IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 9 May 2023, accepted 6 June 2023, date of publication 16 June 2023, date of current version 13 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3286305

RESEARCH ARTICLE

Assessing the Soldier Survivability Tradespace Using a Single IMU

MATTHEW P. MAVOR[1], VICTOR C. H. CHAN[1], KRISTINA M. GRUEVSKI[2], LINDA L. M. BOSSI[2], THOMAS KARAKOLIS[2], AND RYAN B. GRAHAM[1]

[1]Faculty of Health Sciences, School of Human Kinetics, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[2]Defence Research and Development Canada, Toronto Research Centre, Government of Canada, Toronto, ON M3K 3C9, Canada

Corresponding author: Ryan B. Graham (ryan.graham@uottawa.ca)

This work was supported by the Canadian Department of National Defence under Contract W7719-195460/001/TOR.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Ottawa and Defence Research and Development Canada Research Ethics Boards under Approval Nos. H-06-18-721 and 2019-026, and performed in line with the Declaration of Helsinki.

ABSTRACT Soldier burden is influenced by the environment, metabolic demands, equipment properties, and psychological stressors; however, much of our knowledge of soldier burden is in the context of body-borne load mass in controlled laboratory environments. Thus, to further our understanding of how all aspects of soldier burden affect the survivability tradespace (i.e., performance, health, and susceptibility to enemy action), field-based motion capture methods are needed. We developed a human activity recognition method using the deep convolutional long short-term memory neural network architecture, trained using a single inertial measurement unit on the upper back, to identify eleven tactical movement patterns commonly performed by soldiers. Using a two-step logical algorithm, real-world constraints are forced, and class labels are expanded to 19 movements. Presented are three models based on Indoor, SectionAttack (outdoors), and a General approach. Across all three approaches, we obtained an average accuracy of 90.0%. Further, we used these predictions to calculate meaningful tradespace metrics, which had an excellent agreement with calculations using the true labels. Military leaders and defence scientists can use this approach to quantify tradespace metrics in the field, as a preprocessing tool to supplement other technology, and make data-driven decisions that can help improve performance, decrease susceptibility, and increase overall mission success.

INDEX TERMS Activity recognition, performance, LSTM, military, wearables, DNN.

I. INTRODUCTION

Military personnel perform a variety of tactical movement patterns during training and deployment. Typically, these movements are performed while subjected to various aspects of soldier burden: environmental conditions, metabolic demands, equipment properties (i.e., mass, mass distribution, coverage, bulk, stiffness, breathability, and thermal resistance), and psychological stressors [1]. However, during biomechanical research investigations, soldiers often perform movement protocols in a controlled indoor motion capture laboratory with a narrow focus on the mass of the body-borne load. Using this data collection framework, previous literature has unanimously identified many negative kinematic [2], [3], kinetic [4], [5], [6], and performance [7], [8], [9], [10], [11], [12] outcomes due to heavy body-borne load mass. Although these investigations are essential to our fundamental understanding of the underlying phenomena that affect a soldier's movement pattern, they do not fully reflect the movement patterns of soldiers in the field as influenced by the mission, threat, environment, and commander's intent. Thus, to truly understand a soldier's movements within the operational environment and the impact soldier burden has on these movements, a motion capture method that can efficiently and reliably collect meaningful data in the field is needed.

The associate editor coordinating the review of this manuscript and approving it for publication was Razi Iqbal.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

One technology that is suitable for field use are inertial measurement units (IMUs). IMUs are small wearable devices that typically include an accelerometer, gyroscope, and magnetometer. These devices are highly transportable, can collect over long periods of time, and can be combined with a biomechanical model to calculate joint angles [13], [14]. Previous work by Mavor et al. (2022) validated a 17-sensor IMU suit [15] and used data from this suit to morph movement patterns to represent intermediary body-borne loads and personal characteristics. While such a suit can provide a tremendous amount of biomechanical information and drive many essential research questions, it is not practical for widespread usage as it is costly, must be set up and calibrated correctly by trained individuals, and is susceptible to damage in an operational environment (e.g., wires can get pulled/frayed, the onboard computer can breakdown if struck, and one of the many sensors can move, nullifying the calibration). An alternative method is to use a single IMU that does not need to be calibrated to a biomechanical model and can be easily outfitted on the participant. Although it is impossible to determine gross kinematic differences caused by soldier burden, a single IMU is field-deployable, is relatively inexpensive, enables the measurement of multiple participants simultaneously, and could be used for gathering performance and susceptibility metrics. Previous military investigations using a single IMU have been able to determine performance degradations due to body-borne load and personal characteristics [17], [18] and have been used to recognize operationally-relevant movement patterns in the field [19], [20].

Human activity recognition (HAR) is a domain of machine learning research that typically uses deep neural networks (DNNs) to recognize movement patterns based on data obtained from various wearable sensors: a single IMU setup [21], [22], multiple IMU setup [23], IMUs + heart rate monitor [19], IMU + EMG [24], optical motion capture systems [25], and smartphones [26], [27], among others. Researchers have applied HAR to these various data types to recognize hand gestures [23], activities of daily living [21], [23], [27], non-specific sporting movements [25], workplace postures [28] and tasks [24], and military activities [19], [20], [29], among many others. Using the knowledge obtained from the HAR algorithms, researchers have been able to perform workplace ergonomic assessments [28], [30] and at-home monitoring of activities of daily living [22]. In a military context, by using a single IMU to recognize what activities soldiers are performing, with or without combining other wearable technology, there is potential to advance other fields of study within defence science, such as fatigue research (e.g., knowing what activities lead up to sustained heart rate or the evaluation of performance decline over time), ergonomics (e.g., task frequency, rest, interaction with objects), team-based interactions, training (i.e., digitize drills to visualize), and survivability (i.e., firing at digitized avatars, calculation of exposure time), among others.

This work aimed to investigate the efficacy of using a single IMU sensor to recognize key military movements and calculate tradespace metrics. Presented are a series of DNNs, trained to recognize operationally-relevant movement patterns using data from a single IMU, obtained from data collected in an indoor laboratory and during an outdoor two-person 60-metre simulated section attack drill. Three types of DNN architectures, 1) convolutional neural network (CNN), 2) fully connected DNN (FC), and 3) deep convolutional long short-term memory neural network (DeepConvLSTM), were trained using three approaches: indoor-specific, section attack-specific, and general (i.e., both indoor and section attack data). Predictions made by the DNNs were processed through a two-step logical algorithm to ensure predictions adhered to real-world constraints and to expand the number of class labels from 11 to 19. We present our models' performance using traditional machine learning metrics (i.e., accuracy and F1-scores) and compare the survivability tradespace metrics (i.e., exposure time, susceptibility to enemy action) calculated using the predicted labels versus the true labels for our test participants. It should be declared that this paper reuses content from thesis [31] with permission.

II. EXPERIMENTAL DESIGN

A. PARTICIPANTS

Seventeen male dismounted combat arms reserve force soldiers (consisting of Privates, Corporals, and Master Corporals) from the Canadian Armed Forces (CAF) were recruited. Participants' mean height, mass, age, and years served were 1.83 m (±8.9), 81.7 kg (±9.2), 26.6 years (±7.2), and 3.8 years of service (±2.2), respectively. All experimental protocols followed the Declaration of Helsinki and were approved by the University of Ottawa (H-06-18-721) and Defence Research and Development Canada (2019-026) Research Ethics Boards. All participants provided free and informed consent prior to participating in these protocols.

B. EQUIPMENT

Participants performed a series of movements under four load conditions: 1) ***Slick*** ∼5.5 kg, comprising of a CAF helmet, a replica C7A2 rifle, and the temperate combat clothing ensemble (T-shirt, combat shirt, combat trousers, combat boots); 2) ***Full Fighting Order (FFO)*** ∼22 kg, comprising Slick + CAF issued fragmentation protective vest (with 2 in-service ballistic plates, front and back) + a fully loaded tactical vest (2 replica grenades, 2 replica smoke grenades, 4 replica C7 magazines, a simulated canister of C9 ammunition, 2 field dressings, and a water-filled 1 L canteen); 3) ***Backpack*** 38 kg, comprising FFO + a loaded CAF daypack weighing 16 kg; and 4) ***Pockets*** 38 kg, comprising of FFO + 16 kg worth of non-ferrous plates added to the fragmentation vest (section attack only). Participants' whole-body movement patterns were recorded at 240 Hz for all conditions using an Xsens IMU suit (MVN Link, Xsens, Netherlands)

validated for military-based tasks [15]. Concurrently, movement patterns were collected at 100 Hz using a single IMU (S5, Catapult Sports Pty Ltd., Australia) designed for high-performance activities. The IMU was placed on the back of the upper torso using a specialized sports pinnie.

C. MOVEMENT PROTOCOL

1) INDOOR LABORATORY

Twelve soldiers performed two repetitions of eight military-based movements under the Slick, FFO, and Backpack load conditions described above: run, run-to-kneel (RTK), run-to-prone (RTP), kneel-to-run (KTR), prone-to-run (PTR), kneel-to-prone (KTP), prone-to-kneel (PTK), and walk. Trial order and weight conditions were randomized for each participant; a total of 48 movement trials (8 movements $*$ 2 repetitions $*$ 3 conditions $=$ 48 movements) were performed. All movements requiring forward progression were done in a straight line measuring ten metres, and all movements began/ended with the participant in standing.

2) OUTDOOR SECTION ATTACK

Ten soldiers performed two repetitions of fireteam pairs within the context of a simulated 60-metre section attack under the four load conditions (i.e., Slick, FFO, Backpack, and Pockets). For each repetition, team members alternated between fireteam leader and follower roles. Fireteam pairs began by patrolling (walking and watching for threats) for at least 10 m in open grassy terrain. At the command of a Section Commander (non-participant), fireteam pairs began advancing towards the threat and adopted tactical movements (i.e., bounding rushes, wherein one soldier dashes from one point of cover to another while simulated covering fire is provided by the fireteam partner and vice versa) until the desired distance was reached ($\sim$40 metres). Soldiers then completed a simulated assault on the objective for the remaining distance ($\sim$20 metres). These tactical movements replicate the individual movements and postures captured during the laboratory data collection (run, RTK, KTR, PTR, KTR, PTK, and walk).

III. DATA PREPERATION

A. PREPROCESSING

Xsens IMU data were processed in Xsens MVN Analyze 2019.2.1 (Xsens, Netherlands) using the high-definition (HD) reprocessing tool. Data were visually inspected, and segment contact data were manually adjusted when necessary. Files were exported in both .mvnx and .c3d file formats.

Catapult IMU data were downloaded and processed through Sprint 5.1.7 (Catapult, Australia) to produce comma-separated variable (CSV) files hosting the sensor's triaxial acceleration, triaxial gyroscope, and triaxial magnetometer data; GPS data were available outdoors but were not used for this investigation. The output data were a continuous stream hosting the entire data collection protocol; thus, data were cropped and recombined to only include data when the Xsens suit was collecting. This was done so that all input data for the DNNs could be accurately labelled with ground truth labels.

Before each load condition, a jumping task was performed to synchronize the Catapult and Xsens IMU systems. Specifically, the peak vertical acceleration from the Catapult and Xsens (T8 segment) systems were used to align the data in time using the serial date number calculated through Matlab's (MathWorks, USA) datenum function.

B. GROUND TRUTH LABELS

Ground truth labels were visually identified using Matlab 2016b by a single labeller who plotted the Catapult IMU's vertical accelerometer and yaw gyroscope data, along with Xsens' vertical virtual marker position of the T8 spinal process and used Matlab's ginput function to select the bounds of each movement type. The labeller also visualized the whole-body movement patterns using the .c3d files to gain additional information on the movements.

The labeller manually identified two class label sets: original and expanded. The original set was used as ground truth labels for the DNNs and consisted of eleven class labels: stand, jump, descend, ascend, prone, kneel, stand-to-run (STR), run, walk, crawl, and null. By processing the DNNs' predictions through a two-step logical algorithm, they were expanded to give context to the ascend and descend classes, resulting in 19 class labels (Table 1). The null class was used as a catch-all for any movement the participant performed that did not fit into any other movement; examples include, but are not limited to bending over, shuffling feet while standing, and turning.

TABLE 1. Original and expanded class labels.

Original	Expanded	
Stand	Stand	
Jump	Jump	
Descend	Run-to-prone	Run-to-kneel
	Run-to-kneel	Stand-to-prone
	Stand-to-kneel	Kneel-to-prone
Ascend	Prone-to-run	Kneel-to-run
	Prone-to-stand	Kneel-to-stand
	Prone-to-kneel	
Kneel	Kneel	
Prone	Prone	
Stand-to-run	Stand-to-run	
Run	Run	
Walk	Walk	
Crawl	Crawl	
Null	Null	

Note: The original labels were predicted by the DNNs. The expanded labels were generated through a two-step algorithm to give context to the descend and ascend class labels.

C. DATA ORGANIZATION

The data were divided into training, validation, and test subsets by participant (i.e., samples from one participant were always kept together in the same subset), shown in Table 2. The same subsets for a given approach were used across all models (i.e., CNN, FC, and DeepConvLSTM) to allow for

TABLE 2. Datasets used for each approach.

	Total samples	Training subset	Validation subset	Test subset
Indoor	628,063 (12 participants)	477,929 (76.1%) (9 participants)	47,866 (7.6%) (1 participant)	102,268 (16.3%) (2 participants)
Section Attack	1,032,177 (10 participants)	812,402 (78.7%) (8 participants)	109,889 (10.7%) (1 participant)	109,886 (10.7%) (1 participant)
General	1,660,240 (22 participants)	1,337,335 (80.6%) (18 participants)	157,755 (9.5%) (2 participants)	165,150 (9.9%) (2 participants)

Note: The general dataset comprised the indoor and section attack datasets, and both indoor and section attack data were represented in the validation and test subsets. Seventeen unique participants were recruited for this study; however, 5 of the same participants performed both the indoor and section attack protocols. Each sample consisted of 9 features (3*accelerometer, 3*gyroscope, 3*magnetometer).

direct comparisons. Across all approaches, the validation and test participants were not exposed during training to provide an unbiased estimate of model generalizability. In the General approach, indoor and section attack data were represented in the validation and test subsets with a distribution similar to the training subset.

For all approaches, a single matrix was created for each subset by concatenating data from all movement trials and conditions performed by the participants in that subset. Normalization was conducted on a feature-by-feature basis by subtracting the mean and dividing by the standard deviation of all data frames across participants and movements in the corresponding subset [32]. A sliding window approach was used to divide the data into segments containing an equal number of data frames (i.e., 64 or 128), where the step size was set to 1/4 of the window size. Each segment was assigned a label according to the movement performed for the majority of that segment. The length and width of the input data was the window size and nine features (x, y, and z component of the linear accelerometer, gyroscope, and magnetometer data from the Catapult IMU), respectively.

IV. DEEP NEURAL NETWORK-BASED HUMAN ACTIVITY RECOGNITION

A hold-out testing approach was employed to train and evaluate the DNNs for activity recognition. Three approaches to model development were employed: Indoor-specific, Section Attack-specific, and General. The Indoor and Section Attack approaches were used to develop models to classify only indoor or section attack data, respectively. The General approach was used to develop models to classify activities from either scenario. For each approach, three neural network architectures were employed (3 approaches $*$ 3 architectures $=$ 9 models).

A. MODEL ARCHITECTURES, HYPERPARAMETER OPTIMIZATION, AND TRAINING

To perform activity recognition, three neural network architectures based on previous work were implemented in Python 3.8.12 [33] using Tensorflow 2.3.0 [34]: shallow CNN [35], FC [36], and DeepConvLSTM [23]. All machine learning development was synchronously distributed using a mirrored strategy across three Nvidia Titan RTX GPUs (Nvidia, USA) with 24 GB of GDDR6 memory each. This was performed on a server running Ubuntu 18.04 powered by two Intel Xeon Gold 6248 CPUs @ 2.50GHz (Intel, USA) and 384 GB of DDR4 RAM.

A combination of reliance on previous work and hyperparameter optimization was used to select the hyperparameters for each model. Hyperparameter optimization was performed using the indoor dataset and KerasTuner library [37] with the objective of maximizing validation accuracy. Bayesian hyperparameter optimization was implemented to search the hyperparameter spaces defined for each model [38], presented in Table 3; bolded values were the optimized hyperparameters that were used to train the models for all approaches.

All models were trained using a batch size of 100 samples, a maximum epoch of 500, and early stopping was implemented via monitoring the training accuracy with a patience of three. Training data were randomly shuffled before each epoch to reduce the risk of overfitting. The categorical cross-entropy function was used to compute loss for all models and was minimized using either stochastic gradient descent (SGD), or root mean squared propagation (RMSProp). When training was stopped, the weights from the epoch with the smallest validation loss were retained and used to classify all samples in the test set. Loss, accuracy, and weighted average F1-score were computed to evaluate test performance. The architectures and hyperparameter optimization procedures for each model are described in the following sections.

B. CONVOLUTIONAL NEURAL NETWORK

The CNN was comprised of an input layer, a convolutional layer with a rectified linear unit (ReLU) activation function, and a dense layer with a softmax activation function (Fig. 1). During training, the model parameters were optimized using SGD. The SGD hyperparameters (i.e., learning rate and momentum), the number of convolutional filters, and the convolutional kernel size were tuned during hyperparameter optimization (Table 3).

C. FULLY CONNECTED NEURAL NETWORK

The FC neural network was comprised of an input layer, three dense (i.e., fully-connected) layers with ReLU activation

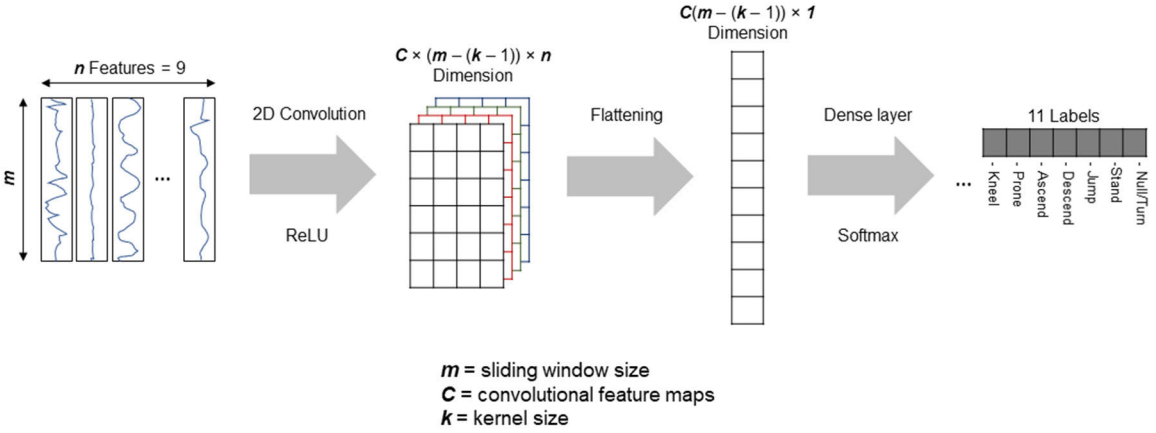

FIGURE 1. Convolutional neural network architecture. The nine features were the x, y, and z components of the accelerometer, gyroscope, and magnetometer signals from the inertial measurement unit. The sliding window size, the number of convolutional feature maps (i.e., number of filters), and the kernel size were tuned during hyperparameter optimization. ReLU = rectified linear unit.

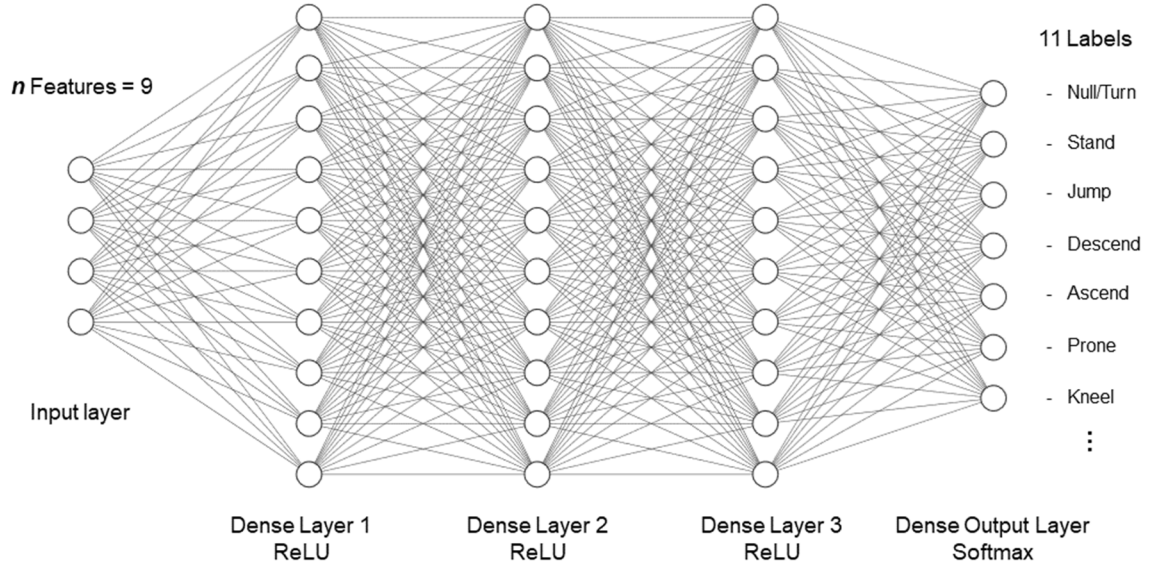

FIGURE 2. Fully connected neural network architecture. The nine features were the x, y, and z components of the accelerometer, gyroscope, and magnetometer signals from the inertial measurement unit. The sliding window size and the number of input units in each dense (i.e., fully connected) layer were tuned during hyperparameter optimization. ReLU = rectified linear unit.

functions, and a dense output layer with a softmax activation function (Fig 2). As a form of regularization, a drop-out operator with a probability set to $p = 0.5$ was implemented on the output of the dense layers. Model parameters were optimized during training using SGD. The SGD hyperparameters (i.e., learning rate and momentum) and the number of input units in the dense layers were tuned during hyperparameter optimization (Table 3).

D. DEEP CONVOLUTIONAL LONG SHORT-TERM NEURAL NETWORK

The DeepConvLSTM model was comprised of an input layer, four convolutional layers with ReLU activation functions, two long short-term memory recurrent layers with hyperbolic tangent (tanh) activation functions, and a dense layer with a softmax activation function (Fig 3). A drop-out operator was implemented on the inputs of the recurrent layers with a probability set to $p = 0.5$.

Following procedures outlined by Ordóñez and Roggen (2016), the model parameters were optimized during training by minimizing the cross-entropy loss function using batch gradient descent and the RMSProp update rule. For RMSProp, the decay factor was set to $\rho = 0.9$, and the momentum was set to 0.0. The RMSProp learning rate, the number of convolutional filters, the convolutional kernel size, and the number of long short-term memory (LSTM) cells in

TABLE 3. The hyperparameter search spaces for each architecture.

	Hyperparameters
CNN	Window size = [64, **128**] CNN channels = [**32**, 64, 96] Kernel size = [5, **7**, 9, 11] SGD learning rate = [**0.0001**, 0.001, 0.01, 0.1, 1] SGD momentum = [0.7, 0.9, **0.95**, 0.98]
FC	Window size = [**64**, 128] Input units = [64, 128, 192, 256, **320**] SGD learning rate = [**0.0001**, 0.001, 0.01, 0.1, 1] SGD momentum = [0.7, 0.9, 0.95, **0.98**]
DeepConvLSTM	Window size = [64, **128**] CNN channels = [32, 64, **96**] Kernel size = [5, **7**, 9, 11] LSTM cells = [64, 128, **192**] RMSProp learning rate = [**0.0001**, 0.001, 0.01, 0.1, 1]

Note: The values for the CNN channels, input units, and LSTM cells were held constant across layers within a model. CNN = convolutional neural network; DeepConvLSTM = deep convolutional LSTM neural network; FC = fully connected neural network; SGD = stochastic gradient descent; LSTM = long short-term memory; RMSProp = root mean squared propagation. Bolded values represent the optimized hyperparameters that were used to train the models for all approaches.

the LSTM recurrent layers were tuned during hyperparameter optimization (Table 3).

V. PREDICTION PROCESSING

A. LOGICAL ALGORITHM

The predictions from the DNNs were processed through a two-step logical algorithm implemented in Matlab 2016b using an Acer Laptop with Intel®Core™i7-8750H CPU @ 2.20GHz 2.21 GHz, 16 GB of RAM. In step one, the initial predictions from the DNN were processed to update the original label set to ensure that the predictions adhere to real-world constraints. Examples of constraints are: no standing after a descend, no standing before an ascend, kneel/prone must be preceded by a descend if the previous label is a stand, walk, or run, etc. The input to the logical algorithm was a 3 × 5 matrix where the columns were the class labels (rows 1 and 2 corrected; rows 3-5 predicted), the number of windows for each class label (i.e., multiple windows can contribute to the class labels), and the mean probability for those windows for that class. The current step for the logical algorithm was row three such that the algorithm could consider the previous two and future two class labels. Considering real-world constraints and the DNN's confidence in those labels, the logical algorithm altered the current (i.e., row 3) class label, which was fed into the next step. If the step failed, the logical algorithm was given the last and next five labels, window sizes, and probabilities to make a decision and could alter any previous label if real-world constraints were violated.

After completing the first step, the second step iterates through the updated original label set to generate the expanded label set (Table 1). To expand the labels, the algorithm was shown the similarly structured 3 × 5 matrix described above; however, the algorithm only made changes to the ascend and descend classes based on what the previous and next class labels were. For example, if the current class was a descend with a previous class of running and a next class of kneeling, the current class descend would be altered to a run-to-kneel.

B. SURVIVABILITY TRADESPACE

Using the expanded label set, the survivability tradespace metrics were calculated: exposure time, the number of enemy shots fired, and susceptibility to enemy action. For the current investigation, the exposure time for each movement was calculated with the assumption that while the participant was on the ground (i.e., kneel, prone, crawl, KTP, PTK), they were considered to be not exposed. For all other movements, participants were considered exposed to enemy action.

Due to the nature of the protocol (especially during the indoor collections), participants stood at the starting and finish line for a variable amount of time. Therefore, all laboratory-based movement trials were cropped to exclude the standing portion at the start or end of the movement (i.e., calculations started at the first movement following a stand and ended at the last movement before a stand). Using the same rationale, the section attack data were cropped between the first descend into a kneel/prone after the initial patrol and the last ascend before the simulated assault. For the indoor movements, the exposure time was calculated for each movement and averaged within each load condition. For the section attack, the exposure time was individually calculated at each instance of exposure between two periods when the participant was unexposed (i.e., one instance of exposure can be represented as the series KTR, run, RTK, and kneel, for example).

The calculated exposure time for each movement was used to determine the number of enemy shots that could be fired at the participant using Equation 1 [17], [39]:

$$Shots = (Exposure\ Time - Reaction\ Time) * Shooting\ Cadence \quad (1)$$

where reaction time was set to one second and shooting cadence was set to 1.3 shots per second [17], [39]. Given the number of shots fired, susceptibility was calculated with a shooting accuracy of 10% (Equation 2 [17], [39]):

$$Susceptibility = \left(1 - (1 - Accuracy)^{Shots}\right) * 100 \quad (2)$$

where susceptibility is represented as a percentage chance of being hit by the number of shots calculated in Equation 1.

VI. RESULTS

Model performance results represent the whole time series, while the tradespace metrics represent the cropped labels as described in section V-B. Model performance was evaluated using the windowed true labels described above, while the

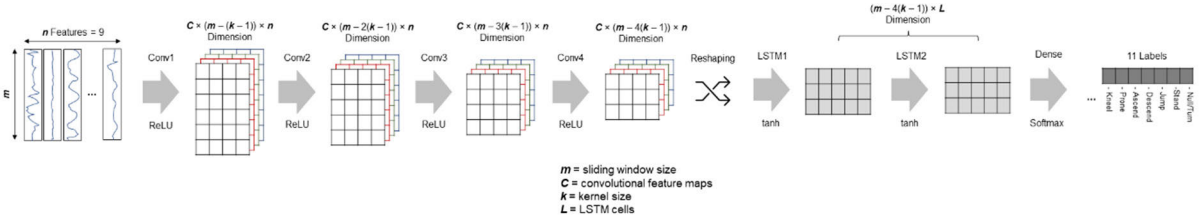

FIGURE 3. Deep convolutional long short-term memory neural network architecture. The nine features were the x, y, and z components of the accelerometer, gyroscope, and magnetometer signals from the inertial measurement unit. The sliding window size, the number of convolutional feature maps (i.e., number of filters), the kernel sizes, and the number of long short-term memory (LSTM) cells in each recurrent layer were tuned during hyperparameter optimization. Conv = convolutional layer; ReLU = rectified linear unit; tanh = hyperbolic tan function.

TABLE 4. Human activity recognition performance for all architectures across all approaches.

	Indoor	Outdoor	General
CNN	Loss = 0.688 Accuracy = 79.86% F1-score = 79.52%	Loss = 0.657 Accuracy = 82.10% F1-score = 82.97%	Loss = 0.638 Accuracy = 81.48% F1-score = 81.41%
FC	Loss = 0.610 Accuracy = 82.26% F1-score = 81.86%	Loss = 0.509 Accuracy = 83.58% F1-score = 84.80%	Loss = 0.481 Accuracy = 84.81% F1-score = 84.58%
DeepConv LSTM	Loss = 0.486 Accuracy = 85.87% F1-score = 86.11%	Loss = 0.406 Accuracy = 87.06% F1-score = 87.36%	Loss = 0.459 Accuracy = 85.88% F1-score = 86.14%
DeepConv LSTM after Logical Algorithm	Accuracy = 91.72% F1-Score = 91.79%	Accuracy = 90.64% F1-Score = 91.11%	Accuracy = 87.70% F1-Score = 87.86%

CNN = convolutional neural network, FC = fully connected neural network, DeepConvLSTM = deep convolutional long short-term memory neural network, F1-score = weighted average F1-score.

TABLE 5. Training and test set prediction times on the general dataset for all three deep neural network architectures.

	Epochs before early stopping	Total training time	Average time per epoch	Prediction time on test set
CNN	59	73.93 s (1.23 min)	1.25 s	0.867 s
FC	99	588.65 s (9.81 min)	5.94 s	1.768 s
DeepConv LSTM	99	944.68 s (15.74 min)	9.54 s	1.932 s

CNN = convolutional neural network, FC = fully connected neural network, DeepConvLSTM = deep convolutional long short-term memory neural network. Times are based on a mirrored strategy across three, Nvidia Titan RTX GPUs (Nvidia, USA) with 24 GB of GDDR6 memory each, performed on a server running Ubuntu 18.04 powered by two, Intel Xeon Gold 6248 CPUs @ 2.50GHz (Intel, USA) and 384 GB of DDR4 RAM.

original true labels were used for tradespace calculations (i.e., the true labels manually identified by the research team before windowing).

A. DEEP NEURAL NETWORK PERFORMANCE

The performance of the DNNs for activity recognition is presented for the original label set only. Across all approaches, the DeepConvLSTM architecture had less loss, greater accuracy, and greater weighted average F1-score compared to the CNN and FC models when evaluated using the test subsets (Table 4). The training and prediction time on the general dataset (i.e., indoor and section attack data combined) were longest for the DeepConvLSTM (15.74 minutes), followed by the FC (9.81 minutes) and CNN (1.23 minutes) models (Table 5). For the general dataset, the logical algorithms took 9.54 seconds to process 165,150 frames of data. The remainder of the results are computed using the output from the DeepConvLSTM models.

TABLE 6. Accuracy and weighted average F1-scores across load conditions.

Approach	Metric	Slick	FFO	Backpack	Pockets
Indoor	Accuracy	90.91	92.86	87.87	-
	F1-Score	90.88	92.93	87.90	-
Section Attack	Accuracy	85.72	90.01	93.02	93.03
	F1-Score	87.13	90.61	93.20	93.26
General	Accuracy	82.26	91.50	85.49	93.15
	F1-Score	82.89	91.47	85.47	93.76
Mean	Accuracy	86.29	91.46	88.79	93.09
SD	F1-Score	4.355	1.423	3.852	0.080

All values are expressed as a percentage. The indoor data collection did not include a Pockets condition. SD = standard deviation.

B. LOGICAL ALGORITHM

Once the predictions were analyzed through the logical algorithm, on average, across all three model approaches, accuracy and weighted F1-scores were 90.0% and 90.3%, respectively (Table 4). This represents an average increase in accuracy and weighted F1-scores of 3.75% and 3.72%, respectively. The General model had a slightly lower overall accuracy (87.7%) compared to the Indoor (91.7%) and Section Attack (90.6%) models. Load condition did not appear to greatly influence the performance of the models (Table 6).

The worst classified load condition was the General model classifying the Slick condition (accuracy: 82.3%; F1: 82.9%), while the Pocket condition classified by the General model was the best (accuracy: 93.1%; F1: 93.8%).

C. TRADESPACE METRICS

For the four unique test participants, all models identified an increase in exposure time with an increase in body-borne load. On average, across all trials/instances of exposure,

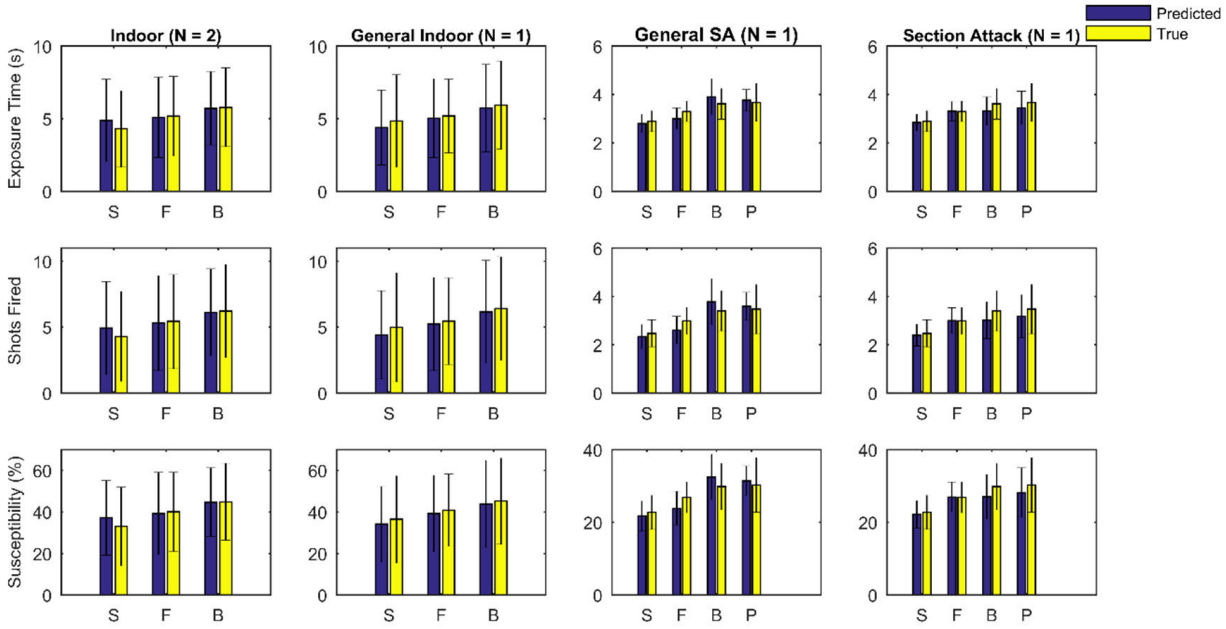

FIGURE 4. Tradespace results for four unique test participants across all approaches. S = Slick, F = Full Fighting Order (FFO), B = Backpack, P = Pockets, General SA = the General model's results on the section attack data, General Indoor = the General model's results of the indoor data, N = the number of test participants included in each comparison. Error bars represent standard deviation of the mean for each trial performed indoors or each instance of exposure during the section attack.

compared to the true labels, the predicted labels had an absolute difference in exposure time of 0.17 s, 0.13 s, and 0.16 s for the Indoor, Section Attack, and General models, respectively. This difference in exposure time between true and predicted labels across all trials/instances of exposure led to an average absolute difference in the number of shots fired of 0.17, 0.16, and 0.21 for the Indoor, Section Attack, and General models, respectively. Ultimately, these differences between the predicted and true labels led to an average absolute difference in susceptibility of 1.25%, 1.20%, and 1.09% for the Indoor, Section Attack, and General models, respectively, across all trials/instances of exposure. Differences in tradespace metrics between the true and predicted labels can be visualized in Fig 4.

VII. DISCUSSION

To further our knowledge into how survivability tradespace metrics are impacted by soldier burden, methods are needed to collect soldiers' movement patterns in their operational environment. These methods need to use technology that is simple (i.e., does not need extensive expertise to operate), cost-effective (i.e., so multiple soldiers can be collected at once), highly transportable (i.e., can be deployed anywhere), and provide meaningful information that can be disseminated to military leaders. Current methods of calculating soldier survivability metrics include timing gates [10], and GPS + accelerometer values on pre-setup courses [17], [18]. While these studies have helped establish foundational knowledge, these methods are suitable for targeted data collections rather than in-the-wild soldier monitoring. In the current work, we have shown how we can address the needs of the military and build on previous work by using a single IMU that is placed on the upper back combined with a DNN to perform HAR.

Our investigation found that the DeepConvLSTM architecture was best at predicting our eleven movement classes. This was likely influenced by the data type (i.e., time series) and the node architecture allowing LSTM nodes to retain memory from previous data windows [23], whereas the CNN and FC models interpret data windows as unique inputs. A benefit of employing models with less complex architectures (i.e., CNN and FC) is shorter training and prediction times. However, considering the intended use cases for this analytical framework, model performance is valued over training time, and the DeepConvLSTM prediction time (i.e., ∼2 s) is still suitable for field-based processing. Thus, the remainder of this discussion will focus on results from the DeepConvLSTM models. Through this investigation, we used three approaches to model development: Indoor-specific, Section Attack-specific, and General. Although not substantial, the Indoor and Section Attack-specific models were more accurate than the General model (90.9% and 90.2% vs. 87.9%, respectively) for classifying the movements. Since data distribution can have adverse effects on a DNN's accuracy [40], the reduced performance in the General model may be caused by the distribution of the Indoor and Section Attack data within the General model's training set (62% Section Attack, 38% Indoor) and that the data collection and preparation methods between Indoor and Section Attack were different (i.e., a structured protocol

cropped into individual movements compared to a continuous movement protocol).

To ensure that our prediction adhered to real-world movement constraints, we processed the original predictions through a two-step logical algorithm, which, on average, increased our models' accuracies by 3.45% for the original labels (i.e., 86.4% vs. 89.8%). When expanding the original labels to the expanded label set (i.e., all 19 labels to give meaning to ascend and descend), accuracy dropped slightly to 87.0% on average, which still represents a small accuracy improvement of 0.6% from the original DNN predictions, with the additional contextual benefit of eight labels. With this additional context, combined with the time dependencies, we could produce digital reconstructions of the movement patterns (e.g., [16]), which can be imported into battle simulators, animation software, and musculoskeletal models.

Soldiers can be tasked to carry a variety of body-borne loads; therefore, to increase generalizability, we trained our models on three common body-borne load conditions (Slick, FFO, Backpack) and one hypothetical load condition (Pockets). All approaches were sensitive enough to identify an increase in exposure time due to body-borne load, a well-established phenomenon [41]. Across all approaches, there were no meaningful performance differences between load conditions; however, across all approaches, the Pocket (93.1%) and FFO (91.5%) conditions tended to have higher accuracies than the Slick (86.3%) and Backpack (88.8%) conditions. Given that the IMU was placed on the upper back, the differences seen could possibly be due to the movement patterns in the FFO and Pockets condition being in a ''sweet spot'' between an upright posture in the Slick and a more flexed posture in the Backpack condition.

Previous literature using HAR models for military movements have reported higher overall accuracies than the presented work. These investigations used a single IMU [29], a single triaxial accelerometer and GPS data [20], and a uniaxial accelerometer with a heart rate monitor [19]; the authors collectively report overall accuracies ranging from 87.5% to 98.5%. The accuracy differences between previous literature and the present work are primarily attributable to the amount and types of movements studied (i.e., walking, running, jumping, and loaded marching), representing a small subset of the movements that soldiers typically perform and fewer than the number of movements included in the present study. Wyss and Mäder (2010) included manual materials handling tasks (lifting and lowering, carrying, and digging) in their HAR model; however, their model only achieved an average accuracy of 54% for these tasks. Across our three approaches, the DeepConvLSTM architecture trained on data from a single IMU produced movement-specific accuracies of 89.1% for walking and 88.6% for running following our two-step logical algorithm that are comparable to previous literature. Aside from Null (66%), the least accurate movement class was Descend at 81.2%, while the most accurate label was Standing at 95.2%. This tight range of accuracies between class labels highlights that by using a DeepConvLSTM architecture, all of our movements are identifiable with a consistently high level of accuracy, building on the large range of accuracies previously reported.

Although the presented results are promising, there are several limitations to the current work. The sliding window method was employed to create segments of 128 frames to train the DNNs. This method reduced our resolution, as movements that were performed in less than 65 frames (i.e., the majority of 128-frame window) are not captured. This primarily affected descend movements during the section attack, which negatively affected the model performance metrics. That is, when the models correctly predicted descend, they were evaluated as errors when the majority of the windowed true label was not a descend. To overcome this limitation, model predictions used to calculate the tradespace metrics (computed after the logical algorithm) were compared to the un-windowed true labels to ensure that all movements were captured. The ''null'' class label was used as a catch-all for all movements that did not fit into other classes thus, there was a large amount of variability for this label, which had a negative impact on performance (i.e., accuracy for this label was 66%). Future work should aim to collect greater samples of these underrepresented movements to avoid the need for a null label. In the present work, we used a DNN approach; however, there are less computationally expensive approaches that have been deployed to HAR. A DNN approach was used to avoid the need for feature engineering to simplify the utilization of our methods for future end users, who may not have the necessary expertise in signal processing/domain knowledge.

The goal of the current work was to display the efficacy of using a single IMU for recognizing tactical movement patterns. As such, there is much more work to be done to realize the full potential of this field of research. Future work should aim to collect more movements as part of realistic situations (i.e., collecting during drills/simulated combat), diversify the movement patterns (e.g., manual materials handling, stair/ladder climbing, window vaulting, etc.), diversify the type of soldier burden experienced by the soldier (e.g., environmental conditions, equipment properties, etc.), diversify the operational environment (e.g., terrain, weather, proximity to operationally-relevant equipment), and diversify the training population (i.e., sex, occupation, rank, experience, etc.). By diversifying the data collection and integrating real-world tactical situations, it will make any developed HAR algorithm more generalizable to the diverse soldier population who operate in various environments around the world. In order to provide defence scientists with recommendations on the optimal number and location of IMUs to use to perform HAR for military movements, future work should focus on a sensitivity analysis using different combinations of all sensors from the Xsens whole-body IMU suit, uncoupled from a biomechanical model.

Presented is a HAR method for identifying military-based movements during a variety of body-borne loads using a single deployable IMU that is being widely used by the Canadian

Armed Forces and allied nations. We used predictions from DeepConvLSTM models to calculate tradespace metrics to produce meaningful information with a high degree of accuracy. Defence scientists and military leaders can employ our method to directly calculate tradespace metrics, use it as a preprocessing tool to supplement other technologies (e.g., heart rate monitor, temperature probes, motion capture), perform workplace assessments (e.g., work-to-rest ratios), identify how the same soldier burden is affecting different soldiers, or as a monitoring tool during deployment, among other applications. The ability to collect meaningful information in the field will accelerate our understanding of the survivability tradespace (i.e., soldier burden factors and their influence on soldier performance, vulnerability, and operational effectiveness). Filling knowledge gaps in this tradespace will also inform the development and employment of modular, scalable protection systems, and data-driven decision support tools to aid military leaders with risk-based trade-off decisions during mission planning.

REFERENCES

[1] N. Armstrong and L. L. M. Bossi, *Introduction*. Brussels, Belguim, 2021, ch. 1.

[2] R. L. Attwells, S. A. Birrell, R. H. Hooper, and N. J. Mansfield, "Influence of carrying heavy loads on soldiers' posture, movements and gait," *Ergonomics*, vol. 49, no. 14, pp. 1527–1537, Nov. 2006.

[3] T. N. Brown, M. O'Donovan, L. Hasselquist, B. D. Corner, and J. M. Schiffman, "Body borne loads impact walk-to-run and running biomechanics," *Gait Posture*, vol. 40, no. 1, pp. 237–242, May 2014.

[4] J. W. Ramsay, C. L. Hancock, M. P. O'Donovan, and T. N. Brown, "Soldier-relevant body borne loads increase knee joint contact force during a run-to-stop maneuver," *J. Biomech.*, vol. 49, no. 16, pp. 3868–3874, Dec. 2016.

[5] B. X. Liew, S. Morris, and K. Netto, "Joint power and kinematics coordination in load carriage running: Implications for performance and injury," *Gait Posture*, vol. 47, pp. 74–79, Jun. 2016.

[6] B. X. W. Liew, S. Morris, and K. Netto, "The effects of load carriage on joint work at different running velocities," *J. Biomech.*, vol. 49, no. 14, pp. 3275–3280, Oct. 2016.

[7] T. Karakolis, B. A. Sinclair, A. Kelly, P. Terhaar, and L. L. M. Bossi, "Determination of orientation and practice requirements when using an obstacle course for mobility performance assessment," *Hum. Factors*, vol. 59, no. 4, pp. 535–545, Jun. 2017.

[8] C. E. Pandorf, E. A. Harman, P. N. Frykman, J. F. Patton, R. P. Mello, and B. C. Nindl, "Correlates of load carriage and obstacle course performance among women," *Work*, vol. 18, no. 2, pp. 179–189, 2002.

[9] R. V. Vitali, S. M. Cain, L. V. Ojeda, M. V. Potter, A. M. Zaferiou, S. P. Davidson, M. E. Coyne, C. L. Hancock, A. Mendoza, L. A. Stirling, and N. C. Perkins, "Body-worn IMU array reveals effects of load on performance in an outdoor obstacle course," *PLoS One*, vol. 14, no. 3, pp. 1–30, 2019.

[10] A. K. L. Treloar and D. C. Billing, "Effect of load carriage on performance of an explosive, anaerobic military task," *Mil. Med.*, vol. 176, no. 9, pp. 1027–1031, Sep. 2011.

[11] L. L. M. Bossi, A. Morton, A. Sy, A. Keefe, T. Karakolis, and M. Jones, "Understanding the trade-offs between protection, performance and integrated survivability," *J. Sci. Med. Sport*, vol. 20, p. S139, Nov. 2017.

[12] L. L. M. Bossi, M. L. H. Jones, A. Kelly, and D. W. Tack, "A preliminary investigation of the effect of protective clothing weight, bulk and stiffness on combat mobility course performance," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, 2016, pp. 701–705.

[13] D. Laidig, T. Schauer, and T. Seel, "Exploiting kinematic constraints to compensate magnetic disturbances when calculating joint angles of approximate Hinge joints from orientation estimates of inertial sensors," in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, Jul. 2017, pp. 971–976.

[14] M. Schepers, M. Giuberti, and G. Bellusci, "Xsens MVN: Consistent tracking of human motion using inertial sensing," *Xsens Technol.*, vol. 1, no. 8, pp. 1–8, 2018.

[15] M. P. Mavor, G. B. Ross, A. L. Clouthier, T. Karakolis, and R. B. Graham, "Validation of an IMU suit for military-based tasks," *Sensors*, vol. 20, no. 15, pp. 1–14, 2020.

[16] M. P. Mavor, K. M. Gruevski, G. B. Ross, M. Akhavanfar, A. L. Clouthier, L. L. Bossi, T. Karakolis, and R. B. Graham, "A data-driven framework for assessing soldier performance, health, and survivability," *Appl. Ergon.*, vol. 104, Oct. 2022, Art. no. 103809.

[17] D. C. Billing, A. J. Silk, P. J. Tofari, and A. P. Hunt, "Effects of military load carriage on susceptibility to enemy fire during tactical combat movements," *J. Strength Conditioning Res.*, vol. 29, no. 11, pp. S134–S138, 2015.

[18] A. P. Hunt, P. J. Tofari, D. C. Billing, and A. J. Silk, "Tactical combat movements: Inter-individual variation in performance due to the effects of load carriage," *Ergonomics*, vol. 59, no. 9, pp. 1232–1241, Sep. 2016.

[19] T. Wyss and U. Mäder, "Recognition of military-specific physical activities with body-fixed sensors," *Mil. Med.*, vol. 175, no. 11, pp. 858–864, Nov. 2010.

[20] C. M. Clements, M. J. Buller, A. P. Welles, and W. J. Tharion, "Real time gait pattern classification from chest worn accelerometry during a loaded road March," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug./Sep. 2012, pp. 364–367.

[21] B. Russell, A. McDaid, W. Toscano, and P. Hume, "Moving the lab into the mountains: A pilot study of human activity recognition in unstructured environments," *Sensors*, vol. 21, no. 2, pp. 1–14, 2021.

[22] T. E. Lockhart, R. Soangra, J. Zhang, and X. Wu, "Wavelet based automated postural event detection and activity classification with single IMU," *Biomed. Sci. Instrum.*, vol. 49, pp. 224–233, Aug. 2013.

[23] F. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, Jan. 2016.

[24] S. S. Bangaru, C. Wang, S. A. Busam, and F. Aghazadeh, "ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors," *Autom. Construct.*, vol. 126, Jun. 2021, Art. no. 103653.

[25] A. L. Clouthier, G. B. Ross, and R. B. Graham, "Sensor data required for automatic recognition of athletic tasks using deep neural networks," *Front. Bioeng. Biotechnol.*, vol. 7, p. 473, Jan. 2020.

[26] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.*, vol. 59, pp. 235–244, Oct. 2016.

[27] Ankita, S. Rani, H. Babbar, S. Coleman, A. Singh, and H. M. Aljahdali, "An efficient and lightweight deep learning model for human activity recognition using smartphones," *Sensors*, vol. 21, no. 11, pp. 1–17, 2021.

[28] A. Abobakr, D. Nahavandi, M. Hossny, J. Iskander, M. Attia, S. Nahavandi, and M. Smets, "RGB-D ergonomic assessment system of adopted working postures," *Appl. Ergon.*, vol. 80, Apr. 2018, pp. 75–88, 2019.

[29] N. Papadakis, K. Havenetidis, D. Papadopoulos, and A. Bissas, "Employing body-fixed sensors and machine learning to predict physical activity in military personnel," *BMJ Mil. Health*, vol. 2, pp. 152–156, Oct. 2020.

[30] M. F. Antwi-Afari, H. Li, W. Umer, Y. Yu, and X. Xing, "Construction activity recognition and ergonomic risk assessment using a wearable insole pressure system," *J. Construct. Eng. Manage.*, vol. 146, no. 7, Jul. 2020, Art. no. 04020077.

[31] M. P. Mavor, "Novel approaches for investigating the soldier survivability tradespace," Ph.D. dissertation, Faculty Health Sci., Dept. Hum. Kinetics, Univ. Ottawa, Ottawa, ON, Canada, 2022.

[32] A. L. Clouthier, G. B. Ross, M. P. Mavor, I. Coll, A. Boyle, and R. B. Graham, "Development and validation of a deep learning algorithm and open-source platform for the automatic labelling of motion capture markers," *IEEE Access*, vol. 9, pp. 36444–36454, 2021.

[33] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA, USA: CreateSpace, 2009.

[34] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, and M. Kudlur, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Conf. Operating Syst. Design Implement.*, 2016, pp. 265–283.

[35] A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Appl. Soft Comput. J.*, vol. 62, pp. 915–922, Jan. 2018.

[36] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newslett.*, vol. 12, no. 2, pp. 74–82, Mar. 2011.

[37] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, and L. Invernizzi, "Keras tuner," 2019. [Online]. Available: https://github.com/keras-team/keras-tuner/issues/264

[38] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1–9.
[39] E. M. Blount, S. I. Ringleb, A. Tolk, M. Bailey, and J. A. Onate, "Incorporation of physical fitness in a tactical infantry simulation," *J. Defense Model. Simul., Appl., Methodol., Technol.*, vol. 10, no. 3, pp. 235–246, Jul. 2013.
[40] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, pp. 1–54, Dec. 2019.
[41] P. W. Sanderson, S. Faulkner, G. Fordy, A. Garcia-Vicencio, S. Malgoyre, and T. Reilly, *Work Package 2: Literature Review on Soldier Load*. 2021, ch. 4.

MATTHEW P. MAVOR received the B.P.H.E. degree from Nipissing University, North Bay, ON, Canada, in 2015, and the M.Sc. and Ph.D. degrees in human kinetics from the University of Ottawa, Ottawa, ON, in 2018 and 2022, respectively.

He is currently a Postdoctoral Fellow with the University of Ottawa. From 2015 to 2022, he was a Research Assistant with the Movement Biomechanics and Analytics Laboratory, University of Ottawa. His research interests include using wearable sensors for characterizing and understanding movement patterns.

VICTOR C. H. CHAN received the B.Kin. degree and the M.Sc. degree in exercise science from the University of Toronto, Toronto, ON, Canada, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree in human kinetics with the University of Ottawa, Ottawa, ON.

From 2015 to 2018, he was a Research Assistant with the Musculoskeletal Biomechanics and Injury Prevention Laboratory, University of Toronto. From 2019 to 2023, he was a Research Assistant with the Movement Biomechanics and Analytics Laboratory, University of Ottawa. His primary research interests include the development of fatigue modeling frameworks, using wearable sensors and machine learning techniques.

KRISTINA M. GRUEVSKI received the B.Sc. degree from McMaster University, the M.Sc. degree in kinesiology from the University of Waterloo, and the Ph.D. degree in kinesiology from the University of Waterloo, in 2018. She completed a Postdoctoral Training with the University of Ottawa, from 2019 to 2020. Currently, she is a Defense Scientist with the Defence Research and Development Canada (DRDC), Toronto Research Centre. Her research interests include musculoskeletal injury prevention, biomechanics, and performance.

LINDA L. M. BOSSI received the B.Sc. degree in human kinetics from the University of Guelph, ON, Canada, in 1983, and the M.Sc. degree in ergonomics from Loughborough University, in 1993. As a uniformed Bioscience Officer (highest rank Lieutenant-Colonel), she led the development, testing, and evaluation of Canadian Aircrew Chemical Biological and Radiological Protection (1984–1991) and DRDC's Applied Research in support of soldier systems human factors (1993–2010). Since 2010, she has been a Civilian Defense Scientist with the DRDC, Toronto Research Centre, specializing in soldier systems integration. Her current research interests include the development and validation of objective measures, metrics, methods, and tools to inform the requirements for specification/design, selection, and employment of soldier clothing, protection, individual equipment, and small arms that improve soldier health, safety, mobility, lethality performance, and dismount team effectiveness.

She received the Canadian Decoration, in 1989, 1999, and 2009; the Chief of Defense Staff Commendation, in 1992; the Ergonomics Society Ulf Aberg Award, in 1995; the Human Factors and Ergonomics Society Best Ergonomics in Design Article Award, in 2005; two NATO Research and Technology Organization Scientific Achievement Awards, in 2010 and 2019; and the Distinguished Service Award from the Technical Cooperation Program, in 2021.

THOMAS KARAKOLIS received the B.Eng. degree and the M.A.Sc. degree in mechanical engineering from McMaster University, Hamilton, ON, Canada, in 2007 and 2009, respectively, and the Ph.D. degree in applied health sciences from the University of Waterloo, Waterloo, ON, in 2022. He is currently a Defense Scientist with the Defence Research Development Canada, an agency of the Canadian Department of National Defence, Toronto Research Centre. His research interests include characterizing physical performance and injury prevention in a military population.

RYAN B. GRAHAM received the B.Sc. degree in biology, the M.Sc. degree in biomechanics and ergonomics, and the Ph.D. degree in biomechanics and ergonomics from Queen's University, Kingston, ON, Canada, in 2005, 2008, and 2012, respectively. He is currently an Associate Professor with the Faculty of Health Sciences, School of Human Kinetics, University of Ottawa, Ottawa, ON, and the University Research Chair of biomechanics and data science for human health and performance. He is also cross-appointed to the Ottawa-Carleton Institute for Biomedical Engineering. His research program blends biomechanics, ergonomics, motor control, and computer science to improve musculoskeletal disorder understanding and prevention; develop and test new technologies and algorithms to objectively score movement quality in clinical, ergonomic, and sport settings; and assess whole-body movement stability, control, and loading during functional movements. He received the David Winter Promising Young Investigator Award from the Canadian Society for Biomechanics; and the Ontario Early Researcher Award, in 2018. In 2019, he was a recipient of the University of Ottawa Faculty of Health Sciences Dean's Award for Excellence in Research.