone. This is because the fusion of motion and muscle intensity data provides distinctive feature patterns for the model to learn quickly. For example, carrying a guardrail (5 Lbs.) and baseboard (33 Lbs.) can have the same motion, but the muscle intensities are different, helping to detect the oxygen uptake levels. It was previously proved that handling different weights results in different EMG signal patterns (Ho et al. 2017). Similarly, the muscle activity is the same for carrying scaffold frame (38 Lbs.) and baseboard (33 Lbs.), but the motion pattern is different. Since the motion and muscle intensity pattern changes with the physical activity level, the EMG+IMU model performance is higher than others. As suggested by previous studies, sensor data fusion has improved model performance compared to individual models (Ahn et al. 2019; Bangaru et al. 2021; Koskimaki and Siirtola 2016; Lu et al. 2018).

Table 6.6 compares this study results with similar studies related to oxygen uptake predictions using wearable sensors and machine learning algorithms. For construction applications, it is essential to use a minimum number of sensors to prevent ongoing work

obstruction. In this study, low-cost wearable armband sensors were used to collect forearm EMG and IMU data that can be worn for the entire workday without any discomfort or obstruction to work (Bangaru et al. 2021; Visconti et al. 2018). Except for Shandhi et al. (2020) study, all the previous studies used multiple sensors for data acquisition. However, Shandhi et al. (2020) evaluated oxygen consumption only for treadmill walking, and the chest strap electrocardiogram sensor might require gel for signal conductivity, which is impractical on construction sites (Bao et al. 2020). Moreover, most previous studies have investigated the oxygen consumption for daily living activities, walking, and cycling, light to moderate intensity. This is the first study to investigate oxygen consumption using wearable sensors for construction activities. Due to the dynamic nature of the construction activities, different intensity activities are performed in a short interval of time (Jebelli et al. 2019). The use of forearm muscle activity and motion data in this study helped capture the complex movements performed quickly.

Table 6.6. Comparison of activities and sensor of the current and previous studies on oxygen uptake prediction

Study	Activities	Activity Type	# Sensors	Sensor Signals	Sensor Location
This Study	Scaffold Building	Light - Heavy	1	IMU EMG	Forearm
Zignoli et al. (2020)	Cycling Exercise	Light - Moderate	2	Heart Rate, Garmin Vector Power Meter	Foot
Shandhi et al. (2020)	Treadmill Walking	Light - Moderate	1	Seismocardiogram Electrocardiogram Atmospheric Pressure	Mid-Sternum
Borror et al. (2019)	Treadmill	Light - Moderate	2	Heart Rate Garmin Vector Power Meter	Chest Foot
Lu et al. (2018)	Office Painting Postal Delivery Meat Cutting Lifting Tasks	Light - Heavy	8	Electrocardiogram Accelerometer	Chest Wrist Thigh
Beltrame et al. (2017)	Daily Living Activities Controlled Walking	Light - Moderate	3	Electrocardiogram Accelerometer Respiratory Bands	Chest Hip
Altini et al. (2015)	Daily Living Activities	Light - Moderate	4	Electrocardiogram, Accelerometer	Chest

Table 6.7 compares the current and previous studies' machine learning model performance related to oxygen consumption using wearable sensors. The application of BiLSTM for oxygen uptake prediction is a novel aspect of this study. The BiLSTM cell's ability to preserve the past and future timesteps information helped model complex data (Meyer et al. 2020). Moreover, the use of BiLSTM and MSE loss function can handle extreme values compared to LSTM and GRU. The coefficient of determination of the proposed model ( $R^2 = 0.80$ ) is lower than Zignoli et al. (2020) ( $R^2 = 0.89$ ) and Borror et al. (2019) ( $R^2 = 0.91$ ). It should be noted that those studies have only evaluated cycling and treadmill activities which are less complex. Lu et al. (2018) model had

achieved RMSE of 1.69, 2.36, 1.62, and 3.88 mL/kg/min for the painting, postal delivery, meat cutting, and lifting activities which are much higher compared to RMSE of the proposed model (i.e., 1.26 mL/kg/min). From Table 6.7, it can be observed that as the complexity of the activities increases, the performance of the previous models goes down. Except for Shandhi et al. (2020), previous studies used raw sensor signals as model input data. The feature engineering process proposed in this study helped to achieve improved model performance on the unseen data.

Table 6.7. Comparison of model performance of the current and previous studies on oxygen uptake prediction

Study	Activities	Model	R <sup>2</sup>	RMSE
This Study	Scaffold Building	BiLSTM	0.80	1.26
Zignoli et al. (2020)	Cycling Exercise	LSTM	0.89	N/A
Shandhi et al. (2020)	Treadmill Walking	Xgboost	Treadmill - 0.77 Outdoor Walk - 0.64	Treadmill (3.68 ± 0.98) Outdoor Walk (4.30 ± 1.47)
Borror et al. (2019)	Treadmill	ANN	0.91	N/A
Lu et al. (2018)	Office Painting Postal Delivery Meat Cutting Lifting Tasks	MLP	N/A	Office - 0.86 Painting - 1.69 Postal Delivery - 2.36 Meat Cutting - 1.62 Lifting - 3.88
Beltrame et al. (2017)	Daily Living Activities Controlled Walking	Random Forest	Daily Living Activities – 0.75 Random Walking - 0.48	N/A
Altini et al. (2015)	Daily Living Activities	Linear, Exponential Logistic	N/A	4.38 ± 0.80

Considering the proposed study's experimental conditions, the average measured and estimated oxygen consumption for building one scaffold unit is 0.77 and 0.79 L/min, which can

be classified as moderate work based on the published work severity guidelines (Astrand et al. 1986). However, continuous monitoring of physical activity is required to evaluate the activity's workload (Abdelhamid and Everett 2002). The statistical analysis has shown that the measured and estimated average oxygen consumption for building one scaffold unit is not significantly different, providing an opportunity to use the forearm-based EMG and IMU sensor instead of an expensive metabolic analyzer. The accurate oxygen consumption prediction using a low-cost wearable sensor helps evaluate the activities' physical workload and quantify the direct impacts of physical workload on construction safety and productivity. Additionally, determining physiological demands by measuring oxygen uptake helps design the construction activities with ergonomic interventions to prevent musculoskeletal disorders.

Even though the proposed model has achieved acceptable performance on unseen datasets, some of the limitations of this study include the oxygen uptake was monitored only for a short time (i.e., for building one scaffold unit), only right-handed participants, and an indoor warehouse environment. The future work includes implementation of the proposed model on the real construction site for the entire workday to monitor oxygen uptake, to understand the influence of sensor position on the model performance, investigate the frequency and time domain features for model building and training, extend the proposed framework for other construction trade activities, and implement time-series data augmentation techniques to improve model performance.

## **6.5.Conclusion**

The study concludes that the proposed BiLSTM RNN model can predict second-by-second oxygen uptake using forearm EMG and IMU features. The model has achieved a coefficient of correlation and RMSE of 0.90 and 1.257 mL/kg/min using the ten participants' data of scaffold building activities. The results show that the data fusion of EMG+IMU ( $R = 0.90$ ) has yielded the

highest performance compared to IMU alone ( $R = 0.88$ ) and EMG alone (0.81). Moreover, the average oxygen consumption for building one scaffold unit is estimated to be 0.77 L/min, whereas the measured oxygen consumption is 0.79 L/min which has no significant difference. The main advantages of the proposed system over previous studies are the use of low-cost sensor, complex construction activities with varying intensities in a short interval of time, use of fully automated framework, feature engineering process to improve performance on the unseen dataset, and use of BiLSTM and appropriate hyperparameters to handle complex time-series data. The proposed model can be embedded into the wearable sensor to provide real-time workload assessment and the worker's physiological status. The continuous monitoring of oxygen uptake helps evaluate the physical workload of various construction activities, physiological status of the worker, and early detection of the potential hazards on the construction site.

## **CHAPTER 7. FATIGUE MONITORING SYSTEM FEASIBILITY VALIDATION AND PERFORMANCE EVALUATION**

### **7.1.Introduction**

The proposed fatigue monitoring framework involves six steps: data acquisition using forearm-based EMG and IMU armband sensor, workers' activity recognition, maximum aerobic capacity (MAC) estimation, oxygen uptake prediction, and aerobic fatigue threshold (AFT) monitoring, and fatigue level assessment, as shown in Figure 3.1. The first four steps of the proposed framework were tested for feasibility and evaluated for the performance in the previous chapters. This chapter focuses on feasibility validation and performance evaluation using AFT for fatigue monitoring and fatigue assessment using the proposed system.

### **7.2.Aerobic Fatigue Threshold – Feasibility Validation and Performance Evaluation**

The proposed fatigue monitoring system uses continuous monitoring of aerobic fatigue threshold to assess the construction workers' fatigue level. The aerobic fatigue threshold or exercise intensity is defined as the ratio of average oxygen consumption to the activity-specific maximum aerobic capacity, as shown in Equation 7.1 (Sharp et al. 1988). The MAC values used for AFT calculation obtained using submaximal experiments are shown in Table 7.1. For simplicity, the fourteen scaffold building activities are assumed to fall into one of these categories. For example, the carrying scaffold, crossbars, guardrail, and baseboard use MAC value of carrying activity.

$$\text{Aerobic Fatigue Threshold} = \frac{\text{Average VO}_2}{\text{Activity Specific MAC}} \quad (7.1)$$

Table 7.1. MAC values for different construction activities.

Sub-Maximal Activities	MAC (L/min)	Scaffold Activities
Walking	2.946	WALK
Carrying	2.552	CPSF, CLJ, CC, CDB, CG
Lifting	2.816	IBDL
Combined	2.304	IALJ, IC, HAM, WRE, DG, IG, GUDVL

The National Institute for Occupational Safety and Health (NIOSH) guidelines recommend that an average aerobic fatigue threshold or exercise intensity should not exceed 33% for an 8h workday (Sharp et al. 1988). Previous studies used aerobic fatigue threshold or exercise intensity to assess the worker capabilities or task workload evaluation (Abut and Akay 2015; Fredericks et al. 2005; Iridiastadi and Aghazadeh 2005). Therefore, it is essential to evaluate the use of AFT for workers' fatigue monitoring. Figure 7.2 presents the data analysis protocol to evaluate the feasibility and performance of AFT for fatigue monitoring. The best classifier was developed to predict the fatigue level on the unseen dataset using suitable features.

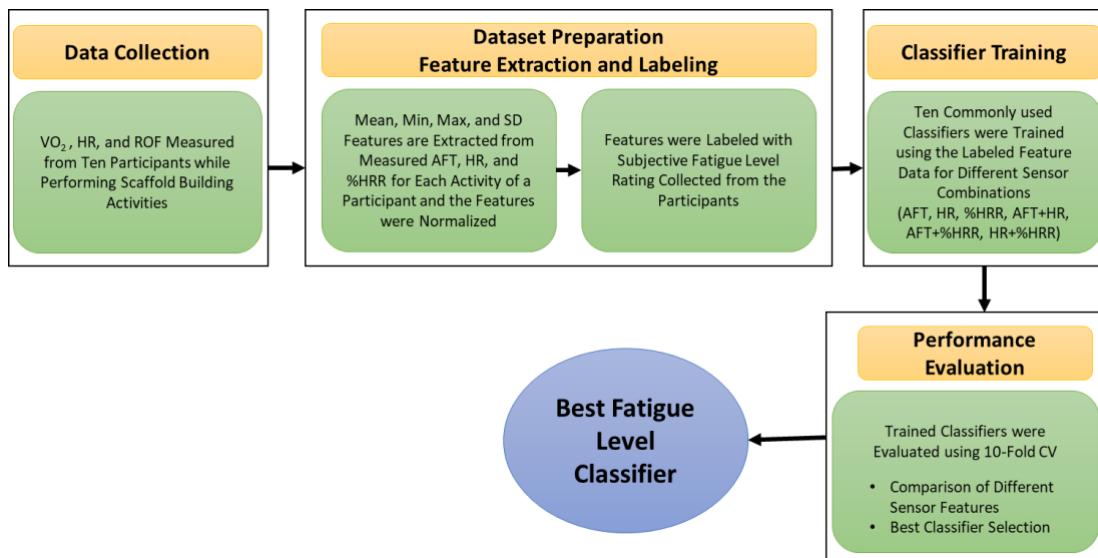


Figure 7.1.Data analysis protocol to evaluate the feasibility and performance of using aerobic fatigue threshold for fatigue monitoring

### 7.1.1. Data Analysis Protocol

The oxygen uptake ( $\text{VO}_2$ ) and heart rate (HR) data collected using  $\text{VO}_2$  metabolic analyzer and Tickr chest strap HR monitor from ten participants (age:  $27 \pm 1.70$  years, weight:  $76.70 \pm 8.25$  kg, and height:  $171.7 \pm 4.13$  cm) while performing scaffold building activities was used for this analysis. The oxygen uptake and heart rate data were recorded at 1Hz frequency. Each participant performed fourteen scaffold building activities listed in Table 7.2 for a maximum of five minutes or until they were exhausted. All the activities were randomized for each participant. In addition to  $\text{VO}_2$  and HR data, the participants' rating of fatigue (ROF) was collected before and after each activity. The ROF was reported using the rating of fatigue scale (0-10) and verbal anchors shown in Table 7.3 (Aryal et al. 2017; Kim et al. 2010). Using the rating of fatigue, the level of fatigue was assigned to one of the labels, i.e., none, low, moderate, high, and very high. In order to compare the performance of AFT to other fatigue assessment metrics such as heart rate (HR) (Aryal et al. 2017) and percentage of HR reserve (%HRR) (Hwang and Lee 2017), the %HRR was calculated using the Equation 7.2 for every one second.

$$\text{Percentage of HR Reserve } (\%) = \frac{\text{HR}_{\text{Working}} - \text{HR}_{\text{Resting}}}{\text{HR}_{\text{Maximum}} - \text{HR}_{\text{Resting}}} \quad (7.2.)$$

Where  $\text{HR}_{\text{Working}}$  = average working heart rate [bpm];  $\text{HR}_{\text{Resting}}$  = resting heart rate [bpm]; and  $\text{HR}_{\text{Maximum}}$  = maximum heart rate is estimated using 220-age [bpm] (Robergs and Landwehr 2002; Verschuren et al. 2011).

Table 7.2. Scaffold building activities description and activity ID.

SL. No.	Activities	Activity ID
1	Walking	WALK
2	Carrying / Positioning Scaffold Frame	CPSF
3	Carrying Leveling Jacks	CLJ
4	Inserting / Adjusting Leveling Jacks	IALJ
5	Carrying Crossbars	CC
6	Installing Crossbars	IC
7	Hammering	HAM
8	Wrenching	WRE
9	Carrying / Dragging Baseboard	CDB
10	Installing Baseboard on Different Level	IBDL
11	Carrying Guardrail	CG
12	Dragging Guardrail	DG
13	Installing Guardrail	IG
14	Going Up / Down Vertical Ladder	GUDVL

Table 7.3. Fatigue rating scale along with verbal anchors and level of fatigue.

Fatigue Rating	Verbal Anchors	Fatigue Level & Labels
0	I am not tired; this is similar to resting	0 - None
1		
2	I am not tired; this is similar to walking	1 - Low
3		
4	I feel fine to continue	2 - Moderate
5		
6		
7	I am getting tired, but I can continue	3 - High
8		
9	I am exhausted; I have to push myself to continue	4 - Very High
10		

The AFT, HR, and %HRR data obtained from ten participants for every one second were used to extract statistical features such as mean, minimum, maximum, and standard deviation for the duration of activity. In total, there twelve features and 140 samples (10 participants \* 14 activities). Since all the features are in different units, the features were normalized to scale all the

features between 0 and 1. Later, the feature data was labeled with the level of fatigue (i.e., none, low, moderate, high, and very high) for each activity performed by the participant. There were no data samples with the "none" label.

The labeled feature data was further used to train ten commonly used machine learning-based classifiers such as Random Forest (RF), Decision Trees (DT), Naïve Bayes (NB), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Support Vector Machine (SVM), Ada Booster (ADA), Logistic Regression (LR), K Nearest Neighbors (KNN), and Multilayer Perceptron (MLP). The classification analysis was performed using PyCaret – an open-source, low-code machine learning library in python (Ali 2020). The models were evaluated using a 10-fold cross-validation technique, and the performance of the models was assessed using accuracy, precision, recall, and F1 Score. To evaluate the feasibility and performance of using AFT for fatigue level assessment, the classifier's performance was compared for different feature combinations such as AFT, HR, %HRR, AFT+HR, AFT+%HRR, HR+%HRR, and AFT+HR+%HRR. Besides, the best fatigue level classifier was selected and used for further analysis on the unseen dataset.

### **7.1.2. Analysis and Results**

First, the average AFT value for each fatigue level was estimated using the ten participants' data, as shown in Table 4. The average AFT is above 33% for the high and very high fatigue levels where an individual is getting tired or very tired, which aligns with the NIOSH recommendation that an individual cannot sustain if AFT exceeds 33%. Moreover, Figure 3 shows the average AFT value for each activity where activities CPSF, CDB, and GUDVL are above the 33% threshold and align with the subjective fatigue rating rated as high or very high fatigue level activities. This shows that the subjective fatigue ratings are reliable for further analysis.

Table 7.4. Average AFT for each fatigue level.

Fatigue Level	Average AFT
Low	22.03%
Moderate	28.91%
High	36.00%
Very High	43.63%

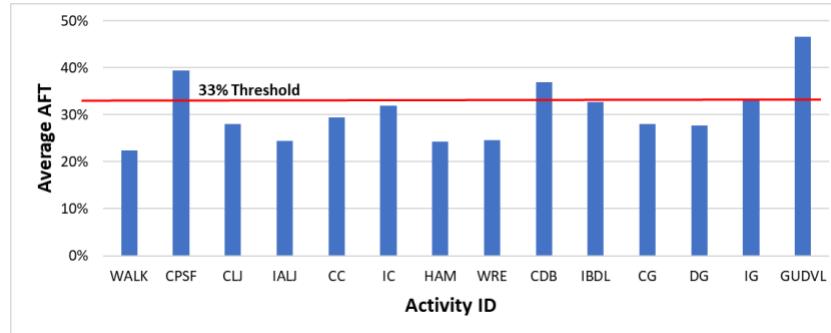


Figure 7.2. Average AFT for each activity.

Further, the classification accuracy for the tested machine learning algorithms for different feature combinations is shown in Table 7.5. The results show that the highest classification accuracy was obtained for AFT features using the decision tree classifier algorithm. Also, the classification accuracy is highest for the features in combination with AFT (i.e., AFT+%HRR = 91.45%, AFT+HR+%HRR = 90.60, and AFT+HR = 90.60%). The highest classification accuracy for different feature combinations is highlighted in Table 7.5. The low classification accuracies were observed for HR (51.28%) and %HRR (50.43%) features. Similarly, the F1 Scores for the AFT (92.40%) are highest compared to other features shown in Table 7.6. Moreover, Figure 7.5 shows the confusion matrix for the decision tree classifier using AFT features. It is observed that the model is classifying all the four levels with 90% accuracy with a high misclassification rate of 11% between high and very high levels. From accuracy, F1 Scores, and confusion matrix, the performance of the classifiers using AFT features is highest compared to HR and %HRR features. See Appendix C for recall and precision results of the classifiers for different feature combinations.

This analysis concludes that AFT features are highly suitable for assessing all four fatigue levels compared to %HRR and HR. The decision tree classifier built using the AFT features was further used for assessing fatigue levels of unseen data.

Table 7.5. Classification accuracies for different feature combinations.

Model	AFT	%HRR	HR	AFT+ %HRR	AFT+ HR	HR+ %HRR	AFT+HR+ %HRR
RF	90.60%	44.44%	41.88%	89.74%	90.60%	44.44%	90.60%
DT	92.31%	41.03%	44.44%	91.45%	88.03%	34.19%	90.60%
NB	82.91%	46.15%	47.01%	79.49%	78.63%	42.74%	74.36%
LDA	87.18%	46.15%	51.28%	88.89%	83.76%	42.74%	82.91%
QDA	85.47%	38.46%	36.75%	79.49%	77.78%	37.61%	67.52%
SVM	72.65%	46.15%	44.44%	72.65%	66.67%	50.43%	68.38%
ADA	52.14%	46.15%	47.86%	52.14%	52.14%	46.15%	52.14%
LR	86.32%	50.43%	46.15%	79.49%	82.05%	45.30%	79.49%
KNN	81.20%	36.75%	38.46%	72.65%	72.65%	35.04%	64.96%
MLP	91.45%	46.15%	45.30%	88.89%	87.18%	46.15%	88.03%

Table 7.6. Classification F1 Score for different feature combinations.

Model	AFT	%HRR	HR	AFT+ %HRR	AFT+ HR	HR+ %HRR	AFT+HR+ %HRR
RF	90.60%	44.60%	42.10%	89.80%	90.60%	44.50%	90.70%
DT	92.40%	40.50%	43.20%	91.50%	88.10%	33.90%	90.60%
NB	82.80%	43.90%	42.20%	79.30%	78.20%	38.20%	73.90%
LDA	87.30%	44.80%	47.40%	88.90%	83.80%	41.60%	82.80%
QDA	85.30%	38.40%	36.30%	79.20%	77.30%	37.60%	65.50%
SVM	68.10%	46.20%	44.44%	70.70%	62.40%	50.40%	66.20%
ADA	52.10%	46.20%	47.90%	52.10%	52.10%	46.20%	52.10%
LR	86.30%	49.00%	42.60%	79.50%	82.10%	44.44%	79.60%
KNN	81.30%	36.20%	38.40%	72.90%	73.10%	35.30%	65.20%
MLP	91.50%	44.10%	43.50%	88.90%	87.20%	45.20%	88.20%

Predicted Level	Low	94%	6%	0%	0%	94.12%	2.40%
	Moderate	4%	91%	4%	0%	91.11%	4.20%
	High	0%	5%	95%	0%	94.74%	4.10%
	Very High	0%	0%	11%	89%	89.47%	0.00%
		Low	Moderate	High	Very High	<b>TP Rate</b>	<b>FP Rate</b>

Figure 7.3. Confusion matrix for decision tree classifier using AFT features.

### 7.3.Fatigue Monitoring System – Feasibility Validation and Performance Evaluation

The proposed fatigue monitoring system uses the forearm EMG and IMU data to assess fatigue levels. The individual components of the proposed framework have been evaluated for feasibility and performance in the previous chapters. It is essential to evaluate the feasibility and performance evaluation of the system to assess the workers' fatigue level. Figure 7.5 presents the data analysis protocol to evaluate the feasibility and performance of the proposed fatigue monitoring framework. We will use the trained BiLSTM activity recognition model, MAC values for construction-specific activities, trained BiLSTM oxygen prediction model, and fatigue level classifier obtained in the previous chapters and sections for the evaluation.

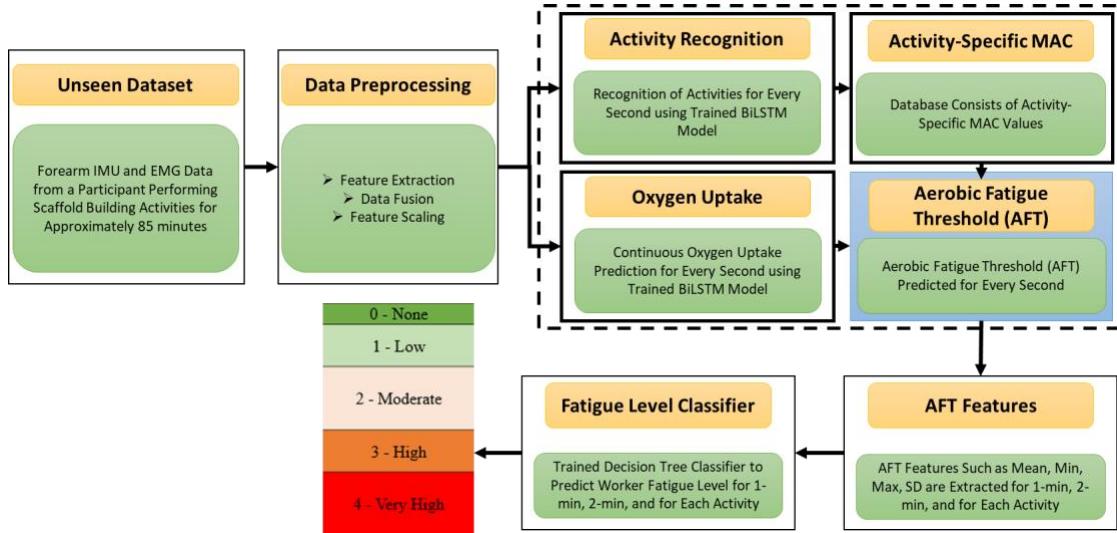


Figure 7.4. Data analysis protocol to evaluate the feasibility and performance of the proposed fatigue monitoring system

### 7.1.3. Data Analysis Protocol

For the proposed system feasibility and performance evaluation, EMG and IMU data were collected from a participant (age = 29 years, height = 168 cm, weight = 75 kg, and Resting HR = 96) performing all fourteen scaffold building activities for approximately 85 minutes (5088 samples). The sequence of the activities and duration of the activities are shown in Table 7.7. Some of the IALJ, CDB, and IG activities are performed for a longer duration because they involve multiple tasks. For example, IALJ involves two tasks installing and adjusting leveling jacks performed for five minutes each continuously. In contrast, the activities such as GUDVL were performed for a short duration because the participant was completely exhausted after 1.40 minutes. In addition to armband data, VO<sub>2</sub> and HR data were continuously recorded for the entire session. Moreover, a rating of fatigue level was collected for every 1-minute.

First, the EMG and IMU data were preprocessed for activity recognition and oxygen uptake prediction using the framework proposed in Chapter IV and V, respectively. The oxygen uptake and heart rate data recorded for every second were synchronized with the preprocessed EMG and

IMU features. Also, the features dataset was labeled with actual activity ID for ground truth. Once the dataset was prepared, the trained activity recognition and oxygen uptake models developed in the previous chapters were implemented on unseen datasets to recognize activities and oxygen uptake for every one second. Using the model predictions and MAC values from Chapter V, we determined AFT for every one second on the unseen dataset. The actual and predicted AFT values were analyzed using linear regression analysis to see how well the proposed system monitored AFT for one second, 1-min, 2-min, and over the entire activity duration.

For the fatigue level assessment, the predicted AFT values for every second were used to extract statistical features (i.e., mean, minimum, maximum, and standard deviation) for every 1-min, 2-min, and activity. The extracted AFT features were labeled with the subjective rating of fatigue level for ground truth. The best fatigue level classifier obtained in the previous section was used to predict the fatigue level of the unseen dataset. The actual and predicted fatigue levels were analyzed to see how well the predicted AFT can recognize workers' fatigue levels compared to HR and %HRR.

Table 7.7. Activity sequence and duration of unseen dataset

Sequence	Activity ID	Duration (Min)
1	WALK	3.27
2	CG	4.98
3	CLJ	4.95
4	IALJ	10.03
5	CPSF	5.00
6	HAM	5.02
7	WRE	5.02
8	CC	5.00
9	CPSF	4.97
10	IC	5.02
11	CDB	10.10
12	IBDL	5.02
13	DG	5.02
14	IG	10.02
15	GUDVL	1.40

#### 7.1.4. Analysis and Results

The actual and predicted AFT on the unseen dataset for 1s, 1-min, 2-min, and each activity over the entire experiment duration are shown in Figures 7.6, 7.7, and 7.8, respectively. The graphs show a peak in the trend for the high-intensity activities for all window sizes. The value of AFT varied based on the intensity and complexity of the activity. Figure 7.9 shows the average actual and predicted AFT for each activity where the highest value was observed for GUDVL and lowest for HAM activities. These results match with the participants' subjective fatigue rating. A linear correlation analysis was performed between actual and predicted AFT for 1s, 1-min, 2-min, and each activity. The correlation results show that the highest coefficient of determination ( $R^2$ ) and root mean square error (RMSE) of 0.85 and 0.027 were observed for both 1-min and 2-min AFT. The lowest correlation was observed for 1s i.e.,  $R^2 = 0.71$  and RMSE = 0.040. The goodness of fit curves for 1s and 5min AFT is shown in Figure 7.10. The correlation analysis shows that the predicted AFT values have achieved a good fit and the variation in the AFT value for different activities shows the feasibility of using predicted AFT values for fatigue assessment. Also, the use of 1-min or 2-min average AFT helps minimize the error and improves the fatigue prediction accuracy.

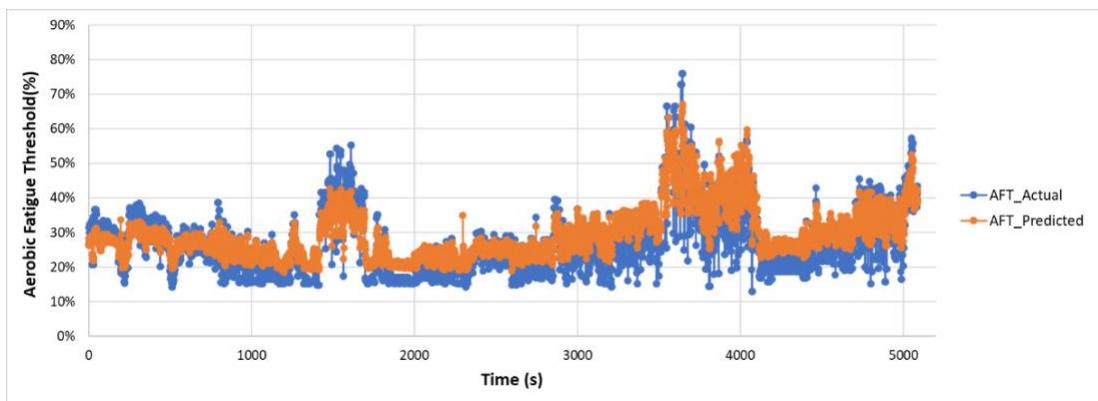


Figure 7.5. Actual and predicted AFT for every 1-second on the unseen dataset

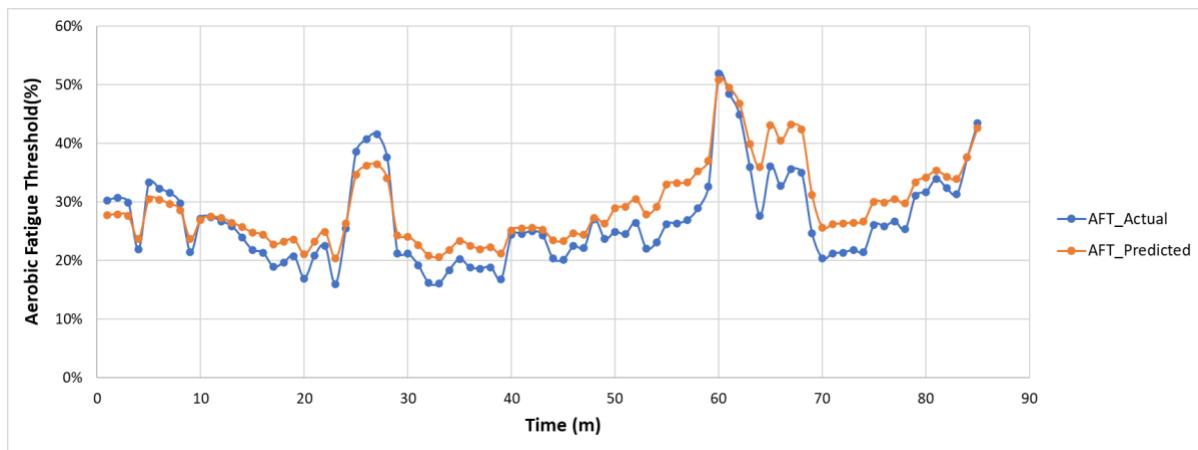


Figure 7.6. Actual and predicted AFT for every 1-minute on the unseen dataset

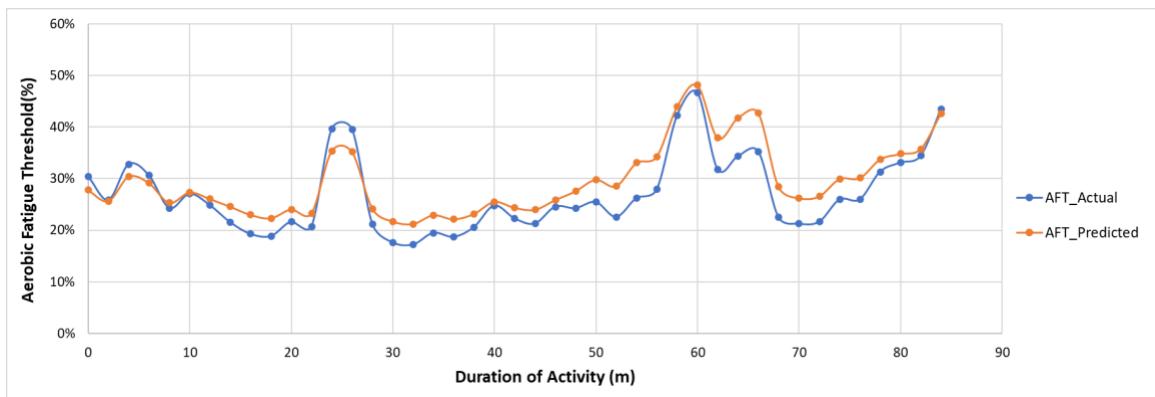


Figure 7.7. Actual and predicted AFT for every 2-minute on the unseen dataset

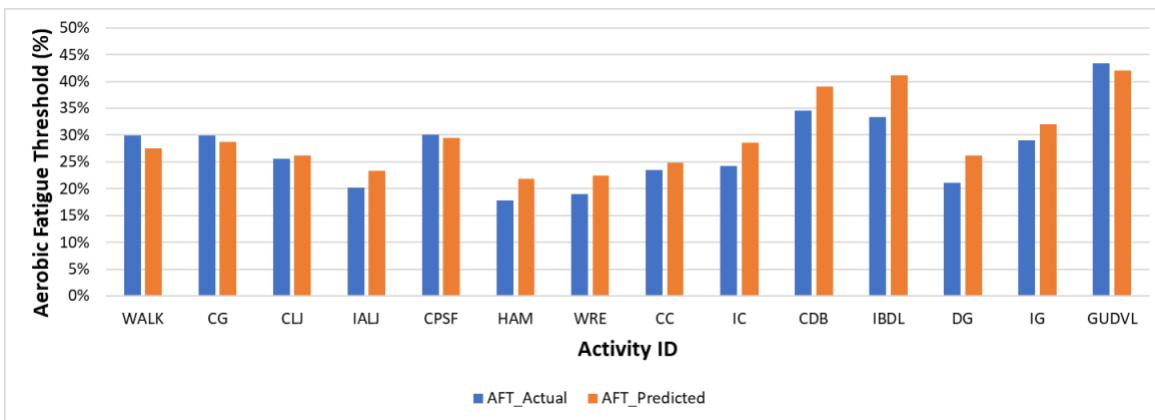


Figure 7.8. Actual and predicted AFT for each activity on the unseen dataset

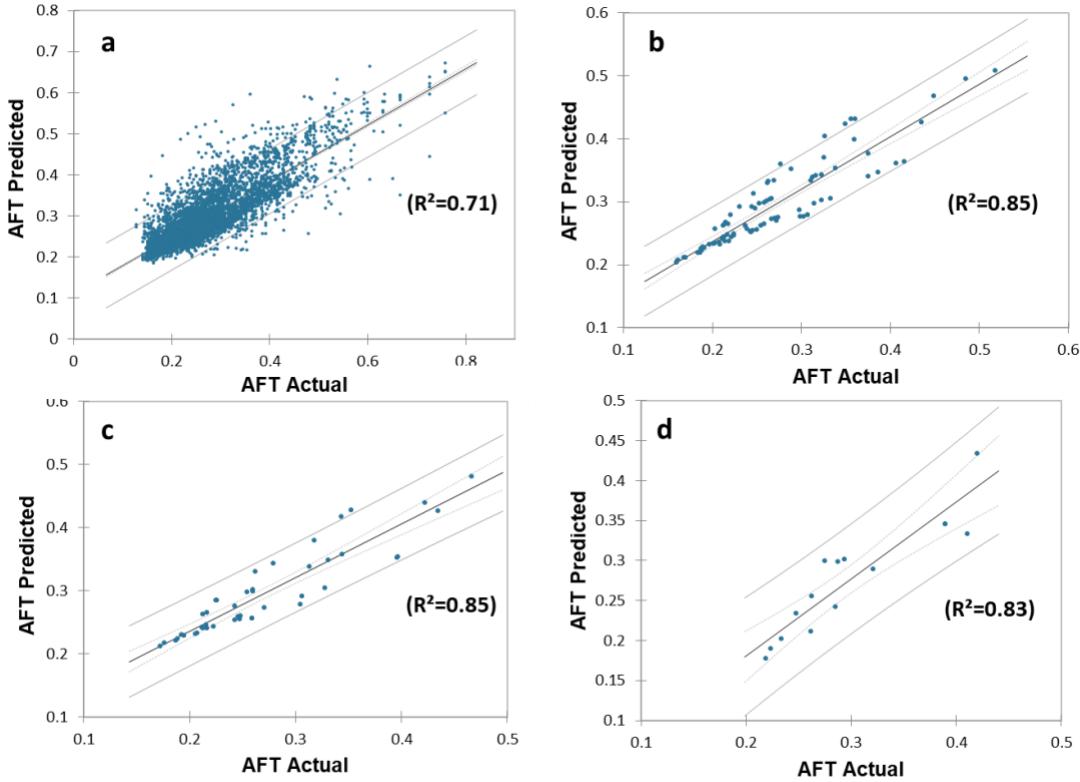


Figure 7.9. Correlation analysis for actual and predicted AFT values on the unseen dataset for (a) every 1-second, (b) average of 1-min, (c) average of 2-min, and (d) average for each activity

Further, the predicted AFT features for 1-min, 2-min, and each activity are used for fatigue assessment to evaluate the performance of the proposed fatigue monitoring system. The classifier performance on predicted AFT features for 1-min, 2-min, and each activity is shown in Table 7.8. The results show that the overall performance of the fatigue assessment is better for the predicted AFT values for 2-min (accuracy = 76.74%) compared to 1-min (accuracy = 71.05%) or for each activity (accuracy = 71.05%). Whereas the classification accuracy using HR (35.71%) and %HRR (35.71%) features of the unseen dataset is very low compared to AFT features. Figure 7.11 shows that the high rate of misclassification when using AFT features was observed between low and moderate fatigue levels.

Table 7.8. Fatigue level assessment using predicted AFT features for every 1-min, 2-min, and each activity

	Accuracy	Recall	Precision	F1 Score
Average Predicted AFT for 1-Minute	71.05%	71.10%	86.10%	72.40%
Average Predicted AFT for 2-Minute	76.74%	76.70%	86.10%	76.10%
Average Predicted AFT for Each Activity	71.43%	71.40%	90.50%	73.80%
HR Features for Each Activity	35.71%	35.70%	82.70%	36.20%
%HRR Features for Each Activity	35.71%	35.70%	40.50%	33.33%

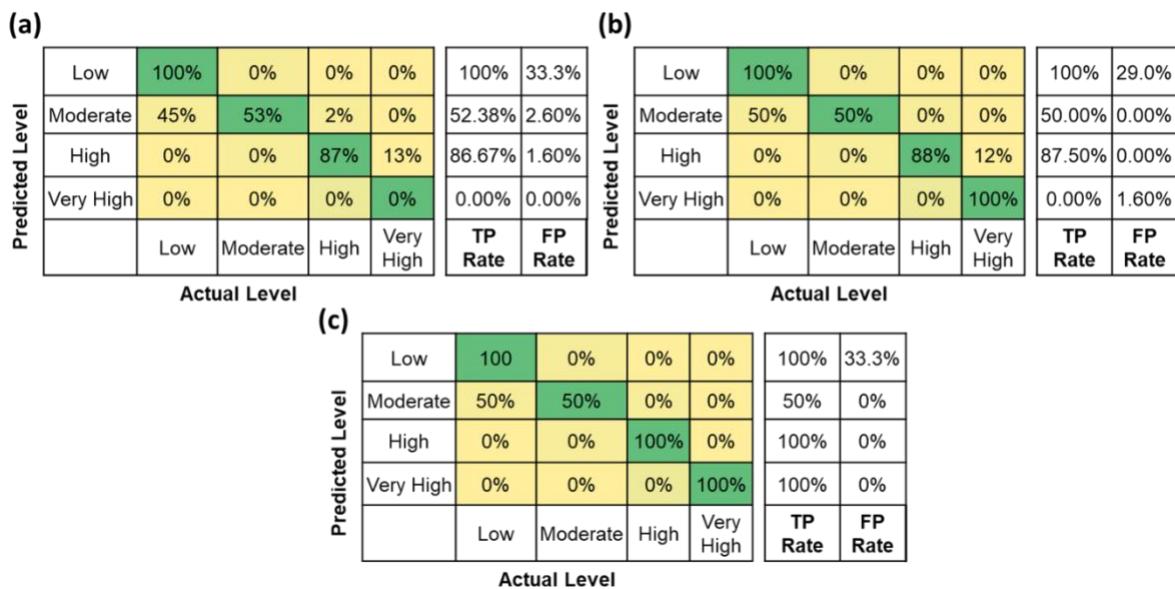


Figure 7.10. Confusion matrix of fatigue level assessment using predicted AFT features (a) for 1-min, (b) for 2-min, and (c) for each activity

Figure 13 and 14 shows the actual and predicted fatigue level over the entire duration of the unseen dataset where most of the time, the low and moderate fatigue levels were misclassified. From Figure 15, it can be observed that fatigue level is accurately predicted for high-intensity activities such as GUDVL, CDB, and IBDL. Some of the low-intensity activities such as ALJ, CLJ, and DG are misclassified as low instead of moderate. Moreover, it can be observed that the misclassifications of lower fatigue levels were less for 2-min compared to 1-min average AFT.

This shows that the longer window sizes are suitable for low-intensity activities. This shows that the proposed system can be used for continuous monitoring of fatigue levels.

Further, hypothesis testing was performed using a Chi-Squared test between actual and predicted fatigue levels for each window size (i.e., 1-min, 2-min, and average over each activity). The null hypothesis was assumed that the actual fatigue levels are not related to predicted fatigue levels. Whereas the alternate hypothesis is that the actual fatigue levels are related to predicted fatigue levels. The results show that the p-value for all the window sizes is less than 0.05, therefore we reject the null hypothesis and accept the alternate hypothesis, which concludes that there exists a significant relationship between actual and predicted fatigue levels for three window sizes (Table 7.9).

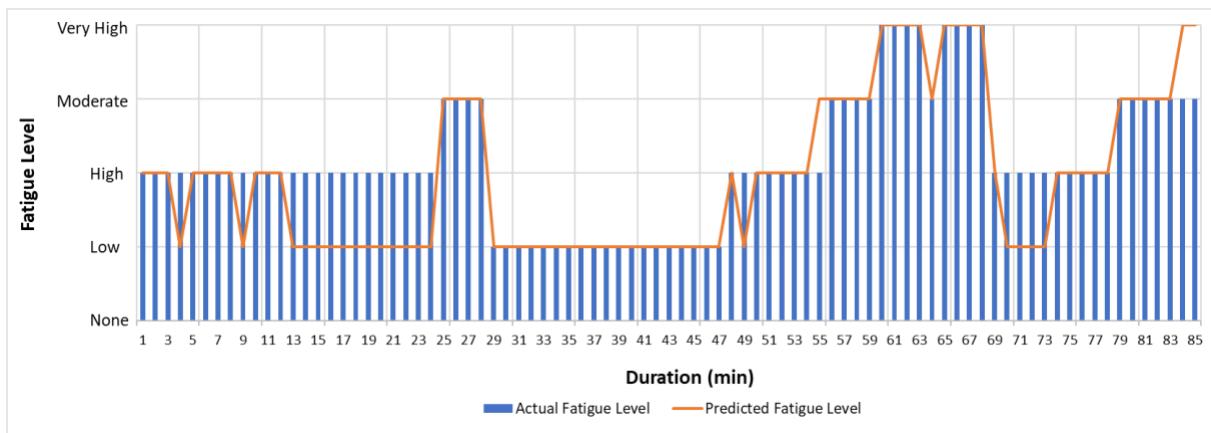


Figure 7.11. Predicted and actual fatigue level for every 1-min on the unseen dataset

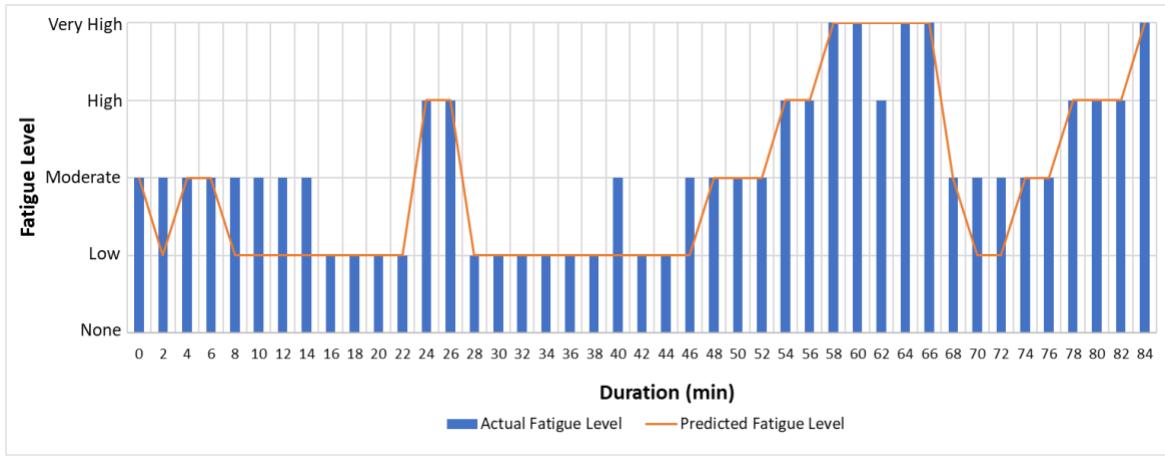


Figure 7.12. Predicted and actual fatigue level for every 2-min on the unseen dataset

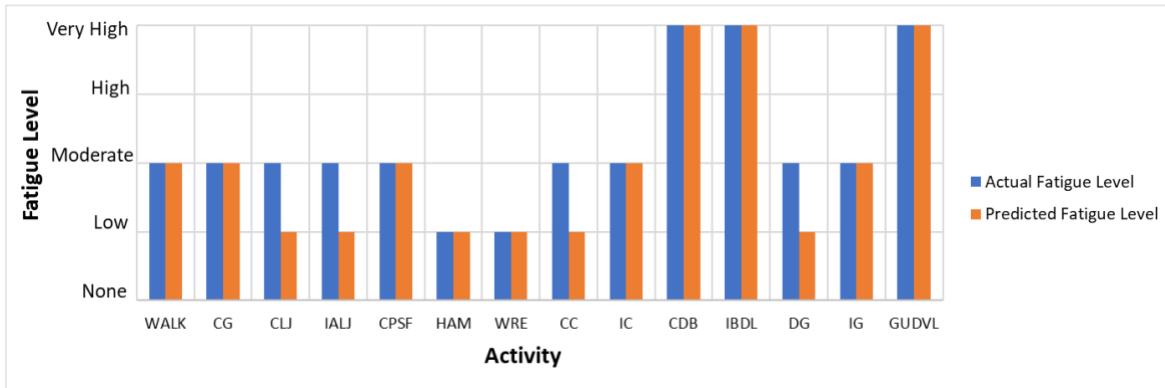


Figure 7.13. Predicted and actual fatigue level for each activity on the unseen dataset

Table 7.9. Chi-Squared test between actual and predicted fatigue levels for different window sizes

	Test statistic	df	p-value
1-Min	160.93	9	2.20E-16
2-Min	86.64	9	7.67E-15
Average Activity	30.333	9	0.000385

#### 7.4.Discussion

This study confirmed that the proposed fatigue monitoring system could continuously assess workers' fatigue levels. The use of the aerobic fatigue threshold (92.31%) to monitor the fatigue level has achieved high classification accuracy compared to the previous studies, which

used energy expenditure (90%) (Jebelli et al. 2019) and skin temperature (80.60%) (Aryal et al. 2017). The better performance of the system is due to the high correlation of the AFT feature with the fatigue levels. Comparing classification accuracy for different feature combinations shows that the AFT features have high performance compared to HR features, which agrees with the previous study (Aryal et al. 2017). This shows that the HR features alone are not suitable for fatigue level assessment for the activities performed in a short interval of time. Previous studies have considered 2-min (Aryal et al. 2017), 35 s (Jebelli et al. 2019), and 30-min (Hwang and Lee 2017) window to assess the worker fatigue level. Jebelli et al. (2019) suggested that the higher window size is required to recognize the physical demand compared, the comparison of classification accuracy for different window sizes identified that 2-min average AFT had achieved high classification accuracy compared to 1-min and over the duration of activity. This study recommends using a higher window size for low-intensity activities and a smaller window size for high-intensity activities.

The AFT values for the high and very high fatigue level activities such as CDB, IBDL, and GUDVL are above 33%, which agrees with published guidelines for oxygen uptake (Abdelhamid and Everett 2002; Åstrand et al. 2003). Moreover, the classification accuracy was higher for high fatigue level activities which shows that AFT is a suitable metric to assess the fatigue level of high-intensity activities. The predicted AFT values using forearm motion and muscle activity data are highly correlated with actual values, showing that oxygen consumption is highly influenced by the type of activity performed in a short interval. The continuous monitoring of AFT can assess activity work severity classification based on published guidelines for oxygen uptake.

The proposed system is highly suitable for construction applications because it uses one single armband sensor to capture EMG and IMU data and AFT metrics, dependent on the activity.

Since the EMG and IMU signals are dependent on activity, which helps recognize complex activities performed in a short interval of time, the proposed system can be used for any trade. Moreover, the performance of the proposed system on the unseen dataset has shown the feasibility of using the system for complex high-intensity activities.

The previous studies classify the physical demand or fatigue level based on the physiological signal data from the worker. Unlike previous studies, the proposed system continuously measures the aerobic fatigue threshold using forearm EMG and IMU data that provides an opportunity to quantify direct impacts of fatigue to accidents, evaluate the worker capabilities, and assess the workload evaluation of the task.

The study has successfully demonstrated forearm EMG and IMU data to monitor the workers' fatigue level continuously; however, it has some limitations. The proposed system was validated and evaluated using only scaffold building tasks. Even though the tasks were complex and highly physically demanding, different tasks should be studied for real-world fatigue monitoring applications. The study was conducted in a warehouse environment at 72°F; however, other environmental and site conditions were not considered. Only one participant data was used to assess the performance of the proposed fatigue monitoring system. An extensive experiment needs to be performed with multiple participants of different ages, ethnicity, work experience, and physical health to evaluate the performance of the proposed system. The combined MAC value was used for hammering, wrenching, and dragging activities instead of individual activity-specific MAC value, which is one of the limitations of the study.

The future research focuses on improving the robustness and performance of the system by training the models using data from workers with different characteristics (i.e., age, work experience, ethnicity, and health conditions), trades, and working conditions. Moreover,

investigate the deep learning algorithms to classify the fatigue levels. To reduce the possible bias due to the subjective fatigue level rating, the use of AFT value ranges from extensive population data to train fatigue level classifier.

## **7.5.Conclusions**

This study proposes an automated framework to monitor the worker fatigue level continuously using forearm-based EMG and IMU sensors by measuring the aerobic fatigue threshold. The system validation and performance evaluation confirmed that the forearm EMG and IMU data could recognize complex construction activities, instantaneous oxygen uptake, continuous aerobic fatigue threshold, and classify fatigue level. The results conclude that AFT features could classify the fatigue levels with high accuracy of 92.31% accuracy compared to HR (51.28%) and %HRR (50.43%). Moreover, the results show that the AFT is greater than 33% for high and very high fatigue levels, which agrees with the NIOSH exercise intensity threshold. The proposed system is highly suitable for construction applications because the entire framework is dependent on the activity performed by the worker. Since the proposed framework is dependent on the activity, the system can be adaptable for any trade and site conditions. The continuous monitoring of fatigue level helps assess the worker's physiological status, evaluate the physical workload of the activity, quantify the direct impacts of the fatigue level on the accidents, and early detection of risk.

## **CHAPTER 8. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH**

### **8.1. Summary of Research**

The construction industry often involves high labor-intensive and repetitive tasks, which results in worker physical fatigue. About 40% of the US construction workforce experiences high-level fatigue daily, which leads to poor judgment, increased risk of injuries, a decrease in productivity, and a lower quality of work. Physical fatigue is a predominant risk factor for injuries and illnesses in the construction industry. Therefore, it is essential to monitor fatigue to reduce the adverse effects and long-term health problems. Since overall physical fatigue demonstrates itself in complex processes, there is no standard measurement for determining the level of fatigue. Moreover, overall physical fatigue results from local (muscular fatigue) interaction and central factors (such as metabolic, cardiovascular, and thermoregulatory). Therefore, fatigue quantification typically involves a combination of kinematic and kinetic measurements, which is often supplemented or substituted by physiological (body temperature, heart rate, or muscle activity) and subjective measures (perceived exertion or discomfort) measures. In the construction industry, subjective measurement methods are currently adopted where the safety staff records the workers' level of fatigue using a fatigue scale, which is invasive and cumbersome. In recent years, few researchers have proposed objective techniques by monitoring physiological responses such as heart rate, skin temperature, electrodermal activity, and photoplethysmogram of the worker to assess the worker's work severity or physical workload. However, these methods have limitations such as unable identify the physical demand of the worker with different characteristics, the physiological responses used are influenced by external factors which may not yield reliable results on construction sites, not capable of continuous workers' fatigue level monitoring for multiple

tasks performed in a short interval of time, and do not help in quantifying direct impacts of the fatigue level on the safety performance or accidents.

To address these challenges, this research aims to promote safe, healthy, and productive workers by developing a system for continuous workers' fatigue monitoring by measuring aerobic fatigue threshold (AFT) using forearm motion and muscle activity data. The forearm motion and muscle activity data were acquired using a low-cost, noninvasive, wearable sensor. The proposed fatigue monitoring system consists of multiple measurable frameworks with five objectives: (1) assess the data quality and reliability of forearm motion and muscle activity data, (2) to develop and validate the construction workers' activity recognition framework using forearm muscle activity and motion data, (3) to estimate construction activity-specific maximum aerobic capacity, (4) to develop and validate continuous oxygen uptake prediction framework using forearm muscle activity and motion data, and (5) to develop fatigue level classifier using AFT features and validate the proposed fatigue monitoring system. This study applies state-of-the-art machine learning techniques and works physiology methodologies to understand and analyze forearm motion and muscle activity patterns to assess worker fatigue while performing complex construction activities such as scaffold building. To achieve these objectives, five inter-related studies were conducted. A summary of these studies and implications are as follows:

### **8.1.1. System Development, Data Quality, and Reliability Assessment**

Chapter 3 evaluated the data quality and reliability of forearm electromyography (EMG) and inertial measurement unit (IMU) data from a low-cost off-the-shelf wearable sensor. This study proposed different experiments to test the data quality and reliability for various conditions on construction sites. Some of the experiments include evaluating the data quality for "at-rest" and "in-motion" activities, the effect of armband sensor position on EMG and IMU data quality, the

reliability of the data while performing construction activities, reliability of activity classification results, and effect of activity intensity on classification results. The experiments confirmed that there was a minimal effect of environment, confounding factors (communication device, power tools, other sensors, and smartwatches), and inter-device variability. Moreover, the reliability tests confirm a strong relative and absolute reliability for the data and classification results. This study concludes that the armband sensor data and classification results are highly reliable, and the scientific evaluation of the armband sensor builds trustworthiness to use the sensor for various construction applications in noisy environments.

### **8.1.2. Automated Construction Workers' Activity Recognition**

Chapter 4 developed an automated framework to recognize construction workers' activity using forearm EMG and IMU data and bidirectional long short-term memory (BiLSTM) model. The feature extraction and data fusion techniques introduced in this chapter have helped to achieve high accuracy. The proposed activity recognition framework was validated using the scaffold building case study since it requires extensive manual efforts and involves different body parts movement and various motions. The leave-one-subject-out cross-validation results using ten participants' data showed that the proposed model achieved an overall accuracy of 99.77%. Moreover, the predictions of the trained model on an unseen dataset collected from a participant performing scaffold building activities in random order was 87.18%. Further, the activity recognition using different sensor features showed that the EMG and IMU (87.18%) data fusion had yielded high accuracy compared to IMU (85.85%) and EMG (48.99%) alone on the unseen dataset. Through a case study, the proposed framework can recognize complex construction activities performed in a short interval of time. Since the proposed activity recognition framework

is fully automated, scalable, robust, and adaptable, the system can be used for the feasibility of workload assessment, fatigue monitoring, and productivity assessment.

### **8.1.3. Maximum Aerobic Capacity for Construction Activities**

Chapter 5 developed maximum aerobic capacity (MAC) estimation models for construction activities using submaximal experiment protocol. The submaximal experiment protocol involves performing an activity at three different intensities, and a regression model was built using measured heart rate and oxygen uptake. This study developed MAC estimation models for walking, carrying, lifting, and combined (carrying, dragging, hammering, lifting, and wrenching). The results show that the average MAC value for walking, carrying, lifting, and combined activities are 2.95 l/m, 2.55 l/m, 2.82 l/m, and 2.30 l/m, respectively. The MAC value of combined construction activities is significantly lower than walking (28%), carrying (11%), and lifting (22%). The study recommends using construction activity-specific exercise protocol for evaluating job or worker to prevent a mismatch between job demands and worker capabilities. The estimation models help design the construction work, which will create a safer work environment and enhance productivity.

### **8.1.4. Oxygen Uptake Prediction during Construction Activities**

Chapter 6 developed an automated framework for continuous oxygen uptake prediction using BiLSTM and Wearable EMG and IMU sensor data. The oxygen uptake prediction framework involves data fusion, feature extraction, and data normalization techniques. The proposed framework was validated for scaffold building activities using ten participants' data. The results show a strong correlation between the EMG and IMU features and oxygen uptake ( $R = 0.90$ ,  $RMSE = 1.257 \text{ mL/kg/min}$ ). Moreover, the measured ( $9.18 \pm 1.97 \text{ mL/kg/min}$ ) and predicted ( $9.22 \pm 0.09 \text{ mL/kg/min}$ ) average oxygen consumption to build one scaffold unit are significantly

the same. The study concludes that the fusion of EMG and IMU features resulted in high model performance compared to EMG and IMU alone. The main advantages of the proposed framework over previous studies are the use of low-cost sensor, complex construction activities with varying intensities in a short interval of time, use of fully automated framework, feature engineering process to improve performance on the unseen dataset, and use of BiLSTM and appropriate hyperparameters to handle complex time-series data. The proposed framework helps evaluate the physical demands of the construction activity, monitor the worker's physiological status, and detect the risk early.

#### **8.1.5. Fatigue Monitoring System Feasibility Validation and Performance Evaluation**

Chapter 7 assessed the proposed fatigue monitoring system feasibility and performance evaluation. First, the study evaluated the feasibility and performance of the AFT features for fatigue level assessment using the oxygen uptake, heart rate, and fatigue rating recorded from ten participants. Moreover, the best fatigue level classifier was build using the AFT features. Second, the performance of the fatigue monitoring system was evaluated using an unseen dataset. The results show that the AFT (92.31%) features have outperformed in assessing the workers' fatigue level compared to heart rate (51.28%) and percentage heart rate reserve (50.43%) features. Among ten commonly used machine learning-based classification algorithms, the decision tree classifier was best in classifying fatigue levels. The correlation analysis has shown that the average actual and predicted AFT values for 1-min and 2-min are highly correlated with  $R^2 = 0.85$ . The continuous monitoring of AFT can assess activity work severity classification based on published guidelines for oxygen uptake. The overall performance of the proposed fatigue monitoring system on unseen data using average 2-min AFT features was 76.74%. Moreover, the results show that the model performed very well for high-intensity activities. The unseen dataset results confirm the

potential of the proposed framework in assessing the workers' fatigue. The proposed system is highly suitable for construction applications because the entire framework is dependent on the activity performed by the worker. Moreover, the physiological measures of oxygen uptake were determined using the forearm EMG and IMU data, which are dependent on activities. The proposed activity-based fatigue monitoring system using noninvasive armband sensor to assess workers' physical fatigue level helps in assessing worker's physiological status, evaluating the physical workload of the activity, quantify the direct impacts of the fatigue level on the accidents, enhancing the workers' safety, health, and productivity through early detection of risk.

## **8.2. Recommendations for Future Research**

This dissertation proposed an automated framework for a continuous workers' fatigue monitoring system, which consists of multiple steps which still warrant further attention in future research efforts to improve these steps and explore other possibilities. Some of the recommendations for future research are as follows:

- This research investigated various confounding factors on the EMG and IMU signal quality in residential construction settings. However, further research is required to assess the quality and reliability of the sensor data in various noisy construction environments and investigate various signal processing techniques to eliminate the signal artifacts and improve signal quality.
- This study explored the signal level and feature level fusion for activity recognition. However, further research efforts are required to investigate the influence of various levels of sensor data fusion (i.e., signal level, feature level, and decision level) on overall fatigue monitoring.

- The current state-of-the-art deep learning models can understand the latent features in the signal data and improve the performance. However, the feature engineering techniques help prevent overfitting and improve the performance on the real-time dataset. This study investigated the performance of statistical features over 1s second interval for activity recognition and oxygen uptake prediction. Further studies are required to investigate the frequency-domain and time-domain features from the sensor data and determine the optimal window size. Also, future studies can explore time series data augmentation techniques to improve the model performance since collecting extensive data for model training on construction sites is challenging.
- This study has considered only one combined activities sequence to build a MAC estimation model for combined construction activities. However, future studies must investigate how different activity sequences impact the MAC value for combined activities.
- This study has evaluated the feasibility and performance of simulated scaffold building activities performed in a warehouse environment for a limited time. Further studies are required to evaluate the system on real job sites for different trades with a diverse population.
- Since this study has proved the feasibility of using wearable sensors for continuous fatigue monitoring, it is beneficial to explore workers' fatigue or workload on the workers' health, job safety, productivity, and quality of work.
- Further studies are required to investigate the use of the proposed system to determine work-rest intervals for various construction activities and workload evaluation of construction activities, which helps to redesign the activities to prevent WMSDs and improve workers' performance.

- With advancements in sensing technologies and machine learning, future studies can explore the possibility of embedding the proposed models within the wearable sensor to build real-time alert systems to prevent accidents.

## APPENDIX A. IRB APPROVAL FORMS



**TO:** Wang, Chao  
LSUAM | Col of ENGR | Construction  
Management

**FROM:** Alex Cohen  
Chair, Institutional Review Board

**DATE:** 12-Aug-2020

**RE:** IRBAM-20-0112

**TITLE:** Wearable Myo Armband Sensor Data Quality  
Assessment

**SUBMISSION TYPE:** Initial Application

**Review Type:** Expedited Review

**Risk Factor:** Minimal

**Review Date:** 11-Aug-2020

**Status:** Approved

**Approval Date:** 11-Aug-2020

**Approval Expiration Date:** 10-Aug-2021

**Re-review frequency:** (three years unless otherwise stated)

**Number of subjects approved:** 10

**LSU Proposal Number:**

**By:** Alex Cohen, Chairman

**Continuing approval is CONDITIONAL on:**

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects\*
  2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
  3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
  4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
  5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
  6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
  7. Notification of the IRB of a serious compliance failure.
- 8. SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

*\* All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at <http://www.lsu.edu/irb>*

Louisiana State University O 225-578-5833  
131 David Boyd Hall F 225-578-5983  
Baton Rouge, LA 70803 [www.research.lsu.edu](http://www.research.lsu.edu)

O 225-578-5833  
F 225-578-5983  
[www.research.lsu.edu](http://www.research.lsu.edu)

**TO:** Wang, Chao  
 LSUAM | Col of ENGR | Construction Management

**FROM:** Alex Cohen  
 Chair, Institutional Review Board

**DATE:** 04-Nov-2020

**RE:** IRBAM-20-0539

**TITLE:** Prediction of oxygen uptake and maximum aerobic capacity using wearable sensors during construction activities

**SUBMISSION TYPE:** Initial Application

**Review Type:** Expedited Review

**Risk Factor:** Minimal

**Review Date:** 04-Nov-2020

**Status:** Approved

**Approval Date:** 04-Nov-2020

**Approval Expiration Date:** 03-Nov-2021

**Re-review frequency:** Annually

**Number of subjects approved:** 15

**LSU Proposal Number:**

**By:** Alex Cohen, Chairman

**Continuing approval is CONDITIONAL on:**

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects\*
  2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
  3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
  4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
  5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
  6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.
  7. Notification of the IRB of a serious compliance failure.
- 8. SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

\* All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at <http://www.lsu.edu/research>

Louisiana State University  
 131 David Boyd Hall  
 Baton Rouge, LA 70803

O 225-578-5833  
 F 225-578-5983  
<http://www.lsu.edu/research>

## APPENDIX B. SAMPLE RAW ARMBAND SENSOR DATA

timestamp	emg1	emg2	emg3	emg4	emg5	emg6	emg7	emg8
1.61205E+15	-1	0	-1	-1	-1	-1	1	-2
1.61205E+15	-2	1	-1	-1	1	0	-1	-1
1.61205E+15	0	-2	0	1	-1	0	0	-1
1.61205E+15	0	0	-1	1	0	-1	-1	-1
1.61205E+15	-2	-1	-3	-2	0	-1	-1	-1
1.61205E+15	0	-2	-2	-2	0	-2	-1	-1
1.61205E+15	-1	-1	-3	0	-2	-1	0	-1
1.61205E+15	1	0	-2	-1	-2	-1	-2	-1
1.61205E+15	0	0	0	-1	-2	1	-1	-1
1.61205E+15	-3	-1	0	-1	0	0	-1	-1
1.61205E+15	3	0	-1	-2	1	-1	-1	-1
1.61205E+15	-2	-2	1	-1	-1	-2	-1	-1
1.61205E+15	-1	1	-1	0	-1	0	-1	-2
1.61205E+15	-2	0	-1	0	-2	0	-1	-1
1.61205E+15	-5	-2	-1	-1	-1	-1	-1	-1
1.61205E+15	-2	0	0	-1	0	1	0	-1
1.61205E+15	0	0	1	-2	1	1	0	0
1.61205E+15	0	-1	-2	0	-2	-1	-1	0
1.61205E+15	1	-2	-2	-1	-2	-1	-1	-1
1.61205E+15	-4	-1	-2	-1	-1	-2	-3	-1
1.61205E+15	1	-1	-1	0	0	-2	-1	-1
1.61205E+15	1	-2	-1	1	-1	0	-1	-1
1.61205E+15	8	3	1	-1	-1	0	-1	2
1.61205E+15	-4	-6	-2	-1	-2	-2	-1	2

Figure B. 1. Sample raw EMG data from armband sensor

timestamp	ax	ay	az	gx	gy	gz	ox	oy	oz	ow	roll	pitch	yaw
1.61205E+15	1.02979	0.438477	0.233398	-40.375	-24.4375	14.1875	0.141663	-0.52277	-0.11047	0.833313	0.704918	-0.9972	-0.65893
1.61205E+15	1.01514	0.39209	0.285645	-53.75	-33.6875	16.625	0.131958	-0.52887	-0.11353	0.830688	0.696441	-1.0135	-0.66933
1.61205E+15	1.03564	0.303711	0.33252	-45.0625	-43.3125	17.125	0.122437	-0.53601	-0.11554	0.827271	0.690145	-1.03245	-0.68003
1.61205E+15	0.987305	0.300781	0.302246	-25.5625	-47.3125	18.625	0.115723	-0.54395	-0.11578	0.822998	0.692495	-1.05224	-0.69258
1.61205E+15	0.879395	0.362305	0.243652	-7.9375	-47.5625	18.25	0.11261	-0.55206	-0.11505	0.818115	0.706018	-1.0704	-0.70958
1.61205E+15	0.779297	0.410645	0.202637	-4.75	-45	15.125	0.111084	-0.55951	-0.11499	0.813232	0.724933	-1.08537	-0.73062
1.61205E+15	0.727539	0.369629	0.195801	-9.875	-42.3125	10.75	0.10907	-0.56616	-0.11597	0.808777	0.742818	-1.09844	-0.75285
1.61205E+15	0.714355	0.299805	0.209961	-11.8125	-40.8125	7.5	0.106445	-0.57221	-0.11719	0.804688	0.758201	-1.11055	-0.77381
1.61205E+15	0.744629	0.28125	0.252441	-3.6875	-43.5	4.4375	0.104858	-0.57831	-0.11792	0.800415	0.777442	-1.12217	-0.79639
1.61205E+15	0.800781	0.284668	0.277344	0.6875	-48.1875	1	0.104004	-0.58484	-0.11871	0.795654	0.80221	-1.13366	-0.82352
1.61205E+15	0.851563	0.290039	0.289063	-6.375	-52.0625	-2	0.102234	-0.59161	-0.12061	0.790527	0.828664	-1.14501	-0.85524
1.61205E+15	0.875	0.298828	0.298828	-23.0625	-54.6875	-5.125	0.09845	-0.59815	-0.12482	0.785461	0.853675	-1.15566	-0.89178
1.61205E+15	0.872559	0.308594	0.308105	-40.75	-55.8125	-8	0.09259	-0.60407	-0.13129	0.780579	0.875217	-1.16485	-0.9311
1.61205E+15	0.848145	0.30127	0.322266	-52.6875	-56.125	-9.8125	0.085205	-0.60931	-0.1394	0.775879	0.89349	-1.17257	-0.9717
1.61205E+15	0.812012	0.288086	0.314941	-55.8125	-55.9375	-8.625	0.077332	-0.61432	-0.14764	0.77124	0.910999	-1.18035	-1.0124
1.61205E+15	0.777344	0.27832	0.323242	-50.125	-55.25	-6.4375	0.069946	-0.61908	-0.15497	0.766724	0.928698	-1.18806	-1.05125
1.61205E+15	0.755371	0.259766	0.338867	-39.0625	-53.3125	-2.75	0.063599	-0.62354	-0.16058	0.762512	0.94611	-1.19572	-1.08584
1.61205E+15	0.72998	0.249512	0.351074	-24	-51.125	0.1875	0.05896	-0.62787	-0.16419	0.758545	0.965368	-1.20314	-1.11688
1.61205E+15	0.73584	0.25293	0.365234	-5.0625	-50.375	1.875	0.056641	-0.63239	-0.16547	0.754639	0.989274	-1.21024	-1.14587
1.61205E+15	0.754395	0.263184	0.367676	10.125	-47.8125	3.625	0.056274	-0.63696	-0.16492	0.750977	1.01687	-1.21718	-1.17312
1.61205E+15	0.770996	0.285645	0.337891	19	-41.3125	5.25	0.05719	-0.64111	-0.16315	0.747742	1.04402	-1.22296	-1.19631
1.61205E+15	0.782227	0.312988	0.330566	21.5	-34	5.875	0.058838	-0.64435	-0.16113	0.745239	1.06726	-1.22683	-1.21423
1.61205E+15	0.8125	0.330566	0.328613	17.625	-27.1875	6.5	0.06012	-0.64667	-0.15942	0.74353	1.0841	-1.22993	-1.22672
1.61205E+15	0.885254	0.326172	0.355469	12	-21.8125	6.5625	0.060608	-0.64789	-0.15814	0.742676	1.0919	-1.23205	-1.23195

Figure B. 2. Sample raw IMU data from armband sensor

## APPENDIX C. FATIGUE LEVEL CLASSIFIER PERFORMANCE

Table C. 1. Classification recall for different feature combinations

Model	AFT	%HRR	HR	AFT+ %HRR	AFT+ HR	HR+ %HRR	AFT+ HR+%HRR
RF	90.60%	44.44%	41.90%	89.70%	90.60%	44.44%	90.60%
DT	92.70%	41.00%	44.44%	91.50%	88.00%	34.20%	90.60%
NB	82.90%	46.20%	47.00%	79.50%	78.60%	42.70%	74.40%
LDA	87.20%	46.20%	51.30%	88.90%	83.80%	42.70%	82.90%
QDA	85.50%	38.50%	36.80%	79.50%	77.80%	37.60%	67.50%
SVM	72.60%	46.20%	44.44%	72.60%	66.70%	50.40%	68.40%
ADA	52.10%	46.20%	47.90%	52.10%	52.10%	46.20%	52.10%
LR	86.30%	50.40%	46.20%	79.50%	82.10%	45.30%	79.50%
KNN	81.20%	36.80%	38.50%	72.60%	72.60%	35.00%	65.00%
MLP	91.50%	46.20%	45.30%	88.90%	87.20%	46.20%	88.00%

Table C. 2. Classification precision for different feature combinations

Model	AFT	%HRR	HR	AFT+ %HRR	AFT+ HR	HR+ %HRR	AFT+ HR+%HRR
RF	90.70%	45.40%	42.30%	90.11%	90.70%	45.60%	90.90%
DT	92.30%	40.30%	43.40%	91.90%	88.40%	34.00%	91.10%
NB	82.90%	45.80%	42.20%	79.40%	79.00%	40.00%	74.30%
LDA	87.60%	44.40%	44.60%	89.20%	84.10%	40.90%	83.10%
QDA	85.60%	38.40%	36.10%	79.80%	77.80%	37.90%	66.80%
SVM	68.20%	46.20%	44.44%	70.70%	59.10%	50.40%	66.10%
ADA	52.10%	46.20%	47.90%	52.10%	52.10%	46.20%	52.10%
LR	86.40%	49.70%	40.20%	79.70%	82.10%	43.90%	79.80%
KNN	81.60%	37.00%	38.50%	74.60%	74.70%	35.70%	66.66%
MLP	91.50%	44.50%	41.80%	88.90%	87.20%	45.70%	88.50%

## APPENDIX D. COPYRIGHT PERMISSION (CHAPTER 3)

 **Attribution 4.0 International (CC BY 4.0)**

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

**You are free to:**

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.



**Under the following terms:**

 **Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**No additional restrictions** — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

## APPENDIX E. COPYRIGHT PERMISSION (CHAPTER 4)

5/7/2021

Rightslink® by Copyright Clearance Center



**RightsLink®**

Home

?



Email Support

Sign in

Create Account



ANN-based automated scaffold builder activity recognition through  
wearable EMG and IMU sensors

Author: Srikanth Sagar Bangaru, Chao Wang, Sri Aditya Busam, Fereydoun Aghazadeh

Publication: Automation in Construction

Publisher: Elsevier

Date: June 2021

© 2021 Elsevier B.V. All rights reserved.

### Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

© 2021 Copyright - All Rights Reserved | [Copyright Clearance Center, Inc.](#) | [Privacy statement](#) | [Terms and Conditions](#)  
Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)

## LIST OF REFERENCES

- Abdelhamid, T. S., and Everett, J. G. (2002). "Physiological demands during construction work." *Journal of Construction Engineering and Management*, 128(5), 427-437.
- Abut, F., and Akay, M. F. (2015). "Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances." *Medical Devices*, 8, 369.
- Abut, F., Akay, M. F., and George, J. (2016). "Developing new VO<sub>2</sub>max prediction models from maximal, submaximal, and questionnaire variables using support vector machines combined with feature selection." *Computers in Biology and Medicine*, 79, 182-192.
- Acevedo, E. O. (2012). *The Oxford handbook of exercise psychology*, Oxford University Press.
- Ahn, C. R., Lee, S., Sun, C., Jebelli, H., Yang, K., and Choi, B. (2019). "Wearable sensing technology applications in construction safety and health." *Journal of Construction Engineering and Management*, 145(11), 03119007.
- Akay, M. F., Zayid, E. I. M., Aktürk, E., and George, J. D. (2011). "Artificial neural network-based model for predicting VO<sub>2</sub>max from a submaximal exercise test." *Expert Systems with Applications*, 38(3), 2007-2010.
- Akhavian, R., and Behzadan, A. H. "Productivity analysis of construction worker activities using smartphone sensors." *Proc., 16th International Conference on Computing in Civil and Building Engineering*.
- Akhavian, R., and Behzadan, A. H. (2016). "Smartphone-based construction workers' activity recognition and classification." *Automation in Construction*, 71, 198-209.
- Akhavian, R., and Behzadan, A. H. (2018). "Coupling human activity recognition and wearable sensors for data-driven construction simulation." *Journal of Information Technology in Construction*, 23, 1-15.
- Albert, A., Hallowell, M. R., Kleiner, B., Chen, A., and Golparvar-Fard, M. (2014). "Enhancing construction hazard recognition with high-fidelity augmented virtuality." *Journal of Construction Engineering and Management*, 140(7), 04014024.
- Ali, M. (2020). "PyCaret: An open source, low-code machine learning library in Python."
- Altini, M., Penders, J., and Amft, O. (2015). "Estimating oxygen uptake during nonsteady-state activities and transitions using wearable sensors." *IEEE Journal of Biomedical and Health Informatics*, 20(2), 469-475.

- Álvarez-García, J. A., Barsocchi, P., Chessa, S., and Salvi, D. (2013). "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), 119-132.
- Alwasel, A., Abdel-Rahman, E. M., Haas, C. T., and Lee, S. (2017). "Experience, productivity, and musculoskeletal injury among masonry workers." *Journal of Construction Engineering and Management*, 143(6), 05017003.
- Ann, O. C., and Theng, L. B. "Human activity recognition: a review." *Proc., 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, IEEE, 389-393.
- Arellano, J., Martínez, J., Perez, J., and Alcaraz, J. (2015). "Relationship between workload and fatigue among Mexican assembly operators." *International Journal of Physical Medicine & Rehabilitation*, 3(315), 2.
- Arief, Z., Sulistijono, I. A., and Ardiansyah, R. A. "Comparison of five time series EMG features extractions using Myo Armband." *Proc., 2015 International Electronics Symposium (IES)*, IEEE, 11-14.
- Armstrong, T. J., Franzblau, A., Haig, A., Keyserling, W. M., Levine, S., Strelein, K., Ulin, S., and Werner, R. (2001). "Developing ergonomic solutions for prevention of musculoskeletal disorder disability." *Assistive Technology*, 13(2), 78-87.
- Aryal, A., Ghahramani, A., and Becerik-Gerber, B. (2017). "Monitoring fatigue in construction workers using physiological measurements." *Automation in Construction*, 82, 154-165.
- Astrand, P., Hultman, E., Juhlin-Dannfelt, A., and Reynolds, G. (1986). "Disposal of lactate during and after strenuous exercise in humans." *Journal of Applied Physiology*, 61(1), 338-343.
- Åstrand, P. O., Rodahl, K., Dahl, H. A., and Strømme, S. B. (2003). *Textbook of work physiology: physiological bases of exercise*, Human Kinetics.
- Awolusi, I., Marks, E., and Hallowell, M. (2018). "Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices." *Automation in construction*, 85, 96-106.
- Ayodele, O. A., Chang-Richards, A., and González, V. (2020). "Factors affecting workforce turnover in the construction sector: a systematic review." *Journal of Construction Engineering and Management*, 146(2), 03119010.
- Ayoub, M. (1992). "Problems and solutions in manual materials handling: the state of the art." *Ergonomics*, 35(7-8), 713-728.

Ayoub, M. M. (1989). *Manual materials handling: design and injury control through ergonomics*, CRC Press.

Bagherzadeh, F., Mehrani, M.-J., Basirifard, M., and Roostaei, J. (2021). "Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance." *Journal of Water Process Engineering*, 41, 102033.

Bangaru, S. S., Wang, C., and Aghazadeh, F. "Biomechanical Analysis of Manual Material Handling Tasks on Scaffold." *Proc., Computing in Civil Engineering 2019: Data, Sensing, and Analytics*, American Society of Civil Engineers Reston, VA, 572-579.

Bangaru, S. S., Wang, C., and Aghazadeh, F. (2020). "Data Quality and Reliability Assessment of Wearable EMG and IMU Sensor for Construction Activity Recognition." *Sensors*, 20(18), 5264.

Bangaru, S. S., Wang, C., Busam, S. A., and Aghazadeh, F. (2021). "ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors." *Automation in Construction*, 126, 103653.

Bangaru, S. S., Wang, C., Hassan, M., Jeon, H. W., and Ayiluri, T. (2019). "Estimation of the degree of hydration of concrete through automated machine learning based microstructure analysis—A study on effect of image magnification." *Advanced Engineering Informatics*, 42, 100975.

Bao, L., and Intille, S. S. "Activity recognition from user-annotated acceleration data." *Proc., International Conference on Pervasive Computing*, Springer, 1-17.

Bao, X., Howard, M., Niazi, I. K., and Kamavuako, E. N. "Comparison between Embroidered and Gel Electrodes on ECG-Derived Respiration Rate." *Proc., 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2622-2625.

Barut, O., Zhou, L., and Luo, Y. (2020). "Multitask LSTM Model for Human Activity Recognition and Intensity Estimation Using Wearable Sensor Data." *IEEE Internet of Things Journal*, 7(9), 8760-8768.

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

Beltrame, T., Amelard, R., Wong, A., and Hughson, R. L. (2017). "Prediction of oxygen uptake dynamics by machine learning analysis of wearable sensors during activities of daily living." *Scientific Reports*, 7(1), 1-8.

Benalcázar, M. E., Motoche, C., Zea, J. A., Jaramillo, A. G., Anchundia, C. E., Zambrano, P., Segura, M., Palacios, F. B., and Pérez, M. "Real-time hand gesture recognition using the Myo armband and muscle activity detection." *Proc., 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)*, IEEE, 1-6.

Bengio, Y. "Deep learning of representations: Looking forward." *Proc., International Conference on Statistical Language and Speech Processing*, Springer, 1-37.

BLS (2017). "Injuries, Illnesses, and Fatalities."  
<https://www.bls.gov/iif/oshwc/cfoi/cftb0321.htm>. (Dec 12, 2019).

BLS (2018). "Survey of Occupational Injuries and Illnesses Data." <<https://www.bls.gov/iif/soii-data.htm>>. (Dec 27, 2019).

BLS (2019). "Injuries, Illnesses, and Fatalities." <[https://www.bls.gov/web/osh/summ1\\_00.htm](https://www.bls.gov/web/osh/summ1_00.htm)>. (April 5, 2021).

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue, and drowsiness." *Neuroscience & Biobehavioral Reviews*, 44, 58-75.

Borror, A., Mazzoleni, M., Coppock, J., Jensen, B. C., Wood, W. A., Mann, B., and Battaglini, C. L. (2019). "Predicting oxygen uptake responses during cycling at varied intensities using an artificial neural network." *Biomedical Human Kinetics*, 11(1), 60-68.

Bosch, J., Gracias, N., Rida, P., and Ribas, D. (2015). "Omnidirectional underwater camera design and calibration." *Sensors*, 15(3), 6033-6065.

Bosecker, C., Dipietro, L., Volpe, B., and Igo Krebs, H. (2010). "Kinematic robot-based evaluation scales and clinical counterparts to measure upper limb motor performance in patients with chronic stroke." *Neurorehabilitation and Neural Repair*, 24(1), 62-69.

Bosquet, L., Merkari, S., Arvisais, D., and Aubert, A. E. (2008). "Is heart rate a convenient tool to monitor over-reaching? A systematic review of the literature." *British Journal of Sports Medicine*, 42(9), 709-714.

Bouktif, S., Fiaz, A., Ouni, A., and Serhani, M. A. (2018). "Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches." *Energies*, 11(7), 1636.

Bruton, A., Conway, J. H., and Holgate, S. T. (2000). "Reliability: what is it, and how is it measured?" *Physiotherapy*, 86(2), 94-99.

- Bujang, M. A., and Baharum, N. (2017). "A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review." *Archives of Orofacial Science*, 12(1).
- Cezar, G. (2012). "Activity recognition in construction sites using 3D accelerometer and gyrometer." Accessed Octobar, 10, 2018.
- Chae, J., Jin, Y., Sung, Y., and Cho, K. (2018). "Genetic algorithm-based motion estimation method using orientations and EMGs for robot controls." *Sensors*, 18(1), 183.
- Chaffin, D. B., Andersson, G. B., and Martin, B. J. (2006). *Occupational biomechanics*, John wiley & sons.
- Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., and Wessely, S. W. (1993). "Development of a fatigue scale." *Journal of Psychosomatic Research*, 37, 147-153.
- Chan, A. P., Yi, W., Wong, D. P., Yam, M. C., and Chan, D. W. (2012). "Determining an optimal recovery time for construction rebar workers after working to exhaustion in a hot and humid environment." *Building and Environment*, 58, 163-171.
- Chan, P., and Kaka, A. (2004). "Construction productivity measurement: A comparison of two case studies." *20th Annual ARCOM Conference* Edinburgh, Scotland.
- Chang, F. L., Sun, Y. M., Chuang, K. H., and Hsu, D. J. (2009). "Work fatigue and physiological symptoms in different occupations of high-elevation construction workers." *Applied Ergonomics*, 40(4), 591-596.
- Chen, J., Qiu, J., and Ahn, C. (2017). "Construction worker's awkward posture recognition through supervised motion tensor decomposition." *Automation in Construction*, 77, 67-81.
- Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z., and Liu, Y. (2020). "Deep learning for sensor-based human activity recognition: overview, challenges, and opportunities." *arXiv*, 2001.07416.
- Cheng, C.-F., Rashidi, A., Mark, D., and Anderson, D. V. (2018). "Evaluation of software and hardware settings for audio-based analysis of construction operations." *International Journal of Civil Engineering*.
- Cheng, T., Migliaccio, G. C., Teizer, J., and Gatti, U. C. (2012). "Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers." *Journal of Computing in Civil engineering*, 27(3), 320-335.
- Cheng, T., Teizer, J., Migliaccio, G. C., and Gatti, U. C. (2013). "Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data." *Automation in Construction*, 29, 24-39.

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

Cho, Y. K., Kim, K., Ma, S., and Ueda, J. "A robotic wearable exoskeleton for construction worker's safety and health." *Proc., ASCE Construction Research Congress*, 19-28.

Choi, E., Schuetz, A., Stewart, W. F., and Sun, J. (2017). "Using recurrent neural network models for early detection of heart failure onset." *Journal of the American Medical Informatics Association*, 24(2), 361-370.

Chollet, F. (2015). "Keras: Deep learning library for theano and tensorflow." URL: <https://keras.io/k>, 7(8), T1.

Christie, C. J., and Scott, P. A. (2005). "Comparison of maximal aerobic capacity during running and lifting activities." *Journal of the Ergonomics Society of South Africa*, 17(1), 41-49.

Cook, D., Feuz, K. D., and Krishnan, N. C. (2013). "Transfer learning for activity recognition: a survey." *Knowledge and Information Systems*, 36(3), 537-556.

Craig, B. N., Congleton, J. J., Kerk, C. J., Lawler, J. M., and McSweeney, K. P. (1998). "Correlation of injury occurrence data with estimated maximal aerobic capacity and body composition in a high-frequency manual materials handling task." *American Industrial Hygiene Association Journal*, 59(1), 25-33.

De Brabandere, A., De Beeck, T. O., Schütte, K. H., Meert, W., Vanwanseele, B., and Davis, J. (2018). "Data fusion of body-worn accelerometers and heart rate to predict VO<sub>2</sub>max during submaximal running." *PloS One*, 13(6), e0199509.

De Pasquale, G., and Somà, A. (2010). "Reliability testing procedure for MEMS IMUs applied to vibrating environments." *Sensors*, 10(1), 456-474.

Deb, S., Carruth, D. W., Sween, R., Strawderman, L., and Garrison, T. M. (2017). "Efficacy of virtual reality in pedestrian safety research." *Applied Ergonomics*, 65, 449-460.

Debnath, A. K., Blackman, R., and Haworth, N. (2015). "Common hazards and their mitigating measures in work zones: a qualitative study of worker perceptions." *Safety Science*, 72, 293-301.

Ding, F., Guo, W., Dai, L., and Du, J. (2020). "Attentive batch normalization for lstm-based acoustic modeling of speech recognition." *arXiv preprint arXiv:2001.00129*.

Dittner, A. J., Wessely, S. C., and Brown, R. G. (2004). "The assessment of fatigue: a practical guide for clinicians and researchers." *Journal of Psychosomatic Research*, 56(2), 157-170.

- Düking, P., Fuss, F. K., Holmberg, H.-C., and Sperlich, B. (2018). "Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity." *JMIR mHealth and uHealth*, 6(4), e102.
- Edwards, R. H. "Human muscle function and fatigue." *Proc., Ciba Foundation Symposium 82 - Human Muscle Fatigue: Physiological Mechanisms*, Wiley Online Library, 1-18.
- Esmaeili, B., and Hallowell, M. R. (2011). "Diffusion of safety innovations in the construction industry." *Journal of Construction Engineering and Management*, 138(8), 955-963.
- Fang, D., Jiang, Z., Zhang, M., and Wang, H. (2015). "An experimental method to study the effect of fatigue on construction workers' safety performance." *Safety Science*, 73, 80-91.
- Frank, A. E., Kubota, A., and Riek, L. D. "Wearable activity recognition for robust human-robot teaming in safety-critical environments via hybrid neural networks." *Proc., 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.
- Fredericks, T. K., Choi, S. D., Hart, J., Butt, S. E., and Mital, A. (2005). "An investigation of myocardial aerobic capacity as a measure of both physical and cognitive workloads." *International Journal of Industrial Ergonomics*, 35(12), 1097-1107.
- Future, B. Y. (2020). "Craft Labor Demand." <<https://byf.org/explore/demand-map/>>. (July 5th, 2020).
- Gatti, U. C., Migliaccio, G. C., Bogus, S. M., and Schneider, S. (2014). "An exploratory study of the relationship between construction workforce physical strain and task level productivity." *Construction Management and Economics*, 32(6), 548-564.
- Gatti, U. C., Schneider, S., and Migliaccio, G. C. (2014). "Physiological condition monitoring of construction workers." *Automation in Construction*, 44, 227-233.
- Gholamiangonabadi, D., Kislev, N., and Grolinger, K. (2020). "Deep Neural Networks for Human Activity Recognition With Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection." *IEEE Access*, 8, 133982-133994.
- Gong, J., and Caldas, C. H. (2009). "An intelligent video computing method for automated productivity analysis of cyclic construction operations." *Computing in Civil Engineering* (2009), 64-73.
- Grandjean, E. (1979). "Fatigue in industry." *Occupational and Environmental Medicine*, 36(3), 175-186.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. (2016). "LSTM: A search space odyssey." *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222-2232.

- Guérat, R., Cierro, A., Vanderdonckt, J., and Pérez-Medina, J. L. (2019). "Gesture elicitation and usability testing for an armband interacting with netflix and spotify." *International Conference on Information Technology & Systems*, Springer, 625-637.
- Guillén-Rogel, P., Franco-Escudero, C., and Marín, P. J. (2019). "Test-retest reliability of a smartphone app for measuring core stability for two dynamic exercises." *PeerJ*, 7, e7485.
- Häikiö, J., Kallio, J., Mäkelä, S.-M., and Keränen, J. (2020). "IoT-based safety monitoring from the perspective of construction site workers." *International Journal of Occupational and Environmental Safety*, 4(1), 1-14.
- Hall, S., Wild, F., and Olde Scheper, T. (2019). "Real-time auditory biofeedback system for learning a novel arm trajectory: a usability study." *Perspectives on Wearable Enhanced Learning (WELL)*, Springer, 385-409.
- Halperin, K. M., and McCann, M. (2004). "An evaluation of scaffold safety at construction sites." *Journal of Safety Research*, 35(2), 141-150.
- Hasanzadeh, S., Esmaeili, B., and Dodd, M. D. (2017). "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), 04017024.
- He, S., Yang, C., Wang, M., Cheng, L., and Hu, Z. (2017). "Hand gesture recognition using MYO armband." *2017 Chinese Automation Congress (CAC)*, IEEE, 4850-4855.
- Health, N. I. f. O. S. a. (1981). *Work Practices Guide for Manual Lifting*, US Department of Health and Human Services, Cincinnati, Ohio.
- Ho, B.-J., Liu, R., Tseng, H.-Y., and Srivastava, M. "Myobuddy: detecting barbell weight using electromyogram sensors." *Proc., Proceedings of the 1st Workshop on Digital Biomarkers*, 27-32.
- Hogan, N., and Sternad, D. (2009). "Sensitivity of smoothness measures to movement duration, amplitude, and arrests." *Journal of Motor Behavior*, 41(6), 529-534.
- Hopkins, W. G. (2000). "Measures of reliability in sports medicine and science." *Sports Medicine*, 30(1), 1-15.
- Horrey, W. J., Noy, Y. I., Folkard, S., Popkin, S. M., Howarth, H. D., and Courtney, T. K. (2011). "Research needs and opportunities for reducing the adverse safety consequences of fatigue." *Accident Analysis & Prevention*, 43(2), 591-594.

- Hsu, C. Y., Tsai, Y. S., Yau, C. S., Shie, H. H., and Wu, C. M. (2016). "Test-retest reliability of an automated infrared-assisted trunk accelerometer-based gait analysis system." *Sensors*, 16(8), 1156.
- Hunn, H. M., Lapuma, P. T., and Holt, D. T. (2002). "The Influence of Pre-test Anxiety, Personality and Exercise on VO<sub>2</sub>max Estimation." *Journal of Exercise Physiology Online*, 5(1).
- Hwang, S., Jebelli, H., Choi, B., Choi, M., and Lee, S. (2018). "Measuring workers' emotional state during construction tasks using wearable EEG." *Journal of Construction Engineering and Management*, 144(7), 04018050.
- Hwang, S., and Lee, S. (2017). "Wristband-type wearable health devices to measure construction workers' physical demands." *Automation in Construction*, 83, 330-340.
- Hwang, S., Seo, J., Jebelli, H., and Lee, S. (2016). "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, 372-381.
- Hwang, S., Seo, J., Ryu, J., and Lee, S. "Challenges and opportunities of understanding construction workers' physical demands through field energy expenditure measurements using a wearable activity tracker." *Proc., Construction Research Congress 2016*, 2730-2739.
- Ilmarinen, J. (1992). "Job design for the aged with regard to decline in their maximal aerobic capacity: Part I - guidelines for the practitioner." *International Journal of Industrial Ergonomics*, 10(1-2), 53-63.
- Intille, S. S., Bao, L., Tapia, E. M., and Rondoni, J. "Acquiring in situ training data for context-aware ubiquitous computing applications." *Proc., Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1-8.
- Iridiastadi, H., and Aghazadeh, F. (2005). "Physiological fatigue limit of combined manual materials handling tasks." *Occupational Ergonomics*, 5(3), 141-148.
- Iridiastadi, H., Aghazadeh, F., and Waly, S. (1997). "Maximum aerobic capacity of combined manual materials handling and standard tasks." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications Sage CA: Los Angeles, CA, 670-674.
- Janidarmian, M., Roshan Fekr, A., Radecka, K., and Zilic, Z. (2017). "A comprehensive analysis on wearable acceleration sensors in human activity recognition." *Sensors*, 17(3), 529.

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

Jebelli, H., Choi, B., and Lee, S. (2019). "Application of wearable biosensors to construction sites. II: assessing workers' physical demand." *Journal of Construction Engineering and Management*, 145(12), 04019080.

Jebelli, H., Khalili, M. M., and Lee, S. (2019). "Mobile EEG-based workers' stress recognition by applying deep neural network." *Advances in Informatics and Computing in Civil and Construction Engineering*, Springer, 173-180.

Jebelli, H., and Lee, S. (2019). "Feasibility of wearable electromyography (EMG) to assess construction workers' muscle fatigue." *Advances in Informatics and Computing in Civil and Construction Engineering*, Springer, 181-187.

Jeelani, I., Han, K., and Albert, A. (2018). "Automating and scaling personalized safety training using eye-tracking data." *Automation in Construction*, 93, 63-77.

Jensenius, A. R., Nymoen, K., Skogstad, S. A. v. D., and Voldsgaard, A. (2012). "A study of the noise-level in two infrared marker-based motion capture systems." *Proceedings of the 9th Sound and Music Computing Conference*.

Joshua, L., and Varghese, K. (2010). "Accelerometer-based activity recognition in construction." *Journal of Computing in Civil Engineering*, 25(5), 370-379.

Joshua, L., and Varghese, K. (2011). "Accelerometer-based activity recognition in construction." *Journal of Computing in Civil Engineering*, 25(5), 370-379.

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

Joshua, L., and Varghese, K. (2014). "Automated recognition of construction labour activity using accelerometers in field situations." *International Journal of Productivity and Performance Management*.

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

Juan, S. (2015). "Facts of Safety at Work." <[https://www.ilo.org/wcmsp5/groups/public/-/-dgreports/-/-dcomm/documents/publication/wcms\\_067574.pdf](https://www.ilo.org/wcmsp5/groups/public/-/-dgreports/-/-dcomm/documents/publication/wcms_067574.pdf)>. (Dec 27, 2019).

- Karimi, H., Taylor, T. R., Goodrum, P. M., and Srinivasan, C. (2016). "Quantitative analysis of the impact of craft worker availability on construction project safety performance." *Construction Innovation*.
- Kefer, K., Holzmann, C., and Findling, R. D. (2017). "Evaluating the Placement of Arm-Worn Devices for Recognizing Variations of Dynamic Hand Gestures." *Journal of Mobile Multimedia*, 12(3&4), 225-242.
- Khan, S. H., and Sohail, M. (2013). "Activity monitoring of workers using single wearable inertial sensor." *2013 International Conference on Open Source Systems and Technologies*, IEEE, 60-67.
- Khosrowpour, A., Fedorov, I., Holynski, A., Niebles, J. C., and Golparvar-Fard, M. "Automated worker activity analysis in indoor environments for direct-work rate improvement from long sequences of RGB-D images." *Proc., Construction Research Congress 2014*, 729-738.
- Khusainov, R., Azzi, D., Achumba, I. E., and Bersch, S. D. (2013). "Real-time human ambulation, activity, and physiological monitoring: taxonomy of issues, techniques, applications, challenges and limitations." *Sensors (Basel)*, 13(10), 12852-12902.
- Kim, C., Park, T., Lim, H., and Kim, H. (2013). "On-site construction management using mobile computing technology." *Automation in construction*, 35, 415-423.
- Kim, E., Lovera, J., Schaben, L., Melara, J., Bourdette, D., and Whitham, R. (2010). "Novel method for measurement of fatigue in multiple sclerosis: Real-Time Digital Fatigue Score." *Journal of Rehabilitation Research & Development*, 47(5).
- Kim, H. J., Lee, Y. S., and Kim, D. "Arm motion estimation algorithm using MYO armband." *Proc., 2017 First IEEE International Conference on Robotic Computing (IRC)*, IEEE, 376-381.
- Kim, K., and Cho, Y. K. (2020). "Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition." *Automation in Construction*, 113, 103126.
- Kim, S., Chang, S., and Castro-Lacouture, D. (2019). "Dynamic modeling for analyzing impacts of skilled labor shortage on construction project management." *Journal of Management in Engineering*, 36(1), 04019035.
- Kim, S., Chang, S., and Castro-Lacouture, D. (2020). "Dynamic modeling for analyzing impacts of skilled labor shortage on construction project management." *Journal of Management in Engineering*, 36(1), 04019035.

- Kim, S., and Nussbaum, M. A. (2013). "Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks." *Ergonomics*, 56(2), 314-326.
- Koo, T. K., and Li, M. Y. (2016). "A guideline of selecting and reporting intraclass correlation coefficients for reliability research." *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Koskimaki, H., Huikari, V., Siirtola, P., Laurinen, P., and Roning, J. "Activity recognition using a wrist-worn inertial measurement unit: a case study for industrial assembly lines." *Proc., 2009 17th Mediterranean Conference on Control and Automation*, IEEE, 401-405.
- Koskimaki, H., and Siirtola, P. (2016). "Accelerometer vs. electromyogram in activity recognition." *Advances in Distributed Computing and Artificial Intelligence Journal*.
- Koskimäki, H., Siirtola, P., and Röning, J. "Myogym: introducing an open gym data set for activity recognition collected using myo armband." *Proc., Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 537-546.
- Kubota, A., Iqbal, T., Shah, J. A., and Riek, L. D. "Activity recognition in manufacturing: The roles of motion capture and semg+ inertial wearables in detecting fine vs. gross motion." *Proc., 2019 International Conference on Robotics and Automation (ICRA)*, IEEE, 6533-6539.
- Kwapisz, J. R., Weiss, G. M., and Moore, S. A. (2011). "Activity recognition using cell phone accelerometers." *ACM SIGKDD Explorations Newsletter*, 12(2), 74-82.
- Ładyga, M., and Faff, J. (2005). "Assessment of the accuracy of prediction of the maximal oxygen uptake based on submaximal exercises in the former elite rowers and paddlers." *Biology Sport*, 22(2), 125-134.
- Lara, O. D., and Labrador, M. A. (2012). "A survey on human activity recognition using wearable sensors." *IEEE Communications Surveys & Tutorials*, 15(3), 1192-1209.
- Lee, H., Ahn, C. R., Choi, N., Kim, T., and Lee, H. (2019). "The effects of housing environments on the performance of activity-recognition systems using Wi-Fi channel state information: An exploratory study." *Sensors*, 19(5), 983.
- Lee, K. A., Hicks, G., and Nino-Murcia, G. (1991). "Validity and reliability of a scale to assess fatigue." *Psychiatry Research*, 36(3), 291-298.
- Lim, T.-K., Park, S.-M., Lee, H.-C., and Lee, D.-E. (2016). "Artificial neural network-based slip-trip classifier using smart sensor for construction workplace." *Journal of Construction Engineering and Management*, 142(2), 04015065.

- Lim, T. K., Park, S. M., Lee, H. C., and Lee, D. E. (2015). "Artificial neural network-based slip-trip classifier using smart sensor for construction workplace." *Journal of Construction Engineering and Management*, 142(2), 04015065.
- Lipton, Z. C., Kale, D. C., Elkan, C., and Wetzel, R. (2015). "Learning to diagnose with LSTM recurrent neural networks." *arXiv preprint arXiv:1511.03677*.
- Lu, K., Yang, L., Seoane, F., Abtahi, F., Forsman, M., and Lindecrantz, K. (2018). "Fusion of heart rate, respiration and motion measurements from a wearable sensor system to enhance energy expenditure estimation." *Sensors*, 18(9), 3092.
- Lu, L., Megahed, F. M., Sesek, R. F., and Cavuoto, L. A. (2017). "A survey of the prevalence of fatigue, its precursors and individual coping mechanisms among US manufacturing workers." *Applied Ergonomics*, 65, 139-151.
- Luo, H., Xiong, C., Fang, W., Love, P. E., Zhang, B., and Ouyang, X. (2018). "Convolutional neural networks: Computer vision-based workforce activity assessment in construction." *Automation in Construction*, 94, 282-289.
- Maman, Z. S., Yazdi, M. A. A., Cavuoto, L. A., and Megahed, F. M. (2017). "A data-driven approach to modeling physical fatigue in the workplace using wearable sensors." *Applied Ergonomics*, 65, 515-529.
- McComb, J. J. R., Roh, D., and Williams, J. S. (2006). "Explanatory variance in maximal oxygen uptake." *Journal of Sports Science & Medicine*, 5(2), 296.
- McGraw, K. O., and Wong, S. P. (1996). "Forming inferences about some intraclass correlation coefficients." *Psychological Methods*, 1(1), 30.
- Medsker, L. R., and Jain, L. (2001). "Recurrent neural networks." *Design and Applications*, 5.
- Mendez, I., Hansen, B. W., Grabow, C. M., Smedegaard, E. J. L., Skogberg, N. B., Uth, X. J., Bruhn, A., Geng, B., and Kamavuako, E. N. (2017). "Evaluation of the Myo armband for the classification of hand motions." *2017 International Conference on Rehabilitation Robotics (ICORR)*, IEEE, 1211-1214.
- Meyer, B. M., Tulipani, L. J., Gurchiek, R. D., Allen, D. A., Adamowicz, L., Larie, D., Solomon, A. J., Cheney, N., and McGinnis, R. (2020). "Wearables and deep learning classify fall risk from gait in multiple sclerosis." *IEEE Journal of Biomedical and Health Informatics*.
- Michielsen, H. J., De Vries, J., Van Heck, G. L., Van de Vijver, F. J., and Sijtsma, K. (2004). "Examination of the dimensionality of fatigue." *European Journal of Psychological Assessment*, 20(1), 39-48.

- Mital, A., Foononi-Fard, H., and Brown, M. L. (1994). "Physical fatigue in high and very high frequency manual materials handling: perceived exertion and physiological indicators." *Human Factors*, 36(2), 219-231.
- Mitropoulos, P., and Memarian, B. (2012). "Task demands in masonry work: Sources, performance implications, and management strategies." *Journal of Construction Engineering and Management*, 139(5), 581-590.
- Munguia Tapia, E. (2008). "Using machine learning for real-time activity recognition and estimation of energy expenditure." Massachusetts Institute of Technology.
- Muppalla, V., Suraj, N. S. S. K., Reddy, V. Y. S., and Suman, D. (2017). "Performance evaluation of different denoising techniques for physiological signals." *2017 14th IEEE India Council International Conference (INDICON)*, IEEE, 1-6.
- Murcia, H., and Triana, J. "A personal activity recognition system based on smart devices." *Proc., Workshop on Engineering Applications*, Springer, 487-499.
- Nath, N. D., Akhavian, R., and Behzadan, A. H. (2017). "Ergonomic analysis of construction worker's body postures using wearable mobile sensors." *Applied Ergonomics*, 62, 107-117.
- Nath, N. D., Chaspari, T., and Behzadan, A. H. (2018). "Automated ergonomic risk monitoring using body-mounted sensors and machine learning." *Advanced Engineering Informatics*, 38, 514-526.
- Nimbarte, A. D., Aghazadeh, F., Ikuma, L. H., and Harvey, C. M. (2010). "Neck disorders among construction workers: understanding the physical loads on the cervical spine during static lifting tasks." *Industrial Health*, 48(2), 145-153.
- Nymoen, K., Haugen, M. R., and Jensenius, A. R. (2015). "Mumyo—evaluating and exploring the myo armband for musical interaction." *International Conference on New Interfaces For Musical Expression*, 215-218.
- Nymoen, K., Voldskund, A., Skogstad, S. A. v. D., Jensenius, A. R., and Tørresen, J. (2012). "Comparing motion data from an iPod touch to a high-end optical infrared marker-based motion capture system." *International Conference on New Interfaces for Musical Expression*, 88-91.
- Ordonez, F. J., and Roggen, D. (2016). "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition." *Sensors (Basel)*, 16(1), 115.
- Ordóñez, F. J., and Roggen, D. (2016). "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition." *Sensors*, 16(1), 115.

OSHA (2018). "OSHA, Commonly Used Statistics." <<https://www.osha.gov/data/commonstats>>. (Dec 27, 2019).

OSHA (2020). "Scaffolding eTool." <<https://www.osha.gov/SLTC/etools/scaffolding/index.html>>. (July 6th, 2020).

Phinyomark, A., N Khushaba, R., and Scheme, E. (2018). "Feature extraction and selection for myoelectric control based on wearable EMG sensors." *Sensors*, 18(5), 1615.

Porta, M., Kim, S., Pau, M., and Nussbaum, M. A. (2021). "Classifying diverse manual material handling tasks using a single wearable sensor." *Applied Ergonomics*, 93, 103386.

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

Ravi, N., Dandekar, N., Mysore, P., and Littman, M. L. "Activity recognition from accelerometer data." *Proc., American Association for Artificial Intelligence*, 1541-1546.

Raynor, B. (2020). "U.S. Construction Industry - Statistics & Facts." <<https://www.statista.com/topics/974/construction/>>. (January 12, 2021).

Regterschot, G. R. H., Zhang, W., Baldus, H., Stevens, M., and Zijlstra, W. (2014). "Test-retest reliability of sensor-based sit-to-stand measures in young and older adults." *Gait & Posture*, 40(1), 220-224.

Robergs, R. A., and Landwehr, R. (2002). "The surprising history of the" HRmax = 220-age" equation." *Journal of Exercise Physiology Online*, 5(2), 1-10.

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

Romain, M., Watson, P., Hasegawa, H., Roelands, B., and Piacentini, M. (2006). "Central fatigue: the serotonin hypothesis and beyond." *Sports Medicine*, 36(10), 881-909.

Rubio-Romero, J. C., Gámez, M. C. R., and Carrillo-Castrillo, J. A. (2013). "Analysis of the safety conditions of scaffolding on construction sites." *Safety Science*, 55, 160-164.

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

- Ryu, J., Seo, J., Jebelli, H., and Lee, S. (2019). "Automated action recognition using an accelerometer-embedded wristband-type activity tracker." *Journal of Construction Engineering and Management*, 145(1), 04018114.
- Sadeghniiat-Haghghi, K., and Yazdi, Z. (2015). "Fatigue management in the workplace." *Industrial Psychiatry Journal*, 24(1), 12.
- Sathiyanarayanan, M., and Rajan, S. (2016). "MYO armband for physiotherapy healthcare: A case study using gesture recognition application." *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*, IEEE, 1-6.
- Schuster, M., and Paliwal, K. K. (1997). "Bidirectional recurrent neural networks." *IEEE Transactions on Signal Processing*, 45(11), 2673-2681.
- Seel, T., Raisch, J., and Schauer, T. (2014). "IMU-based joint angle measurement for gait analysis." *Sensors (Basel)*, 14(4), 6891-6909.
- Šerbetar, I. (2015). "Establishing some measures of absolute and relative reliability of a motor test." *Croatian Journal of Education: Hrvatski časopis za odgoj i obrazovanje*, 17(Sp. Ed. 1), 37-48.
- Shandhi, M. M. H., Bartlett, W. H., Heller, J., Etemadi, M., Young, A., Ploetz, T., and Inan, O. (2020). "Estimation of Instantaneous Oxygen Uptake during Exercise and Daily Activities using a Wearable Cardio-Electromechanical and Environmental Sensor." *IEEE Journal of Biomedical and Health Informatics*.
- Shanker, M., Hu, M. Y., and Hung, M. S. (1996). "Effect of data standardization on neural network training." *Omega*, 24(4), 385-397.
- Sharp, M., Harman, E., Vogel, J., Knapik, J., and Legg, S. (1988). "Maximal aerobic capacity for repetitive lifting: comparison with three standard exercise testing modes." *European Journal of Applied Physiology and Occupational Physiology*, 57(6), 753-760.
- Sherafat, B., Ahn, C. R., Akhavian, R., Behzadan, A. H., Golparvar-Fard, M., Kim, H., Lee, Y.-C., Rashidi, A., and Azar, E. R. (2020). "Automated methods for activity recognition of construction workers and equipment: state-of-the-art review." *Journal of Construction Engineering and Management*, 146(6), 03120002.
- Shoaib, M., Bosch, S., Incel, O., Scholten, H., and Havinga, P. (2016). "Complex human activity recognition using smartphone and wrist-worn motion sensors." *Sensors*, 16(4), 426.
- Skogstad, S. A. v. D., Nymoen, K., and Høvin, M. E. (2011). "Comparing inertial and optical mocap technologies for synthesis control." *Proceedings of SMC 2011 8th Sound and Music Computing Conference "Creativity rethinks science"*, 421-426.

- Snook, S., and Irvine, C. (1968). "Maximum frequency of lift acceptable to male industrial workers." *American Industrial Hygiene Association Journal*, 29(6), 531-536.
- Søgaard, K., Christensen, H., Jensen, B., Finsen, L., and Sjøgaard, G. (1996). "Motor control and kinetics during low level concentric and eccentric contractions in man." *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, 101(5), 453-460.
- St-Amant, Y., Rancourt, D., and Clancy, E. A. (1998). "Influence of smoothing window length on electromyogram amplitude estimates." *IEEE Transactions on Biomedical Engineering*, 45(6), 795-799.
- Stančin, S., and Tomažič, S. (2011). "Angle estimation of simultaneous orthogonal rotations from 3D gyroscope measurements." *Sensors*, 11(9), 8536-8549.
- Stanton, N., Salmon, P. M., and Rafferty, L. A. (2013). *Human factors methods: a practical guide for engineering and design*, CRC Press.
- Sun, G., Jiang, C., Wang, X., and Yang, X. (2020). "Short-term building load forecast based on a data-mining feature selection and LSTM-RNN method." *IEEJ Transactions on Electrical and Electronic Engineering*, 15(7), 1002-1010.
- Tao, W., Lai, Z.-H., Leu, M. C., and Yin, Z. (2018). "Worker activity recognition in smart manufacturing using IMU and sEMG signals with convolutional neural networks." *Procedia Manufacturing*, 26, 1159-1166.
- Tixier, A. J.-P., Hallowell, M. R., Rajagopalan, B., and Bowman, D. (2016). "Application of machine learning to construction injury prediction." *Automation in Construction*, 69, 102-114.
- Tsai, M.-K. (2014). "Automatically determining accidental falls in field surveying: A case study of integrating accelerometer determination and image recognition." *Safety Science*, 66, 19-26.
- Valero, E., Sivanathan, A., Bosché, F., and Abdel-Wahab, M. (2017). "Analysis of construction trade worker body motions using a wearable and wireless motion sensor network." *Automation in Construction*, 83, 48-55.
- Van Dieën, J. H., Toussaint, H. M., Maurice, C., and Mientjes, M. (1996). "Fatigue-related changes in the coordination of lifting and their effect on low back load." *Journal of Motor Behavior*, 28(4), 304-314.
- Vandersmissen, G., Verhoogen, R., Van Cauwenbergh, A., and Godderis, L. (2014). "Determinants of maximal oxygen uptake (VO<sub>2</sub> max) in fire fighter testing." *Applied Ergonomics*, 45(4), 1063-1066.

Vepakomma, P., De, D., Das, S. K., and Bhansali, S. (2015). "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, 1-6.

Verschuren, O., Maltais, D. B., and Takken, T. (2011). "The 220-age equation does not predict maximum heart rate in children and adolescents." *Developmental Medicine & Child Neurology*, 53(9), 861-864.

Vigliensoni, G., and Wanderley, M. M. (2012). "A Quantitative Comparison of Position Trackers for the Development of a Touch-less Musical Interface." *The International Conference on New Interfaces for Musical Expression*.

Visconti, P., Gaetani, F., Zappatore, G., and Primiceri, P. (2018). "Technical features and functionalities of Myo armband: an overview on related literature and advanced applications of myoelectric armbands mainly focused on arm prostheses." *International Journal on Smart Sensing and Intelligent Systems*, 11(1), 1-25.

Visconti, P., Gaetani, F., Zappatore, G., and Primiceri, P. (2018). "Technical features and functionalities of Myo armband: An overview on related literature and advanced applications of myoelectric armbands mainly focused on arm prostheses." *Int. J. Smart Sens. Intell. Syst*, 11(1), 1-25.

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

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

Wang, T. (2019). "U.S. Construction Industry - Statistics & Facts."  
<<https://www.statista.com/topics/974/construction/>>. (July 5th, 2020).

Wang, X., Dong, X. S., Choi, S. D., and Dement, J. (2017). "Work-related musculoskeletal disorders among construction workers in the United States from 1992 to 2014." *Occupational and Environmental Medicine*, 74(5), 374-380.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.

Wong, D. P.-l., Chung, J. W.-y., Chan, A. P.-c., Wong, F. K.-w., and Yi, W. (2014). "Comparing the physiological and perceptual responses of construction workers (bar benders and bar fixers) in a hot environment." *Applied Ergonomics*, 45(6), 1705-1711.

- Wu, W., Yang, H., Chew, D. A., Yang, S.-h., Gibb, A. G., and Li, Q. (2010). "Towards an autonomous real-time tracking system of near-miss accidents on construction sites." *Automation in Construction*, 19(2), 134-141.
- Xi, Y., Russell, J., Zhang, Q., Wang, Y. Y., Zhang, J., and Zhao, W. H. (2019). "Validity and Reliability of the Wristband Activity Monitor in Free-living Children Aged 10-17 Years." *Biomedical and Environmental Sciences*, 32(11), 812-822.
- Xu, Y., Yang, C., Liang, P., Zhao, L., and Li, Z. (2016). "Development of a hybrid motion capture method using MYO armband with application to teleoperation." *2016 IEEE International Conference on Mechatronics and Automation*, IEEE, 1179-1184.
- Yang, C. C., and Hsu, Y. L. (2010). "A review of accelerometry-based wearable motion detectors for physical activity monitoring." *Sensors (Basel)*, 10(8), 7772-7788.
- Yang, J. (2009). "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, 1-10.
- Yang, J., Shi, Z., and Wu, Z. (2016). "Vision-based action recognition of construction workers using dense trajectories." *Advanced Engineering Informatics*, 30(3), 327-336.
- Yang, K., Ahn, C. R., Vuran, M. C., and Aria, S. S. (2016). "Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit." *Automation in Construction*, 68, 194-202.
- Yang, K., Aria, S., Ahn, C. R., and Stentz, T. L. (2014). "Automated detection of near-miss fall incidents in iron workers using inertial measurement units." *Construction Research Congress 2014*, 935-944.
- Yang, K., Jebelli, H., Ahn, C., and Vuran, M. (2015). "Threshold-based approach to detect near-miss falls of iron workers using inertial measurement units." *Computing in Civil Engineering*, 148-155.
- Yang, L., and Shami, A. (2020). "On hyperparameter optimization of machine learning algorithms: theory and practice." *arXiv:2007.15745*, 415, 295-316.
- Yang, Z., Yuan, Y., Zhang, M., Zhao, X., and Tian, B. (2019). "Assessment of construction workers' labor intensity based on wearable smartphone system." *Journal of Construction Engineering and Management*, 145(7), 04019039.
- Yi, W., Chan, A. P., Wang, X., and Wang, J. (2016). "Development of an early-warning system for site work in hot and humid environments: a case study." *Automation in Construction*, 62, 101-113.

- Yung, M., and Wells, R. (2016). "How do we measure neuromuscular fatigue at the workplace? The Toronto CRE-MSD workshop." <<https://uwaterloo.ca/centre-of-research-expertise-for-the-prevention-of-musculoskeletal-disorders/resources/position-papers/how-do-we-measure-neuromuscular-fatigue-workplace-toronto>>. (Dec 31st, 2019).
- Zhang, L., Diraneyya, M. M., Ryu, J., Haas, C. T., and Abdel-Rahman, E. (2018). "Assessment of jerk as a method of physical fatigue detection." *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers Digital Collection.
- Zhang, M., Chen, S., Zhao, X., and Yang, Z. (2018). "Research on construction workers' activity recognition based on smartphone." *Sensors*, 18(8), 2667.
- Zhang, M., Murphy, L., Fang, D., and Caban-Martinez, A. J. (2015). "Influence of fatigue on construction workers' physical and cognitive function." *Occupational Medicine*, 65(3), 245-250.
- Zhang, M., Sparer, E. H., Murphy, L. A., Dennerlein, J. T., Fang, D., Katz, J. N., and Caban-Martinez, A. J. (2015). "Development and validation of a fatigue assessment scale for US construction workers." *American Journal of Industrial Medicine*, 58(2), 220-228.
- Zhang, T., Lee, Y.-C., Scarpiniti, M., and Uncini, A. "A supervised machine learning-based sound identification for construction activity monitoring and performance evaluation." *Proc., Construction Research Congress*, 358-366.
- Zhang, W., Regterschot, G. R. H., Schaabova, H., Baldus, H., and Zijlstra, W. (2014). "Test-retest reliability of a pendant-worn sensor device in measuring chair rise performance in older persons." *Sensors*, 14(5), 8705-8717.
- Zhang, Z. (2018). "Improved adam optimizer for deep neural networks." *2018 IEEE/ACM 26th International Symposium on Quality of Service*, IEEE, 1-2.
- Zignoli, A., Fornasiero, A., Ragni, M., Pellegrini, B., Schena, F., Biral, F., and Laursen, P. B. (2020). "Estimating an individual's oxygen uptake during cycling exercise with a recurrent neural network trained from easy-to-obtain inputs: a pilot study." *PloS One*, 15(3), e0229466.

## **VITA**

Srikanth Bangaru was born in Mahabubnagar, Telangana, India. He earned a bachelor's degree in Civil Engineering from the National Institute of Technology, Warangal, India, in 2015 and a master's degree in Civil Engineering from Oklahoma State University, Stillwater, OK, USA, in 2017. Later, he moved to Louisiana State University (LSU) to pursue a master's degree in Construction Management and a doctoral degree in Engineering Science. Srikanth earned his master's degree in Construction Management from LSU in 2020. During his doctoral study at LSU, he worked as a graduate research and teaching assistant. As a graduate assistant, he worked on various research projects funded by state and federal agencies. His research findings were published in multiple journals and conference proceedings. He was a teaching assistant and lab instructor for construction courses such as construction scheduling and cost control and advanced construction productivity. Also, he was the instructor for building information modeling and construction materials and methods II courses. For his outstanding teaching abilities and service to students, he was awarded LSU Alumni Association Teaching Assistant Award - 2021. Apart from research and teaching, he served as a president for student organizations such as Construction Management Graduate Student Association, Indian Students' Association, and Tiger Cricket Club at LSU. He was honored with LSU Graduate Student Leader of the Year – 2020 for his significant contribution to the quality of student life at LSU.