



Worker-centric heat strain analysis: Integrating physiological signals with ensemble learning and domain adaptation

Amit Ojha^a, Shayan Shayesteh^a, Ali Sharifironizi^b, Yizhi Liu^c, Houtan Jebelli^{a,*}

^a Department of Civil and Environmental Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States of America

^b Robotic, Automation, and Intelligent Sensing (RAISe) Laboratory, United States of America

^c Department of Civil and Environmental Engineering, Syracuse University, Syracuse, NY 13244, United States of America

ARTICLE INFO

Keywords:

Occupational heat strain risk
Data-driven health monitoring
Heat-vulnerable occupations
Ensemble learning
Domain adaptation
Immersive environments

ABSTRACT

Despite advances in utilizing physiological sensors and machine learning (ML) algorithms, accurately and consistently monitoring heat strain levels of field workers on job sites remains a challenge not fully addressed by previous research. Existing frameworks often fail to adapt to the diverse physiology of workers and their varied working conditions, leading to concerns about stability, reliability, and accuracy. To address these limitations, a worker-centered heat strain monitoring framework was introduced, leveraging physiological data from wearable biosensors for a more accurate and consistent estimation of heat strain risks. A high-fidelity virtual reality (VR) environment was developed to simulate heat-vulnerable occupations for quality data collection. Building on this foundation, EnsmTrBoost, a physiological sensing framework integrating ensemble learning and domain adaptation, was developed. This framework exhibited a remarkable prediction accuracy exceeding 93%. This paper advances heat strain monitoring and supports the development of early warning systems for heat-related fatalities at work sites.

1. Introduction

Numerous workers in field-oriented industries, such as construction, agriculture, forestry, and fishing, struggle with serious work-related injuries and health problems [1–3]. One of the critical health issues in this context is workers' exposure to heat stress, which can be associated with the prolonged hours of intensive physical field operations in the high ambient temperature [4–6]. Heat stress can abruptly raise the core body temperature above a safe threshold, shut down the temperature-regulating system, and adversely affect workers' overall safety, health, and productivity, leading to heat-related illness and fatalities [7–10]. The long-term consequence of heat stress also includes a myriad of chronic health problems, such as cardiovascular complications [11–14], or kidney diseases [15,16], as well as respiratory disorders [8,17]. Further, heat stress can negatively affect brain activities, such as cognition [18] and concentration [9]. Notably, in field-oriented industries, particularly construction, agriculture, forestry, and fishing, workers are more vulnerable to heat stress [7]. In 2018, the Bureau of Labor Statistics reported that 9262 workers were diagnosed with heat related illness [19]. Regrettably, owing to anticipated climate changes and the detrimental impacts of global warming, along with more

frequent and intense heatwaves, it is expected that these assertions will see a rise in the coming years [20].

Current practices for preventing traumatic heat-related illnesses and fatalities mainly rely on monitoring of environmental conditions of the workplace [1,21–23]. Such techniques attempt to evaluate the likelihood of heat stress exposure based on environmental parameters such as temperature, humidity, wind velocity, and cloud cover [24]. Since these methods extrapolate environmental parameters to estimate the likelihood of heat stress exposure universally for all workers with a singular value, they fail to consider individual differences in personal characteristics [25]. Another common technique for estimating the workers' likelihood of exposure to heat stress is the use of perception-based self-assessment techniques, such as the Heat Strain Score index (HSSI; [26]). Such methods are concerned with the human thermo-perception or how individuals perceive heat [26–28]. In this regard, implementing these perception-based self-assessment techniques continuously in the field can be challenging due to their inherently intrusive nature [29,30]. These techniques require workers to actively engage in self-assessment and provide feedback on their thermal comfort or discomfort. Such involvement can disrupt the workflow of workers, potentially causing interruptions or distractions from the primary tasks. Furthermore,

* Corresponding author.

E-mail addresses: amito2@illinois.edu (A. Ojha), shayan@illinois.edu (S. Shayesteh), yliu580@syr.edu (Y. Liu), hjebelli@illinois.edu (H. Jebelli).

several proactive strategies such as development of work-rest schedules [31,32], and the use of cooling vests [33–35] are also being adopted to prevent heat-related illness and fatalities in the construction industry. Work-rest schedules are designed to limit heat exposure by alternating work and rest periods according to the physiological and metrological parameters, thereby reducing the duration of exposure to extreme temperatures [31,32]. While creating work-rest schedules for workers helps manage the amount of heat stress exposure by providing regular breaks and reducing heat stress exposure during peak times, they alone may not be sufficient due to the fluctuating environmental conditions and diverse job demands at construction sites. Cooling vests, on the other hand, provide direct cooling to the body, aiming to maintain core temperature within safe limits during heat exposure [33–35]. While effective in reducing core temperature, cooling vests cannot be worn for prolonged periods of time due to their tendency to become cumbersome and less effective as their cooling properties diminish over time [33]. In this vein, the objective, personalized, and timely monitoring of heat strain risks is necessary to guide the use of such interventions for optimal worker safety.

Recently, few studies have examined the potential of physiological responses to assess heat strain, a non-invasive, objective, and more accurate approach to protecting workers from heat stress exposure [28,36,37]. Several researchers have leveraged the non-invasive physiological sensors in conjunction with prevalent machine learning (ML) algorithms, to identify workers' bodily responses to heat – heat strain – based on their physiological signals [38–40]. While these investigations presented the potential to identify the risk of heat strain among the workers, their application for real-world safety application in the field is associated with several challenges. Firstly, the current studies leverage invasive techniques such as calorimetry, to elicit physiological responses from the users [40]. The use of invasive techniques is bound to interfere with the workers' ongoing work and thus is infeasible to be applied in the field for heat strain assessment. Secondly, the current studies use limited physiological responses to predict the likelihood of heat strain among the workers. For instance, Yi et al. [38] leveraged heart rate as a physiological metric to evaluate heat strain in occupational settings. While the use of heart rate allowed to effectively predict heat strain on job sites, the reactions to a stimulus can be represented through numerous physiological variations. Different individuals respond differently to the same stressors; even the same worker can react differently to similar stressors during multiple exposures. A specific understanding of physiological responses (i.e. using only heart rate to assess the risk of heat strain) is insufficient to provide an informative assessment of risk of heat strain exposure on field settings. Thirdly, most of the current studies account for the individual differences by solely considering personal biometric parameters such as height, weight, age, body mass index, etc. However, this reliance on static data does not adequately address the real-time physiological changes experienced by workers, thereby diminishing the effectiveness and relevance of the predictive models for personalized risk assessments. Lastly, the traditional ML classifiers leveraged by the current studies in tandem with the physiological signals, have difficulty in effectively decoding human physiological signals with high accuracy and stability due to intra- and inter-subject variability [41,42]. This variability challenges the foundational assumptions of traditional ML algorithms, which expect consistent distributions between training and test data. The shift in data distribution due to individual differences often leads to models that cannot consistently or accurately translate real-time physiological signals into actionable insights regarding heat strain [43].

To bridge these gaps, this study aims to develop an enhanced worker-centered heat strain monitoring framework for consistently and accurately estimating the likelihood of heat strain among the field workers using the multimodal physiological data collected from non-invasive wearable biosensors. In this regard, a high-fidelity immersive virtual reality (VR)-based environment was developed to simulate the common physically demanding tasks in heat-vulnerable occupations. More

specifically, a VR-based environment with a tangible user interface was developed to offer life-like experiences in a risk-free setting. The developed high-fidelity VR environments closely replicated the physical and sensory aspects of the actual workplace. This level of realism is vital because it ensures that the physiological responses of workers in the virtual environment closely mirror what they would experience in the real world [44]. To that end, the developed environment offered a virtual testbed for collecting high volumes of multimodal physiological signals, which include EDA, PPG, ST, as well as respiratory signals. This multimodal approach is vital for capturing the complex and individualized physiological responses of workers, addressing the limitations of previous studies that relied on limited physiological signals or static biometric data. Then, the collected multimodal physiological signals were used to train a physiological-sensing framework, EnsmTrBoost, which integrates ensemble learning and domain adaptation to accurately identify workers' heat strain from captured physiological signals with variability issues. Ensemble learning combines predictions from multiple ML classifiers, thereby reducing the likelihood of overfitting to specific data characteristics and increasing the overall accuracy. Domain adaptation, on the other hand, allows the framework to generalize better across different individuals and environmental conditions, a crucial aspect in the dynamic and varied real-world job site scenarios. This integration is a key differentiator from current heat strain prediction frameworks, contributing to a more robust and accurate prediction model. By being designed to work with multimodal physiological signals and their inherent variabilities, the EnsmTrBoost framework can make more accurate predictions by inferring physiological signals with variability issues.

2. Research background

2.1. AI-driven physiological sensing-based heat strain assessment

Extreme heat triggers physiological responses that are indicative of clinical syndromes arising from the body's inefficiency in dissipating heat. Physical fatigue, high perspiration, elevated body temperature, increased heart rate, and severe blood pressure reduction can be linked to a noticeable decrease in bodily fluid during heat stress exposure [45,46]. In such a case, the physiological symptoms manifest as excessive perspiration, increased heartbeat, and reduced blood volume [47]. Based on these clinical symptoms of heat stress, several physiological responses such as EDA [48], which measures skin's electrical changes related to sweat production, PPG [49], which tracks blood volume changes due to heart activity, and ST [50], which monitors skin temperature, have been studied. If interpreted correctly, these physiological signals can become proper indicators of the worker's heat-strain vulnerability [8,51–54]. In the study by Yi et al. [38], heart rate was utilized as a physiological metric to evaluate heat strain in occupational settings. The study introduced a conventional multilayer perceptron model for an early-warning system to detect heat strain. This system used Artificial Neural Networks (ANN), to estimate workers' Rate of Perceived Exertion (RPE). Additionally, a contingent process was implemented to integrate the heart rate data with the estimated RPE, enabling the timely issuance of heat strain warnings to workers. However, the study's limitation lies in its sole reliance on heart rate as the metric for assessing heat strain at job sites. Given that responses to heat stress can manifest through a range of physiological changes, understanding heat strain merely through heart rate data may not yield a fully comprehensive assessment of the risks associated with exposure in the field. Further, a prominent constraint associated with traditional multilayer perceptron model lies in their inherent "black box" characteristic, which can yield predictions seemingly at odds with established domain expertise [55]. In another study, Chan et al. [40] developed a multiple linear model (MLR) to predict worker's heat strain based on the diverse physiological, work-related, environmental, and personal parameters. While the study presented potential to identify the risk of heat

strain among the workers, the study leverages invasive calorimetry sensors to elicit workers physiological responses, which includes, oxygen consumption, minute ventilation, respiratory exchange ratio, metabolic equivalent, energy expenditure, and heart rate during their ongoing work [3]. The use of invasive techniques is bound to interfere with the workers' ongoing work and thus is infeasible for real-world application. In another study, Shakerian et al. [39] proposed an AI-driven physiological sensing-based framework to evaluate workers' heat strain continuously. In the developed process, the researchers leveraged off-the-shelf wearable sensing devices to capture EDA, PPG, and ST signals and concurrently decoded the captured signals through several prevalent ML algorithms for assessing workers' heat strain in the job site [39]. While the investigation established an AI-driven physiological sensing-based heat strain assessment framework to prevent heat-related illnesses by identifying workers' heat strain, the data collection of physiological readings from the human subjects for developing the heat strain assessment framework occurred in a lab-controlled environment with hypothetical task scenarios. However, the lack of realism in task scenarios is certain to impact the human subjects' behavior and affect the physiological readings [44,56,57]. In this vein, the framework's ability to provide accurate results and predictions may be compromised, given its reliance on data collected under conditions that do not faithfully replicate real-world circumstances. To that end, there is a critical need to collect a high volume of physiological data from workers under realistic working conditions to develop robust predictive models of heat strain.

Likewise, to ensure the efficiency and accuracy of physiological sensing-based data-driven framework, a fundamental requirement is to accurately decode physiological signals into meaningful information. Towards that end, several studies (including the study by [39], of developing an AI-driven heat strain assessment framework) leveraged prevalent ML classifiers, including, ANN, Support Vector Machines (SVM), k-Nearest Neighbor (kNN), random forest (RF), etc. as decoders in physiologically based data-driven models for objectively discerning different levels of workers' physiological states [39,58–62]. In this regard, the accuracy of data-driven models heavily relies on the classification accuracy of the ML classifier. However, such ML classifiers have difficulty in effectively decoding human physiological signals with high accuracy and stability due to intra- and inter-subject variability [41,42]. The intra-subject variability refers to the variation in distribution of the elicited physiological signal over time due to its non-stationary nature [41]. Likewise, the inter-subject variations refer to the variation in distribution among signals elicited from different subjects due to the inherent individual differences in anatomy and physiology [63]. Taken together, the variability issues in physiological signals violate the foundations of the traditional ML algorithm [42,64,65]. More specifically, the training data on which the ML classifier is trained, and test data, on which the ML classifier is consequently applied should have the same distribution. However, the intra- and inter-subject variability causes a shift in the distribution of test data from that of the training data. In this regard, the ML classifiers leveraged in the previous studies for developing heat strain assessment framework are unable to decode workers' real-time physiological signals consistently and accurately into meaningful information for workers' heat strain. Such decoding problems are bound to impede the accuracy of heat strain assessment in the field.

2.2. Potential of domain adaptation for addressing variability issues in physiological sensing-based data-driven models

In physiological sensing-based data-driven models, domain adaptation can improve the performance and generalization of the models by reducing the variability in the data and making the model more robust to changes in the data. In the context of this study, "domain" refers to diverse workers and conditions under which they operate, encompassing factors like individual differences (age and body mass index),

different field works, and varying working environments. These factors shape an individual's unique physiological response to heat stress, thereby constituting different "domains". Domain adaptation involves using a model trained on one dataset (the "source domain") and adapting it to work on a different dataset (the "target domain") [66]. Currently, three major approaches to achieve domain adaptation are widely explored, which include: (a) incorporating domain specific information [67]; (b) adversarial training [68]; and (c) transfer learning [69].

The first method to achieve domain adaptation is by incorporating domain-specific information, such as demographic details, medical history, or task-specific information. This approach can allow physiological sensing-based data-driven models to be more attuned to the nuances of physiological data, thereby improving the model's performance in the target domain. However, the addition of domain-specific information can inadvertently introduce or amplify biases present in the data, potentially leading to misdiagnoses [70]. Specifically, if the incorporated information is not fully representative, the model may perform poorly on underrepresented groups [70]. Furthermore, adding more information increases the model's complexity, which can lead to overfitting, especially if the additional data is not carefully curated and relevant [71]. Likewise, the incorporation of domain-specific information, such as demographic details and medical history, to achieve domain adaptation in physiological sensing-based data-driven models is bound to raise concerns about data privacy.

The second method of achieving domain adaptation is through adversarial training [72]. Adversarial training uses a loss function for a model to learn the domain-invariant features of the data by training a domain discriminator in parallel with the main model and using it to distinguish between the source and target domains [68,73]. Adversarial training can thus impart robustness to data-driven models based on physiological sensing against intra- and inter-subject variability by learning domain-invariant features. However, physiological data often contains complex, non-linear patterns [74], and adversarial training might struggle to learn the complex relationships between the source and target domains, potentially leading to suboptimal performance [75]. Specifically, the training process can be unstable, particularly with complex physiological data, leading to difficulties in model convergence. Furthermore, adversarial training requires a large amount of labeled data from the target domain [76], especially challenging when dealing with physiological data, where collecting large amounts of labeled datasets is difficult.

As aforementioned, both approaches to achieving domain adaptation—incorporating domain-specific information and adversarial training—face certain limitations. In contrast, the third method, transfer learning can be an effective method for addressing variability issues in physiological sensing-based data-driven models. In transfer learning, a model trained in one domain is fine-tuned for a new domain [69]. For instance, a model initially trained on a physiological dataset with a specific statistical distribution can be fine-tuned on another dataset with varying statistical distributions. By adapting the model to the unique characteristics of the new domain, transfer learning techniques can manage variability issues, improving the performance and generalization of data-driven models, and enhancing their robustness to data variations [72]. Transfer learning involves fine-tuning a pre-trained model on a smaller, target dataset. Consequently, the model is less likely to overfit the specificities of the target dataset compared to one trained from scratch and incorporating domain-specific information. Moreover, transfer learning entails a more stable and straightforward fine-tuning process [66], a significant advantage when dealing with complex physiological signal data. Additionally, transfer learning generally requires fewer labeled examples from the target domain to be effective [77]. More specifically, transfer learning enables the model to leverage knowledge from a source domain, which can be useful to improve the model performance in a new domain, even if the labeled data is not available. In summary, transfer learning can significantly

improve generalization to new individuals and environments and address data variability issues. This is particularly crucial in physiological sensing-based data-driven models, where data can be highly variable due to individual and environmental differences.

3. Research methodology

The research methodology is mainly orchestrated through two major key modules. In the first module, an interactive, immersive virtual testbed was generated for a realistic simulation of heat-vulnerable field occupation in a risk-free setting. The developed immersive environment offered a virtual testbed for collecting EDA, PPG, ST, and respiratory signals and consisted of two steps: 1) immersive virtual environment and scenario development, in which a common field task (logging task) with a chainsaw was simulated; and 2) tangible user interface development, in which an actual chainsaw was integrated with the simulated immersive scenario to provide a life-like fieldwork experience. In the second module, a physiological-sensing framework, EnsmTrBoost, was developed to identify workers' heat strain from collected physiological signals with variability issues. Notably, the developed network architecture featured two steps: 1) the ensemble learning step, in which the collected physiological signals were processed to remove artifacts, extract informative features, and generate an initial ensemble classifier; and 2) the domain adaptation step, in which the transfer learning technique is employed to calibrate the initial ensemble classifier, enabling the classifier to decode physiological signals whose statistical distribution differed from the data used to train and generate the initial ensemble classifier, and accurately identify the workers' heat strain. Fig. 1 provides an overview of the research methodology. Detailed explanations of the various steps within this methodology are elaborated upon in the subsequent subsections.

3.1. Virtual testbed for realistic simulation of heat-vulnerable field occupation

3.1.1. Immersive environment and scenario development

In the first module of the developed framework, a high-fidelity virtual testbed was developed to simulate common physically demanding tasks in heat-vulnerable occupations to provide a high-quality physiological dataset from workers under realistic working conditions. To that end, an immersive virtual environment was leveraged to simulate common physically demanding tasks (logging task) that field workers may experience on the job site. Logging task was specifically chosen due to its significant relevance to timber construction, which constitutes

about 92% of the U.S. residential construction sector [78], and its embodiment of the intense work experienced across various heat-vulnerable field occupations. This simulation involved the design and development of a user-centered, simulated workspace to safely examine different field-oriented tasks. Therefore, an outdoor workspace was simulated and rendered in VR using the Unity game engine, as shown in Fig. 2-A. The creation of various components within the scene was done using commercially available 3D modeling tools (i.e., Blender) as well as game engine built-in functions. In addition, the simulation incorporated various 3D elements obtained from the Unity Asset Store. These components were treated as Game Objects within the Unity Editor, facilitating the inclusion of collision meshes, scripts, and physical properties to different elements. The game engine also provided the capability to enhance the scene with realistic textures, sounds, and lighting effects. Furthermore, a virtual representation of a chainsaw was integrated into the scene to serve as the primary tool for interaction within the designed scenario. To enhance the realism of the physical logging tasks, the simulation integrated haptic feedback to replicate the resistance and vibration felt during logging, using the virtual chainsaw as the primary tool for interaction within the virtual setting. Furthermore, different components of the virtual workspace, such as the terrain, tools, and equipment, were combined to create a more immersive user experience. To enhance the user's sense of presence and spatial awareness in the scene, several soundtracks related to different situations (e.g., idle working of tools, cutting woods) were added to different Game Objects that can be triggered in certain conditions. Additionally, spatial audio techniques were employed to simulate the ambient sounds of the work environment and noises of a logging site, such as the roar of chainsaws and noise generated from cutting the wood, further immersing the user in the environment. This auditory experience, coupled with tactile feedback, considers the resistance, vibration, and noise typical to logging tasks. The representative nature of simulated task ensures the study addresses a broad spectrum of field workers susceptible to heat stress, enhancing its applicability and relevance. In addition to the immersive environment, a tangible user interface was developed for user interaction within the developed scenario. More details about the interface are elaborated in the following section.

3.1.2. Tangible Interface development

Considering the physical demand of the selected field-oriented task, an actual chainsaw (without the chain and powered off) was integrated with the simulated immersive scenario through a tangible user interface. The tangible interface setup provides both visual and kinesthetic feedback that enhances the realism of the simulation, allowing users to experience its weight and resistance mimicking the handling of an operational tool. As shown in Fig. 2-B, a motion-tracking sensor (i.e., Vive tracker) was attached to the physical chainsaw that is synchronously rendered in the virtual environment. The motion-tracking sensor of choice features an IMU sensor that can roughly calculate the location of the device. The selected VR system (i.e., HTC Vive Pro) also included two laser emitters that can compensate for any errors made by IMU sensors using the variations in response times of the photodiodes on the tracker. The position and orientation of the virtual chainsaw in the virtual setting matched the corresponding ones of the physical chainsaw in the real world, which is essential for accurate replication of physical interactions within the VR environment. Additional vibration feedback was programmed to simulate the interaction between the chainsaw and the wood, thereby mimicking the operational vibrations typical to real-world scenarios. The provided visual and force feedback enabled the users to intuitively interact with the chainsaw and feel its weight, resistance, and vibration throughout the task performance.

To simulate the cutting effect, a mesh slicing technique was applied in the Unity game engine. More specifically, a script based on the collision detection between the cutting tool (i.e., chainsaw) and an object (i.e., wooden log) was developed in this study. The mesh slicing technique, implemented through Unity game engine scripts, simulated

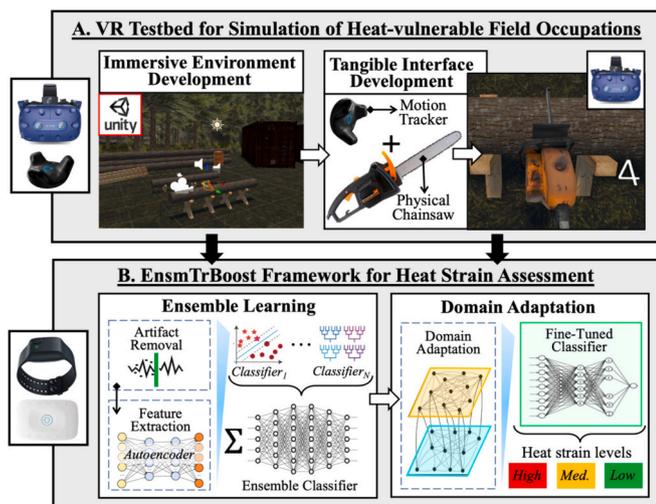


Fig. 1. Research methodology.

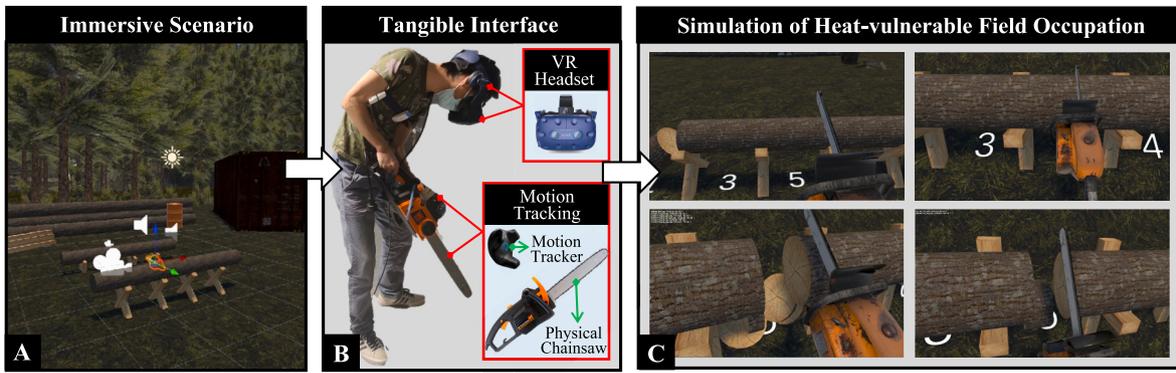


Fig. 2. Development of virtual testbed: (A) Immersive scenario development; (B) Tangible interface development; and (C) Realistic simulation of heat-vulnerable field occupation.

the cutting effect, offering a realistic tactile response that varied by wood density, in tandem with the real-world physical feedback. The cutting starts with the intersection of the chainsaw with the mesh triangles of the object. To that end, an invisible plane was attached to the cutting tool which hit the other object. Once the cut happened, the mesh triangles that are not intersected were divided into two groups and copied into one of the resulting objects from the cut. For intersected triangles, the plane divided all intersected edges into two groups, each on one side of the plane. As a result, three new vertices (one triangle with an original vertex and two new vertices and two triangles with two original vertices and a new vertex) were generated from each intersected triangle. Furthermore, since the newly generated meshes were hollow at the cutting point, new triangles were created from new vertices and the average position of all new vertices (center) and added to both resulting objects. The cutting could be done multiple times on an object, allowing the user to repeat the task for extended periods of time. This process integrated the digital work environment with the physical tools to create an immersive testbed where the risk of using actual tools can be mitigated. Fig. 2-C represents the generated simulations of the virtual

testbed for collecting EDA, PPG, ST, and respiratory signals.

3.2. EnsmTrBoost framework for heat strain assessment

To enhance the performance of machine learning (ML) classifiers in identifying workers' heat strain from captured physiological signals (EDA, PPG, ST, and respiratory signals) with variability issues, the authors developed a physiological sensing framework, EnsmTrBoost. The developed framework was featured in two steps, presented schematically in Fig. 3. The first step, the ensemble learning step, was developed to train an ensemble ML classifier to generate an initial estimation of workers' heat strain from physiological signals. In this step, the physiological signals were first processed to remove signal noises and accordingly extract informative features. Then, the extracted features were used to train and generate an initial ensemble ML classifier for preliminary identification of workers' heat strain. The second step of EnsmTrBoost, the domain adaptation step, was developed to employ the transfer learning technique to calibrate the initial ensemble ML classifier, enabling the classifier to decode physiological signals whose

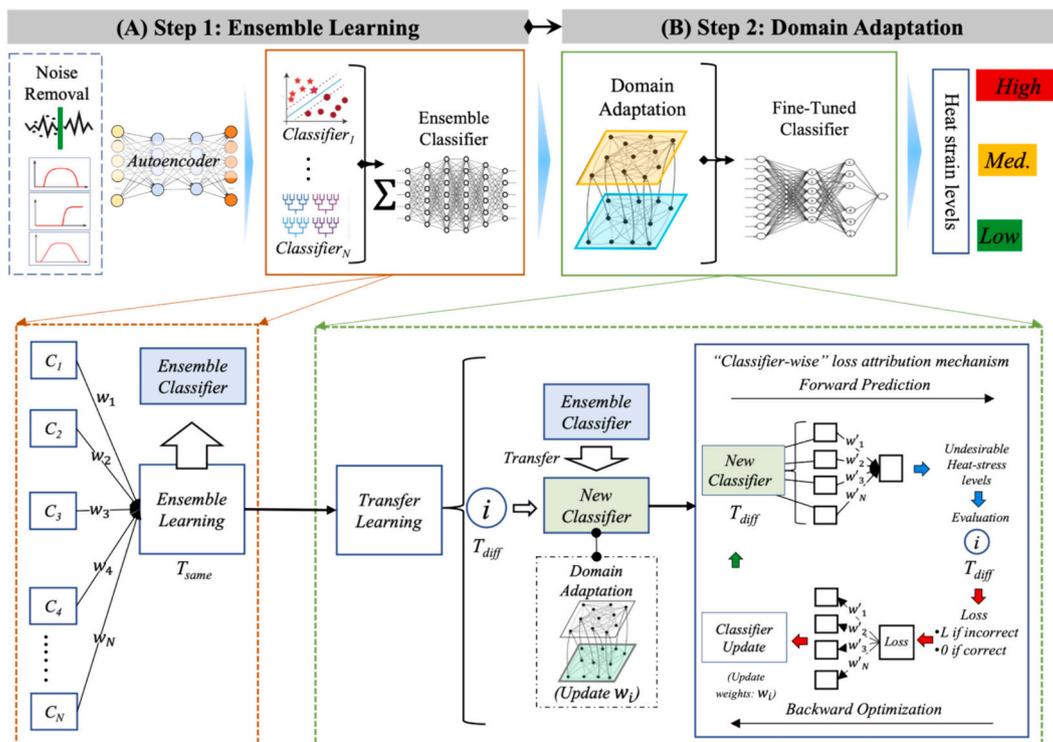


Fig. 3. Overview of the developed EnsmTrBoost framework: (A) Step 1: Ensemble learning; and (B) Step 2: Domain adaptation.

statistical distribution differed from the data used to train and generate the classifier (due to individual differences). In the developed EnsmTrBoost, the ensemble algorithm was leveraged to improve the performance of the ML-based classifier in decoding physiological signals with inter-subject variability issues [79,80]. The transfer learning technique was implemented to enable the ML-based classifier to resolve the intra-subject variability issues in physiological signals, as demonstrated in [41,65]. Using these two steps, the developed EnsmTrBoost can generate an ensemble ML classifier to enhance the performance in identifying workers' heat strain from their physiological signals with variability issues. The EnsmTrBoost framework specifically accounts for individual differences by calibrating the ML classifier to recognize and adapt to the physiological variations among workers, ensuring personalized and accurate assessments of heat strain. The details of each step will be presented in the following subsections. The pseudo-code of the EnsmTrBoost paired with the developed EnsmTrBoost framework is shown in Algorithm 1. Furthermore, to generate the ensemble classifier using the EnsmTrBoost framework, the dataset used in the first step of EnsmTrBoost is denoted as T_{same} , where all data for each category of physiological signals (EDA, PPG, ST, or respiratory signal) have the same statistical distribution. x_i ($i = \{n + 1, \dots, n + m\}$) indicates a data point in T_{same} . T_{diff} represents the dataset used in the second step, in which the distribution of the data is different from the data in T_{same} . x_i ($i = \{1, \dots, n\}$) represents a data point in T_{diff} . The acquisition of datasets T_{same} and T_{diff} will be introduced in Section 4.

3.2.1. Step 1 of the EnsmTrBoost framework – Preprocessing and ensemble learning step towards the development of the initial ensemble ML classifier

3.2.1.1. Noise removal.

The physiological signals (i.e., EDA, PPG, ST, and respiratory signals) collected from the wearable biosensor can be contaminated by various types of noise in field settings. These noises include the electrode noises caused by contact between electrodes and skin [81], the motion artifacts caused by a myriad of manual activities performed by the workers [82], electromagnetic noises caused by the ambient electromagnetic fields of the applied biosensors [39], and the thermal noises and electro-surgical noises [83]. Accordingly, the presence of noises will adversely affect the quality of signal collections and the following interpretations of the signals using ML-based classifiers [82,84,85]. To reduce these signal noises and thus improve the performance of the ML-based classifier in identifying workers' heat strain, the authors implemented several noise-removal techniques, substantiated by the previous investigations [58–62,86]. For EDA signals, a lowpass filter with a 1.5 Hz cutoff frequency was leveraged to target the typical low-frequency range of EDA [87]. The chosen frequency for noise removal has been documented for its success in eliminating artifacts in EDA signals [58–62,86]. ST signals were processed using a high pass filter with a 0.05 Hz cutoff frequency to eliminate low-frequency artifacts, a technique validated in previous studies [58–62,86]. Additionally, PPG signals, which typically range from 0.5 to 5 Hz PPG [49], was leveraged with a bandpass filter of corresponding cutoff frequencies. Likewise, additional techniques, including double median filtering and adaptive filtering techniques (i.e., Fourier Linear Combiner), were employed to remove motion artifacts from PPG signals. Such methods have been validated by previous research for significantly improving signal quality and stability [88–90]. For the respiratory signal, informed by the previous investigation, the authors applied an adaptive filter with a 0.16–1 Hz bandpass filter to remove noises, including motion artifacts, electro-surgical noise, and electrode noise [83]. Further, to remove the motion artifacts in the signals, sparse recovery method was utilized [91]. The efficacy of this sparse recovery method in enhancing signal quality has been demonstrated in prior studies [86,91].

3.2.1.2. Feature extraction.

Once physiological signals have been denoised, the authors leveraged an autoencoder network to extract informative features from these signals to train an ML-based heat strain identification model [92–94]. In contrast to conventional feature extraction methods utilized in authors' previous studies [95,96], this autoencoder approach can (1) automatically extract features to train ML-based classifiers without requiring researchers to manually extract features; and (2) reduce the computational resources of the feature extraction [97]. Structurally, the autoencoder applied in this study was comprised of an encoder and a decoder – the encoder was used to extract features from the input signals, and the decoder was designed to evaluate the feature extraction performance of the encoder. The encoder of the autoencoder consisted of a three-layer 1-dimensional (1D) convolutional neural network (CNN). The stride of each 1D-CNN layer was set to 2; the activation function of each layer was LeakyReLU. The kernel size for each of the three layers was 16×1 , 8×1 , and 4×1 , respectively. The kernel number corresponding to each layer was 32, 64, and 128, respectively. The structure of the decoder was comprised of three deconvolutional neural network layers. Like the encoder, for each deconvolutional layer, the stride size was 2; the activation function was LeakyReLU. In addition, the kernel size of each deconvolutional layer was 4×1 , 8×1 , and 16×1 , respectively. The corresponding kernel numbers were 128, 64, and 32, respectively. Relying on this structure, the applied autoencoder was trained, and the output of the encoder of the trained autoencoder was regarded as the features to train the ML-based classifiers. Details of training the autoencoder (objective function, hyperparameters) can be found in the authors' previous research [97].

3.2.1.3. Initial ensemble classifier.

For each captured physiological signal in the dataset T_{same} – EDA, PPG, ST, or respiratory signal in T_{same} – once its features were extracted through the above autoencoder network, the features were first fed into a single ML classifier for heat strain identification of workers. As shown in Algorithm 1, C_1 , C_2 , C_3 , and C_4 represent the single classifier for the EDA, PPG, ST, and respiratory signals, respectively. Notably, in this study, the single classifier was selected from the following traditional ML classifiers, including kNN, SVM, SVM with Gaussian kernel, SVM with Polynomial kernel, Logistic Regression, Quadratic Discriminant Analysis (QDA), and MLP Neural Network. For a specific kind of physiological signal (EDA or PPG or ST or respiratory signal), the classifier (among these traditional ML classifiers) with the optimal performance in identifying the heat strain of workers was selected as the corresponding single classifier. After obtaining all single classifiers (C_s), the authors applied the ensemble learning algorithm to combine these classifiers to generate an ensemble classifier. The steps of generating the initial ensemble classifier were shown in the first four lines of the pseudo-code of the EnsmTrBoost framework (see Algorithm 1). As shown in lines 2 and 3, for each selected single classifier, $C_s; s \in \{1, 2, 3, 4\}$, the developed EnsmTrBoost first generated the corresponding ensemble weight, $w_s; s \in \{1, 2, 3, 4\}$. Line 3 of Algorithm 1 also showed these ensemble weights – w_1, w_2, w_3, w_4 – were normalized before being applied to the corresponding single ML classifier. Then, based on the ensemble weights ($w_s; s \in \{1, 2, 3, 4\}$), an initial ensemble classifier was generated from the single classifiers ($C_s; s \in \{1, 2, 3, 4\}$) using the equation shown in line 4 of Algorithm 1. As demonstrated in [79,98,99], this integration can reduce the bias or prediction variance of the single ML classifier and produce a robust classifier with low-variance prediction capability to identify heat strain across workers, thereby allowing the classifier to decode physiological signals with inter-subject variability issues.

Algorithm 1. Ensemble for transfer learning.

3.2.2. Step 2 of the EnsmTrBoost framework – Transfer learning for fine-tuning the ensemble classifier to accurately decode physiological signals with subject variability issues

Once the initial ensemble classifier was obtained, the next step of the EnsmTrBoost was to endow the classifier with the capability of decoding the physiological signals with intra-subject variability issues. To this end, the authors applied the domain adaptation technique, transfer learning, to update the initial ensemble classifier. The applied transfer learning rule is shown in lines 5 to 10 of Algorithm 1. For each transfer learning iteration, the initial ensemble classifier was employed to classify all data points x_i in T_{same} and T_{diff} datasets. As mentioned, x_i ($i = \{n+1, \dots, n+m\}$) indicates a data point belongs to T_{same} used to train and generate the initial ensemble classifier; x_i ($i = \{1, \dots, n\}$) indicates a data point in T_{diff} whose statistical distribution is different from the data point in T_{same} due to the variability issues. The classification error, ϵ_k , was calculated by using the equation shown in line 7 of Algorithm 1— where $h_k(x_i)$ represents the label (worker's heat strain) identified by the ensemble classifier, $c(x_i)$ is the true label of the data point x_i , and t_i^k indicates the weight of the difference between $c(x_i)$ and $h_k(x_i)$, initially set to 1. After calculating the classification error ϵ_k ($0 < \epsilon_k < 0.5$), ϵ_k was leveraged to generate a domain-adaptation parameter $\beta_k = \frac{1 - \epsilon_k}{\epsilon_k}$ (line 8). β_k was used to update the weight (t_i^k), as presented in line 9 of Algorithm 1. Then, t_i^k was used to multiply by the corresponding data point x_i . In this manner, the weights of the misclassified data points ($x_i; i = \{1, \dots, n\}$) in T_{diff} could be increased by $\beta_k^{|h_k(x_i) - c(x_i)|}$. Conversely, the weights of the misclassified data points ($x_i; i = \{n+1, \dots, n+m\}$) in T_{same} could be decreased by $\beta_k^{|h_k(x_i) - c(x_i)|}$. After N iterations of transfer learning (N is the user-defined parameter), the weights of the misclassified data points in T_{diff} would be larger than

those of the misclassified data points in T_{same} . Subsequently, the misclassified data points in T_{diff} and their weights would be integrated into T_{same} to generate an updated dataset, $T_{combine}$. From the ensemble classifier's perspective, $T_{combine}$ could train the classifier to focus more on learning from the data points with larger weights [64,65,80]. In other words, using $T_{combine}$, the ensemble classifier can learn more from data points that have different statistical distributions than the data used to train and generate it in the first step (Section 3.2.1). As such, the robustness of the ensemble classifier in identifying data points with shifts in the statistical distribution can be enhanced. In addition, the above-mentioned domain adaptation process to obtain $T_{combine}$ is shown in Fig. 4.

After obtaining the updated dataset $T_{combine}$, this dataset was applied to optimize the initial ensemble classifier by fine-tuning the initial ensemble weights, w_i ($i \in \{1, 2, 3, 4\}$). Lines 11 to 15 of Algorithm 1 illustrate the fine-tuning process. As expressed in line 11, the updated dataset was employed to allow the initial ensemble classifier featured with the initial ensemble weights (w_1, w_2, w_3, w_4) to calculate the overall classification loss, L . Based on each data point in $T_{combine}$, this loss function leveraged norm_2 distance to measure the difference between the identification result generated by the ensemble classifier and the ground truth of the data. Subsequently, the overall loss, denoted as L , was allocated to each single classifier ($C_i; i \in \{1, 2, 3, 4\}$) of the ensemble classifier, based on the performance of each classifier in classifying the datapoint in $T_{combine}$. This allocation mechanism is highlighted in line 12 of Algorithm 1. To elaborate further, in cases where a datapoint is misclassified, the corresponding loss is uniformly distributed among the single classifiers that inaccurately classify such datapoint (in line 12, the parameter δ was used to prevent the denominator from becoming zero). Conversely, single classifiers that accurately classify the datapoint are exempt from being allocated the loss

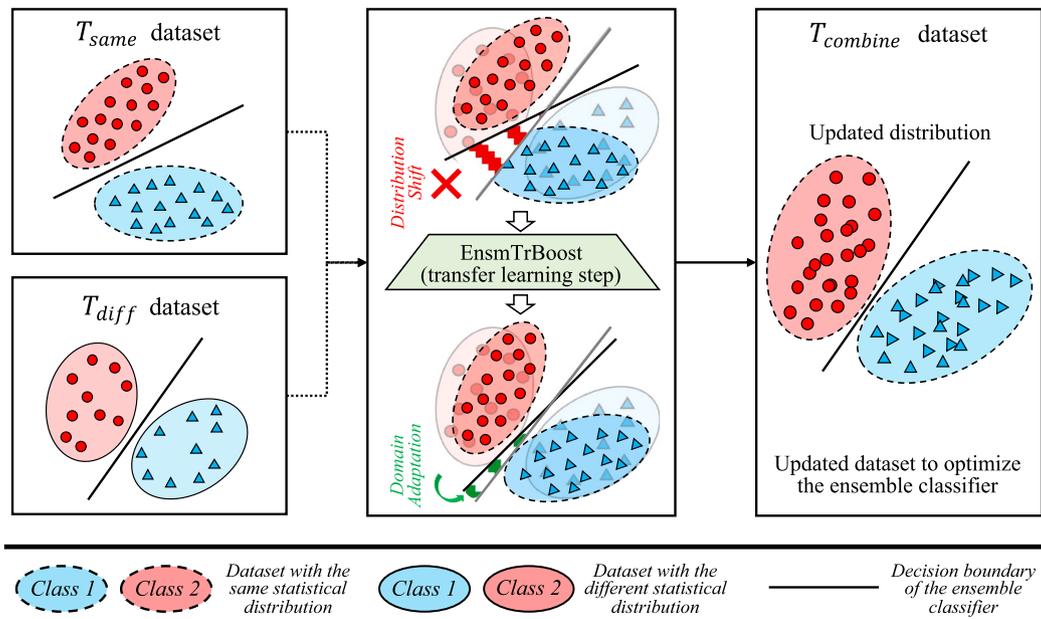


Fig. 4. Illustration of the domain adaptation process.

related to such datapoint. According to the loss assigned to each single classifier, the fast proximal gradient method (FPGM) was applied to optimize the ensemble weight of every single classifier (line 13). More details of the FPGM optimization method can be found in [100]. The optimized weights w'_1, w'_2, w'_3, w'_4 were regarded as the new ensemble weights to update the ensemble classifier (lines 14 and 15). Notably, after obtaining the updated ensemble classifier ($w^{p+1} = [w'_s, s = 1, 2, 3, 4]; p + 1$ stands for the new iteration), the EnsmTrBoost framework returned to line 2 of Algorithm 1 and started a new iteration. For a new iteration, the initial weights of the ensemble classifier were the ensemble weights w'_1, w'_2, w'_3, w'_4 updated from the previous iteration and normalized using the

equation shown in line 3. Furthermore, the iteration will be repeated until each ensemble weight of the ensemble classifier converges to a constant value and the overall loss of the ensemble classifier no longer changes. The ensemble classifier generated from the final iteration is expected to have enhanced performance in decoding the physiological signals (EDA, PPG, ST, and respiratory signals) with intra- and inter-subject variability issues. The performance of the above EnsmTrBoost framework will be examined in the next section.

4. Performance assessment of EnsmTrBoost framework

In order to assess the effectiveness of the developed worker-centered

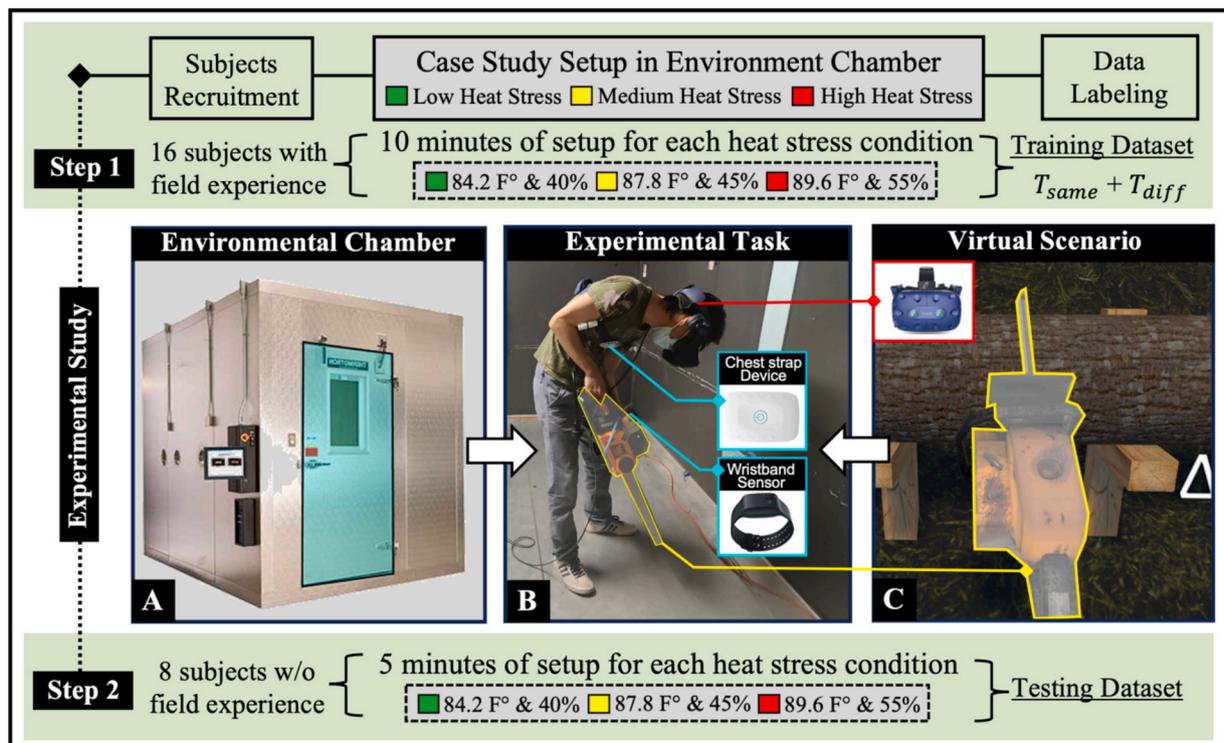


Fig. 5. Schematic diagram of experimental study.

heat strain monitoring framework, an experiment was devised and implemented within a controlled environmental chamber. This experiment comprised two distinct steps. The first step was designed to collect the workers' physiological data during a physically demanding task (forestry task) in the immersive VR setting under three heat stress exposure conditions. Concurrently, Kullback-Leibler (KL) divergence [101,102] was utilized to assess the similarity or dissimilarity in statistical distribution within the collected datasets. Based on the KL-divergence score, the original dataset was divided into datapoints with a particular statistical distribution (T_{same}), and dataset with a different statistical distribution than that of T_{same} (T_{diff}). The obtained dataset T_{same} and T_{diff} were leveraged to train the developed heat strain monitoring framework. Likewise, the second step of the case study setup was to collect the test dataset to evaluate the performance of the EnsmTr-Boost framework. Fig. 5 illustrates the schematic diagram of the experimental study.

4.1. Step 1: Collecting training dataset to generate the ensemble classifier

4.1.1. Subjects and device information

Sixteen able-bodied subjects (mean age: 25; SD of age: 2.6; mean weight: 190 lb.; SD of weight: 27 lb.; mean height: 5' 11"; SD of height: 1.7") with proper experience of field occupation were recruited to participate in the experiments. Before the experiments, all the subjects were provided with informed consent forms explaining the purpose of this study, a comprehensive explanation of the data collection process, confidentiality of the data, and subjects' rights. Prior to the experiments, the authors ensured that participants had no medical history that might impact their performance in the immersive virtual environment within a controlled environmental chamber. Additionally, subjects were instructed to be adequately hydrated before the commencement of the experimental trial.

After the informed consent forms were signed, subjects were asked to wear a wristband-type wearable sensor (E4 Empatica™) and a chest-strap-type wearable device (bioPLUX™ Inductive Respiration Sensor). While the EDA, PPG, and ST signals show promise in accurately deducing and monitoring workers' heat strain, it is important to note that the body's response to different stimuli is not solely confined to these indicators. When an individual is exposed to heat stress conditions, the body employs thermoregulation as a protective mechanism, facilitating heat dissipation [103]. In order to uphold its proper functioning, the body necessitates the maintenance of a core temperature of 37 °C, ideally within a range of ± 1 °C [104]. To maintain a balanced body temperature, which hinges on the exchange of heat between the body and its surroundings, adjustments in breathing patterns can lead to heat loss [105]. Such respiratory patterns can be monitored through respiration sensors. To that end, this study attempts to incorporate respiratory signals along with the EDA, PPG, and ST signals in the developed framework. Physiological responses are multifaceted and different workers may exhibit unique response patterns even under similar environmental conditions. The use of a multimodal physiological data method is crucial for capturing individual differences in physiological responses, ensuring personalized and accurate prediction of heat strain risks.

The wristband-type wearable sensor was equipped with a PPG sensor, an EDA sensor, and an infrared thermopile sensor (for skin temperature), collecting physiological signals at frequencies of 64, 4, and 4 Hz, respectively. Likewise, the chest-strap-type wearable device included an inductive-type respiration sensor and triaxial accelerometer to collect the respiratory signals with the frequency of 100 Hz. For further analysis, EDA and ST signals were up-sampled to 32 Hz, whereas PPG and respiratory signals were down-sampled to 32 Hz. Additionally, participants were provided with an HMD VR headset (HTC Vive Pro Eye) and motion tracking sensors (Vive Trackers and Leap Motion) alongside the physiological sensors.

4.1.2. Data acquisition

After the setup of wearable physiological sensors and the VR system, participants were tasked with performing the logging exercise within the immersive testbed. This took place in a controlled environmental chamber under three distinct climatic conditions. To start, participants were given a brief period to familiarize themselves with the simulated immersive environment, acquainting themselves with features like the interaction mechanism and the task at hand. Once participants felt ready, they were then instructed to commence the logging task within the virtual testbed, all while still within the controlled environmental chamber. The logging tasks consist of subjects cutting lumber using a chainsaw. To emulate the physically demanding nature of the selected field-oriented task, subjects were required to emulate the cutting-a-lumber task through an actual chainsaw (without the chain and powered off), which was integrated with the virtual testbed through a tangible user interface. Participants were tasked with performing the logging exercise under three distinct conditions, which were controlled by varying levels of ambient temperature and humidity.

To mitigate selection bias and account for any potential trial-order effects, subjects were grouped in sets of three, and each set was assigned a unique sequence of climatic conditions (resulting in six distinct sequences for three conditions). Environmental parameters such as humidity and ambient temperature were regulated within a climate-controlled chamber room to ensure consistency across data collection.

As previously mentioned, the three different climatic conditions were emulated by regulating both the temperature and humidity levels within the chamber room. According to OSHA guidelines for heat stress, three distinct climatic conditions were simulated by adjusting the temperature and humidity of the chamber room to 84.2 °F and 40% for low, 87.8 °F and 45% for medium, and 89.6 °F and 55% for high heat stress, respectively [25]. All experimental trials were conducted in a chamber room. For each climatic condition, subjects were asked to perform the logging tasks for 10 min with 10 min break in between. The duration of the tasks across the climatic conditions was chosen based on the capability to induce heat strain within a safe and controlled timeframe. Research indicates that the short exposures to high temperature and humidity also can significantly increase heart rate and core body temperature, cumulatively leading to heat strain and dehydration risks [106–110]. To that end, the duration of the experimental tasks was carefully formulated to collect the necessary physiological data while minimizing the health risks to participants. Additionally, a medical consultant with extensive expertise in human physiology was present throughout the experimental setup. The consultant confirmed that the task duration was sufficient to elicit measurable physiological responses indicative of medium and high heat strain without jeopardizing participant well-being. To ensure that the experimental duration could induce different heat strain levels, a reliable heat strain assessment technique called the Heat-Strain Score Index (HSSI) questionnaire [26] was used. HSSI evaluates the heat strain score index, providing a classification of the subjects' perceived heat strain. This classification includes low heat stress exposure ($HSSI < 13.5$), medium heat stress exposure ($13.6 < HSSI < 18$), and high heat stress exposure ($18 < HSSI$). After each trial (and during the break in between), subjects were asked to fill out the HSSI questionnaire to indicate their perceived heat strain level for the task. The responses from the questionnaire consistently indicated that subjects experienced medium to high levels of heat strain. To that end, the responses corroborated the intended effects of the experimental setup, which was designed to expose subjects to different heat strain intensities, ranging from low to high. Throughout the tasks, the subjects' wearable biosensors continuously recorded all physiological data, including PPG, EDA, skin temperature (ST), and respiratory signals. This data was subsequently uploaded to a cloud server for further analysis.

4.1.3. Data labeling

Throughout all the sessions, the physiological data collected were categorized based on three distinct climatic conditions, each designed to

Table 1
Overview of participants' collected data and the labeled data.

Subject	Temperature (Humidity)	HSSI score ¹	Data Size ²	Label ³	Subject	Temperature (Humidity)	HSSI score ¹	Data Size ²	Label ³
1	84.2 °F (40%)	12.6	76,800	L	9	84.2 °F (40%)	13.1	76,800	L
	87.8 °F (45%)	17.4	76,800	M		87.8 °F (45%)	17.5	76,800	M
	89.6 °F (55%)	20.1	76,800	H		89.6 °F (55%)	20.8	76,800	H
2	84.2 °F (40%)	11.1	76,800	L	10	84.2 °F (40%)	12.2	76,800	L
	87.8 °F (45%)	16.8	76,800	M		87.8 °F (45%)	16.4	76,800	M
	89.6 °F (55%)	22.2	76,800	H		89.6 °F (55%)	22.7	76,800	H
3	84.2 °F (40%)	13.2	76,800	L	11	84.2 °F (40%)	12.9	76,800	L
	87.8 °F (45%)	18.2	76,800	Omit		87.8 °F (45%)	17.2	76,800	M
	89.6 °F (55%)	21.7	76,800	H		89.6 °F (55%)	21.8	76,800	H
4	84.2 °F (40%)	12.8	76,800	L	12	84.2 °F (40%)	11.7	76,800	L
	87.8 °F (45%)	15.9	76,800	M		87.8 °F (45%)	16.3	76,800	M
	89.6 °F (55%)	21.9	76,800	H		89.6 °F (55%)	22.4	76,800	H
5	84.2 °F (40%)	13.4	76,800	L	13	84.2 °F (40%)	12.5	76,800	L
	87.8 °F (45%)	17.2	76,800	M		87.8 °F (45%)	17.3	76,800	M
	89.6 °F (55%)	19.9	76,800	H		89.6 °F (55%)	21.9	76,800	H
6	84.2 °F (40%)	12.8	76,800	L	14	84.2 °F (40%)	13.4	76,800	L
	87.8 °F (45%)	16.8	76,800	M		87.8 °F (45%)	17.7	76,800	M
	89.6 °F (55%)	22.7	76,800	H		89.6 °F (55%)	21.6	76,800	H
7	84.2 °F (40%)	12.3	76,800	L	15	84.2 °F (40%)	13.1	76,800	L
	87.8 °F (45%)	15.9	76,800	M		87.8 °F (45%)	16.7	76,800	M
	89.6 °F (55%)	23.1	76,800	H		89.6 °F (55%)	20.9	76,800	H
8	84.2 °F (40%)	11.9	76,800	L	16	84.2 °F (40%)	12.9	76,800	L
	87.8 °F (45%)	15.4	76,800	M		87.8 °F (45%)	17.4	76,800	M
	89.6 °F (55%)	21.6	76,800	H		89.6 °F (55%)	20.6	76,800	H

¹ HSSI score label: (HSSI <13.5): low heat strain; (13.6 > HSSI >18): medium heat strain; (18 < HSSI): high heat strain.

² Data Size: [32 datapoints/s]: 19,200 datapoints for each physiological signal.

³ L: low heat strain; M: medium heat strain; H: high heat strain.

represent varying levels of heat stress exposure (low, medium, and high). To ensure accuracy, HSSI [26] was used in post-processing phase to identify and rectify any potential mislabeling of datasets.

After each trial, subjects were asked to fill out the HSSI questionnaire to indicate their perceived heat strain level for the task. Each previously labeled dataset underwent a thorough examination to ensure consistent results. This additional verification step significantly enhanced the accuracy of the labeling process. Any dataset that did not yield similar outcomes was omitted from analysis. Table 1 provides an overview of the data size, HSSI scores, and final labels for the 16 subjects. The authors applied the KL divergence to quantitatively evaluate the dissimilarity between the collected datasets, considering their respective statistical distributions. For this purpose, probability distributions for each physiological signal under each label were computed based on mean, variance, and higher-order moments. Simultaneously, KL divergence was employed to compare the resulting probability distributions, leading to the classification of the dataset into two distinct classes; one with a particular statistical distribution (T_{same}) and another with dissimilar statistical distributions compared to T_{same} (T_{diff}).

4.2. Step 2: Collecting test dataset to examine the EnsmTrBoost framework

Eight able-bodied subjects with no experience of field occupation were recruited to participate in the second step. Each subject was equipped with the same set of equipment, including a wristband-type wearable sensor, chest-strap-type wearable device, HMD VR headset,

Table 2
Heat stress classification accuracy (testing) for single classifiers.

	EDA	PPG	ST	Respiration
kNN (k = 3)	54.9%	51.9%	67.3%	66.9%
Logistic Regression	68.1%	67.5%*	67.7%	73.5%
SVM (linear kernel)	60.6%	60.1%	63.8%	71.9%
SVM with Gaussian	70.3%*	66.9%	73.1%*	75.6%
SVM with polynomial	66.7%	63.2%	68.4%	72.2%
QDA	65.4%	57.6%	68.4%	71.7%
MLP	70.1%	67.0%	72.2%	80.2%*

and motion tracking sensors (Vive Trackers and Leap Motion). Subjects were instructed to perform the same logging task within a developed immersive testbed in a controlled environmental chamber. This task was carried out under three distinct climatic conditions, as in the first step. However, unlike the previous step, the subjects were only required to perform the logging tasks for 5 min, with a 5-min break in between. Throughout the tasks, the wearable biosensors continuously recorded all physiological data, which was subsequently uploaded to a cloud server. Additionally, the collected physiological data from all the sessions were categorized based on three different climatic conditions, specifically designed to induce low, medium, and high levels of heat stress. To ensure the varying levels of climatic conditions could induce different heat strain levels, the subjects were asked to complete the HSSI questionnaire after each trial. This questionnaire allowed the participants to indicate their perceived level of heat strain during the task. The results of the questionnaire suggested that the subjects experienced medium, and high levels of heat strain, validating the effectiveness of the experimental design in simulating different levels of heat strain.

5. Results

The authors applied the developed EnsmTrBoost framework to the data collected from the subjects. T_{same} was used in the first step of the EnsmTrBoost to generate the ensemble classifier. Since the training dataset has three labels, low, medium, and high heat stress, the baseline accuracy of each classifier was 33.3%. Correspondingly, the performance of the trained ML classifier was evaluated with the collected test dataset. Table 2 demonstrates the accuracy of the ML classifier in estimating the likelihood of workers' heat strain levels for each physiological signal trained using the five-folder cross-validation technique. As depicted in Table 2, SVM with Gaussian emerges as the leading classifier for EDA signal in predicting the likelihood of heat strain level with an accuracy of 70.3%. Likewise, Logistic Regression has the best accuracy of 67.5% for predicting the likelihood of heat strain from a PPG signal. Similarly, SVM with Gaussian is the top classifier for ST signal in predicting the likelihood of heat strain level with an accuracy of 73.1%. Furthermore, MLP has the best accuracy of 80.2% for predicting the likelihood of heat strain from respiration signals.

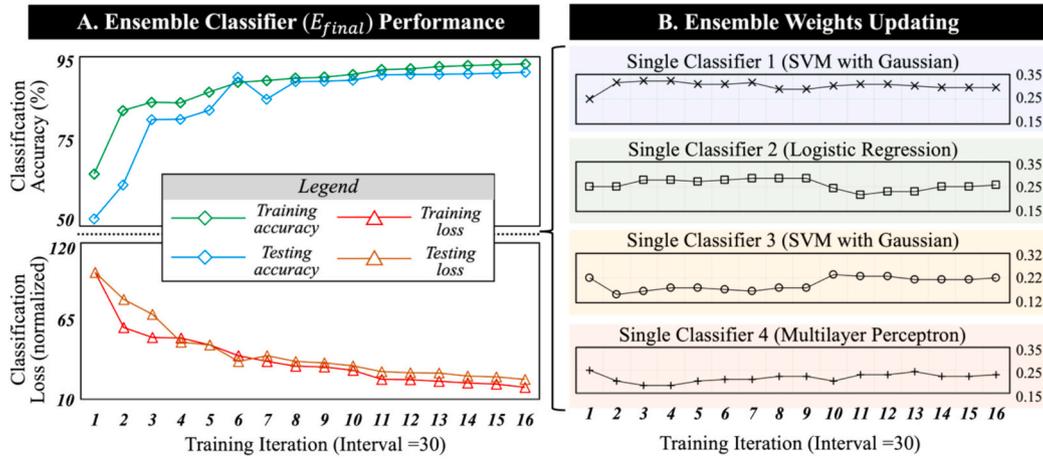


Fig. 6. Results of the experimental setup: (A) Ensemble classifier (E_{final}) performance; and (B) Updated weights of selected single classifiers over training iterations.

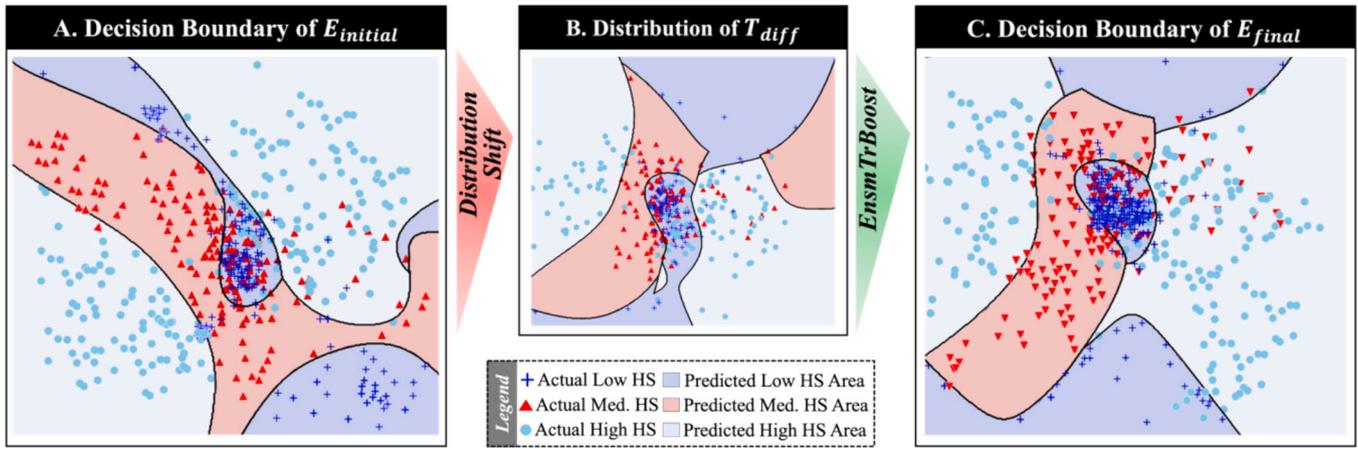


Fig. 7. Performance of EnsmTrBoost framework in addressing distribution shift issues: (A) Decision boundary of initial ensemble classifier ($E_{initial}$); (B) Statistical distribution of T_{diff} ; and (C) Decision boundary of final ensemble classifier (E_{final}).

For each category of the physiological signals, the authors selected the optimal single classifier for each signal to generate the initial ensemble classifier ($E_{initial}$) by following the procedure mentioned in Section 3.2. Using the ensemble rule, these single classifiers were used to generate $E_{initial}$. Consequently, the $E_{initial}$ was trained to develop an ensemble classifier (E_{final}) using the T_{diff} . The performance of E_{final} was evaluated using the test dataset. Fig. 6-A shows the training and testing performance of the E_{final} . The classification accuracy of the E_{final} is 93.2%, and the testing accuracy of the E_{final} is 91.7%. Fig. 6-A also shows the learning graph curves (training versus testing loss) where the red lines represent training loss, and the orange lines represent the testing loss. Training loss is an important indicator for monitoring the training of the E_{final} . The training was considered complete once the disparity between the training losses of two successive training epochs fell below the predefined threshold (i.e., 0.01). Every 30 iterations, the training loss and testing loss were tested and visualized to determine the stability of the training process. As seen in Fig. 6-A, the training loss and testing loss have decreased in the training process, whereas the testing loss is higher than the training loss. Moreover, the testing losses are in good agreement with their training losses, which means that the E_{final} fits the training dataset closely without overfitting. Further, the authors have also reported the ensemble weights of the selected single classifiers. Fig. 6-B illustrates the updated weights of the single classifier over the training iterations.

The authors also demonstrated the performance of EnsmTrBoost in calibrating the initial ensemble classifier to decode the physiological

signals with different statistical distributions. Fig. 7-A visualizes the decision boundary of $E_{initial}$. Likewise, Fig. 7-B demonstrates the statistical distribution of T_{diff} which statistical distribution differed from the T_{same} . More specifically, the data points collected in three pre-determined environment conditions and the corresponding statistical distribution of T_{diff} is represented in Fig. 7-B. Notably, these statistical distributions were depicted by leveraging Principal Component Analysis (PCA) to project the data into a two-dimensional space. A significant difference in distribution can be seen between the datasets in Fig. 7-A and Fig. 7-B caused by variability issues in the elicited physiological signals, as corroborated by KL divergence. As such, the $E_{initial}$ trained on data with a particular statistical distribution cannot accurately decode the dataset with a different statistical distribution. Following the procedure mentioned in Section 3.2, the $E_{initial}$ was trained to develop E_{final} . Fig. 7-C visualizes the decision boundary of E_{final} . The decision boundary is calibrated to match the statistical distribution of the T_{diff} by employing EnsmTrBoost framework. As shown in Fig. 7-C, the E_{final} has high potential in accurately identifying workers' heat strain from captured physiological signals with variability issues.

To assess the consistency in the performance of EnsmTrBoost for the dataset with different statistical distributions, the authors leveraged traditional ML algorithms alongside EnsmTrBoost to T_{diff} . The selected ML algorithms included SVM with Gaussian, Logistic Regression, and MLP, due to their effective performance for specific physiological signals, as demonstrated in Table 2. The authors evaluated the consistency by comparing the mean and standard deviation (SD) of the prediction

accuracy from the leveraged algorithms across the T_{diff} for each test subject. SVM with Gaussian achieved a testing accuracy of $78.6\% \pm 4.71\%$. Logistic Regression demonstrated a testing accuracy of $77.3\% \pm 6.29\%$. MLP achieved a testing accuracy of $84.3\% \pm 4.3\%$. In this vein, compared to the traditional ML approaches, EnsmTrBoost with a testing accuracy of $91.7\% \pm 1.43\%$ lead to consistent performance across datasets with different statistical distributions.

6. Discussion

This paper presents an enhanced worker-centered heat strain monitoring framework that utilizes physiological signals obtained from wearable biosensors to reliably and precisely estimate the likelihood of heat strain among field workers. Firstly, a high-fidelity immersive virtual reality environment with a tangible user interface was developed to stimulate life-like experiences of a common physically demanding task in heat-vulnerable occupations and collect physiological signals, which included EDA, PPG, ST, and respiratory signals. The developed virtual environment provided a safe and controlled environment for reducing the risk of injury and allowing for the collection of a large amount of data. Furthermore, the integration of a tangible interface (with an actual chainsaw) in the virtual environment provided a life-like experience for the participants, making the collected data more representative of real-world conditions. In sum, the developed simulated experience in a risk-free setting provided a better avenue in eliciting the life-like human subjects' behavior and the corresponding physiological signals, which facilitated the development of a robust framework for accurate heat strain assessment. Notably, the developed approach holds promise for future studies that involve developing AI-driven models for monitoring and assessing human health in extreme environments.

Secondly, a physiological-sensing framework, EnsmTrBoost, was developed to identify workers' heat strain from collected physiological signals. The framework successfully non-intrusively captured EDA, PPG, ST, and respiratory signals from the human subject, enabling the prompt

estimation of their heat stress exposure with an accuracy of 93.2%. This advancement is particularly significant when considered in the context of previous research where physiological sensors were coupled with prevalent ML algorithms to identify heat strain based on workers' physiological signals [38–40]. Notably, Chan et al. [40] developed an MLR to predict worker's heat strain with a mean absolute percent value (MAPE) of 5.6%, but its reliance on invasive calorimetry, which disrupts ongoing work, limits real-world applicability. Additionally, MLR often struggles to accurately predict the complex interplay of personal, environmental, and work-related factors [111]. Similarly, Yi et al. utilized ANN to estimate worker's RPE with a low MAPE [38], and concurrently, integrated the estimation with heart rate data to issue early heat strain alerts. However, RPE is a subjective measure, and relying solely on heart rate to predict workers' RPE may not fully capture the varied physiological responses to heat stress. Further, Shakerian et al. [39] leveraged off-the-shelf wearable sensing devices to capture physiological signals and used traditional ML algorithm to predict the risk of heat strain with 92% accuracy [41]. The traditional ML algorithms have difficulty in objectively discerning physiological signals due to intra- and inter-subject variability [41,42]. Such variability issues fundamentally challenge the traditional ML algorithms' assumptions and necessitate a uniform distribution between training and test data—an aspect often overlooked in previous studies [60,112–117]. Contrary to the limitations in these state-of-the-art heat strain models, the EnsmTrBoost framework designed to enhance the performance of ML classifiers in identifying workers' heat strain from the multimodal physiological signals with variability issue, exhibited a testing accuracy of over 90%. Upon closer examination, the EnsmTrBoost with a low SD of 0.0143 had consistent performance across the dataset with different statistical distribution, as compared with the traditional classifiers. By integrating ensemble learning and domain adaptation, EnsmTrBoost not only improves the performance and generalization of data-driven models but also becomes more robust to variations in the data. Furthermore, its design, which accommodates multimodal physiological signals and their

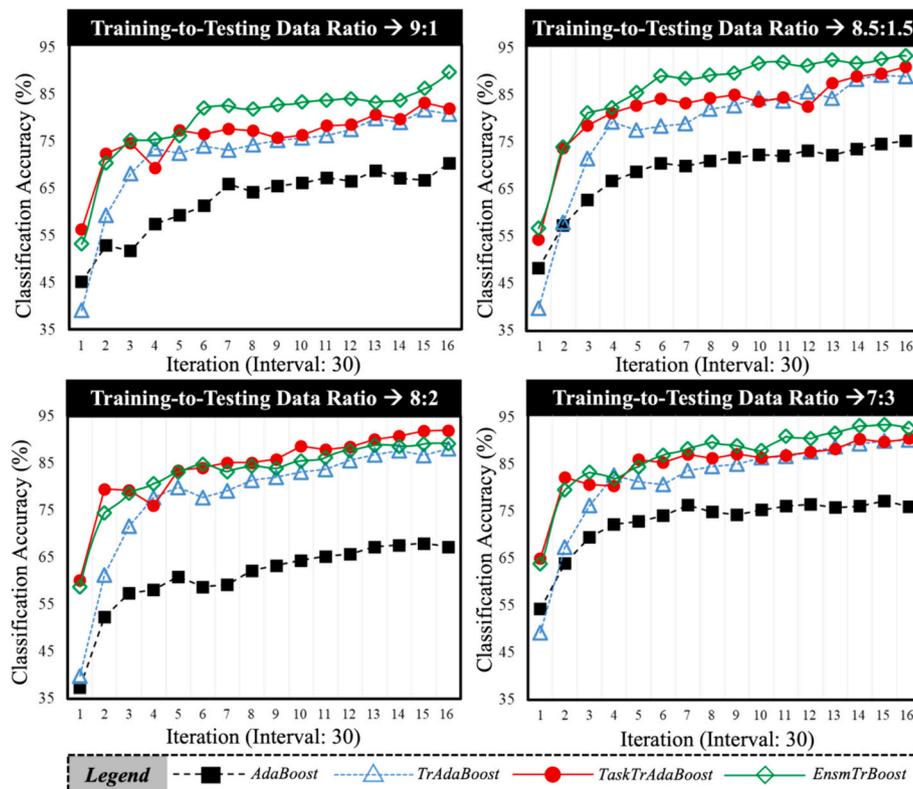


Fig. 8. Performance of the EnsmTrBoost framework over the state-of-art transfer learning approaches.

inherent variabilities, allows us to make more accurate predictions by effectively inferring physiological signals with variability issues. This key advantage enables the framework to accurately predict the likelihood of heat strain in new individuals under real-world job site conditions.

To further examine the efficiency of the developed framework in identifying workers' heat strain from collected physiological signals with variability issues, the authors compared the classification accuracy of the developed EnsmTrBoost with other prevalent transfer learning approaches, namely Adaboost, TrAdaBoost, and TaskTrAdaBoost under four distinct scenarios, each time by dividing the physiological data collected from the subjects into the different ratio of training data and testing data. The ratio of training data and testing data were 9:1, 8.5:1.5, 8:2, and 7:3, respectively. Fig. 8 shows the competitive performance of the developed framework over the state-of-the-art transfer learning approaches.

As seen, the EnsmTrBoost framework had the highest classification accuracy in the scenarios with the 9:1, 8.5:1.5, and 7:3 training-to-testing data ratios. In the scenario where the ratio was 8:2, the performance of TaskTrAdaBoost was slightly better than the EnsmTrBoost. One plausible explanation for such competitive performance of EnsmTrBoost is the capability of the developed framework to adapt the statistical properties of the data to the properties of the data with different statistical distribution through inner loop domain adaptation. The physiological signals collected from different individuals or different environments can have different distributions, making it more challenging for a model trained on one domain to generalize to another. In this vein, the EnsmTrBoost framework can robustly handle distributional differences by updating the model parameters to better fit the target domain data and achieve better performance. As shown in Fig. 8, AdaBoost had the lowest accuracy among different transfer learning approaches in all scenarios. This might be because AdaBoost is more focused on modifying the decision boundary rather than the distribution itself, which cannot effectively address the distribution shift issues, subsequently having less accurate prediction results.

The developed framework is expected to promote a deeper understanding of human thermoregulation and assist in preventing heat stress in field workers. The developed approach provides a systematic and continuous assessment of heat stress exposure, allowing for timely recognition and prevention of occupational health and safety risks. By detecting workers' heat strain in a timely manner, appropriate preventive measures can be implemented to avoid the undesirable consequences of heat stress, ultimately enhancing workplace safety. While the scope of this study was limited to estimating workers' heat strain from EDA, PPG, ST, and respiratory signals with variability issues, the framework can also be leveraged for other non-stationary physiological signals to monitor diverse physiological parameters of field workers accurately and consistently. For instance, the developed framework can be leveraged to accurately estimate muscular fatigue from electromyogram (EMG) signals, cognitive performance, vigilance, and alertness from electroencephalogram (EEG) signals, and overexertion from electrocardiogram (ECG) signals. In this context, the introduction of a domain-adaptation-based approach capable of addressing variability concerns in physiological signals holds the potential for enhancing physiological-sensing-based health monitoring across a spectrum of sectors, including sports, military, mining, agriculture, and firefighting.

While the results of this study hold significant promise for the creation of a sturdy heat strain monitoring system for field workers, it is crucial to acknowledge certain limitations that could be explored in forthcoming research endeavors. First, the study used commercially available off-the-shelf wearable biosensors to estimate the likelihood of heat strain. While wearable biosensors endow for objective, noninvasive, and continuous monitoring of heat strain among the field workers, the accessibility, ease of use, and cost involved with the sensor are bound to limit its ubiquitous implementation on the field. The body's response to various stimuli is reflected as several physiological

alterations, which cannot be measured through a single commercially available wearable sensor. A wearable sensor that captures the contributory biosignals is currently unavailable. Future research can develop and leverage a wearable sensing system for accurate and noninvasive measurement of workers' diverse contributory physiological responses to stressors on the field. Second, the study only used four physiological signals (EDA, PPG, ST, and respiratory signals) to estimate heat strain. It would be interesting to include other factors such as hydration level, acclimation status, and individual differences to assess heat strain comprehensively. Third, the developed virtual environment does not consider the effects of powder in logging tasks, which can also exacerbate workers' heat strain. Future research could explore safe and practical methods to incorporate the effect of powder in the virtual environment, aiming to enhance the realism of logging tasks using advanced immersive technologies or alternative simulation strategies. In future research, the authors recommend examining the robustness of EnsmTrBoost model with a larger sample size for longer tasks duration in varying environmental conditions with diverse field tasks. Further, the developed framework can be validated in the naturalistic field for continuous heat strain monitoring and proactive safety interventions. The predicted results can be sent to the worker as a safety intervention to prompt the worker to rest, self-pace or stop physical activity, and replenish body nutrients using dietary supplements, such as electrolyte tablets. The implementation of such a proactive safety intervention mechanism will ultimately serve to enhance the safety and overall well-being of field workers.

7. Conclusion

This paper presented and examined an enhanced worker-centered heat strain monitoring framework for accurately and consistently estimating the likelihood of heat strain among field workers based on their physiological signals. To achieve this, a high-fidelity immersive virtual reality environment with a tangible user interface was developed to stimulate life-like experiences of a common physically demanding task in heat-vulnerable occupations and collect physiological signals, including EDA, PPG, ST, and respiratory signals. A physiological-sensing framework, EnsmTrBoost, was developed to identify workers' heat strain from the collected physiological signals, particularly concerning the variability issues that often arise in real-world settings. The EnsmTrBoost framework employed ensemble learning and transfer learning techniques to improve the performance and generalization of the data-driven models, making them more robust to variations in the physiological data. The investigation revealed that EnsmTrBoost could promptly estimate the likelihood of heat strain with an accuracy of 93.2%. The developed heat strain assessment framework is expected to promote a deeper understanding of human thermoregulation and assist in preventing heat stress in field workers. Swift detection of workers' heat strain allows for timely implementation of preventive measures, averting the adverse effects of heat stress and thereby augmenting workplace safety. As time progresses, the implementation of this safety monitoring system has the potential to greatly improve the safety and overall health of field workers, especially those working in extremely hot weather conditions. The encouraging outcomes of this study hold significant promise for the development of a robust heat strain monitoring system for field workers. Nevertheless, it is important to acknowledge certain limitations that could be addressed in future research endeavors. For example, this study was conducted with a relatively small sample size, comprising individuals with limited field-work experience. Subsequent investigations could benefit from recruiting a larger and more diverse pool of subjects, ideally with extensive prior experience in fieldwork, to offer a more comprehensive assessment of the developed framework's performance. Likewise, future research can be conducted to evaluate the performance of the developed system in naturalistic field conditions. Moreover, future studies may explore the potential impact of employing a more precise and continuous heat stress

assessment method, such as utilizing ingestible pills for monitoring core body temperature. This can lead to further refinement and accuracy in heat strain evaluation.

CRedit authorship contribution statement

Amit Ojha: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Shayan Shayesteh:** Visualization, Validation, Methodology. **Ali Sharifronizi:** Visualization, Validation. **Yizhi Liu:** Writing – review & editing, Formal analysis. **Houtan Jebelli:** Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Houtan Jebelli reports financial support was provided by National Institute of Occupational Safety and Health. Houtan Jebelli reports financial support was provided by National Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This material is based upon work supported by the National Institute of Occupational Health and Safety (NIOSH) under Grant No. 1 R21 OH012220-01A1 and the National Science Foundation (NSF) Award No. 2401745. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NIOSH and NSF.

References

- [1] G.M. Waehrer, X.S. Dong, T. Miller, E. Haile, Y. Men, Costs of occupational injuries in construction in the United States, *Accid. Anal. Prev.* 39 (2007) 1258–1266, <https://doi.org/10.1016/j.aap.2007.03.012>.
- [2] D.L. Hard, J.R. Myers, Fatal work-related injuries in the agriculture production sector among youth in the United States, 1992–2002, *J. Agromed.* 11 (2006) 57–65, https://doi.org/10.1300/J096v11n02_09.
- [3] D.L. Lucas, S.L. Case, Work-related mortality in the US fishing industry during 2000–2014: new findings based on improved workforce exposure estimates, *Am. J. Ind. Med.* 61 (2018) 21–31, <https://doi.org/10.1002/ajim.22761>.
- [4] L. Bai, G. Ding, S. Gu, P. Bi, B. Su, D. Qin, G. Xu, Q. Liu, The effects of summer temperature and heat waves on heat-related illness in a coastal city of China, 2011–2013, *Environ. Res.* 132 (2014) 212–219, <https://doi.org/10.1016/j.envres.2014.04.002>.
- [5] B. Tawatsupa, L.-Y. Lim, T. Kjellstrom, S. Seubsmann, A. Sleight, The association between overall health, psychological distress, and occupational heat stress among a large national cohort of 40,913 Thai workers, *Glob. Health Action* 3 (2010) 5034, <https://doi.org/10.3402/gha.v3i0.5034>.
- [6] K. Lundgren, K. Kuklane, C. Gao, I. Holmér, Effects of heat stress on working populations when facing climate change, *Ind. Health* 51 (2013) 3–15, <https://doi.org/10.2486/indhealth.2012-0089>.
- [7] H.-A. Rother, J. John, C.Y. Wright, J. Irlam, R. Oosthuizen, R.M. Garland, Perceptions of occupational heat, sun exposure, and health risk prevention: a qualitative study of forestry workers in South Africa, *Atmosphere* 11 (2019) 37, <https://doi.org/10.3390/atmos11010037>.
- [8] M.A. Febbraio, R.J. Snow, C.G. Stathis, M. Hargreaves, M.F. Carey, Effect of heat stress on muscle energy metabolism during exercise, *J. Appl. Physiol.* 77 (1994) 2827–2831, <https://doi.org/10.1152/jappl.1994.77.6.2827>.
- [9] K.K. Zander, W.J.W. Botzen, E. Oppermann, T. Kjellstrom, S.T. Garnett, Heat stress causes substantial labour productivity loss in Australia, nature, *Climate Change* 5 (2015) 647–652, <https://doi.org/10.1038/NCLIMATE2623>.
- [10] M. Al-Bouwarthan, M.M. Quinn, D. Kriebel, D.H. Wegman, Assessment of heat stress exposure among construction workers in the hot Desert Climate of Saudi Arabia, *Ann. Work Exposures Health* 63 (2019) 505–520, <https://doi.org/10.1093/annweh/wxz033>.
- [11] W.L. Kenney, A review of comparative responses of men and women to heat stress, *Environ. Res.* 37 (1985) 1–11, [https://doi.org/10.1016/0013-9351\(85\)90044-1](https://doi.org/10.1016/0013-9351(85)90044-1).
- [12] L.R. Leon, B.G. Helwig, Heat stroke: role of the systemic inflammatory response, *J. Appl. Physiol.* 109 (2010) 1980–1988, <https://doi.org/10.1152/jappphysiol.00301.2010>.
- [13] C.G. Crandall, J. González-Alonso, Cardiovascular function in the heat-stressed human, *Acta Physiol.* 199 (2010) 407–423, <https://doi.org/10.1111/j.1748-1716.2010.02119.x>.
- [14] F.W. Hsu, C.J. Lin, Y.H. Lee, H.J. Chen, Effects of elevation change on mental stress in high-voltage transmission tower construction workers, *Appl. Ergon.* 56 (2016) 101–107, <https://doi.org/10.1016/j.apergo.2016.03.015>.
- [15] J. Glaser, J. Lemery, B. Rajagopalan, H.F. Diaz, R. García-Trabanino, G. Taduri, M. Madero, M. Amarasinghe, G. Abraham, S. Anutrakulchai, V. Jha, P. Stenvinkel, C. Roncal-Jimenez, M.A. Lanasa, R. Correa-Rotter, D. Sheikh-Hamad, E.A. Burdmann, A. Andres-Hernando, T. Milagres, I. Weiss, M. Kanbay, C. Wesseling, L.G. Sánchez-Lozada, R.J. Johnson, Climate change and the emergent epidemic of CKD from heat stress in rural communities: the case for heat stress nephropathy, *Clin. J. Am. Soc. Nephrol.* 11 (2016) 1472–1483, <https://doi.org/10.2215/CJN.13841215>.
- [16] C. Herath, C. Jayasumana, P.M.C.S. De Silva, P.H.C. De Silva, S. Siribaddana, M. E. De Broe, Kidney diseases in agricultural communities: a case against heat-stress nephropathy, *Kidney Int. Rep.* 3 (2018) 271–280, <https://doi.org/10.1016/j.ekir.2017.10.006>.
- [17] D.G.C. Rainham, K.E. Smoyer-Tomic, The role of air pollution in the relationship between a heat stress index and human mortality in Toronto, *Environ. Res.* 93 (2003) 9–19, [https://doi.org/10.1016/S0013-9351\(03\)00060-4](https://doi.org/10.1016/S0013-9351(03)00060-4).
- [18] P.A. Hancock, I. Vasmatazidis, P.A. Hancock, I. Vasmatazidis, Effects of heat stress on cognitive performance: the current state of knowledge, *Int. J. Hypertherm.* 19 (2003) 355–372, <https://doi.org/10.1080/0265673021000054630>.
- [19] CFOI, Census of Fatal Occupational Injuries (CFOI) - Current and Revised Data, <https://www.bls.gov/news.release/pdf/cfoi.pdf>, 2018 (accessed July 30, 2024).
- [20] T. Kjellstrom, I. Holmer, B. Lemke, Workplace heat stress, health and productivity – an increasing challenge for low and middle-income countries during climate change, *Glob. Health Action* 2 (2009), <https://doi.org/10.3402/GHA.V2I0.2047>.
- [21] J. Xiang, P. Bi, D. Pisainello, A. Hansen, Health impacts of workplace heat exposure: an epidemiological review, *Ind. Health* 52 (2014) 91–101, <https://doi.org/10.2486/indhealth.2012-0145>.
- [22] R. Rameezdeen, A. Elmualim, The impact of heat waves on occurrence and severity of construction accidents, *Int. J. Environ. Res. Public Health* 14 (2017), <https://doi.org/10.3390/ijerph14010070>.
- [23] CPWR, The U.S. Construction Industry and its Workers, <https://www.bls.gov/news.release/pdf/cfoi.pdf>, 2018 (accessed July 30, 2024).
- [24] G.M. Budd, Wet-bulb globe temperature (WBGT)-its history and its limitations, *J. Sci. Med. Sport* 11 (2008) 20–32, <https://doi.org/10.1016/j.jsams.2007.07.003>.
- [25] B. Jacklitsch, W. Williams, K. Musolin, A. Coca, J.-H. Kim, N. Turner, NIOSH Criteria for a Recommended Standard: Occupational Exposure to Heat and Hot Environments, US Department of Health and Human Services, 2016 <https://www.cdc.gov/niosh/docs/2016-106/pdfs/2016-106.pdf?id=10.26616/NIOSH-PUB2016106> (accessed July 30, 2024).
- [26] H. Dehghan, S.B. Mortzavi, M.J. Jafari, M.R. Maracy, Development and validation of a questionnaire for preliminary assessment of heat stress at workplace, *J. Res. Health Sci.* 15 (2015) 175–181, <https://doi.org/10.34172/JRHS152067>.
- [27] P. Habibi, H. Dehghan, S. Rezaei, K. Maghsoudi, Physiological and Perceptual Heat Strain Responses in Iranian Veiled Women under Laboratory Thermal Conditions, <http://www.ijhse.ir/index.php/IJHSE/article/view/48>, 2014.
- [28] P. Tikuisis, T.M. McLellan, G. Selkirk, Perceptual versus physiological heat strain during exercise-heat stress, *Med. Sci. Sports Exerc.* 34 (2002) 1454–1461, <https://doi.org/10.1097/00005768-200209000-00009>.
- [29] R.E. Rabeji, Evaluation of indoor heat stress on workers of bakeries at Assiut City, Egypt, *Int. J. Environ. Sci. Technol.* 16 (2019) 2637–2642, <https://doi.org/10.1007/s13762-018-1839-z>.
- [30] J.B.M. Malchaire, Occupational heat stress assessment by the predicted heat strain model, *Ind. Health* 44 (2006) 380–387, <https://doi.org/10.2486/INDHEALTH.44.380>.
- [31] W. Yi, A.P.C. Chan, Optimal work pattern for construction workers in hot weather: a case study in Hong Kong, *J. Comput. Civ. Eng.* 29 (2015) 05014009, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000419/ASSET/B560B3A5-8F0F-4794-8E2F-83ECA798C6CF/ASSETS/IMAGES/LARGE/FIGURE3.JPG](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000419/ASSET/B560B3A5-8F0F-4794-8E2F-83ECA798C6CF/ASSETS/IMAGES/LARGE/FIGURE3.JPG).
- [32] W. Yi, A.P.C. Chan, Optimizing work–rest schedule for construction rebar workers in hot and humid environment, *Build. Environ.* 61 (2013) 104–113, <https://doi.org/10.1016/J.BUILDENV.2012.12.012>.
- [33] Y. Zhao, W. Yi, A.P.C. Chan, D.P. Wong, Impacts of cooling intervention on the heat strain attenuation of construction workers, *Int. J. Biometeorol.* 62 (2018) 1625–1634, <https://doi.org/10.1007/S00484-018-1562-Y/FIGURES/6>.
- [34] P. Roelofsens, K. Jansen, Comfort and performance improvement through the use of cooling vests for construction workers, *Int. J. Cloth. Sci. Technol.* 35 (2023) 152–161, <https://doi.org/10.1108/IJCS-08-2021-0104/FULL/PDF>.
- [35] W. Yi, Y. Zhao, A.P.C. Chan, E.W.M. Lam, Optimal cooling intervention for construction workers in a hot and humid environment, *Build. Environ.* 118 (2017) 91–100, <https://doi.org/10.1016/J.BUILDENV.2017.03.032>.
- [36] A. Frank, M. Belokopytov, Y. Shapiro, Y. Epstein, The cumulative heat strain index – a novel approach to assess the physiological strain induced by exercise-heat stress, *Eur. J. Appl. Physiol.* 84 (2001) 527–532, <https://doi.org/10.1007/s004210000368>.

- [37] D.S. Moran, A. Shitzer, K.B. Pandolf, A physiological strain index to evaluate heat stress, *Regulat. Integrat. Comp. Physiol.* 275 (1998) 129–134, <https://doi.org/10.1152/ajpregu.1998.275.1.R129>.
- [38] W. Yi, A.P.C. Chan, X. Wang, J. Wang, Development of an early-warning system for site work in hot and humid environments: a case study, *Autom. Constr.* 62 (2016) 101–113, <https://doi.org/10.1016/j.autcon.2015.11.003>.
- [39] S. Shakerian, M. Habibnezhad, A. Ojha, G. Lee, Y. Liu, H. Jebelli, S.H. Lee, Assessing occupational risk of heat stress at construction: a worker-centric wearable sensor-based approach, *Saf. Sci.* 142 (2021) 105395, <https://doi.org/10.1016/J.SSCI.2021.105395>.
- [40] A.P.C. Chan, M.C.H. Yam, J.W.Y. Chung, W. Yi, Developing a heat stress model for construction workers, *J. Facil. Manag.* 10 (2012) 59–74, <https://doi.org/10.1108/14725961211200405/FULL/XML>.
- [41] S. Saha, M. Baumert, Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: a review, *Front. Comput. Neurosci.* (2020), <https://doi.org/10.3389/fncom.2019.00087>.
- [42] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, F. Yger, A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update, *J. Neural Eng.* 15 (2018), <https://doi.org/10.1088/1741-2552/aab2f2>.
- [43] A. Ojha, Y. Liu, H. Jebelli, H. Cheng, M. Kiani, Improving health monitoring of construction workers using physiological data-driven techniques: an ensemble learning-based framework to address distributional shifts, in: *Computing in Civil Engineering 2023: Resilience, Safety, and Sustainability - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2024*, 2023, pp. 631–638, <https://doi.org/10.1061/9780784485248.076>.
- [44] R. Zheng, T. Wang, J. Cao, P.P. Vidal, D. Wang, Multi-modal physiological signals based fear: a high analysis in virtual reality scenes, *Biomed. Sign. Proc. Control.* 70 (2021) 102988, <https://doi.org/10.1016/J.BSPC.2021.102988>.
- [45] G.P. Bates, V.S. Miller, Sweat rate and sodium loss during work in the heat, *J. Occupation. Med. Toxicol. (London, England)*. 3 (2008) 4, <https://doi.org/10.1186/1745-6673-3-4>.
- [46] M.-L. Chen, C.-J. Chen, W.-Y. Yeh, J.-W. Huang, I.-F. Mao, Heat stress evaluation and worker fatigue in a steel plant 64 (2010) 352–359, <https://doi.org/10.1080/15428110308984827>.
- [47] C.L. Lim, K. Suzuki, Systemic inflammation mediates the effects of Endotoxemia in the mechanisms of heat stroke, *Biol. Med.* 9 (2016), <https://doi.org/10.4172/0974-8369.1000376>.
- [48] L.A. Pugh, C.R. Oldroyd, T.S. Ray, M.L. Clark, Muscular effort and electrodermal responses, *J. Exp. Psychol.* 71 (1966) 241–248, <https://doi.org/10.1037/h0022834>.
- [49] S. Bagha, L. Shaw, A real time analysis of PPG signal for measurement of SpO2 and pulse rate, *Int. J. Comput. Appl.* 36 (2011), <https://doi.org/10.5120/4537-6461>.
- [50] P. Egenberger, B.A. MacRae, S. Kemp, M. Bürgisser, R.M. Rossi, S. Annaheim, Prediction of Core body temperature based on skin temperature, heat flux, and heart rate under different exercise and clothing conditions in the heat in young adult males, *Front. Physiol.* 9 (2018), <https://doi.org/10.3389/fphys.2018.01780>.
- [51] H.F. Posada-Quintero, N. Reljin, C. Mills, I. Mills, J.P. Florian, J.L. VanHeest, K. H. Hon, Time-varying analysis of electrodermal activity during exercise, *PLoS One* 13 (2018) 1–12, <https://doi.org/10.1371/journal.pone.0198328>.
- [52] M. Krishnamurthy, P. Ramalingam, K. Perumal, L.P. Kamalakannan, J. Chinnadurai, R. Shanmugam, K. Srinivasan, V. Venugopal, Occupational Heat Stress Impacts on Health and Productivity in a Steel Industry in Southern India, 2017, <https://doi.org/10.1016/j.shaw.2016.08.005>.
- [53] R.K. Sinha, Electro-encephalogram Disturbances in Different Sleep-wake States Following Exposure to High Environmental Heat, 2004, <https://doi.org/10.1007/bf02344701>.
- [54] Z. Zamanian, Z. Sedaghat, M. Hemehrezaee, F. Khajehnasiri, Evaluation of environmental heat stress on physiological parameters, *J. Environ. Health Sci. Eng.* 15 (2017) 1–8, <https://doi.org/10.1186/s40201-017-0286-y>.
- [55] P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, Explainable AI: a review of machine learning interpretability methods, *Entropy* 23 (2021) 18, 23 (2020) pp. 18, <https://doi.org/10.3390/E23010018>.
- [56] I. Tarnanas, W. Schlee, M. Tsolaki, R. Müri, U. Mosimann, T. Nef, Ecological validity of virtual reality daily living activities screening for early dementia: longitudinal study, *JMIR Serious Games* 1 (2013), <https://doi.org/10.2196/GAMES.2778>.
- [57] N. Chaytor, M. Schmitter-Edgecombe, The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills, *Neuropsychol. Rev.* 13 (2003) 181–197, <https://doi.org/10.1023/B:NERV.000009483.91468.FB>.
- [58] Y. Liu, M. Habibnezhad, H. Jebelli, S. Asadi, S.H. Lee, Ocular artifacts reduction in EEG Signals acquired at construction sites by applying a dependent component analysis (DCA), in: *Construction Research Congress 2020: Computer Applications - Selected Papers from the Construction Research Congress 2020*, American Society of Civil Engineers (ASCE), 2020, pp. 1281–1289, <https://doi.org/10.1061/9780784482865.135>.
- [59] G. Lee, B. Choi, H. Jebelli, C. Ryan Ahn, S. Lee, Noise reference signal-based denoising method for EDA collected by multimodal biosensor wearable in the field, *J. Comput. Civ. Eng.* 34 (2020) 04020044, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000927](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000927).
- [60] H. Jebelli, S. Hwang, S. Lee, EEG signal-processing framework to obtain high-quality brain waves from an off-the-shelf wearable EEG device, *J. Comput. Civ. Eng.* 32 (2018), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000719](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000719).
- [61] H. Jebelli, B. Choi, S. Lee, Application of wearable biosensors to construction sites. II: assessing workers' physical demand, *J. Constr. Eng. Manag.* 145 (2019) 04019080, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001710](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001710).
- [62] H. Jebelli, B. Choi, S. Lee, Application of wearable biosensors to construction sites. I: assessing workers' stress, *J. Constr. Eng. Manag.* 145 (2019) 04019079, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001729](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001729).
- [63] C.S. Wei, Y.P. Lin, Y. Te Wang, C.T. Lin, T.P. Jung, A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection, *NeuroImage* 174 (2018) 407–419, <https://doi.org/10.1016/j.neuroimage.2018.03.032>.
- [64] W. Dai, Q. Yang, X. Gui-Rong, Y. Yong, Boosting for transfer learning, in: *Proceedings of the 24th International Conference on Machine Learning*, 2007, pp. 193–200, <https://doi.org/10.1145/1273496.1273521>.
- [65] Y. Yao, G. Doretto, Boosting for transfer learning with multiple sources, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1855–1862, <https://doi.org/10.1109/CVPR.2010.5539857>.
- [66] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (2010) 1345–1359, <https://doi.org/10.1109/TKDE.2009.191>.
- [67] M. Walker, A. Stent, F. Mairesse, R. Prasad, Individual and domain adaptation in sentence planning for dialogue, *J. Artif. Intell. Res.* 30 (2007) 413–456, <https://doi.org/10.1613/JAIR.2329>.
- [68] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, in: *Advances in Computer Vision and Pattern Recognition*, 2017, pp. 189–209, https://doi.org/10.1007/978-3-319-58347-1_10/FIGURES/8.
- [69] K. Weiss, T.M. Khoshgoftaar, D.D. Wang, A survey of transfer learning, *J. Big Data* 3 (2016) 1–40, <https://doi.org/10.1186/S40537-016-0043-6/TABLES/6>.
- [70] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on Bias and fairness in machine learning, *ACM Comp. Surv. (CSUR)*. 54 (2021), <https://doi.org/10.1145/3457607>.
- [71] M.M. Bejani, M. Ghatte, A systematic review on overfitting control in shallow and deep neural networks, *Artif. Intell. Rev.* 54 (2021) 6391–6438, <https://doi.org/10.1007/S10462-021-09975-1>.
- [72] Y. Gautam, Y. Liu, H. Jebelli, Unsupervised adversarial domain adaptation in wearable physiological sensing for construction workers' health monitoring using Photoplethysmography, *construction research congress 2024*, CRC 2024 (1) (2024) 339–348, <https://doi.org/10.1061/9780784485262.035>.
- [73] Y. Ji, H. Hou, J. Chen, N. Wu, Adversarial training for unknown word problems in neural machine translation, in: *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)* 19, 2019, <https://doi.org/10.1145/3342482>.
- [74] J. Gao, J. Hu, W.W. Tung, E. Blasch, Multiscale analysis of biological data by scale-dependent Lyapunov exponent, *Front. Physiol.* 2 JAN (2012) 18738, <https://doi.org/10.3389/FPHYS.2011.00110/BIBTEX>.
- [75] Y. Ganin, V. Lempitsky, Unsupervised domain adaptation by backpropagation, in: *32nd International Conference on Machine Learning, ICMML 2015 2*, 2014, pp. 1180–1189, <https://doi.org/10.48550/arxiv.1409.7495>.
- [76] M. Wang, W. Deng, Deep visual domain adaptation: a survey, *Neurocomputing* 312 (2018) 135–153, <https://doi.org/10.1016/J.NEUCOM.2018.05.083>.
- [77] K. Feng, T. Chaspari, A review of generalizable transfer learning in automatic emotion recognition, *Front. Comp. Sci.* 2 (2020) 488594, <https://doi.org/10.3389/FCOMP.2020.00009/BIBTEX>.
- [78] The Share of Wood-Framed Homes Increased in 2021, NAHB Eye on Housing: National Association of Home Builders Discusses Economics and Housing Policy. <https://eyeonhousing.org/2022/07/the-share-of-wood-framed-homes-increased-in-2021/>, 2021.
- [79] Y. Liu, M. Habibnezhad, H. Jebelli, Brain-computer interface for hands-free teleoperation of construction robots, *Autom. Constr.* 123 (2021) 103523, <https://doi.org/10.1016/J.AUTCON.2020.103523>.
- [80] G. Rätsch, T. Onoda, K.R. Müller, Soft margins for AdaBoost, *Mach. Learn.* 42 (2001) 287–320, <https://doi.org/10.1023/A:1007618119488>.
- [81] S. Taylor, N. Jaques, W. Chen, S. Fedor, A. Sano, R. Picard, Automatic identification of artifacts in electrodermal activity data, in: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Institute of Electrical and Electronics Engineers Inc.*, 2015, pp. 1934–1937, <https://doi.org/10.1109/EMBC.2015.7318762>.
- [82] Y. Gautam, H. Jebelli, Autoencoder-based Photoplethysmography (PPG) signal reliability enhancement in construction health monitoring, *Autom. Constr.* 165 (2024) 105537, <https://doi.org/10.1016/J.AUTCON.2024.105537>.
- [83] M. Sadat-Mohammadi, S. Shakerian, Y. Liu, S. Asadi, H. Jebelli, Non-invasive physical demand assessment using wearable respiration sensor and random forest classifier, *J. Build. Eng.* 44 (2021) 103279, <https://doi.org/10.1016/J.JOBE.2021.103279>.
- [84] K.T. Sweeney, T.E. Ward, S.F. McLoone, Artifact removal in physiological signals-practices and possibilities, *IEEE Trans. Inf. Technol. Biomed.* 16 (2012) 488–500, <https://doi.org/10.1109/TTTB.2012.2188536>.
- [85] F. Mohd-Yasin, D.J. Nagel, C.E. Korman, Noise in MEMS, *Meas. Sci. Technol.* 21 (2010) 012001, <https://doi.org/10.1088/0957-0233/21/1/012001>.
- [86] H. Jebelli, B. Choi, H. Kim, S. Lee, Feasibility study of a wristband-type wearable sensor to understand construction workers' physical and mental status, in: *Construction Research Congress 2018: Construction Information Technology - Selected Papers from the Construction Research Congress 2018*, American Society of Civil Engineers (ASCE), New Orleans, Louisiana, 2018, pp. 367–377, <https://doi.org/10.1061/9780784481264.036>.

- [87] J.J. Braithwaite, D. Watson, R. Jones, M. Rowe, A guide for Analysing Electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments, *Psychophysiology* 49 (2013) 1017–1034, <https://doi.org/10.1111/j.1469-8986.2012.01384.x>.
- [88] S.C. Kim, E.J. Hwang, D.W. Kim, Noise reduction of PPG signal during free movements using adaptive SFLC (Scaled fourier linear combiner), in: *IFMBE Proc*, Springer Verlag, 2007, pp. 1191–1194, https://doi.org/10.1007/978-3-540-36841-0_287.
- [89] S. Li, L. Liu, J. Wu, B. Tang, D. Li, Comparison and noise suppression of the transmitted and reflected Photoplethysmography signals, *Biomed. Res. Int.* 2018 (2018) 4523593, <https://doi.org/10.1155/2018/4523593>.
- [90] Y. Liu, Y. Gautam, S. Shayesteh, H. Jebelli, M. Mahdi Khalili, Towards an Efficient Physiological-Based Worker Health Monitoring System in Construction: An Adaptive Filtering Method for Removing Motion Artifacts in Physiological Signals of Workers, *Computing in Civil Engineering 2023: Resilience, Safety, and Sustainability - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2023, 2024*, pp. 483–491, <https://doi.org/10.1061/9780784485248.058>.
- [91] M. Kelsey, R.V. Palumbo, A. Urbaneja, M. Akcakaya, J. Huang, I.R. Kleckner, L. F. Barrett, K.S. Quigley, E. Sejdic, M.S. Goodwin, Artifact detection in electrodermal activity using sparse recovery, in: *Compressive Sensing VI: From Diverse Modalities to Big Data Analytics* 10211, 2017, p. 102110D, <https://doi.org/10.1117/12.2264027>.
- [92] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, *Commun. ACM* 63 (2014) 139–144, <https://doi.org/10.1145/3422622>.
- [93] J. Li, Z. Struzik, L. Zhang, A. Cichocki, Feature learning from incomplete EEG with denoising autoencoder, *Neurocomputing* 165 (2015) 23–31, <https://doi.org/10.1016/j.neucom.2014.08.092>.
- [94] M.-P. Hosseini, D. Pompili, K. Elisevich, H. Soltanian-Zadeh, Optimized deep learning for EEG big data and seizure prediction BCI via internet of things, *IEEE Trans. Big Data* 3 (2017) 392–404, <https://doi.org/10.1109/tbdata.2017.2769670>.
- [95] Y. Liu, M. Habibnezhad, H. Jebelli, Brainwave-driven human-robot collaboration in construction, *Autom. Constr.* 124 (2021) 103556, <https://doi.org/10.1016/j.autcon.2021.103556>.
- [96] H. Jebelli, S. Hwang, S. Lee, EEG-based workers' stress recognition at construction sites, *Autom. Constr.* 93 (2018) 315–324, <https://doi.org/10.1016/j.autcon.2018.05.027>.
- [97] Y. Liu, A. Ojha, S. Shayesteh, H. Jebelli, S. Lee, Human-centric robotic manipulation in construction: generative adversarial networks (GAN) based physiological computing mechanism to enable robots to perceive workers' cognitive load, *Can. J. Civ. Eng.* (2022), <https://doi.org/10.1139/cjce-2021-0646>.
- [98] W. Tu, S. Sun, A subject transfer framework for EEG classification, *Neurocomputing* 82 (2012) 109–116, <https://doi.org/10.1016/j.neucom.2011.10.024>.
- [99] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.R. Müller, C. Grozea, Subject-independent mental state classification in single trials, *Neural Netw.* 22 (2009) 1305–1312, <https://doi.org/10.1016/j.neunet.2009.06.003>.
- [100] I. Lemhadri, F. Ruan, R. Tibshirani, A neural network with feature sparsity, *J. Mach. Learn. Res.* 22 (2021) 1–29, <https://doi.org/10.48550/arXiv.1907.12207>.
- [101] A. Clim, R.D. Zota, G. Tinica, The Kullback-Leibler divergence used in machine learning algorithms for health care applications and hypertension prediction: a literature review, *Procedia Comp. Sci.* 141 (2018) 448–453, <https://doi.org/10.1016/J.PROCS.2018.10.144>.
- [102] W. Ha, E.Y. Sidky, R.F. Barber, T.G. Schmidt, X. Pan, Estimating the spectrum in computed tomography via Kullback-Leibler divergence constrained optimization, *Med. Phys.* 46 (2019) 81–92, <https://doi.org/10.1002/MP.13257>.
- [103] J.D. Périard, T.M.H. Eijssvogels, H.A.M. Daanen, Exercise under heat stress: thermoregulation, hydration, performance implications, and mitigation strategies, *Physiol. Rev.* 101 (2021) 1873–1979, https://doi.org/10.1152/PHYSREV.00038.2020/ASSET/IMAGES/LARGE/PHYSREV.00038.2020_F025.JPEG.
- [104] C.L. Lim, Fundamental concepts of human thermoregulation and adaptation to heat: a review in the context of global warming, *Int. J. Environ. Res. Public Health* 17 (2020) 1–33, <https://doi.org/10.3390/IJERPH17217795>.
- [105] K. Blazejczyk, J. Baranowski, A. Blazejczyk, Heat stress and occupational health and safety - spatial and temporal differentiation, *Miscellanea Geograph.* 18 (2014) 61–67, <https://doi.org/10.2478/MGRSD-2014-0011>.
- [106] C. Demoury, R. Aerts, B. Vandenninden, B. Van Schaebroeck, E.M. De Clercq, Impact of short-term exposure to extreme temperatures on mortality: a multi-city study in Belgium, *Int. J. Environ. Res. Public Health* 19 (2022) 3763, <https://doi.org/10.3390/IJERPH19073763/S1>.
- [107] R.J. Maughan, H. Otani, P. Watson, Influence of relative humidity on prolonged exercise capacity in a warm environment, *Eur. J. Appl. Physiol.* 112 (2012) 2313–2321, <https://doi.org/10.1007/S00421-011-2206-7/TABLES/2>.
- [108] J. González-alonso, C.G. Crandall, J.M. Johnson, The cardiovascular challenge of exercising in the heat, *J. Physiol.* 586 (2008) 45, <https://doi.org/10.1113/JPHYSIOL.2007.142158>.
- [109] J.D. Périard, S. Racinais, M.N. Sawka, Adaptations and mechanisms of human heat acclimation: applications for competitive athletes and sports, *Scand. J. Med. Sci. Sports* 25 (2015) 20–38, <https://doi.org/10.1111/SMS.12408>.
- [110] A. Ojha, H. Jebelli, L. Alexander, J.R. Loeffert, Quantifying the implications of humidity and temperature on heat stress exposure of construction workers: a worker-centric physiological sensing approach, *construction research congress 2024, CRC 2024 (1)* (2024) 196–205, <https://doi.org/10.1061/9780784485262.021>.
- [111] S. Rowlinson, A. Yunyanjia, B. Li, C. Chuanjingju, Management of climatic heat stress risk in construction: a review of practices, methodologies, and future research, *Accid. Anal. Prev.* 66 (2014) 187–198, <https://doi.org/10.1016/J.AAP.2013.08.011>.
- [112] H. Jebelli, B. Choi, Feasibility study of a wristband-type wearable sensor to understand construction workers, in: *Physical and Mental Status*, in: *Construction Research Congress*, New Orleans, Louisiana, 2018, pp. 367–377, <https://doi.org/10.1061/9780784481264.036>.
- [113] H. Jebelli, S. Lee, Feasibility of wearable electromyography (EMG) to assess construction workers' muscle fatigue, in: *Advances in Informatics and Computing in Civil and Construction Engineering*, Springer International Publishing, 2019, pp. 181–187, https://doi.org/10.1007/978-3-030-00220-6_22.
- [114] H. Jebelli, M.M. Khalili, S. Lee, Mobile EEG-based workers' stress recognition by applying deep neural network, in: *Advances in Informatics and Computing in Civil and Construction Engineering*, Springer International Publishing, 2019, pp. 173–180, https://doi.org/10.1007/978-3-030-00220-6_21.
- [115] M.F. Antwi-Afari, H. Li, Fall risk assessment of construction workers based on biomechanical gait stability parameters using wearable insole pressure system, *Adv. Eng. Inform.* 38 (2018) 683–694, <https://doi.org/10.1016/J.AEI.2018.10.002>.
- [116] K. Yang, C.R. Ahn, Inferring workplace safety hazards from the spatial patterns of workers' wearable data, *Adv. Eng. Inform.* 41 (2019) 100924, <https://doi.org/10.1016/J.AEI.2019.100924>.
- [117] W. Fang, D. Wu, P.E.D. Love, L. Ding, H. Luo, Physiological computing for occupational health and safety in construction: review, challenges and implications for future research, *Adv. Eng. Inform.* 54 (2022) 101729, <https://doi.org/10.1016/J.AEI.2022.101729>.