Deep learning-based construction equipment operators' mental fatigue classification using wearable EEG sensor data

Authors

Imran Mehmood[1*], Ph.D. Student, (Email: imran.mehmood@connect.polyu.hk)

Heng Li[1], Ph.D., (Email: heng.li@polyu.edu.hk)

Yazan Qarout[2], Ph.D., (Email: yazan.qarout@gmail.com)

Waleed Umer[3], Ph.D., (Email; waleed.umer@northumbria.ac.uk)

Shahnawaz Anwer[1], Ph.D., (Email; shah-nawaz.anwer@polyu.edu.hk)

Haitao Wu[1], Ph.D. Student, (Email: haitao.wu@connect.polyu.hk)

Mudasir Hussain[4], Ph.D. Student, (Email: mudasir.hussain@connect.polyu.hk)

Maxwell Fordjour Antwi-Afari[5**], Ph.D., (Email; m.antwiafari@aston.ac.uk)

Affiliation

1. Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR
2. The Manufacturing Technology Centre Ltd, Ansty Park, Coventry, CV7 9JU, United Kingdom.
3. Department of Architecture and Built Environment, Northumbria University, NE1 8ST, Newcastle upon Tyne, United Kingdom
4. Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR

5. Department of Civil Engineering, College of Engineering and Physical Sciences, Aston University, Birmingham, B4 7ET, United Kingdom

Corresponding Author*

Imran Mehmood[1*], Ph.D. Student, (Email: imran.mehmood@connect.polyu.hk)

Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR

Corresponding Author**

Maxwell Fordjour Antwi-Afari[5**], Ph.D., (Email; m.antwiafari@aston.ac.uk)

Department of Civil Engineering, College of Engineering and Physical Sciences, Aston University, Birmingham, B4 7ET, United Kingdom

Abstract

Operator attention failure due to mental fatigue during extended equipment operations is a common cause of equipment-related accidents that result in catastrophic injuries and fatalities. As a result, tracking operators' mental fatigue is critical to reducing equipment-related accidents on construction sites. Previously, several strategies aimed at recognizing mental fatigue with adequate accuracy, such as machine learning utilizing EEG-based wearable sensing systems, have been proposed. However, the ability to track operators' mental fatigue for its implementation on an actual construction site is still an issue. For instance, the mobility and systemic instability of EEG sensors necessitate their application in laboratory settings rather than on actual construction sites. Furthermore, while the machine learning classifiers achieved acceptable accuracy, their input is limited to manually developed EEG features, which may compromise the models' performance on real construction sites. Accordingly, the current research proposes the viability of a construction site strategy that uses flexible headband-based sensors for acquiring raw EEG data and deep learning networks to recognize operators' mental fatigue. To serve this purpose, a one-hour excavator operation by fifteen operators was conducted on a construction site. The NASA-TLX score was used as the ground truth of mental fatigue, and brain activity patterns were recorded using a wearable EEG sensor. The raw EEG data was then used to develop deep learning-based classification models. Finally, the performance of deep learning models, i.e., long short-term memory, bidirectional LSTM, and one-dimensional convolutional networks, was investigated using accuracy, precision, recall, specificity, and an F1-score. The findings indicate that the Bi-LSTM model outperforms the other deep learning models with a high accuracy of 99.941% and F1-score between 99.917% and 99.993%. These findings demonstrate the feasibility of

applying the Bi-LSTM model and contribute to wearable sensor-based mental fatigue recognition and classification, thus enhancing on-site health and safety operations.

Keywords: mental fatigue, deep learning networks, electroencephalography, construction equipment operators, construction safety

1 Introduction

The construction industry has significantly contributed to the development of countries, with over 350 million workers around the world (Birhane et al., 2022). Besides, it is also regarded as one of the most hazardous industries (Hinze and Teizer, 2011) and is distinguished by its complexity, uncertainty, and disorderliness. In particular, construction projects are carried out in an uncertain and unpredictable environment (Choi et al., 2020, Laitinen and Päivärinta, 2010). Almost every day, the activities of workers, materials, and construction equipment create an environment that is unique and dynamic (Zhu et al., 2016). Such an environment makes construction workers more vulnerable to accidents compared to other occupational industries (Albert et al., 2020). As a result, accidents at work happen often in the construction industry (Koc and Gurgun, 2022). These accidents not only cause serious injuries and deaths but also halt the flow of work at sites (Sarkar et al., 2020). Among these, construction equipment-related accidents constitute a significant proportion and are unarguably one of the most frequent types of construction accidents (Li et al., 2021, Li et al., 2017b). For instance, in the United States construction industry, construction equipment is a major cause of work-related fatalities and injuries (Vahdatikhaki et al., 2019). According to Hinze and Teizer (2011), one in four construction industry fatalities is caused by accidents involving equipment. Therefore, reducing the occurrence of equipment-related incidents at construction sites is crucial to reducing fatalities and injuries in the construction industry.

One of the leading causes of construction equipment-related accidents is human behavior (Ma et al., 2021), which is highly influenced by fatigue states (Behrens et al., 2023, Yang et al., 2021, Molan and Molan, 2021, Bucsuházy et al., 2020). According to Bai and Qian (2021), over 65% of all accidents can be attributed to human error.

According to Brown (1994), fatigue is "a state of energy depletion" that leads to "difficulties in maintaining task-directed efforts and a loss of attentiveness." Fatigue is a risk to both workers' health and safety (Williamson et al., 2011), resulting in reduced energy levels and increased fatigue during and after work (Frone and Tidwell, 2015). There are two primary types of fatigue: physical and mental (Villani et al., 2022). Physical fatigue is the feeling of tiredness, weakness, or lack of energy that results from physical activity, such as exercise or manual labor (Alghadir and Anwer, 2015). It poses a risk for construction accidents and occupational injuries due to poor worker judgment in dynamic environments (Wu et al., 2018, Umer et al., 2018, Adane et al., 2013, Chan, 2011). Similarly, mental fatigue is the outcome of the brain's engagement in intellectually demanding tasks for an extended period of time and can lead to decreased behavioral and cognitive performance (Borragán et al., 2016, Boksem and Tops, 2008, van der Linden et al., 2003). In addition, mental fatigue is significant in occupations that demand workers to be cognitively active and vigilant, such as long-distance driving (Hu and Lodewijks, 2020), airport baggage screening (Chavaillaz et al., 2019), and nurses working prolonged shifts (Farag et al., 2022). Both physical and mental fatigue can have negative impacts on performance, safety, and well-being (Chen and Hsu, 2020). The fundamental distinction between physical and mental fatigue is the source of fatigue. Nonetheless, the symptoms of physical and mental fatigue can be similar, such as decreased energy, decreased motivation, and impaired performance (Behrens et al., 2023, Van Cutsem et al., 2017).

In the construction industry, construction equipment is utilized to execute several challenging tasks, such as excavation, material lifting, and compaction. These tasks are cognitively demanding and require the equipment operators to maintain a significant level of sustained effort and alertness (Li et al., 2020b). According to Wagstaff and Sigstad Lie (2011), such protracted construction operations and attentive tasks lead to mental fatigue among construction equipment operators, which results in an inability to maintain equipment operations requiring sustained attention. Their judgment and focus are impaired (Das et al., 2020), resulting in a decrease in productivity and performance (Masullo et al., 2020). This renders equipment operators more susceptible to equipment-related incidents, resulting in injuries and fatalities on the site. Therefore, preventing the inattention of construction equipment operators is crucial for improving site safety (Han et al., 2019). As a result, it is imperative that the mental fatigue of construction equipment operators be monitored constantly, so that safety personnel can respond immediately if necessary.

To prevent accidents and ensure the safety and health of construction equipment operators, proactive safety management has become a critical component of construction safety (Hallowell et al., 2013, Carbonari et al., 2011). Previously, several studies were conducted to monitor and analyze mental fatigue on construction sites, either by using psychological or physiological techniques. Initially, the operators' mental fatigue was subjectively assessed using questionnaires, where NASA-TLX was the most widely utilized assessment tool (Li et al., 2019b). However, this assessment is intrusive in nature and time-consuming (Umer et al., 2020). Further, it lacks accuracy as it is prone to bias (Han et al., 2019). As a result, researchers were motivated to develop a more objective assessment of mental fatigue. Hence, wearable sensors have gained significant attention from researchers in recent years owing to technological developments that allow more objective monitoring of mental fatigue on construction sites. As a result, research efforts were conducted to assess mental fatigue by evaluating the workers' physiological signals, e.g., electrodermal activity (EDA) (Lee et al., 2021, Choi et al., 2019); electroencephalogram (EEG) (Ke et al., 2021b, Xing et al., 2020); electrocardiograph (ECG) (Umer, 2022); and eye-tracking (Li et al., 2020b). According to the literature, there is a substantial association between these signals and workers' mental states, and they can be employed to reliably identify construction workers' fatigue. Recently, geometric measurements of facial features have also been used to identify mental fatigue during on-site construction operations (Mehmood et al., 2022). However, among these technologies, EEG has emerged as one of the fastest-growing ones that has attracted significant attention from researchers for assessing workers' cognitive and mental states under dynamic workplace conditions (Zhang et al., 2019b). As a result, the overarching purpose has been to enhance safety performance on construction projects in order to ensure that construction sites are safer for workers.

EEG is an electrophysiological monitoring system that records the electrical activities generated by cortical neurons (Sanei and Chambers, 2013). It is recognized as a potent technique in the field of construction research since it detects brain activity rapidly, cost-effectively, with a high temporal resolution, and in a portable manner (Saedi et al., 2022). There has been extensive research into the construction industry's use of EEG data gathered from wearable devices for the purpose of analyzing distinct mental states among construction workers, such as fatigue (Tehrani et al., 2021, Xing et al., 2020, Li et al., 2019a), stress (Lee and Lee, 2022, Jebelli et al., 2019a), distraction (Ke et al., 2021b, Ke et al., 2021a), workload (Chen et al., 2017), vigilance (Wang et al., 2019), emotion (Xing et al., 2019, Hwang et al., 2018), and hazard identification (Wang et al., 2022, Liao et al., 2022, Jeon and Cai, 2022). In the aforementioned studies, different mental states were analyzed and computed using either statistical methods or machine learning. Several statistical significance tests, including the Kruskal-Wallis test, the analysis of variance (ANOVA), the Mann-Whitney U test, the Wilcoxon signed-rank test, the Spearman rank-order correlation test, and the paired sample t-test, have been used to draw conclusions between experimental and control groups in EEG-based studies to compute the cognitive status of construction workers (Ke et al., 2021b, Chae et al., 2021, Xing et al., 2020). The purpose of these analyses was to ascertain whether or not there was a statistically significant relationship between construction workers' EEG signals and their performance on the task pertaining to their mental states. Even if these assessments fared well, they have significant shortcomings when it comes to drawing reasonable and trustworthy inferences about the mental wellbeing of construction workers. Cheng et al. (2022) reported these limitations: that conventional statistical approaches are inadequate for modeling complex mapping, whereas the relationships between EEG patterns and cognitive ability are rather complex. This is mostly because the sample data used to test these statistical approaches is typically subject to stringent requirements. Therefore, the researchers turned to machine learning methods, which offer a high degree of adaptability (Rajula et al., 2020).

Machine learning may be utilized to compute the mental state of construction workers using their EEG signals. Various machine learning models have been developed by researchers to estimate the mental state of construction workers. For instance, using a supervised learning algorithm, Jebelli et al. (2019a) proposed a framework that can identify stress levels among construction workers and achieved an accuracy of 84.5%. Aryal et al. (2017) predicted the fatigue of construction workers with an accuracy of 82% using a boosted tree classifier. Furthermore, Hwang et al. (2018) demonstrated that two aspects of construction workers' emotional states (arousal and valence) could be measured and quantified using EEG signals as they performed various construction-related tasks. To identify mental stress in construction workers, Jebelli et al. (2018a) compared the efficacy of K-nearest neighbors (KNNs), support vector machines (SVM), and gaussian discriminant analysis (GDA). When compared to other methods, they discovered the highest accuracy of 80.32% with SVM. In another study, Ke et al. (2021b) proposed a distraction monitoring method for construction workers, and validation was done using SVM classifier. Similarly, Jeon and Cai (2022) explored multi-class classification for hazard identification in construction workers using EEG signals in a virtual reality environment and achieved 82.3% accuracy. Selecting a suitable model and optimizing its hyperparameters are key phases in machine learning for achieving optimal results. Regardless of these advances, robust and accurate detection of construction workers' cognitive performance by EEG remains a challenge. The EEG-based studies conducted in the construction industry using machine learning were conducted offline by first measuring the data and then downloading the raw electroencephalography data for analysis (Cheng et al., 2022). In such a case, the model development may be suitable for implementation on real-time monitoring of the mental states of construction workers while they are facing dynamic site conditions (Cesa-Bianchi and Orabona, 2021). Furthermore, it is generally understood that EEG manifestations are very non-stationary and change over time within and between subjects (Thodoroff et al., 2016). Hence, recognizing overarching trends in EEG data is difficult since the signals are constantly changing (Zeng et al., 2018).

According to Türk and Özerdem (2021) and Li et al. (2020a), the ability of deep learning to analyze raw data and identify key features is its major strength. Deep learning approaches are actively applicable to different signal processing because they can learn the features from raw data and have cutting-edge performance and robust skills in creating trustworthy features in time-series data analysis (Rastgoo et al., 2019, Liu et al., 2017, Zheng et al., 2014). They have been utilized in several fields, including computer vision, natural language processing, and speech recognition (LeCun et al., 2015). Subsequently, construction-related research domains have recently shown a significant deal of interest in deep learning networks due to their outstanding performance in a variety of research areas, such as image classification (Yeşilmen and Tatar, 2022, Duan et al., 2022, Del Savio et al., 2022, Zhong et al., 2020, Yang et al., 2018), object identification and recognition (Wu et al., 2021, Fang et al., 2018b, Fang et al., 2018a), natural language processing (Wu et al., 2022, Moon et al., 2022, Ding et al., 2022, Zhong et al., 2020, Zhang et al., 2019a), and recognition of work-related risk factors (Zhao et al., 2022, Antwi-Afari et al., 2022, Zhao and Obonyo, 2021, Wang et al., 2021, Seo and Lee, 2021, Zhao and Obonyo, 2020, Yang et al., 2020, Lee et al., 2020, Kim and Cho, 2020, Yu et al., 2019, Zhang et al., 2018). Although EEG analysis and decoding of data with deep learning algorithms have become hot research topics in recent years, unfortunately, EEG-based classification of mental fatigue using deep learning approaches has not previously been investigated for construction equipment operators on real construction sites.

Therefore, the objective of the current research is to evaluate the feasibility of using deep learning techniques to classify construction equipment operators' mental fatigue using raw EEG data and has two major contributions. The present study represents the first attempt to acquire and analyze EEG data from construction equipment operators in real-world construction sites, thus demonstrating the feasibility and applicability of the proposed method for construction site settings. This approach enabled the authors to collect data in a natural environment, providing a more authentic and realistic context. Moreover, the study is likely to have higher external validity, which refers to the extent to which the findings of a study can be generalized. Many prior investigations on mental fatigue have been carried out in controlled laboratory settings with student participants, as exemplified by studies conducted by Li et al. (2020b) and Li et al. (2019b). However, such laboratory experiments may face challenges related to generalization and validity since they lack the dynamics and complexity of actual construction sites (Xing et al., 2020). Therefore, the current study collected EEG data from construction equipment operators during an on-site excavation operation to support the study's findings, resulting in more comprehensive, accurate, and realistic results.

Secondly, the current study evaluates the usefulness and performance of deep learning models in detecting and classifying mental fatigue in construction equipment operators using EEG sensor data. Hypothetically, these models are more suitable for time-dependent data such as EEG signals, as they account for temporal dependencies and trends that cannot be captured using traditional classification machine learning algorithms. To the best of the authors' knowledge, no previous research in the construction industry (Cheng et al., 2022) has demonstrated the innovative approach of using deep learning models and EEG signals for detecting and classifying mental fatigue in construction workers. This is attributed to the difficulty in collecting EEG data in the field due to various factors such as noise, motion artifacts, and safety concerns, as highlighted in previous studies (Ke et al., 2021b, Ahn et al., 2019). Furthermore, the limited availability of large EEG datasets in the construction industry, as observed in studies by Wang et al. (2019) and Jebelli et al. (2018a), may constrain the training and validating of deep learning models. These challenges have hindered the application of deep learning models in the construction industry. To overcome these challenges, the current study collected an EEG dataset using a four-channel EEG sensor and recorded one hour of EEG data from each equipment operator. This resulted in more than 18 million data points for the entire experiment, enabling the effective application of deep learning models. This gap was also filled by the current study.

There is a plethora of deep learning architectures to choose from in the literature; nevertheless, choosing the right one is crucial for EEG data processing. Recent studies by Nakagome et al. (2022), Roy et al. (2019), and Craik et al. (2019) have examined the latest trends in EEG research and identified that convolutional neural networks (CNN) and recurrent neural networks (RNN) are gaining popularity for processing EEG data. According to Nakagome et al. (2022), more than half of EEG studies used CNN or RNN, particularly with raw EEG data as input, to analyze EEG data end-to-end, eliminating the need for time-consuming feature extraction processes. Moreover, both these deep learning architectures have been effectively used in studies involving individuals exposed to external stimuli (Nakagome et al., 2022). Subsequently, this study employed and investigated the performance of three deep learning techniques, i.e., long short-term memory, bidirectional long short-term memory, and one-dimensional convolutional networks, for mental fatigue recognition in construction equipment operators. Therefore, the findings of this study are expected to provide a better understanding of the application of electroencephalography technology for mental fatigue detection in construction equipment operators in real construction scenarios based on field tests. Furthermore, using this approach, operators of construction equipment might have their mental fatigue continuously monitored without having to be observed or watched by a supervisor. Having said that, this study will also contribute to classifying construction equipment operators' mental fatigue using raw EEG data, without any human intervention for manual crafting of features. As a result, the suggested method has the potential to improve the standardization of safety management within the construction industry.

2 Methodology

Figure 1 shows an overview of the research process. It demonstrates the proposed method for detecting mental fatigue in construction equipment operators by analysing brain activity patterns acquired using an EEG device. The research process consists of four steps. In the first step, an experiment was conducted to acquire relevant data. A headband was mounted on the head of construction equipment operators to obtain EEG data, and data related to

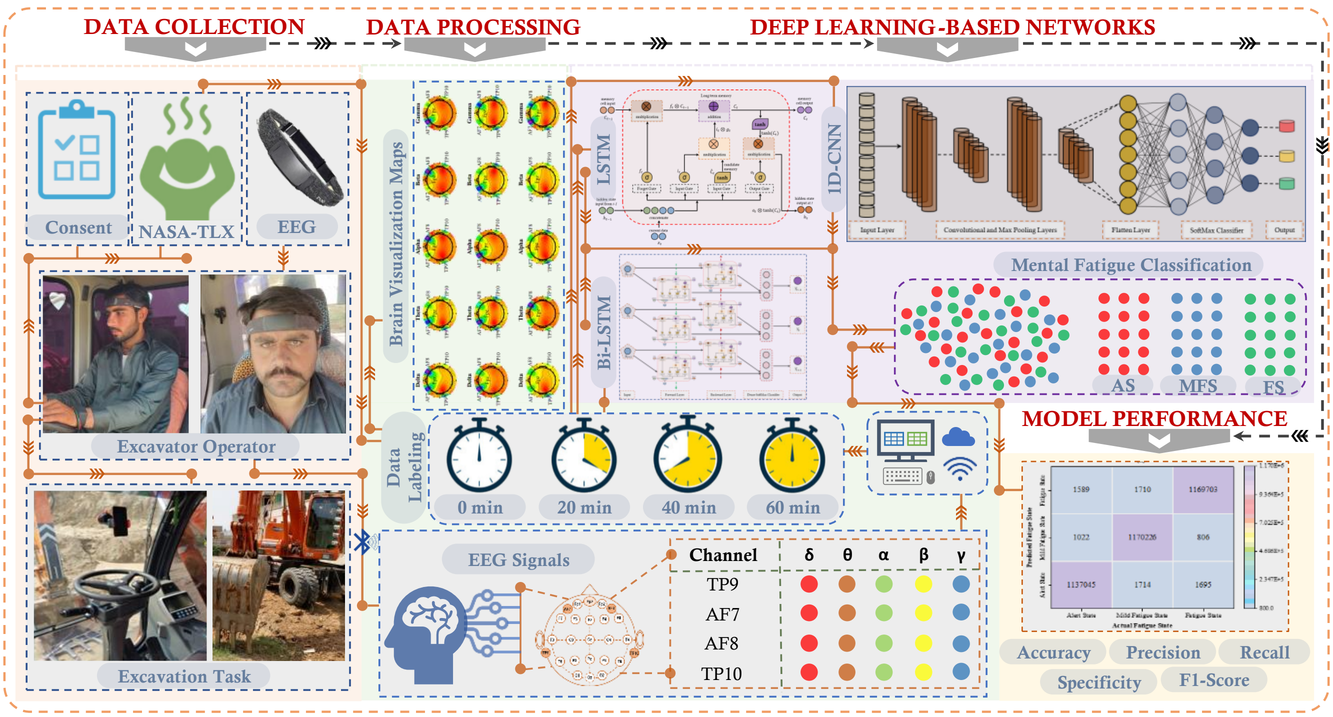

Figure 1: Overview of the research process

subjective feelings of mental fatigue was gathered using a questionnaire. In the second step, the EEG data was analysed and labelled into mental fatigue levels using subjective scores, artifacts were removed, and the data was down sampled. In the third step, detection of multiple mental fatigue levels in construction equipment operators was done based on deep learning techniques. Each deep learning model was trained using raw EEG data from an EEG device as input data. In the last step, the performance of each deep learning architecture was assessed using metrics.

2.1 Experiment design and data collection

2.1.1 Participants

Fifteen male construction equipment operators were voluntarily recruited to participate in the experiments. The operators' mean age was 33.07 years (SD = 3.95). Construction equipment operators were recruited and participated in this study because excavation operation tasks (ground excavation and moving the material from the pits to the transport vehicles) are repetitive, cognitively demanding, and often involve prolonged working hours, which require the operators to have a significant level of sustained attention (Li et al., 2020b). Furthermore, we determined the sample size of excavator operators to recruit for our research investigations based on sample sizes from previous studies. In earlier studies with similar purposes, 12 crane operators (Das et al., 2020), 12 excavator operators (Li et al., 2019b), 11 drivers (Ahn et al., 2016), 6 excavator operators (Li et al., 2020b), and 5 crane operators (Liu et al., 2021a) were recruited. Considering previous research in the literature, we decided that more than fifteen operators would be sufficient for our investigation and to justify our results. All the participants were excavator operators with prior experience in excavator operations at construction sites. All the excavator operators had slept at least eight hours the previous night and abstained from alcoholic drinks for at least 24 hours before experimentation. The operators were required to directly come for experiments on their designated days, and they were not involved in any other tasks or activities before the start of the experiment. In addition, we ensured that each operator remained fully engaged during the length of the task. The experimental protocol for data collection was reviewed and approved by the ethics subcommittee of the Hong Kong Polytechnic University (Reference Number: HSEARS20210927008) and conducted in accordance with the Declaration of Helsinki. In addition, written consent was obtained from each participant after a verbal explanation of the experimental procedures. Table 1 provides the demographic information of the construction equipment operators who participated in the study.

Table 1: Construction equipment operators' demographic information.

	Mean	**SD**	**Range (Min-Max)**
Age (Years)	33.07	3.95	13 (26-39)
Job Experience (Years)	7.27	2.58	8 (3-11)
Height (cm)	175.87	5.32	18 (166-184)
Weight (kg)	77.86	7.72	22 (68-90)
Body Mass Index (kg/m²)	25.16	2.06	7.48 (21.91-29.39)

2.1.2 Subjective Assessment

The NASA-TLX score was used for the labeling of construction equipment operators by assessing their individual subjective feelings of mental fatigue. The subjective assessment was used as a ground truth for construction equipment operators' mental fatigue levels. It has been widely used in various research investigations since its development, and its reliability and sensitivity have been tested in a consistent number of independent tests. The NASA-TLX is intended to measure operators' perceived workload in six dimensions: mental demand, physical demand, effort, own performance, temporal demand, and frustration. An overall NASA-TLX score was computed by adding the scores from each of the six dimensions of the scale. Overall NASA-TLX scores were used, with no weight applied to the individual categories. Adding the subscale scores to calculate an overall score is a common approach to simplifying the original scale (Hart, 2006, Byers, 1989). Additionally, several recent studies by Kaduk et al. (2021), Mehmood et al. (2022), Bitkina et al. (2021), Das et al. (2020), Li et al. (2019b), and Chen et al. (2017) reported that an increase in the NASA-TLX score over time during the same task can serve as a reliable indicator of mental fatigue. Moreover, in the construction industry, studies by Mehmood et al. (2022), Li et al. (2020b), and Li et al. (2019b) have employed the increase in the overall NASA-TLX score for the same task as a subjective indicator of mental fatigue. Likewise, in our study, an increase in the NASA-TLX score was the result of an increase in mental fatigue.

2.1.3 Electroencephalogram (EEG) Recording

To capture EEG signals, we employed the Muse headband, a flexible and user-friendly EEG recording device. Dry electrodes are located at AF7, AF8, TP9, and TP10 on a four-channel headband, with the FPz serving as the reference electrode. Electrodes are typically made of silver. The Muse headband has a sampling rate of 256 Hz, which makes it suitable for capturing EEG data. Through a Bluetooth connection, data was transmitted from the Muse headband to a smartphone. The construction equipment operators' EEG data was gathered on a smartphone using an app called Mind Monitor, then transferred to a computer for post-processing. The recorded EEG signals are subjected to artifact removal techniques to remove muscular artifacts, power line noise, and other artifacts. The Muse EEG headband has an on-board noise cancellation mechanism to filter out the noise based on the statistical properties of the data. The statistical properties used by the MUSE headband include amplitude, variance, and kurtosis. An EEG signal is considered clean if its statistical properties are below a predetermined threshold; otherwise, the signal is considered noisy and discarded (Arsalan et al., 2019). Although the on-board noise cancellation method has been successful in various fields, including research by Cannard et al. (2021) and Arsalan et al. (2019), construction site tasks are demanding and dynamic (Xing et al., 2020). It involves the continual body movement of workers to perform these tasks on construction sites (Mehmood et al., 2022). Hence, it is crucial to remove artifacts that cause noise in the acquired EEG data. Therefore, the acquired data underwent further preprocessing techniques for artifact removal, including the third-order one-dimensional median filter (Krauss et al., 1994) and the Savitzky-Golay (SG) filter (Orfanidis, 1995). The classical SG filter is designed based on the least-squares polynomial approximation phenomenon (Savitzky and Golay, 1964) and is used to remove inappropriate and large spikes in the EEG sensor data. The goal was to smooth the data while retaining the quality of the signal. To achieve this, we applied an overlapping window of 50% (Krauss et al., 1994). Previous studies in the construction industry by Aryal et al. (2017) have effectively used this noise cancellation method to smooth the data while retaining the quality of the acquired EEG data. Once the artifacts were removed from the data, it was down sampled to 128 Hz by selecting each second sample and effectively reducing the number of data points by half. It is a common method to reduce the dimensionality of the data (Frydenlund and Rudzicz, 2015). According to Roy et al. (2019), 72% of the studies employing EEG sensors have used the down sampling technique to preprocess their EEG data. In our investigation, doing down sampling did not affect the data model's predictive power, yet it improved the training time of the models significantly. Figure 2 demonstrates the electrode positioning system on the scalp of construction equipment operators as well as the various views of the EEG device used in the study.

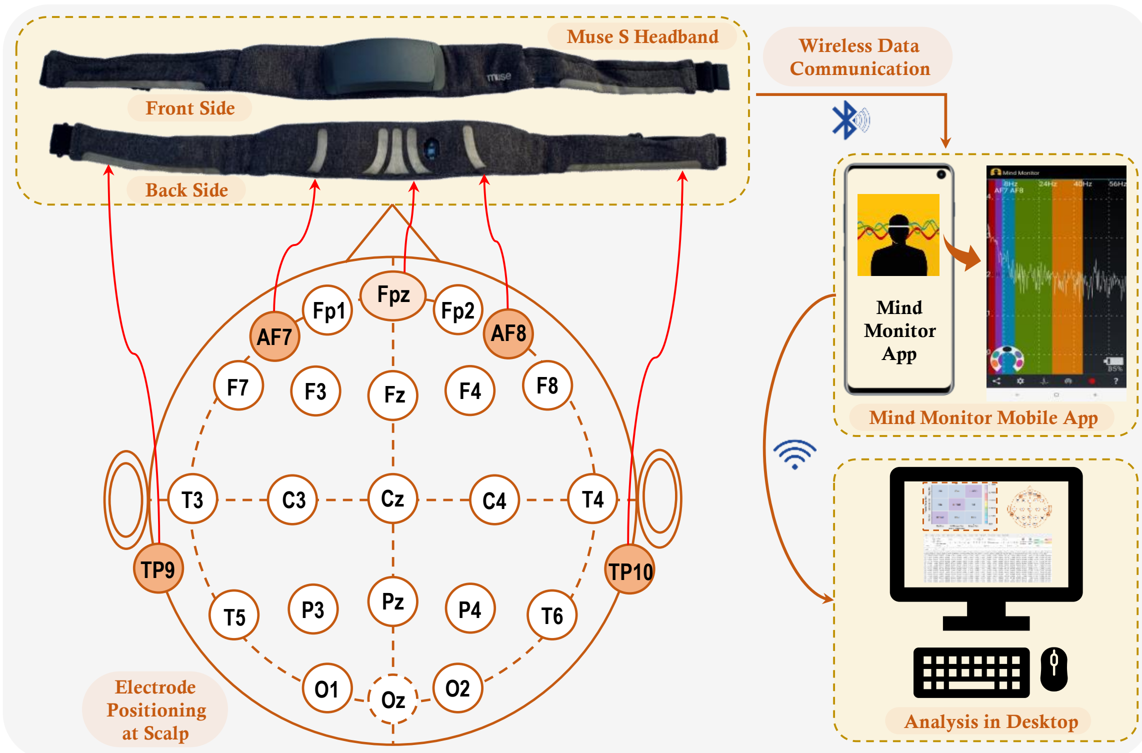

Figure 2: Overview of headband-based EEG device for measuring the activity of the brain, 10-20 system of electrode positioning, and mobile application for acquiring data

2.1.4 Experiment Procedure

Figure 3 shows an overview of the experiment's procedure. At a construction site, an excavator operation experiment was carried out to collect data for detecting the mental fatigue of construction equipment operators. The experiment was carried out on different days at the same time, from 9:00 a.m. to 11:00 a.m. (Li et al., 2019b) in the morning, under similar weather conditions, particularly clear weather on all data collection days. The experiment involved a repetitive and time-consuming excavation and discharge task on a construction site. It was a time-on-task approach, which is a common approach to induce mental fatigue (Li et al., 2020b, Morales et al., 2017, Hopstaken et al., 2016). For an hour, the excavator operators were required to conduct a repetitive and protracted excavation operation that included ground excavation and transporting material from pits to transport vehicles. The conditions for each excavator operator were the same, requiring them to continuously operate the equipment in the manner of a cyclic operation. The amount of earth excavated or moved, as well as the number of vehicles filled, were not fixed since it was a time-on-task experiment. Furthermore, no prior practice session was scheduled for the operators, as they already had experience with excavation operations. During the experiment, the operators were wearing a headband-based EEG device to collect data on their brain activity regarding active brain areas for mental fatigue while doing their tasks. Furthermore, the NASA-TLX score was used to quantify the subjective evaluation of equipment operators' mental fatigue. It has been used in various previous studies to subjectively assess mental fatigue in operators (Das et al., 2020). For the one-hour experiment, the subjective mental fatigue levels were recorded every 20 minutes, i.e., at 20, 40, and 60 min. Accordingly, the acquired EEG data was then labelled as per the subjective assessment into three mental fatigue states, i.e., alert state, mild fatigue state, and fatigue state (Prabaswari et al., 2019, Grier, 2015). There was no practice session included in the experiment because all the operators were professional excavator operators with prior experience in excavation operations. Furthermore, the exact duration of the experiment was not revealed to the operators. The purpose was to avoid the end-spurt effect reactivation that occurs when participants realize the experiment is nearing its conclusion (Morales et al., 2017).

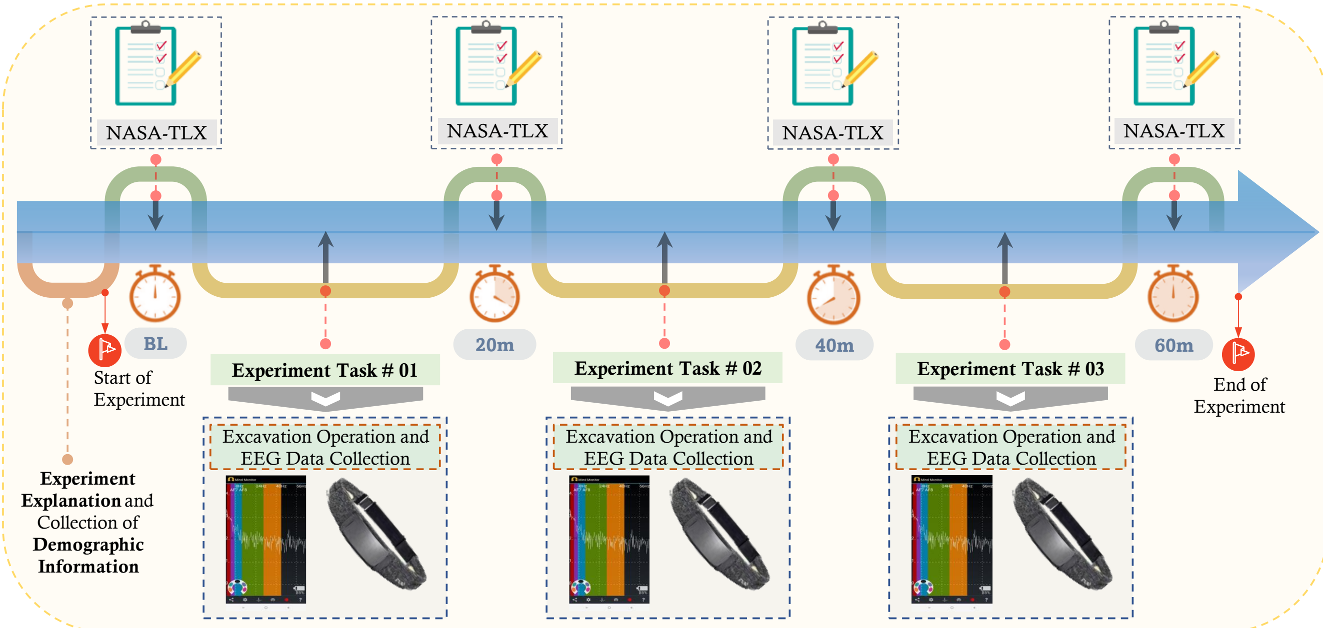

Figure 3: Experimental procedure for temporal assessment through NASA-TLX score and electroencephalography

2.2 Deep Learning-based Networks

The aim of our research was not to develop new and unique models, but rather to evaluate the innovative approach of utilizing deep learning techniques and headband-based wearable EEG sensor data to identify and classify mental fatigue states in construction equipment operators. Hence, the primary purpose was to contribute to the advancement of knowledge in the construction field by providing a deeper understanding of the cognitive processes and mental states of construction workers through a more sophisticated analysis of EEG data. This, in turn, could pave the way for improving safety and productivity, reducing accidents and injuries, and enhancing the overall well-being of construction workers. To achieve this, we employed three types of deep learning models: long short-term memory, bidirectional long short-term memory, and one-dimensional convolutional networks to train raw EEG data acquired by a wearable sensor. The sub-sections explain the details about the structures of the deep learning architectures we adopted in our research.

2.2.1 Long Short-Term Memory (LSTM)

In the last decade of the twentieth century, Hochreiter and Schmidhuber (1997) presented the first examples of LSTMs. These networks have the unique ability to learn long-term dependencies. Since it also has a memory component, it is one of the finest algorithms for processing sequence data. As a result of its memory component, LSTM can recall its prior actions in a process. With just a little structural tweak, it can solve the problem of the vanishing gradient that plagues RNN. The basic layout of an LSTM cell is depicted in Figure 4 (Olah, 2015). Because of this cell state, LSTM can only allow specific sets of information to pass through it. To implement this function, three logic gates are used. Input to these gates is provided by the sigmoid activation function. The first gate to determine what data can be safely erased from the cell is known as the Forget Gate f_t and is described in Eq. 1:

$$f_t = \sigma\left(x_t W^f + h_{t-1} U^f + b_f\right) \qquad Eq. 1$$

The result is either 0 or 1, with 0 indicating forget and 1 indicating keep. The second phase is the input gate, which determines which data will be added to the cell state or saved. As indicated in Eq. 2, the input gate also includes a second sigmoid layer for determining fresh candidate inputs that may be utilized to modify the cell's status.

$$i_t = \sigma\left(x_t W^i + h_{t-1} U^i + b_i\right) \qquad Eq. 2$$

In the following phase of LSTM, the old cell is replaced with a new one. As demonstrated in Eq. 3, the tanh function generates a vector of possible values that could be appended to the state.

$$\hat{C}_t = tanh(x_t W^g + h_{t-1} U^g + b_c) \qquad Eq. 3$$

Then, the new cell state replaces the previous one in C_{t-1} by discarding the information created by the forget gate in Eq. 1. The current cell state, denoted by C_t in Eq. 4, has been modified.

$$C_t = \sigma\left(f_t \text{ x } C_{t-1} + i_t \text{ x } \hat{C}_t\right) \qquad Eq.4$$

Finally, a sigmoid layer and subsequently a $tanh$ layer is employed to classify the output, as stated in Eq. 5 and 6.

$$\sigma_t = \sigma(x_t W^o + h_{t-1} U^o + b_o) \qquad Eq.5$$

$$h_t = tanh(C_t) \text{ x } \sigma_t \qquad Eq.6$$

where, i_t, f_t, and σ_t denotes the input gates, forget gates, and output gates, respectively. W^i, W^f, and W^o denotes the weights for the input gate, forget gate, and output gates at time step t, respectively. W^g is the weight of the candidate layer. U^i, U^f, and U^o are the weights for the input gate, forget gate, and output gates at time step $t-1$. U^g is the weight for the candidate layer. x_t is the input at the current time step t. h_t and h_{t-1} are

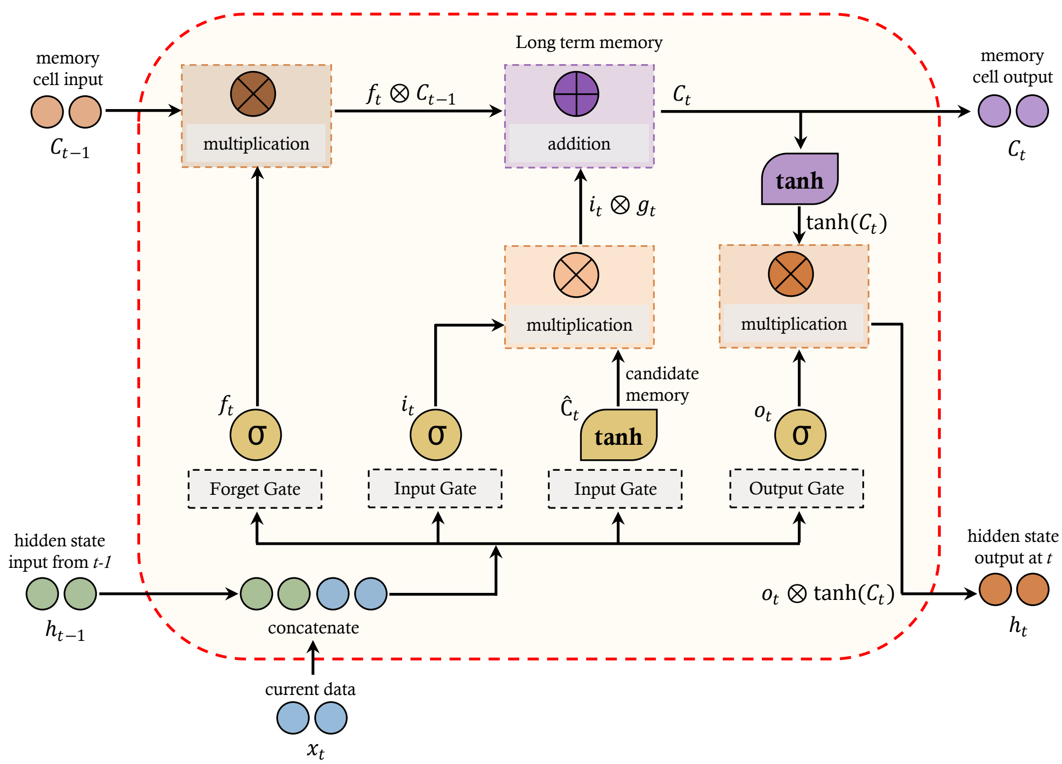

Figure 4: Long short-term memory (LSTM) cell architecture

the cell outputs at the current time step t and the previous time step $t-1$, respectively. C_t and C_{t-1} are the states of the cell at time steps t and $t-1$, respectively. b_i, b_f, and b_o denotes the biases for the input gate, forget gate, and output gates, respectively. b_c is the bias for the candidate layer, and σ is the sigmoid function.

2.2.2 Bidirectional Long Short-Term Memory (Bi-LSTM)

The Bi-LSTM layer structure is shown in Figure 5, and it consists of three independent layers that share the same input sequence and whose outputs are combined and displayed in the sequence. The state cells of a standard LSTM are split into a forward layer that controls the forward time path and a backward layer that controls the backward

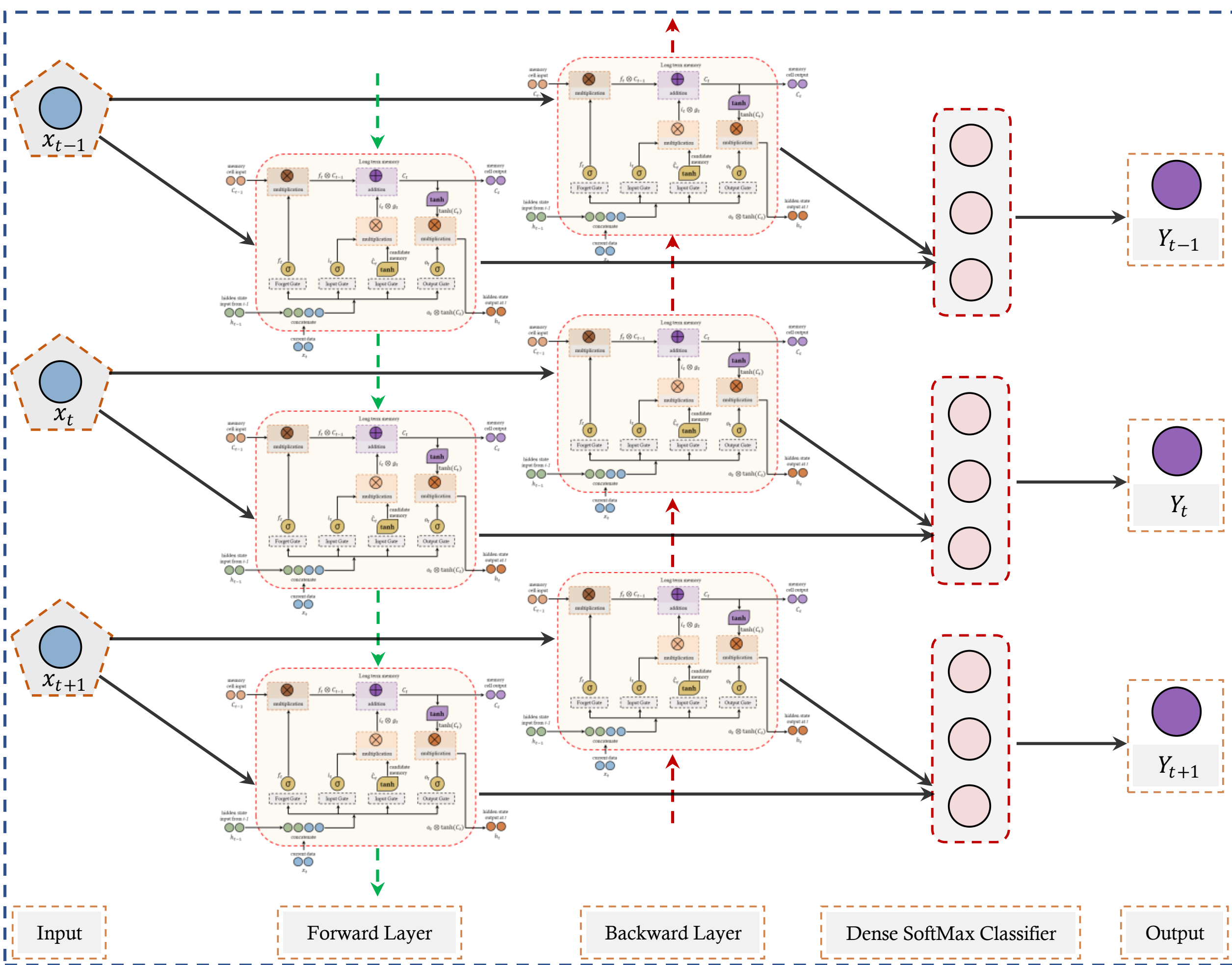

Figure 5: Bidirectional long short-term memory (Bi-LSTM) layer architecture

time direction in a Bi-LSTM model. For each time step forward and backward, information can be obtained by concatenating the outputs of the forward and backward layers. Given the established dependency between adjacent data pairs, this method improves the learning process.

2.2.3 1-Dimensional Convolutional Network

Deep convolutional neural networks, as they have been traditionally described in the literature, were developed with a focus on processing only two-dimensional data, such as images and recordings (Kiranyaz et al., 2021). For this reason, 2D-CNNs have become commonplace. Recently, however, 1D convolutional neural networks (1D-CNN) have been designed to work on one-dimensional data and have been applied to a wide variety of scenarios instead of 2D-CNN, such as by Eren et al. (2019), Kiranyaz et al. (2018), and Abdeljaber et al. (2018). Typically, specialized hardware is required for training deep 2D CNNs (e.g., cloud computing or GPU farms). Conversely, training small 1D CNNs with few hidden layers is practical and can be done quickly on any CPU implementation on a desktop machine (Kiranyaz et al., 2021). As a result of their minimal processing requirements, small 1D CNNs are ideal for real-time and low-cost scenarios (Eren, 2017). The 1D-CNN structure of the time-series prediction models used in this study is depicted in Figure 6. The network has several layers, including input, convolution, pooling, flattening, fully connected, and output layers. The features of the input are passed into a convolution layer. A feature map is generated by filtering an input feature in the convolution layer. The outcomes are then activated using the provided function. To shrink the feature map, the convolution layer's output is fed into a pooling layer. After that, to prepare the merged feature map for further processing, it is given to a flattening layer, which transforms it into a one-dimensional array. The completely linked layer then receives input from the layered-flattened layer. The weights are used in the fully connected layer to process the data. The output layer receives the signal from the layer with all connections made. When it comes to activation functions, ReLU is used in the convolution layer of this research. All other layers are ignored by the activation function (Chaerun Nisa and Kuan, 2021).

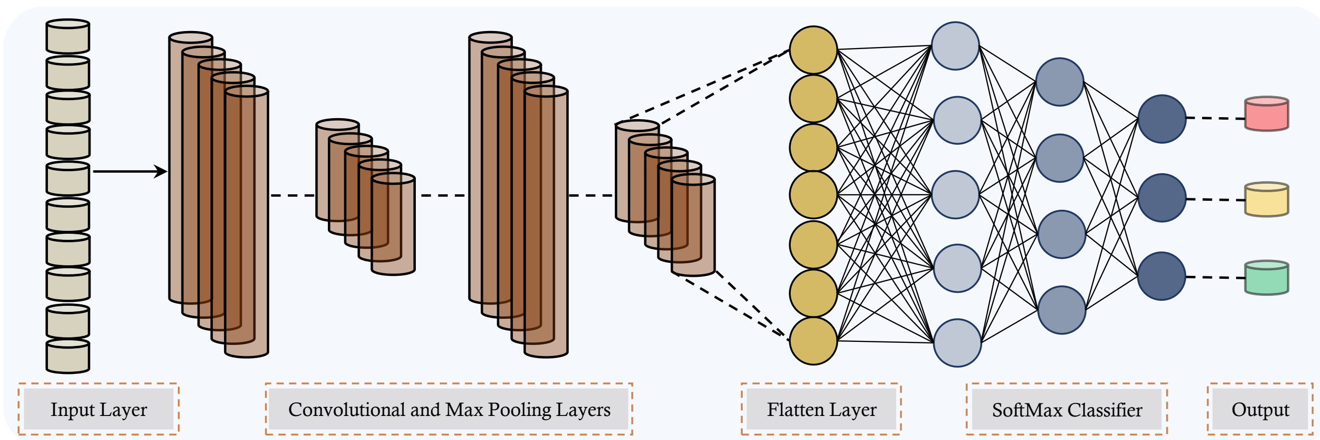

Figure 6: A sample one-dimensional convolutional network layer, flatten layer and SoftMax layer architecture

2.3 Training and performance evaluation of deep learning models

The EEG data of brain activity patterns was trained using three different deep learning techniques in the present study: long short-term memory (LSTM), bidirectional long short-term memory (Bi-LSTM), and one-dimensional convolutional networks (1D-CNN). To ensure consistency among the deep learning models, they were all constructed with the same dataset for training and evaluation. Based on EEG analysis, each designated class represents a single construction equipment operator. The electroencephalography data vector for each excavator operation task done by each operator has a dimensionality of 20 vectors (5 brain waves from each electrode x 4 electrodes of the EEG device) x 256 data samples. As a result, 5120 values serve as data samples in total. Input data for the current investigation consisted of 6,971,010 sample values from each electrode for every brain wave from fifteen equipment operators, since each window size contained 256 data samples and data was gathered for one hour. A sliding window approach was utilized with a window size of nine seconds to split the EEG sensor data into smaller segments, in order to capture long-term dependencies in the data. Overlapping of consecutive windows was then employed to ensure that no relevant data was missed. Specifically, a 50% overlap of adjacent data segment lengths was used in this study, as described by Liu et al. (2021b). However, there is no consensus on the optimal percentage of overlap, as previous studies have reported a range of overlapping percentages from 1% to 95% (Roy et al., 2019). Each deep learning model consists of three layers, with the number of hidden units varying from one hundred to five hundred. A similar architecture was utilized in a previous study, also using 200 hidden units per layer. When assessing the accuracy of our models, we employed a cost function based on the cross-entropy losses (the log loss function). In a classification problem, the loss function is what ultimately decides the model's performance. It is more indicative of reality when the loss value is lower. The optimization function is responsible for making the necessary adjustments to the model's weights and biases. An adaptive form of stochastic gradient descent was utilized for model training (Kingma and Ba, 2014), in addition to the Adam optimization function. Adam is a trustworthy optimizer that provides precise and quick updates to the network's settings. This research utilized the dropout technique (Srivastava et al., 2014), a popular stochastic regularization method, to prevent model overfitting. When the loss function is extremely small on the training data and extremely large on the testing data, overfitting occurs. The primary objective of the dropout method is to inhibit neurons in the system from over-adapting to one another, which leads to poor model generalization.

During the evaluation of the model's performance, the available data was partitioned into two subsets, with 70% being allocated for training purposes and the remaining 30% for testing. The original training dataset was split into two parts, with 80% going to the training phase and 20% going to the validation phase. We used the validation data to fine-tune our hyper-parameters and find the perfect spot for each of our three deep learning models' unit counts. Analogous to earlier research using deep learning networks (Antwi-Afari et al., 2022, Yang et al., 2020, Kim and Cho, 2020), the 10-fold cross-validation method was utilized to evaluate the classification performance of deep learning models. The optimum hyper-parameters can be chosen by 10-fold cross-validation, and the deep learning models can be tested as generalized models that exhibit acceptable classification performance with an unseen dataset. Based on the model, we chose the parameter values that achieved the highest level of accuracy with the least amount of time spent in training. The findings demonstrate that by adjusting the parameters of epoch, dropout, batch size, learning rate, and hidden units to 30, 0.5, 64, 0.001, and 200, respectively, our tuning procedure yielded the best accuracy for the datasets. To run the tests and train the models, we used a computer outfitted with a 2.50 GHz Quad-Core Intel Core i7-9750H CPU, 16 GB of RAM, a 64-bit operating system (Windows 10 Pro), and an Intel Iris Plus Graphics 650, 1,536 MB GPU running MATLAB R2020b. Table 2 displays the fine-tuned hyperparameters of the proposed deep learning models and the detailed dataset.

Table 2: Dataset and hyperparameters of proposed deep learning models

Dataset and Parameters	**Value**
Number of classes	3 (Alert State, Mild Fatigue State, Fatigue State)
Number of EEG sensors	4 (TP9, AF7, AF8, TP10)
Window size	9 s
Overlap of adjacent windows	50 %
Sampling rate	128 Hz

Epoch	30
Dropout	5%
Batch size	1000
Learning rate	0.001 (Adam optimizer: provides adaptive optimization)
Number of sample data	6,971,010 data samples

Accuracy, precision, recall, specificity, and the F1-score were employed to evaluate the three different types of deep learning models' performance in terms of evaluation and classification (Phutela et al., 2022, Attal et al., 2015). Each metric's breakdown for evaluation may be seen in Table 3. The most widely utilized metric to sum up classification performance across all classes is accuracy. Specifically, it is the ratio of instances that were correctly labeled relative to the total number of instances. Precision is the rate at which positive cases are correctly identified as such. In this sense, it is a quantitative indicator of precision. It is the ratio of positive instances that were correctly labeled compared to the total number of positive instances classified. Recall, also referred to as sensitivity, is a measure of how accurately positive examples were identified as such. Correctly classifying positive instances as a percentage of all positive instances is the definition of this metric. Whereas specificity is measured by how many times negative examples are correctly labeled as negative. To put it simply, it is the ratio of false-negatives that were identified compared to the total number of false-negatives. Precision and recall are combined into a single number called the F1-score, which is then used to evaluate the efficacy of the classification model without introducing any systematic bias (Ordóñez and Roggen, 2016). In addition to these measures, the confusion matrix was used to evaluate the performance of each model in particular classes, and the accuracy and loss curves were plotted to determine which model performed the best. The confusion matrix describes the discrepancies between the data's true labels and the model-generated labels. Elements on the diagonal of this matrix represent correctly classified fatigue states, whereas those off the diagonal represent incorrectly classified fatigue states. Furthermore, the Mann-Whitney test was conducted to analyze the results obtained from the deep learning models. While previous studies on EEG data and deep learning models for mental fatigue classification have compared models based on their achieved accuracy or training time, they have not statistically evaluated the difference in accuracy between models. To address this, we chose the Mann-Whitney test as it is a non-parametric test that does not require any assumptions about the distribution (Mat Roni et al., 2021), resulting in more conservative results. Velarde et al. (2022) and Phutela et al. (2022) employed analogous techniques in their investigations of the significance of predicted outcomes for time-series data using deep learning models.. Table 10 shows the inferences from classifiers with a p-value of less than 0.01 were considered significant, while the others were considered insignificant. If the p-value was less than 0.01, it was deduced that the classifier used for analysis is significant; otherwise, it is insignificant.

Table 3: Details of performance evaluation metrics for deep learning models

Performance metric	**Equation**
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$
Precision	$(TP)/(TP+FP)$
Recall	$(TP)/(TP+FN)$
Specificity	$(TN)/(TN+FP)$
F1-Score	$2 \text{ x } \frac{Precision \text{ x } Recall}{Precision + Recall}$

3 Experimental Results

In this section, we describe the findings of our investigations and the data we acquired from the operators. All fifteen construction equipment operators successfully completed the experiment. Therefore, data from all operators was used for analysis.

3.1 *Analysis of ground truth data*

The NASA-TLX score was utilized as a ground truth for recognizing mental fatigue states. Accordingly, Table 4 displays descriptive and analytical statistics derived from the ground truth evaluation. Subjective mental fatigue was significantly higher at the end of the NASA-TLX than at the start, increasing from 11.067 (SD = 2.764) to 64.733 (SD = 4.543). According to Table 3, operators reported increasing mental fatigue as the excavation operation progressed.

Table 4: Subjective assessment as a ground truth of mental fatigue

		Mental Fatigue States		
	Baseline	**Alert State**	**Mild Fatigue State**	**Fatigue State**
Subjective Assessment				
NASA-TLX Score (0-100)	11.25 (2.77)	30.81 (2.99)	45.00 (4.27)	65.25 (4.85)

**The scores for each state are mentioned as mean score (standard deviation)*

3.2 *Analysis of physiological data*

The absolute power for each frequency band of the EEG data acquired from all the channels of the MUSE headband was analyzed using the paired t-test across the three experimental phases (alert, mild fatigue, and fatigue) to make inferences about the underlying physiological processes. T-test results were interpreted in light of a null hypothesis and associated p-value. If the p-value for rejecting the null hypothesis was less than 0.05, then there was a statistically significant difference among the studied fatigue states. Table 5 displays the *t*-Stat for the power spectral density of EEG recordings made from various regions of the brain depicted in Figure 3. These findings indicate a statistically significant difference. For example, the t-test applied to EEG signals revealed that the alpha band was not found to be statistically significant at right frontal channel i.e., AS-MFS (t_{Stat} = 4.991, p < 0.05) and MFS-FS (t_{Stat} = -3.641, p < 0.05), whereas at the left frontal channel the alpha band was statistically significant only for comparison at fatigue states; AS-MFS (t_{Stat} = -4.816, p < 0.05). However, there was an increase in alpha activity as the experiment progressed from AS to FS, as demonstrated in Figure 7. Likewise, a similar trend was also shown for beta band at right frontal channel i.e., AS-MFS (t_{Stat} = 7.172, p < 0.05) and MFS-FS (t_{Stat} = -4.741, p < 0.05). However, this trend was inverse at left frontal channel for beta band. The beta band showed differences that were statistically significant in both the frontal and temporal parts of the brain. Overall, the theta band showed an increasing trend with an increase in mental fatigue from AS to FS in the frontal region of the brain, as demonstrated in Figure 7. The Delta band was found to be statistically significant in the left and right temporal regions. However, it was not significant for AS-MFS and MFS-FS at AF8 (t_{Stat} = 0.523) and AF7 (t_{Stat} = -1.559), respectively. Furthermore, the gamma band was found to be statistically significant between MFS-FS with p < 0.05 at TP9 (t_{Stat} = -3.175), AF7 (t_{Stat} = 4.814), AF8 (t_{Stat} = -4.791) and TP10 (t_{Stat} = -3.553). Table 5 depicts the statistical assessment of all the channels' bands. Previous studies have reported similar findings to our own, such as the study by Zhao et al. (2012) , which found significant theta and beta activity in the frontal regions of the brain. Other studies, including those by Nguyen et al. (2017), Käthner et al. (2014), and Dasari et al. (2010), also reported increased alpha and beta activity in the parietal region of the brain with an increase in mental fatigue. Additionally, Ma et al. (2018) reported increased alpha activity due to mental fatigue. Our investigations into mental fatigue show an increasing trend of frontal theta activity, which is consistent with previous studies by Trejo et al. (2015), Roy et al. (2013), and Dasari et al. (2010). Furthermore, the theta, alpha, and beta bands are the most commonly investigated EEG metrics for measuring mental fatigue. The (θ+α)/β ratio is the most widely used EEG metric for mental fatigue assessment. Higher mental fatigue is associated with an increase in this metric, according to findings by Jap et al. (2009). In our research, we found that time-on-task had a significant increasing effect on the EEG metric (θ+α)/β [$F = 15.011$, $p < 0.05$, $\eta^2 = 0.517$]. The value of the EEG metric (θ+α)/β in the alert state, mild fatigue state, and fatigue state was 1.015, 1.482, and 1.739, respectively, indicating an increase in mental fatigue. These findings are consistent with previous investigations by Ma et al. (2018) and Li et al. (2017a). Construction equipment operators' brain activity was visualized by calculating the power spectral density from their EEG data when they were in the alert state, mild fatigue state and fatigue state, as shown in Figure 7. The red color on the brain maps represents high levels of cortical activity, whereas the orange tint represents low levels. Brain activity in the alpha and beta bands of the frontal AF7 and AF8 channels can be seen to change graphically from the alert state to fatigue state on the brain maps.

Table 5: t-Stat for EEG power spectral densities at different brain regions

Mental	**Channels**	**EEG Bands (*t-Stat*)**

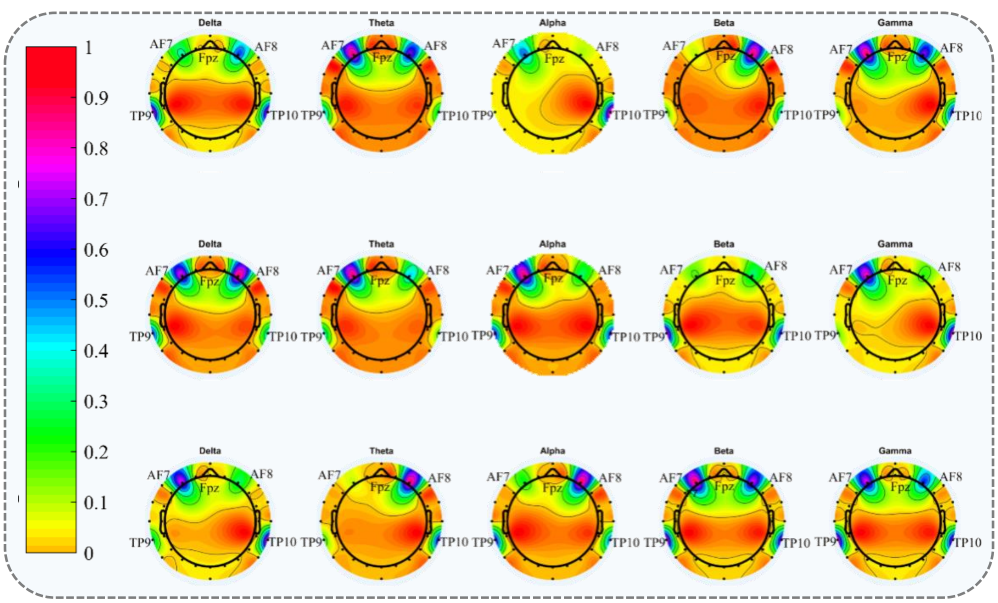

Figure 7: Brain visualization using EEG power spectral densities from different brain regions

Fatigue States		Delta	Theta	Alpha	Beta	Gamma
AS and MFS	TP9	2.526*	0.655	-1.325	7.560*	1.403
	AF7	-5.699*	-0.604	-4.816*	0.299	-1.871
	AF8	0.523	4.389*	4.991*	7.172*	1.574
	TP10	-2.662*	0.020	-1.351	-1.621	-17.812*
MFS and FS	TP9	-3009*	-4.464*	-2.823*	-3.450*	-3.175*
	AF7	-1.559	-0.370	-0.036	3.856*	4.814*
	AF8	-3.306*	-3.543*	-3.641*	-4.741*	-4.791*
	TP10	-2.670*	-5.596*	-1.411	-3.348*	-3.553*

**The t-Stat values are significant at the 0.05 level*

3.3 *Deep learning-based classification results*

Three deep learning models, LSTM, Bi-LSTM, and 1D-CNN were used to classify mental fatigue in construction equipment operators into alert, mild fatigue, and fatigue states. The implementation of LSTM, Bi-LSTM, and 1D-CNN employed cutting-edge parameter values, as demonstrated in Table 2. To limit experimental error, these models were run on the same system. The classification accuracies of the Bi-LSTM and LSTM deep learning models were both greater than 99%. The 1D-CNN deep learning model, on the other hand, only attained a classification accuracy marginally higher than 69%. However, upon evaluating the performance of the three deep learning models in terms of training time, it was found that the average duration for LSTM, Bi-LSTM, and 1D-CNN models was 68 minutes and 21 seconds, 163 minutes and 56 seconds, and 16 minutes and 57 seconds, respectively, as presented in Table 6. The findings show that, when trained on data reflecting operators' brain activity patterns over three increasingly demanding phases of work, the Bi-LSTM model outperformed the other deep learning models investigated in this study in terms of accuracy. Furthermore, LSTM also achieved accuracy slightly lower than Bi-LSTM when trained on the EEG sensor data.

Table 6: Classification accuracy and training time for deep learning models

Deep Learning Models	Accuracy (%)	Training Time
Long short-term memory (LSTM)	99.7063	68 mins 21 seconds
Bidirectional long short-term memory (Bi-LSTM)	99.9410	163 mins 56 seconds
One-dimensional convolutional network (1D-CNN)	69.4726	16 mins 57 seconds

3.3.1 Long short-term memory

Table 7 and Figure 8 illustrate the evaluation metrics and confusion matrix for the LSTM model. In general, the evaluation metrics demonstrated a good level of performance of the LSTM model on EEG-based brain activity data for identifying different mental fatigue levels in construction equipment operators. However, the performance of this model was slightly lower than that of Bi-LSTM. The LSTM model attained classification performance values ranging from 99.556% to 99.963% in terms of precision. FS represented 99.963% of instances of correctly identified fatigue levels. In addition, AS and MFS states exhibited the same effect on the LSTM model compared to FS, i.e., 99.556% and 99.589%, respectively. However, their effects were less than FS. Furthermore, higher recall and precision indicated that the model yielded fewer false negatives and false positives, respectively. Likewise, specificity and F1-score measures have values ranging between 99.761% and 99.818% and 99.681% and 99.718%, respectively. High specificity indicates the true negative rate, i.e., that a person identified as being in a fatigued state was in fact in that fatigued state. Besides, the confusion matrix was utilized to determine whether classes were misclassified or confused with others. As illustrated in Figure 8, each column depicts the actual mental fatigue states, while each row represents the predicted mental fatigue states. The diagonal cells indicate the correct instances for a more comprehensive evaluation of the classification performance at the end of the 30th epoch. The diagonal members of this matrix represent the cases in the dataset for which classification was accurate. Incorrectly classified instances include nondiagonal elements. The high values of the diagonal elements imply that the model correctly distinguishes between the three classifications of mental fatigue. The other cells indicate the incidents that were incorrectly classified. It is also evident that alert and mild fatigue states were misclassified more often than fatigue state. In spite of this, the misclassification rate remains remarkably low when compared to their number of classified instances. Furthermore, AS was confused with MFS and FS in 1299 and 1949 instances, respectively.

Table 7: Performance evaluation metrics for LSTM model

	Testing			
Indicator		**Alert State**	**Mild Fatigue State**	**Fatigue State**
Accuracy	99.7063%			
Precision		99.5569%	99.5898%	99.9634%
Recall		99.8807%	99.7735%	99.4693%
Specificity		99.7613%	99.8185%	99.9808%
F1-score		99.7186%	99.6816%	99.7157%

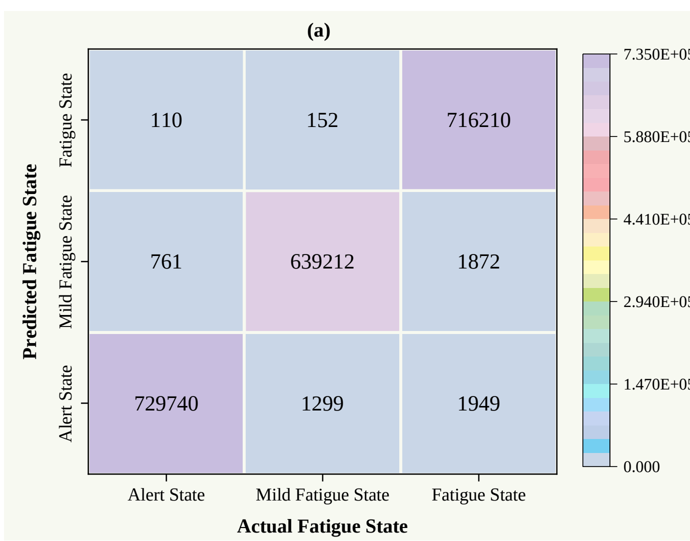

Figure 8: Confusion matrix for LSTM model *(the value in purple cells shows true positive instances)*

3.3.2 Bidirectional long short-term memory

The evaluation matrix and confusion matrix of the bidirectional LSTM model are presented in Table 8 and Figure 9, respectively. Bi-LSTM evaluation measures indicated the highest performance on EEG-based brain activity data for identifying distinct mental fatigue levels in construction equipment operators. This shows that Bi-LSTM is most effective in our construction equipment operation-related task. Results for accuracy-related classification performance for the Bi-LSTM model ranged from 99.840% to 99.995%. The MFS and FS indicated approximately comparable instances of correctly identified fatigue levels with a precision slightly above 99.995%; however, the AS exhibited a little less of an effect on the Bi-LSTM model with a precision of 99.840%. In addition, greater recall and precision indicated that the model produced fewer false negatives and, consequently, false positives. Similarly, specificity measures have values ranging from 99.914% to 99.997%, while the F1-score has values ranging from 99.917% to 99.993%. High specificity demonstrates the true negative rate, i.e., a person identified with any fatigue state was indeed experiencing that fatigue level. According to the confusion matrix in Figure 9, it can be observed that MFS and FS are the most recognized classes, with 640609 and 718872 positive instances, respectively. Furthermore, it is notable that AS was misclassified more frequently than MFS and FS. However, the misclassification rate was exceptionally low in comparison to the number of instances that were correctly identified. The confusion matrix further indicates that the AS was 1141 times confused with the FS. However, the confusion among the remaining states was modest.

Table 8: Performance evaluation metrics for Bi-LSTM model

	Testing			
Indicator		**Alert State**	**Mild Fatigue State**	**Fatigue State**
Accuracy	99.9410%			
Precision		99.8409%	99.9945%	99.9952%

Recall	99.9972%	99.9915%	99.8390%
Specificity	99.9144%	99.9975%	99.9975%
F1-score	99.9190%	99.9930%	99.9170%

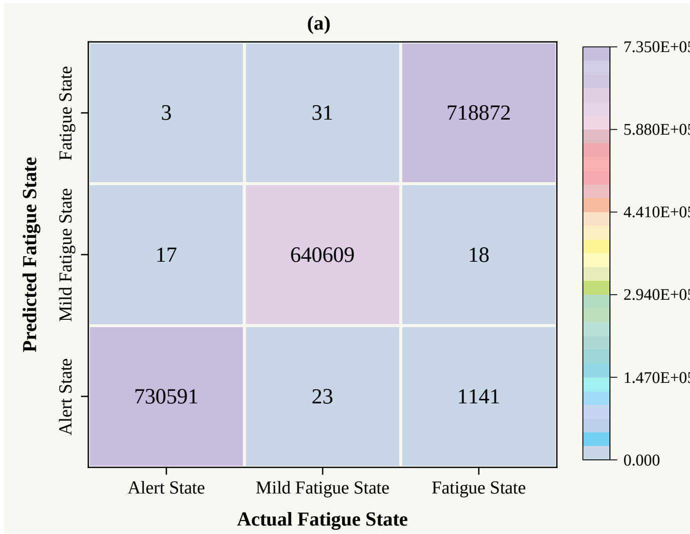

Figure 9: Confusion matrix for Bi-LSTM model *(the value in purple cells shows true positive instances)*

3.3.3 One-dimensional convolutional network

Table 9 and Figure 10 exhibit the evaluation matrix and confusion matrix of the 1-dimensional convolutional network (1DCN) model, with correct classes provided in the diagonal cells for a more detailed evaluation of classification performances at the end of the 30th epoch. When compared to the LSTM and Bi-LSTM models, the evaluation metrics of the 1DCN model achieved the lowest performance. In terms of precision, the 1-dimensional convolutional model produced classification performance values ranging from 54.600% to 84.241%. FS had the highest percentage of accurately classified instances, i.e., 72.545%. Furthermore, AS had the lowest accurately categorized instances, i.e., 65.387%. Moreover, for MFS, the model produced a high number of false negatives and false positives in this state as compared to other fatigue stages, i.e., 227,581 times with AS and 148,916 times with FS. Similarly, specificity measurements range from 74.046% to 92.874%, while the F1-score ranges from

61.607% to 77.957%. These findings reveal that the 1-dimensional convolutional model underperformed the LSTM or Bi-LSTM models based on EEG data in classifying mental fatigue in construction equipment operators. Furthermore, the confusion matrix in Figure 10 indicates that FS was the most recognized class, with 52,2348 affirmative instances.

Table 9: Performance evaluation metrics for 1D-CNN model

	Testing			
Indicator		**Alert State**	**Mild Fatigue State**	**Fatigue State**
Accuracy	69.4726%			
Precision		74.4194%	54.6009%	84.2415%
Recall		65.3874%	70.6780%	72.5452%
Specificity		87.9317%	74.0461%	92.8743%
F1-score		69.6116%	61.6078%	77.9570%

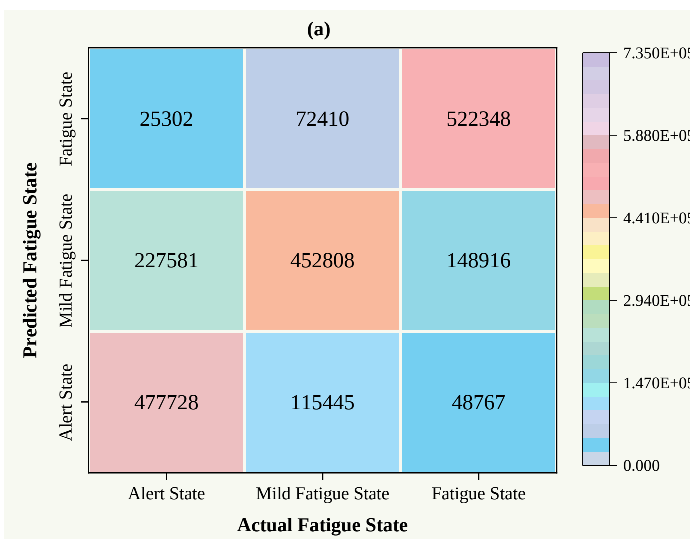

Figure 10: Confusion matrix for 1D-CNN model *(the value in purple cells shows true positive instances)*

3.3.4 Train/test accuracy and loss

The accuracy and loss over iterations curves of the three deep learning models investigated in this study are shown in Figure 11, respectively. The training and validation results for a bidirectional LSTM model show higher accuracy and lower loss, as shown in Figure 11(b). Specifically, the bidirectional LSTM model exhibited the maximum accuracy during training and validation, while the associated loss value was the lowest at the 30th epoch. As a result, the Bi-LSTM model was effectively trained without overfitting the EEG-based brain activity data of construction equipment operators, as demonstrated by the smallest difference between training accuracy and validation accuracy or training loss and validation loss.

3.3.5 Comparison of p-values for deep learning models

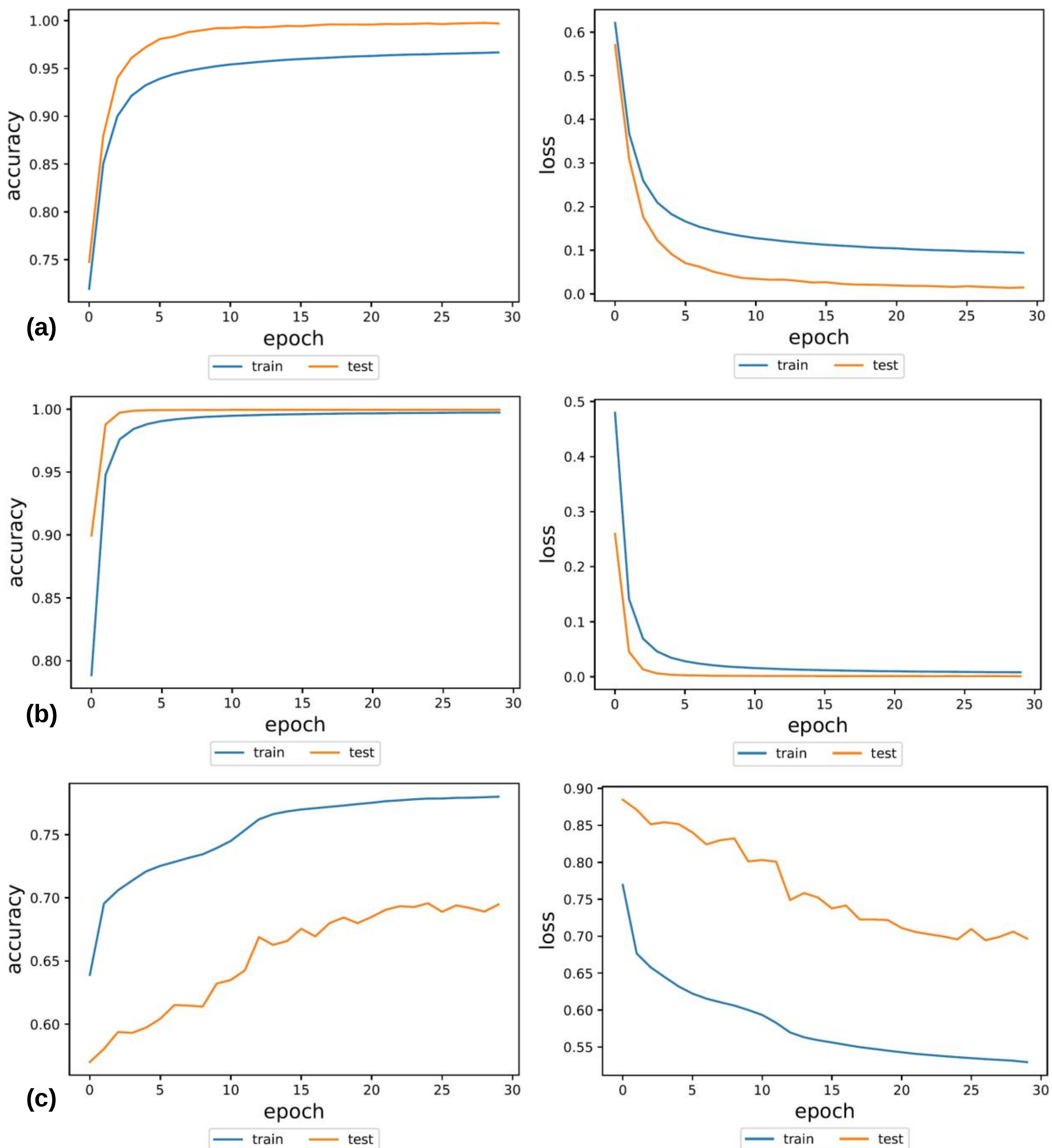

Figure 2: Accuracy and loss over iteration curves with the tuned hyperparameters of (a) LSTM model, (b) Bi-LSTM model, and (c) 1D-CNN model

The p-values of the Mann-Whitney test computed on the results given by the bidirectional long short-term memory

(Bi-LSTM) findings are presented in Table 10. The results demonstrate that the bidirectional LSTM model's accuracy was considerably higher to that of the other two models, i.e., the LSTM and the 1D-CNN, for 10-fold cross-validation.

Table 10: Mann-Whitney test-based comparison of p-values for Bi-LSTM model

Model	Validation method	LSTM		1D-CNN	
		***p*-value**	**Significance**	***p*-value**	**Significance**
Bi-LSTM	10-fold cross validation	0.007994	Sig.	7.22×10^{-30}	Highly Sig.
LSTM	10-fold cross validation			1.4×10^{-28}	Highly Sig.

** Significant at $p < 0.01$*

4 Discussion

4.1.1 Deep learning networks and EEG data for mental fatigue

Construction equipment operations are cognitively demanding and necessitate the operators' undivided attention. Such protracted attention induces mental fatigue in construction equipment operators, which is one of the leading causes of construction-site equipment-related accidents. As a result, it is imperative that the mental fatigue of construction equipment operators be monitored non-invasively to reduce equipment-related incidents and make construction sites safe for workers. Thus, the objective of this study was to assess a novel approach for recognizing and classifying different types of mental fatigue states in equipment operation using deep learning-based networks and wearable EEG data gathered from equipment operators. This study compared three types of deep learning models for training time-series raw EEG data acquired by a wearable headband: long short-term memory (LSTM), bidirectional long short-term memory (bi-LSTM), and one-dimensional convolutional network (1D-CNN). To our knowledge, this study is the first to propose a deep learning-based model for recognizing and classifying alert, mild fatigue, and fatigue states from EEG signals in construction equipment operators under sustained attention. Our results show that mental fatigue can be accurately classified in construction equipment operators with varying mental fatigue levels, i.e., alert state, mild fatigue state, and fatigue state.

Comparing the deep learning models utilized in this study revealed that the bidirectional LSTM model had the highest accuracy of 99.941%. In addition, the results demonstrate that the bidirectional LSTM model achieved precision, recall, specificity, and F1-score metrics ranging from 99.840% to 99.995%, 99.839% to 99.997%, 99.914% to 99.997%, and 99.917% to 99.993%, respectively, when classifying multiple states of mental fatigue in construction equipment operators. Regarding the confusion matrix, it was concluded that the fatigue state (FS) and mild fatigue state (MFS) had the fewest misclassified instances, i.e., 34 and 35, respectively. While alert state (AS) was the most misclassified class, with 1164 instances of misclassification. In addition, the performance of the Bi-LSTM model increases in accuracy and decreases in loss during both training and testing. These findings indicate that the Bi-LSTM model, which is a deep learning network model, could provide a more accurate classification of the mental fatigue states of operators. This finding might be explained from the perspective of the model. Bidirectional LSTM can remember previous time-series patterns and is more effective at processing time-series data. The Bi-LSTM model is a cyclic neural network comprised of two distinct LSTM networks that can collect information not only from past input but also from future input states. Consequently, the concept of bidirectional LSTM's design is to collect the characteristics of a time series at the current time while also having information about the past and the future, resulting in outstanding classification accuracy. In this study, the bidirectional LSTM model surpassed other deep learning-based models in terms of classification accuracy. Similar to previous research, our findings indicate that the bidirectional LSTM model is better than unidirectional LSTM models (Li et al., 2020c), while bidirectional LSTM performs better with time-series data (Siami-Namini et al., 2019). Similarly, Sarkar et al. (2022) reported a high performance of the LSTM architecture compared to CNN when the models were applied to a testing dataset with varied percentages such as 20%, 30%, and 40%. The study established that when dealing with time-series EEG data, LSTM was found to perform better than a convolutional neural network, stating that CNN is more useful for image data. Similarly, Phutela et al. (2022) also reported that LSTM is a promising option for classifying stress-related brain activity data. Likewise, Rastgoo et al. (2019) also stated that long short-term memory has a strong ability to exploit the temporal dependencies in time-series data. Similarly, the results of the study by Cai et al. (2021) and our findings also demonstrated that the proposed approach, based on the three-layer Bi-LSTM prediction model, has a greater conception of the data exhibited in the forecast models, leading to more reliable forecasts of mental fatigue in construction equipment operators. The three-layer Bi-LSTM model learns from its errors during unsupervised training of the EEG-based brain activity patterns, to increase precision while maintaining the original attributes of the input EEG data. As a result, the forecast model of mental fatigue classification we developed in this study is more robust for its application in the construction industry. Hence, the use of bidirectional LSTM and LSTM models is recommended to classify mental fatigue states in construction equipment operators based on EEG data.

In Table 11, we contrast the performance of our approach with other methods found in the literature that are relevant to construction workers. Based on the comparison, it is evident that our classification method employs bidirectional LSTM and LSTM-based deep learning models, is better in performance. Previously, many studies had been conducted in construction to classify the stress or fatigue of construction workers, and some acceptable accuracy had been achieved. However, all these studies used manually crafted features from EEG data and applied machine learning to classify either stress or fatigue. Our approach is significantly different from previous machine learning approaches, where raw EEG data has been directly used without any manual crafting of input features. Although Jebelli et al. (2019c) have used deep learning neural networks to classify mental stress with an accuracy of 86.62%, such a study is significantly different in several ways. For example, their task was not a prolonged task because the focus of the author was to classify mental stress in construction workers. It was a simple task with a short duration that was performed by workers once standing on the ground (low stress) and then on the top of a ladder (high stress). Furthermore, such experimental settings are not suitable to induce mental fatigue in workers, particularly construction equipment operators. Also, it's worth noting that, because of variations in the experimental setup, the nature of tasks performed by operators, the number of subjects, the subjects themselves, etc., direct comparison with the methods is not possible or will be quiet challenging.

Table 6: Comparison of mental fatigue classification accuracies in construction domain

Reference	**No. of subjects**	**No. of electrodes**	**Stress or Fatigue (Levels)**	**Stimulus (Type of data collection settings)**	**Classification Method**	**Accuracy (%)**
(Aryal et al., 2017)	12	Beta 1 channel	Fatigue (4)	Psychomotor Vigilance Task (indoor simulated)	Boosted trees	82.60
(Jebelli et al., 2018a)	11	14	Stress (2)	Working on ladder (construction site)	Fully connected NN	79.26
(Jebelli et al., 2018b)	5	14	Stress (2)	Working on ladder (construction site)	OMTL-VonNeuman	77.60
(Jebelli et al., 2019b)	7	14	Stress (2)	Working on ladder (construction site)	Gaussian support vector machine	80.32
(Jeon and Cai,	30	16	Hazard (3)	Simulated environment	CatBoost	65.2

2022)				(Laboratory setting)	LightGBM	63.7
Current study	15	4	Fatigue (3)	Construction Equipment Operation (Site)	Deep learning (Bi-LSTM)	99.9410
Current study	15	4	Fatigue (3)	Construction Equipment Operation (Site)	Deep learning (LSTM)	99.7076
Current study	15	4	Fatigue (3)	Construction Equipment Operation (Site)	Deep learning (1D CNN)	69.4726

4.1.2 Study implications

The proposed research investigation is expected to impact safety management practices on construction sites in many ways, especially as they pertain to construction equipment operators. It provides researchers and practitioners in the construction industry with pertinent results and practical implications. First, the feasibility of onsite experimental data collection for detecting mental fatigue using a wearable electroencephalography sensor has important practical implications. Collecting wearable sensing data in a real-world construction scenario is extremely challenging for a variety of reasons, including the dynamic nature of the construction site (Chen et al., 2022), the need for enormous resources, and the plethora of onsite risks. However, many prior investigations on mental fatigue monitoring, such as the investigations by Liu et al. (2021a), Li et al. (2020b), and Li et al. (2019a), were conducted in a laboratory under controlled settings. In addition, other investigations were undertaken on construction sites, but they did not focus on mental fatigue, which necessitates the collection of extensive data from workers and operators. These experiments, like those by Lee and Lee (2022), and Jebelli et al. (2019a), only focused on small tasks to detect stress in workers. In contrast, the present research investigated the use of wearable electroencephalography data while construction equipment operators performed an extended excavation task on a construction site. Therefore, the findings are highly valuable to construction industry researchers. As a consequence of protracted equipment operations, mental fatigue is a typical human behavior. The proposed method could be utilized not only for excavation operations (such as excavating earth and then tracking the bucket before and after transferring the excavated material to the trucks) but also for other prolonged repetitive equipment operations in the construction industry, such as crane operations, etc. Secondly, the insights can assist construction managers in establishing a framework for managing worker shifts. In the current study, researchers analyzed changes in the brain activity of construction equipment operators at three levels for one hour. They can be observed by construction managers every 30 to 45 minutes. Breaks between shifts can be introduced to allow equipment operators to recuperate from the mental fatigue they have induced. Thirdly, the proposed method facilitates the recognition and classification of mental fatigue states in construction equipment operators. Using EEG data and deep learning networks, it is possible to recognize and classify the mental fatigue that affects human behavior due to attention failure in construction equipment operators. Identification of mental fatigue is the first step towards proactively preventing attentional failure. As a result, this EEG data and deep learning networks-based approach can serve as a proactive intervention tool for tracking and identifying various brain fatigue states in operators, thereby decreasing mental fatigue-related incidents on construction sites and minimizing accidents. In addition, this method may be effective for assisting workers in many construction industry occupations, such as monitoring the mental states of structural design engineers, who are frequently required to do multiple redesign tasks in a short period of time with sustained attention. Furthermore, the current technique will assist in the development of real-time wearable EEG sensor computing through the use of brain activity pattern performance and a bidirectional LSTM model for the classification of distinct states of mental fatigue. Managers in charge of workplace safety might utilize this data to better protect their workers. The best deep learning models, i.e., Bi-LSTM and LSTM, can learn the brain activity patterns of equipment operators and reliably forecast mental fatigue states, as demonstrated by the performance accuracy of all three deep learning models in this study. However, misclassification of mild fatigue states was revealed to occur more frequently than misclassification of other fatigue states, which could lead to identification problems. Nonetheless, the study's findings can be applied to other cognitive failures, such as mental stress, mental workload, hazard identification, emotions, etc., leading to better incident management for construction workers experiencing cognitive issues.

4.1.3 Limitations and future research

The proposed research is the first study to use deep learning-based models to classify mental fatigue in construction equipment operators using EEG sensor data. Although the findings have expanded our comprehension of mental fatigue monitoring using EEG data of brain activity patterns, they have some limitations that should be acknowledged and addressed in future research investigations. First, the sample size in this study (fifteen equipment operators) was modest, and there were three mental fatigue levels. Despite the fact that we determined the sample size based on the sample sizes employed in previous research of a similar kind, findings with such a limited number of operators may limit the application of the proposed approach to the construction industry. To generalize the results to the entire population of operators, future studies should collect large data sets representing a variety of mental fatigue states. Second, this research exclusively employed excavation operators as equipment operators. Future research may corroborate these findings for operators of other construction equipment, such as crane operators, dozer operators, grader operators, etc. In general, a data set with enough samples from different groups of equipment operators to identify additional mental fatigue states is important for training, testing, and building a general model for construction operations. Thirdly, this study did not define the thresholds for various levels of mental fatigue based on EEG data. Depending on whether or not these mental fatigue thresholds can be set and applied to all operators of construction equipment, future research may use such thresholds to define and classify mental fatigue states using statistical analysis, machine learning, and deep learning. Fourthly, the current research exclusively analyzed electroencephalography data from wearable sensors to classify equipment operators' mental fatigue. Additionally, different types of wearable sensor networks or biosensors exist for gathering data on things like an operator's heart rate, respiration rate, skin conductance, and geometric facial feature measurements. Construction site monitoring and recognition applications could benefit from combining the acquired data from multiple sensors. Therefore, multimodal networks for monitoring and managing the mental fatigue of construction site workers should be developed in future studies that integrate information gathered from various biomarkers. Finally, the current study used only three types of deep learning networks to recognize and classify mental fatigue in construction equipment operators. Nevertheless, deep learning models based on bidirectional LSTM are intended primarily to handle sequence and time-series data. They do, however, come at a higher cost. They require more time to train the model. As a result, future research could combine multiple deep learning networks as a fusion model or use multi-deep learning models to classify mental fatigue in equipment operators.

5 Conclusions

Operators' attention failure as a consequence of mental fatigue due to prolonged equipment operations is a common cause of equipment-related accidents. These incidents lead to serious injuries and fatalities on construction sites.

As a result, the current study proposes a construction site procedure for classifying mental fatigue in construction equipment operators, which is an effective way to decrease the incidence of equipment-related accidents. Given that mental fatigue can cause inattention failures in construction equipment operators, this study evaluated a novel approach that used deep learning-based networks and EEG sensor data to distinguish and classify mental fatigue states. Through subjective assessment, three different types of mental fatigue states, i.e., alert state, mild fatigue state, and fatigue state, were labeled, and subsequent brain activity patterns of fifteen equipment operators were acquired using a wearable headband EEG sensor. The mental fatigue was induced by a monotonous and prolonged excavation task on a real construction site. Accuracy, precision, recall, specificity, and the F1-score were used to evaluate the classification performance of the three deep learning models, i.e., LSTM, bidirectional LSTM, and 1D-CNN. Furthermore, the Mann-Whitney test was conducted to assess the statistical significance of the results obtained from three deep learning models. The experimental findings demonstrate that the bidirectional LSTM model performs better compared to the other deep learning models, with an accuracy of 99.941% and an F1-score of 99.917% to 99.993%. Bidirectional LSTM and LSTM models both outperformed 1D-CNN models; however, the difference between their accuracies was less than 1%. These findings provide support for the notion that the Bi-LSTM model, which is prevalently used for the classification of time-series and sequential data, can be employed to learn sequential brain activity patterns captured by an EEG sensor in order to distinguish and classify different types of mental fatigue states during construction operations. In addition, the current method will aid in the development of real-time wearable EEG sensor computing through the classification of different states of mental fatigue utilizing brain activity pattern performance and a bidirectional LSTM model. Furthermore, it will help to improve the safety and health management on construction sites by allowing safety managers to constantly track the mental fatigue level of construction equipment operators in real-time.

Acknowledgement

The authors acknowledged the following two funding grants: 1. General Research Fund (GRF) Grant (15201621) titled "Monitoring and managing fatigue of construction plant and equipment operators exposed to prolonged sitting"; 2. General Research Fund (GRF) Grant (15210720) titled "The development and validation of a noninvasive tool to monitor mental and physical stress in construction workers"; and 3. Research Institute for Intelligent Wearable System (RI-IWEAR) - Strategic Supporting Scheme (CD47) titled "An Automated Assessment of Construction Equipment Operators' Mental Fatigue based on Facial Expressions".

References

ABDELJABER, O., SASSI, S., AVCI, O., KIRANYAZ, S., IBRAHIM, A. A. & GABBOUJ, M. 2018. Fault detection and severity identification of ball bearings by online condition monitoring. *IEEE Transactions on Industrial Electronics,* 66**,** 8136-8147.

ADANE, M. M., GELAYE, K. A., BEYERA, G. K., SHARMA, H. R. & YALEW, W. W. 2013. Occupational injuries among building construction workers in Gondar City, Ethiopia. *Occupational Medicine & Health Affairs*.

AHN, C. R., LEE, S., SUN, C., JEBELLI, H., YANG, K. & CHOI, B. 2019. Wearable sensing technology applications in construction safety and health. *Journal of Construction Engineering and Management,* 145**,** 03119007.

AHN, S., NGUYEN, T., JANG, H., KIM, J. G. & JUN, S. C. 2016. Exploring Neuro-Physiological Correlates of Drivers' Mental Fatigue Caused by Sleep Deprivation Using Simultaneous EEG, ECG, and fNIRS Data. *Frontiers in Human Neuroscience,* 10.

ALBERT, A., PANDIT, B., PATIL, Y. & LOUIS, J. 2020. Does the potential safety risk affect whether particular construction hazards are recognized or not? *Journal of safety research,* 75**,** 241-250.

ALGHADIR, A. & ANWER, S. 2015. Prevalence of musculoskeletal pain in construction workers in Saudi Arabia. *The Scientific World Journal,* 2015.

ANTWI-AFARI, M. F., QAROUT, Y., HERZALLAH, R., ANWER, S., UMER, W., ZHANG, Y. & MANU, P. 2022. Deep learning-based networks for automated recognition and classification of awkward working postures in construction using wearable insole sensor data. *Automation in Construction,* 136, 104181.

ARSALAN, A., MAJID, M., BUTT, A. R. & ANWAR, S. M. 2019. Classification of Perceived Mental Stress Using A Commercially Available EEG Headband. *IEEE Journal of Biomedical and Health Informatics,* 23, 2257-2264.

ARYAL, A., GHAHRAMANI, A. & BECERIK-GERBER, B. 2017. Monitoring fatigue in construction workers using physiological measurements. *Automation in Construction,* 82, 154-165.

ATTAL, F., MOHAMMED, S., DEDABRISHVILI, M., CHAMROUKHI, F., OUKHELLOU, L. & AMIRAT, Y. 2015. Physical human activity recognition using wearable sensors. *Sensors,* 15, 31314-31338.

BAI, X.-P. & QIAN, C. 2021. Factor validity and reliability performance analysis of human behavior in green architecture construction engineering. *Ain Shams Engineering Journal,* 12, 4291-4296.

BEHRENS, M., GUBE, M., CHAABENE, H., PRIESKE, O., ZENON, A., BROSCHEID, K.-C., SCHEGA, L., HUSMANN, F. & WEIPPERT, M. 2023. Fatigue and Human Performance: An Updated Framework. *Sports Medicine,* 53, 7-31.

BIRHANE, G. E., YANG, L., GENG, J. & ZHU, J. 2022. Causes of construction injuries: a review. *International Journal of Occupational Safety and Ergonomics,* 28, 343-353.

BITKINA, O. V., PARK, J. & KIM, H. K. 2021. The ability of eye-tracking metrics to classify and predict the perceived driving workload. *International Journal of Industrial Ergonomics,* 86, 103193.

BOKSEM, M. A. S. & TOPS, M. 2008. Mental fatigue: Costs and benefits. *Brain Research Reviews,* 59, 125-139.

BORRAGÁN, G., SLAMA, H., DESTREBECQZ, A. & PEIGNEUX, P. 2016. Cognitive fatigue facilitates procedural sequence learning. *Frontiers in human neuroscience,* 10, 86.

BROWN, I. D. 1994. Driver fatigue. *Human factors,* 36, 298-314.

BUCSUHÁZY, K., MATUCHOVÁ, E., ZŮVALA, R., MORAVCOVÁ, P., KOSTÍKOVÁ, M. & MIKULEC, R. 2020. Human factors contributing to the road traffic accident occurrence. *Transportation research procedia,* 45, 555-561.

BYERS, J. C. 1989. Traditional and raw task load index (TLX) correlations: are paired comparisons necessary? *Advances in Industrial Erfonomics and Safety l: Taylor and Francis.*

CAI, C., TAO, Y., ZHU, T. & DENG, Z. 2021. Short-Term Load Forecasting Based on Deep Learning Bidirectional LSTM Neural Network. *Applied Sciences,* 11, 8129.

CANNARD, C., WAHBEH, H. & DELORME, A. Validating the wearable MUSE headset for EEG spectral analysis and Frontal Alpha Asymmetry. 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021. IEEE, 3603-3610.

CARBONARI, A., GIRETTI, A. & NATICCHIA, B. 2011. A proactive system for real-time safety management in construction sites. *Automation in construction,* 20, 686-698.

CESA-BIANCHI, N. & ORABONA, F. 2021. Online Learning Algorithms. *Annual Review of Statistics and Its Application,* 8, 165-190.

CHAE, J., HWANG, S., SEO, W. & KANG, Y. 2021. Relationship between rework of engineering drawing tasks and stress level measured from physiological signals. *Automation in Construction,* 124, 103560.

CHAERUN NISA, E. & KUAN, Y.-D. 2021. Comparative Assessment to Predict and Forecast Water-Cooled Chiller Power Consumption Using Machine Learning and Deep Learning Algorithms. *Sustainability,* 13, 744.

CHAN, M. 2011. Fatigue: The most critical accident risk in oil and gas construction. *Construction Management and Economics,* 29, 341-353.

CHAVAILLAZ, A., SCHWANINGER, A., MICHEL, S. & SAUER, J. 2019. Work design for airport security officers: Effects of rest break schedules and adaptable automation. *Applied ergonomics,* 79, 66-75.

CHEN, C.-F. & HSU, Y.-C. 2020. Taking a closer look at bus driver emotional exhaustion and well-being: evidence from Taiwanese urban bus drivers. *Safety and Health at Work,* 11, 353-360.

CHEN, J., TAYLOR, J. E. & COMU, S. 2017. Assessing Task Mental Workload in Construction Projects: A Novel Electroencephalography Approach. *Journal of Construction Engineering and Management,* 143, 04017053.

CHEN, S., DEMACHI, K. & DONG, F. 2022. Graph-based linguistic and visual information integration for on-site occupational hazards identification. *Automation in Construction,* 137, 104191.

CHENG, B., FAN, C., FU, H., HUANG, J., CHEN, H. & LUO, X. 2022. Measuring and Computing Cognitive Statuses of Construction Workers Based on Electroencephalogram: A Critical Review. *IEEE Transactions on Computational Social Systems*, 1-16.

CHOI, B., JEBELLI, H. & LEE, S. 2019. Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk. *Safety Science,* 115, 110-120.

CHOI, J., GU, B., CHIN, S. & LEE, J.-S. 2020. Machine learning predictive model based on national data for fatal accidents of construction workers. *Automation in Construction,* 110, 102974.

CRAIK, A., HE, Y. & CONTRERAS-VIDAL, J. L. 2019. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering,* 16, 031001.

DAS, S., MAITI, J. & KRISHNA, O. B. 2020. Assessing mental workload in virtual reality based EOT crane operations: A multi-measure approach. *International Journal of Industrial Ergonomics,* 80, 103017.

DASARI, D., CROWE, C., LING, C., ZHU, M. & DING, L. EEG pattern analysis for physiological indicators of mental fatigue in simulated air traffic control tasks. Proceedings of the human factors and ergonomics society annual meeting, 2010. SAGE Publications Sage CA: Los Angeles, CA, 205-209.

DEL SAVIO, A., LUNA, A., CÁRDENAS-SALAS, D., VERGARA, M. & URDAY, G. 2022. Dataset of manually classified images obtained from a construction site. *Data in Brief,* 42, 108042.

DING, Y., MA, J. & LUO, X. 2022. Applications of natural language processing in construction. *Automation in Construction,* 136, 104169.

DUAN, R., DENG, H., TIAN, M., DENG, Y. & LIN, J. 2022. SODA: A large-scale open site object detection dataset for deep learning in construction. *Automation in Construction,* 142, 104499.

EREN, L. 2017. Bearing fault detection by one-dimensional convolutional neural networks. *Mathematical Problems in Engineering,* 2017.

EREN, L., INCE, T. & KIRANYAZ, S. 2019. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. *Journal of Signal Processing Systems,* 91**,** 179-189.

FANG, Q., LI, H., LUO, X., DING, L., LUO, H., ROSE, T. M. & AN, W. 2018a. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction,* 85**,** 1-9.

FANG, W., DING, L., LUO, H. & LOVE, P. E. 2018b. Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction,* 91**,** 53-61.

FARAG, A., SCOTT, L., PERKHOUNKOVA, Y., SAEIDZADEH, S. & HEIN, M. 2022. A human factors approach to evaluate predicators of acute care nurse occupational fatigue. *Applied Ergonomics,* 100**,** 103647.

FRONE, M. R. & TIDWELL, M.-C. O. 2015. The meaning and measurement of work fatigue: Development and evaluation of the Three-Dimensional Work Fatigue Inventory (3D-WFI). *Journal of occupational health psychology,* 20**,** 273.

FRYDENLUND, A. & RUDZICZ, F. Emotional affect estimation using video and EEG data in deep neural networks. Advances in Artificial Intelligence: 28th Canadian Conference on Artificial Intelligence, Canadian AI 2015, Halifax, Nova Scotia, Canada, June 2-5, 2015, Proceedings 28, 2015. Springer, 273-280.

GRIER, R. A. How high is high? A meta-analysis of NASA-TLX global workload scores. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2015. SAGE Publications Sage CA: Los Angeles, CA, 1727-1731.

HALLOWELL, M. R., HINZE, J. W., BAUD, K. C. & WEHLE, A. 2013. Proactive construction safety control: Measuring, monitoring, and responding to safety leading indicators. *Journal of construction engineering and management,* 139**,** 04013010.

HAN, Y., JIN, R., WOOD, H. & YANG, T. 2019. Investigation of Demographic Factors in Construction Employees' Safety Perceptions. *KSCE Journal of Civil Engineering,* 23**,** 2815-2828.

HART, S. G. NASA-task load index (NASA-TLX); 20 years later. Proceedings of the human factors and ergonomics society annual meeting, 2006. Sage publications Sage CA: Los Angeles, CA, 904-908.

HINZE, J. W. & TEIZER, J. 2011. Visibility-related fatalities related to construction equipment. *Safety science,* 49**,** 709-718.

HOCHREITER, S. & SCHMIDHUBER, J. 1997. Long short-term memory. *Neural computation,* 9**,** 1735-1780.

HOPSTAKEN, J. F., VAN DER LINDEN, D., BAKKER, A. B., KOMPIER, M. A. J. & LEUNG, Y. K. 2016. Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology: Human Perception and Performance,* 42**,** 878-889.

HU, X. & LODEWIJKS, G. 2020. Detecting fatigue in car drivers and aircraft pilots by using non-invasive measures: The value of differentiation of sleepiness and mental fatigue. *Journal of safety research,* 72**,** 173-187.

HWANG, S., JEBELLI, H., CHOI, B., CHOI, M. & LEE, S. 2018. Measuring Workers’ Emotional State during Construction Tasks Using Wearable EEG. *Journal of Construction Engineering and Management,* 144**,** 04018050.

JAP, B. T., LAL, S., FISCHER, P. & BEKIARIS, E. 2009. Using EEG spectral components to assess algorithms for detecting fatigue. *Expert Systems with Applications,* 36**,** 2352-2359.

JEBELLI, H., CHOI, B. & LEE, S. 2019a. Application of Wearable Biosensors to Construction Sites. I: Assessing Workers’ Stress. *Journal of Construction Engineering and Management,* 145**,** 04019079.

JEBELLI, H., HWANG, S. & LEE, S. 2018a. EEG-based workers' stress recognition at construction sites. *Automation in Construction,* 93**,** 315-324.

JEBELLI, H., KHALILI, M. M., HWANG, S. & LEE, S. 2018b. A Supervised Learning-Based Construction Workers’ Stress Recognition Using a Wearable Electroencephalography (EEG) Device. *Construction Research Congress 2018.*

JEBELLI, H., KHALILI, M. M. & LEE, S. 2019b. A Continuously Updated, Computationally Efficient Stress Recognition Framework Using Electroencephalogram (EEG) by Applying Online Multitask Learning Algorithms (OMTL). *IEEE Journal of Biomedical and Health Informatics,* 23**,** 1928-1939.

JEBELLI, H., KHALILI, M. M. & LEE, S. Mobile EEG-Based Workers’ Stress Recognition by Applying Deep Neural Network. *In:* MUTIS, I. & HARTMANN, T., eds. Advances in Informatics and Computing in Civil and Construction Engineering, 2019// 2019c Cham. Springer International Publishing, 173-180.

JEON, J. & CAI, H. 2022. Multi-class classification of construction hazards via cognitive states assessment using wearable EEG. *Advanced Engineering Informatics,* 53**,** 101646.

KADUK, S. I., ROBERTS, A. P. J. & STANTON, N. A. 2021. Driving performance, sleepiness, fatigue, and mental workload throughout the time course of semi-automated driving—Experimental data from the driving simulator. *Human Factors and Ergonomics in Manufacturing & Service Industries,* 31**,** 143-154.

KÄTHNER, I., WRIESSNEGGER, S. C., MÜLLER-PUTZ, G. R., KÜBLER, A. & HALDER, S. 2014. Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological psychology,* 102**,** 118-129.

KE, J., DU, J. & LUO, X. 2021a. The effect of noise content and level on cognitive performance measured by electroencephalography (EEG). *Automation in Construction,* 130**,** 103836.

KE, J., ZHANG, M., LUO, X. & CHEN, J. 2021b. Monitoring distraction of construction workers caused by noise using a wearable Electroencephalography (EEG) device. *Automation in Construction,* 125**,** 103598.

KIM, K. & CHO, Y. K. 2020. Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition. *Automation in Construction,* 113**,** 103126.

KINGMA, D. P. & BA, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980.*

KIRANYAZ, S., AVCI, O., ABDELJABER, O., INCE, T., GABBOUJ, M. & INMAN, D. J. 2021. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing,* 151**,** 107398.

KIRANYAZ, S., GASTLI, A., BEN-BRAHIM, L., AL-EMADI, N. & GABBOUJ, M. 2018. Real-time fault detection and identification for MMC using 1-D convolutional neural networks. *IEEE Transactions on Industrial Electronics,* 66**,** 8760-8771.

KOC, K. & GURGUN, A. P. 2022. Scenario-based automated data preprocessing to predict severity of construction accidents. *Automation in Construction,* 140**,** 104351.

KRAUSS, T. P., SHURE, L. & LITTLE, J. 1994. *Signal processing toolbox for use with MATLAB®: user's guide*, The MathWorks.

LAITINEN, H. & PÄIVÄRINTA, K. 2010. A new-generation safety contest in the construction industry–a long-term evaluation of a real-life intervention. *Safety science,* 48**,** 680-686.

LECUN, Y., BENGIO, Y. & HINTON, G. 2015. Deep learning. *Nature,* 521**,** 436-444.

LEE, B. G., CHOI, B., JEBELLI, H. & LEE, S. 2021. Assessment of construction workers' perceived risk using physiological data from wearable sensors: A machine learning approach. *Journal of Building Engineering,* 42**,** 102824.

LEE, G. & LEE, S. 2022. Feasibility of a Mobile Electroencephalogram (EEG) Sensor-Based Stress Type Classification for Construction Workers. *Construction Research Congress 2022.*

LEE, H., YANG, K., KIM, N. & AHN, C. R. 2020. Detecting excessive load-carrying tasks using a deep learning network with a Gramian Angular Field. *Automation in Construction,* 120**,** 103390.

LI, G., LEE, C. H., JUNG, J. J., YOUN, Y. C. & CAMACHO, D. 2020a. Deep learning for EEG data analytics: A survey. *Concurrency and Computation: Practice and Experience,* 32**,** e5199.

LI, H., WANG, D., CHEN, J., LUO, X., LI, J. & XING, X. 2019a. Pre-service fatigue screening for construction workers through wearable EEG-based signal spectral analysis. *Automation in Construction,* 106**,** 102851.

LI, J., LI, H., UMER, W., WANG, H., XING, X., ZHAO, S. & HOU, J. 2020b. Identification and classification of construction equipment operators' mental fatigue using wearable eye-tracking technology. *Automation in Construction,* 109**,** 103000.

LI, J., LI, H., WANG, F., CHENG, A. S. K., YANG, X. & WANG, H. 2021. Proactive analysis of construction equipment operators' hazard perception error based on cognitive modeling and a dynamic Bayesian network. *Reliability Engineering & System Safety,* 205**,** 107203.

LI, J., LI, H., WANG, H., UMER, W., FU, H. & XING, X. 2019b. Evaluating the impact of mental fatigue on construction equipment operators' ability to detect hazards using wearable eye-tracking technology. *Automation in Construction,* 105**,** 102835.

LI, R., SU, W. & LU, Z. 2017a. Physiological signal analysis for fatigue level of experienced and inexperienced drivers. *Traffic Injury Prevention,* 18**,** 139-144.

LI, W., ZHANG, L. & LIANG, W. 2017b. An Accident Causation Analysis and Taxonomy (ACAT) model of complex industrial system from both system safety and control theory perspectives. *Safety science,* 92**,** 94-103.

LI, Y.-H., HARFIYA, L. N., PURWANDARI, K. & LIN, Y.-D. 2020c. Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model. *Sensors,* 20**,** 5606.

LIAO, P.-C., ZHOU, X., CHONG, H.-Y., HU, Y. & ZHANG, D. 2022. Exploring construction workers' brain connectivity during hazard recognition: a cognitive psychology perspective. *International Journal of Occupational Safety and Ergonomics***,** 1-9.

LIU, P., CHI, H.-L., LI, X. & GUO, J. 2021a. Effects of dataset characteristics on the performance of fatigue detection for crane operators using hybrid deep neural networks. *Automation in Construction,* 132**,** 103901.

LIU, X., LI, G., WANG, S., WAN, F., SUN, Y., WANG, H., BEZERIANOS, A., LI, C. & SUN, Y. 2021b. Toward practical driving fatigue detection using three frontal EEG channels: A proof-of-concept study. *Physiological Measurement,* 42**,** 044003.

LIU, Y., CHEN, X., PENG, H. & WANG, Z. 2017. Multi-focus image fusion with a deep convolutional neural network. *Information Fusion,* 36**,** 191-207.

MA, J., GU, J., JIA, H., YAO, Z. & CHANG, R. 2018. The Relationship Between Drivers' Cognitive Fatigue and Speed Variability During Monotonous Daytime Driving. *Frontiers in Psychology,* 9.

MA, L., GUO, H. & FANG, Y. 2021. Analysis of Construction Workers’ Safety Behavior Based on Myers-Briggs Type Indicator Personality Test in a Bridge Construction Project. *Journal of Construction Engineering and Management,* 147**,** 04020149.

MASULLO, M., TOMA, R., PASCALE, A., RUGGIERO, G. & MAFFEI, L. Research methodology used to investigate the effects of noise on overhead crane operator's performances. International Ergonomics Conference, 2020. Springer, 223-231.

MAT RONI, S., DJAJADIKERTA, H. G., MAT RONI, S. & DJAJADIKERTA, H. G. 2021. Non-Parametric Tests. *Data Analysis with SPSS for Survey-based Research***,** 219-260.

MEHMOOD, I., LI, H., UMER, W., ARSALAN, A., SAAD SHAKEEL, M. & ANWER, S. 2022. Validity of facial features' geometric measurements for real-time assessment of mental fatigue in construction equipment operators. *Advanced Engineering Informatics,* 54**,** 101777.

MOLAN, G. & MOLAN, M. 2021. Theoretical model for accident prevention based on root cause analysis with graph theory. *Safety and health at work,* 12**,** 42-50.

MOON, S., LEE, G. & CHI, S. 2022. Automated system for construction specification review using natural language processing. *Advanced Engineering Informatics,* 51**,** 101495.

MORALES, J. M., DÍAZ-PIEDRA, C., RIEIRO, H., ROCA-GONZÁLEZ, J., ROMERO, S., CATENA, A., FUENTES, L. J. & DI STASI, L. L. 2017. Monitoring driver fatigue using a single-channel electroencephalographic device: A validation study by gaze-based, driving performance, and subjective data. *Accident Analysis & Prevention,* 109**,** 62-69.

NAKAGOME, S., CRAIK, A., SUJATHA RAVINDRAN, A., HE, Y., CRUZ-GARZA, J. G. & CONTRERAS-VIDAL, J. L. 2022. Deep learning methods for EEG neural classification. *Handbook of Neuroengineering.* Springer.

NGUYEN, T., AHN, S., JANG, H., JUN, S. C. & KIM, J. G. 2017. Utilization of a combined EEG/NIRS system to predict driver drowsiness. *Scientific reports,* 7**,** 43933.

OLAH, C. 2015. Understanding lstm networks.

ORDÓÑEZ, F. J. & ROGGEN, D. 2016. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors,* 16**,** 115.

ORFANIDIS, S. J. 1995. *Introduction to signal processing*, Prentice-Hall, Inc.

PHUTELA, N., RELAN, D., GABRANI, G., KUMARAGURU, P. & SAMUEL, M. 2022. Stress Classification

Using Brain Signals Based on LSTM Network. *Computational Intelligence and Neuroscience,* 2022**,** 7607592.

PRABASWARI, A. D., BASUMERDA, C. & UTOMO, B. W. The mental workload analysis of staff in study program of private educational organization. IOP Conference Series: Materials Science and Engineering, 2019. IOP Publishing, 012018.

RAJULA, H. S. R., VERLATO, G., MANCHIA, M., ANTONUCCI, N. & FANOS, V. 2020. Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. *Medicina (Kaunas),* 56.

RASTGOO, M. N., NAKISA, B., MAIRE, F., RAKOTONIRAINY, A. & CHANDRAN, V. 2019. Automatic driver stress level classification using multimodal deep learning. *Expert Systems with Applications,* 138**,** 112793.

ROY, R. N., BONNET, S., CHARBONNIER, S. & CAMPAGNE, A. Mental fatigue and working memory load estimation: interaction and implications for EEG-based passive BCI. 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013. IEEE, 6607-6610.

ROY, Y., BANVILLE, H., ALBUQUERQUE, I., GRAMFORT, A., FALK, T. H. & FAUBERT, J. 2019. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering,* 16**,** 051001.

SAEDI, S., FINI, A. A. F., KHANZADI, M., WONG, J., SHEIKHKHOSHKAR, M. & BANAEI, M. 2022. Applications of electroencephalography in construction. *Automation in Construction,* 133**,** 103985.

SANEI, S. & CHAMBERS, J. A. 2013. *EEG signal processing*, John Wiley & Sons.

SARKAR, A., SINGH, A. & CHAKRABORTY, R. 2022. A deep learning-based comparative study to track mental depression from EEG data. *Neuroscience Informatics,* 2**,** 100039.

SARKAR, S., PRAMANIK, A., MAITI, J. & RENIERS, G. 2020. Predicting and analyzing injury severity: A machine learning-based approach using class-imbalanced proactive and reactive data. *Safety science,* 125**,** 104616.

SAVITZKY, A. & GOLAY, M. J. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry,* 36**,** 1627-1639.

SEO, J. & LEE, S. 2021. Automated postural ergonomic risk assessment using vision-based posture classification. *Automation in Construction,* 128**,** 103725.

SIAMI-NAMINI, S., TAVAKOLI, N. & NAMIN, A. S. The Performance of LSTM and BiLSTM in Forecasting Time Series. 2019 IEEE International Conference on Big Data (Big Data), 9-12 Dec. 2019 2019. 3285-3292.

SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. & SALAKHUTDINOV, R. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research,* 15**,** 1929-1958.

TEHRANI, B. M., WANG, J. & TRUAX, D. 2021. Assessment of mental fatigue using electroencephalography (EEG) and virtual reality (VR) for construction fall hazard prevention. *Engineering, Construction and*

Architectural Management, ahead-of-print.

THODOROFF, P., PINEAU, J. & LIM, A. Learning robust features using deep learning for automatic seizure detection. Machine learning for healthcare conference, 2016. PMLR, 178-190.

TREJO, L. J., KUBITZ, K., ROSIPAL, R., KOCHAVI, R. L. & MONTGOMERY, L. D. 2015. EEG-based estimation and classification of mental fatigue. *Psychology,* 6**,** 572.

TÜRK, Ö. & ÖZERDEM, M. S. 2021. The convolutional neural network approach from electroencephalogram signals in emotional detection. *Concurrency and Computation: Practice and Experience,* 33**,** e6356.

UMER, W. 2022. Simultaneous monitoring of physical and mental stress for construction tasks using physiological measures. *Journal of Building Engineering,* 46**,** 103777.

UMER, W., LI, H., LU, W., SZETO, G. P. Y. & WONG, A. Y. 2018. Development of a tool to monitor static balance of construction workers for proactive fall safety management. *Automation in Construction,* 94**,** 438-448.

UMER, W., LI, H., YANTAO, Y., ANTWI-AFARI, M. F., ANWER, S. & LUO, X. 2020. Physical exertion modeling for construction tasks using combined cardiorespiratory and thermoregulatory measures. *Automation in Construction,* 112**,** 103079.

VAHDATIKHAKI, F., EL AMMARI, K., LANGROODI, A. K., MILLER, S., HAMMAD, A. & DOREE, A. 2019. Beyond data visualization: A context-realistic construction equipment training simulators. *Automation in construction,* 106**,** 102853.

VAN CUTSEM, J., MARCORA, S., DE PAUW, K., BAILEY, S., MEEUSEN, R. & ROELANDS, B. 2017. The Effects of Mental Fatigue on Physical Performance: A Systematic Review. *Sports Medicine,* 47**,** 1569-1588.

VAN DER LINDEN, D., FRESE, M. & SONNENTAG, S. 2003. The Impact of Mental Fatigue on Exploration in a Complex Computer Task: Rigidity and Loss of Systematic Strategies. *Human Factors,* 45**,** 483-494.

VELARDE, G., BRAÑEZ, P., BUENO, A., HEREDIA, R. & LOPEZ-LEDEZMA, M. 2022. An Open Source and Reproducible Implementation of LSTM and GRU Networks for Time Series Forecasting. *Engineering Proceedings,* 18**,** 30.

VILLANI, V., GABBI, M. & SABATTINI, L. Promoting operator’s wellbeing in Industry 5.0: detecting mental and physical fatigue. 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 9-12 Oct. 2022 2022. 2030-2036.

WAGSTAFF, A. S. & SIGSTAD LIE, J.-A. 2011. Shift and night work and long working hours – a systematic review of safety implications. *Scandinavian Journal of Work, Environment & Health***,** 173-185.

WANG, D., LI, H. & CHEN, J. 2019. Detecting and measuring construction workers' vigilance through hybrid kinematic-EEG signals. *Automation in Construction,* 100**,** 11-23.

WANG, J., CHEN, D., ZHU, M. & SUN, Y. 2021. Risk assessment for musculoskeletal disorders based on the characteristics of work posture. *Automation in Construction,* 131**,** 103921.

WANG, M., ZHAO, Y. & LIAO, P.-C. 2022. EEG-based work experience prediction using hazard recognition. *Automation in Construction,* 136**,** 104151.

WILLIAMSON, A., LOMBARDI, D. A., FOLKARD, S., STUTTS, J., COURTNEY, T. K. & CONNOR, J. L. 2011. The link between fatigue and safety. *Accident Analysis & Prevention,* 43**,** 498-515.

WU, C., LI, X., GUO, Y., WANG, J., REN, Z., WANG, M. & YANG, Z. 2022. Natural language processing for smart construction: Current status and future directions. *Automation in Construction,* 134**,** 104059.

WU, H., ZHONG, B., LI, H., LOVE, P., PAN, X. & ZHAO, N. 2021. Combining computer vision with semantic reasoning for on-site safety management in construction. *Journal of Building Engineering,* 42**,** 103036.

WU, Y., MIWA, T. & UCHIDA, M. Heart rate based evaluation of operator fatigue and its effect on performance during pipeline work. Advances in Physical Ergonomics and Human Factors: Proceedings of the AHFE 2017 International Conference on Physical Ergonomics and Human Factors, July 17-21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8, 2018. Springer, 446-454.

XING, X., LI, H., LI, J., ZHONG, B., LUO, H. & SKITMORE, M. 2019. A multicomponent and neurophysiological intervention for the emotional and mental states of high-altitude construction workers. *Automation in Construction,* 105**,** 102836.

XING, X., ZHONG, B., LUO, H., ROSE, T., LI, J. & ANTWI-AFARI, M. F. 2020. Effects of physical fatigue on the induction of mental fatigue of construction workers: A pilot study based on a neurophysiological approach. *Automation in Construction,* 120**,** 103381.

YANG, J., YE, G., XIANG, Q., KIM, M., LIU, Q. & YUE, H. 2021. Insights into the mechanism of construction workers' unsafe behaviors from an individual perspective. *Safety Science,* 133**,** 105004.

YANG, K., AHN, C. R. & KIM, H. 2020. Deep learning-based classification of work-related physical load levels in construction. *Advanced Engineering Informatics,* 45**,** 101104.

YANG, X., LI, H., YU, Y., LUO, X., HUANG, T. & YANG, X. 2018. Automatic pixel-level crack detection and measurement using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering,* 33**,** 1090-1109.

YEŞILMEN, S. & TATAR, B. 2022. Efficiency of convolutional neural networks (CNN) based image classification for monitoring construction related activities: A case study on aggregate mining for concrete production. *Case Studies in Construction Materials,* 17**,** e01372.

YU, Y., YANG, X., LI, H., LUO, X., GUO, H. & FANG, Q. 2019. Joint-level vision-based ergonomic assessment tool for construction workers. *Journal of construction engineering and management,* 145**,** 04019025.

ZENG, H., YANG, C., DAI, G., QIN, F., ZHANG, J. & KONG, W. 2018. EEG classification of driver mental states by deep learning. *Cognitive Neurodynamics,* 12**,** 597-606.

ZHANG, F., FLEYEH, H., WANG, X. & LU, M. 2019a. Construction site accident analysis using text mining and natural language processing techniques. *Automation in Construction,* 99**,** 238-248.

ZHANG, H., YAN, X. & LI, H. 2018. Ergonomic posture recognition using 3D view-invariant features from single ordinary camera. *Automation in Construction,* 94**,** 1-10.

ZHANG, Y., ZHANG, M. & FANG, Q. 2019b. Scoping review of EEG studies in construction safety. *International journal of environmental research and public health,* 16**,** 4146.

ZHAO, C., ZHAO, M., LIU, J. & ZHENG, C. 2012. Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. *Accident Analysis & Prevention,* 45, 83-90.

ZHAO, J. & OBONYO, E. 2020. Convolutional long short-term memory model for recognizing construction workers' postures from wearable inertial measurement units. *Advanced Engineering Informatics,* 46, 101177.

ZHAO, J. & OBONYO, E. 2021. Applying incremental Deep Neural Networks-based posture recognition model for ergonomics risk assessment in construction. *Advanced Engineering Informatics,* 50, 101374.

ZHAO, Y. S., JAAFAR, M. H., MOHAMED, A. S. A., AZRAAI, N. Z. & AMIL, N. 2022. Ergonomics Risk Assessment for Manual Material Handling of Warehouse Activities Involving High Shelf and Low Shelf Binning Processes: Application of Marker-Based Motion Capture. *Sustainability,* 14, 5767.

ZHENG, Y., LIU, Q., CHEN, E., GE, Y. & ZHAO, J. L. Time series classification using multi-channels deep convolutional neural networks. International conference on web-age information management, 2014. Springer, 298-310.

ZHONG, B., XING, X., LUO, H., ZHOU, Q., LI, H., ROSE, T. & FANG, W. 2020. Deep learning-based extraction of construction procedural constraints from construction regulations. *Advanced Engineering Informatics,* 43, 101003.

ZHU, Z., PARK, M.-W., KOCH, C., SOLTANI, M., HAMMAD, A. & DAVARI, K. 2016. Predicting movements of onsite workers and mobile equipment for enhancing construction site safety. *Automation in Construction,* 68, 95-101.