Wearable activity recognition for robust human-robot teaming in safety-critical environments via hybrid neural networks

Andrea E. Frank, Alyssa Kubota, and Laurel D. Riek
Computer Science and Engineering, UC San Diego

***Abstract*—In this work, we present a novel non-visual HAR system that achieves state-of-the-art performance on realistic SCE tasks via a single wearable sensor. We leverage surface electromyography and inertial data from a low-profile wearable sensor to attain performant robot perception while remaining unobtrusive and user-friendly. By capturing both convolutional and temporal features with a hybrid CNN-LSTM classifier, our system is able to robustly and effectively classify complex, full-body human activities with only this single sensor. We perform a rigorous analysis of our method on two datasets representative of SCE tasks, and compare performance with several prominent HAR algorithms. Results show our system substantially outperforms rival algorithms in identifying complex human tasks from minimal sensing hardware, achieving F1-scores up to 84% over 31 strenuous activity classes. To our knowledge, we are the first to robustly identify complex full-body tasks using a single, unobtrusive sensor feasible for real-world use in SCEs. Using our approach, robots will be able to more reliably understand human activity, enabling them to safely navigate sensitive, crowded spaces.**

Fig. 1. Activities from the MIT-UCSD Human Motion dataset. From left to right, top to bottom: *Walking*, *Scanning Part*, *Attaching Part*, *Pincer Grasp*, *Palmar Grasp*, *Thumb-3 Fingers Grasp*.

I. Introduction

As the technology becomes more robust, capable, and affordable, robots are increasingly recognized as invaluable assets in dynamic, safety-critical environments (SCEs), such as emergency departments (EDs) and manufacturing plants. Clinicians and manufacturing workers routinely face substantial physical and cognitive burden, placing them among the populations most at-risk of developing work-related injury and burnout [1, 2]. For instance, Welfare et al. [3] found that manufacturing workers often worry about health issues such as poor ergonomics and high physical demands in the workplace. Similarly, burnout among physicians can deteriorate their physical and mental health [4, 5], as well as increase the risk of preventable medical errors, the third leading cause of death in the United States [6]. Robots' consistency, precision, and physical strength make them promising candidates for alleviating or eliminating tasks that contribute to these burdens. However, existing robots lack the situational understanding to navigate in dynamic, uncertain, and risk-laden workspaces around human agents to safely perform complex tasks.

If robots are to improve conditions for workers and clientele in SCEs, they need to autonomously and robustly recognize tasks performed by their human teammates [7, 8]. This topic of human activity recognition (HAR) in real-world, dynamic environments like SCEs is a central problem in robotics [9, 10]. To date, most HAR research has relied on visual sensor data, but visual sensing is not always viable for robots operating in SCEs due to poor lighting, prevalent occlusion, and disruptive installation [10, 11]. In addition, these settings are often privacy-sensitive, involving patients seeking medical attention, or employees handling proprietary materials. Thus, from an ethical engineering perspective, it is important to consider non-visual sensing approaches [12].

Thus, many researchers have shown interest in non-visual HAR (NVHAR). Prior work often leverages wearable sensors (e.g. accelerometers, gyroscopes) to capture body motion [9]. Recently, researchers started exploring surface electromyography (sEMG), a noninvasive technology that measures the electrical activity of skeletal muscles. Classifiers use sEMG data to differentiate subtle differences in fine motion that is difficult to capture with inertial sensors alone (c.f. [13, 14]).

Despite the abundance of techniques studied in NVHAR, only a handful of systems are designed to recognize tasks robots will encounter in real-world SCEs [10]. It is standard in the field to build and evaluate systems by recognizing activities of daily living (ADLs) (c.f. [15]) or a closed set of carefully chosen hand gestures (c.f. [13, 14]). However, the motions represented in such datasets are not generalizable to the complex tasks robots must respond to in SCEs, such as deft manipulation of specialized tools, or specific medical procedures [16, 17]. The question of how best to perform NVHAR in real-world SCEs remains an open research area.

Moreover, despite the potential for sEMG sensors to im-

Research reported in this paper is supported by the National Science Foundation under Grant Nos. IIS-1734482 and IIS-1724982.

Andrea E. Frank, Alyssa Kubota, and Laurel D. Riek are with the Department of Computer Science and Engineering, UC San Diego, La Jolla, CA 92093, USA (email: {a2frank, akubota, lriek}@eng.ucsd.edu)

prove classification accuracy and robustness, to our knowledge no work has explored how sEMG can benefit NVHAR in SCEs. The few projects that investigate NVHAR in SCEs have largely utilized extensive arrays of accelerometers (c.f. [18, 19]), occasionally augmented with nontraditional sensors such as microphones (c.f. [20]). Considering its success in recognizing fine hand motions and other traditionally difficult actions, sEMG presents a promising avenue for improving the accuracy of NVHAR in SCEs.

In this work, we present a novel NVHAR system to support robust human-robot teaming in SCEs. Our system informs the robot of its teammates' actions through inertial and sEMG signals captured by an unobtrusive armband. This raw multimodal input is processed by an hybrid neural network (NN) architecture that leverages the complementary benefits of convolutional and recurrent layers to capture complex spatial and temporal features. We evaluate our work on two datasets representative of tasks performed in SCEs: MIT-UCSD Human Motion [16], which consists of common manufacturing tasks (see Fig. 1), and MyoGym, a dataset of strenuous exercises demonstrating action primitives for manual labor. Evaluation results show that our system achieves state-of-the-art performance when presented with ample training data of relevant SCE tasks.

The contributions of the paper are threefold: First, to our knowledge, we are the first to compare the performance of several prominent NVHAR classifiers in addressing tasks specific to SCEs, rather than everyday activities such as ADLs. Second, we conduct an analysis of the effect of supplementing inertial data with sEMG on the feasibility of classifying whole-body SCE tasks from single-sensor recordings. Finally, we present a novel wearable NVHAR system that leverages hybrid deep learning to improve upon rival algorithms and achieve state-of-the-art human task awareness for robots in SCEs.

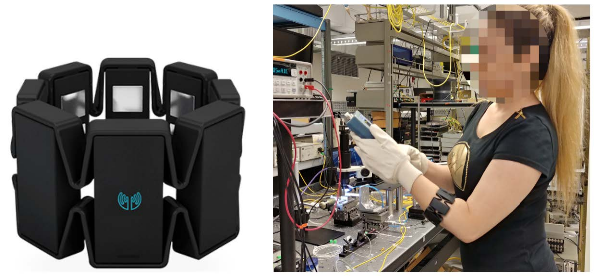
The approach presented in this work will enable robots to fluently understand and collaborate with human partners on complex, strenuous tasks, and confer the numerous benefits of human-robot collaboration to people in SCEs worldwide.

II. Background

A. Wearable NVHAR Sensors

The most common sensors used for NVHAR are IMUs [9, 10]. IMUs measure linear acceleration, rotational acceleration, orientation, or a combination of the three, and are often worn on the limbs or torso. Researchers have also taken advantage of the IMUs found in mobile devices for a variety of studies in-the-wild [9, 10, 21, 22].

A limited number of prior work investigated IMUs for NVHAR in SCEs. Stiefmeier et al. [19] fused 27 IMUs and radio-frequency identification sensors to recognize tasks on a car assembly line. Inoue et al. [18] recorded inertial data from multiple accelerometers to recognize a variety of nursing tasks. In contrast to these past systems, which employ complicated and bulky sensor arrays, our approach uses a single armband sensor in order to recognize activities with minimal encumbrance.



(a) (b)

Fig. 2. (a) The Myo armband can measure sEMG, linear acceleration, and angular acceleration of the wearer's arm movements. (b) Recognizing activities in a manufacturing setting with the Myo.

Recently, researchers have begun to investigate sEMG sensors for NVHAR, either alone or as a supplement to inertial signals. Several groups have employed sEMG in recognition and assessment of ADLs [23], balance [24], and gait [25]. Others have investigated fine hand motions, and have used arm and wrist sEMG to determine hand gestures [26], or to recognize American Sign Language [14]. However, these tasks do not represent the specialized activities or equipment that robots would encounter in SCEs. Furthermore, these systems often utilize numerous obtrusive sensors and are thus not appropriate for use in real-world SCEs.

In contrast, the Myo armband (see Fig. 2) is a compact, arm-worn device that houses an 8-channel sEMG and a 9-axis IMU [27]. Recent studies leverage affordable, unobtrusive sensor for exploring multimodal NVHAR [16, 28–30]. Researchers found that augmenting inertial sensors with sEMG sensors from the Myo considerably improves classification accuracy of strenuous exercises [29] and ADLs [30]. However, approaches such as that presented by Koskimäki et al. [29] still exhibit substandard results (up to 72% accuracy), leaving considerable room for improvement. Totty et al. [30] achieved up to 89.2% accuracy classifying ADL functional groups. However, the approach exhibits several limitations. First, the approach presented is unable to recognize the specific activity performed, but only the high level category (e.g. no activity, functional). In addition, the dataset considered only included basic upper extremity tasks, and does not represent the intensive whole-body tasks relevant to SCEs.

Despite the success of the Myo and of sEMG HAR in general, to our knowledge, there is no work demonstrating a system that can reliably recognize realistic, complex worker tasks performed in real-world SCEs. The complex networks of sensors suggested in studies such as [23] and [19] are cumbersome and delicate, which makes them unfit for use in real-world SCEs. Furthermore, none of these studies explored more than a few basic classifiers on inertial+sEMG data. It remains an open question what classification approach is best suited to decoding these complicated multimodal signals.

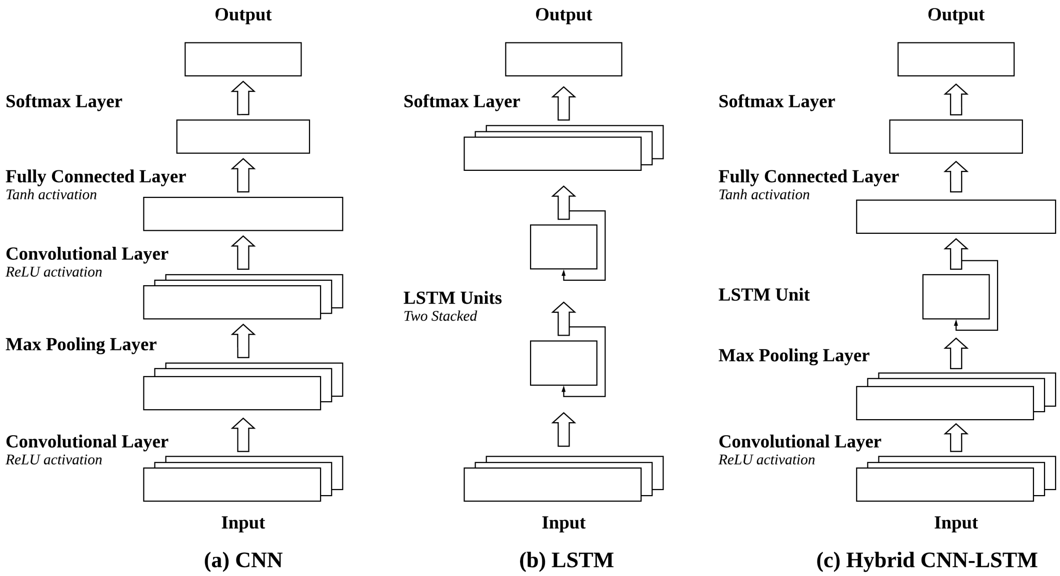


Fig. 3. The architectures of the deep HAR models used in this work. For clarity, the input vectors displayed represent a single time window in a batch. (a) The CNN extracts local patterns among multiple data channels. (b) The LSTM network identifies important temporal features. (c) The hybrid CNN-LSTM identifies temporal patterns using convolutional features.

B. NVHAR Classification

Researchers have employed a variety of classification techniques to address NVHAR. One attractive method is Linear Discriminant Analysis (LDA) as it easily generalizes to multi-class classification and does not require hyperparameter tuning [29]. Another widely-used approach is *k*-Nearest Neighbors (*k*-NN), which classifies new instances as the most common class of the *k* most similar training samples. *k*-NNs produce noteworthy results for inertial NVHAR of ADLs compared to other popular algorithms [30, 31].

Recently, successes in areas such as image and speech recognition have inspired researchers to employ deep learning for NVHAR. Deep learning approaches mitigate the need for hand-crafted features, which are difficult to design for mobile and wearable sensor streams [32]. The most common deep learning approaches for NVHAR are convolutional NNs (CNNs) and recurrent NNs (RNNs) [33]. CNNs construct spatial features from signals by taking convolutions of input channels at each timepoint. On the other hand, RNNs extract temporal features, i.e. how the evolution of the signal over time informs the prediction. Long short-term memory networks (LSTMs), the most widely used RNN, improve upon traditional RNNs by selectively truncating error gradients in backpropagation to allow the network to learn long-term dependencies in input signals [34].

CNNs have become popular in NVHAR for applications such as classifying ADLs (c.f. [35]) and fall detection (c.f. [36]). In addition, the ability to leverage the temporal structure of activity signals makes LSTMs a promising tool for NVHAR of tasks with complicated, time-dependent patterns [33]. Other recent work uses a combination of convolutional and recurrent layers for NVHAR (c.f. [33]). These combined CNN-LSTM architectures capitalize on the CNN layers' ability to extract convolutional features that best represent the state at each timestep, from which the LSTM layers learn the temporal evolution of that state over the input sequence. While CNN-LSTMs are rather new to NVHAR, they can achieve state-of-the-art accuracy on non-visual data [33].

III. Methodology

In this paper, we expand upon past work in the following ways. First, we evaluate the performance of several prominent NVHAR algorithms on realistic worker tasks to determine the most effective techniques for use in SCEs. In particular, we investigate three deep learning approaches (CNN, LSTM, CNN-LSTM) and two machine learning classifiers (*k*-NN, LDA). Second, we investigate the promise of supplementing inertial wearable sensors with sEMG across each dataset, task, and algorithm. Finally, we present and validate a cohesive NVHAR system that employs a single, practical armband sensor to effectively classify SCE tasks, enabling fluent human-robot interaction in these spaces.

A. Datasets and Preprocessing

We evaluate our system on two datasets representative of tasks common in SCEs: MIT-UCSD Human Motion [16] and MyoGym [29]. MIT-UCSD Human Motion consists of 24 trials each containing 13 gross and fine assembly tasks performed in a realistic automobile factory simulation: four gross motion dashboard assembly tasks, and nine fine motion block assembly and grasping tasks. MyoGym includes 10 participants performing 30 strenuous gym exercises, representative of lifting, pushing, and carrying tasks. All data in both datasets were collected by a Myo armband worn on

TABLE I

MEAN F1 SCORES OBTAINED FOR EACH DATA MODALITY ON EACH DATASET FOR EACH CLASSIFIER. ACROSS THE CLASSIFIERS, DATA MODALITY, AND DATASET, WE AVERAGED THE F1 SCORES FROM EVERY TRIAL. A HIGHER F1 SCORE IS BETTER.

	MIT-UCSD Human Motion					**MyoGym**				
	CNN-LSTM	CNN	LSTM	*k*-NN	LDA	CNN-LSTM	CNN	LSTM	*k*-NN	LDA
Inertial+sEMG	**.35**	**.36**	.22	.31	**.33**	**.84**	**.39**	**.28**	.36	**.74**
Only Inertial	.31	.36	**.24**	**.35**	.30	.84	.38	.23	**.39**	.69

the dominant forearm, and contain 6-channel inertial (tri-axial accelerometer, tri-axial gyroscope) and 8-channel sEMG data collected at 50 Hz.

Data were segmented into 50% overlap 1 second and 1.5 second input windows for MIT-UCSD and MyoGym, respectively. We use a shorter window for MIT-UCSD due to shorter task durations. We standardized each input channel of each train set to $\mu = 0$ and $\sigma = 1$ over all training data. To simulate real-time performance, we standardized test data to $\mu = 0$ and $\sigma = 1$ with a moving window of data points in the past 1 second. Because data in MyoGym were collected continuously through all 30 exercises, the *null* class represents approximately 78% of all training data. To discourage the trivial solution (i.e. always predicting the majority class), we reduced the number of *null* class instances through random undersampling of *null* sequences.

B. Classifiers

We built three NN classifiers to perform NVHAR: a CNN, an LSTM, and a hybrid CNN-LSTM. We aimed to minimize variation due to arbitrary hyperparameter choices (e.g. number of layers, activation functions) by making analogous design choices across networks. In this way, we ensure that differences in classification accuracy are more closely tied to the type of network than differences in these hyperparameters. Each network was designed with two convolutional or recurrent layers, a fully connected layer, and a softmax output layer (see Fig. 3). We use two feature-extracting layers for each network to control for layer ordering effects and isolate the effect of layer type on classification.

All kernels use a stride of 20 ms, the sample rate of the Myo. Convolutional layers used a kernel of 500 ms for MIT-UCSD, and a kernel of 1200 ms for MyoGym. We chose these relatively large kernels to simulate temporal memory in convolutional layers. Per convention, we apply a max-pooling layer between convolutional layers. This layer uses a kernel size of 40 ms. LSTM layers contained 64 hidden units, and fully connected layers contained 1000, as chosen by cross-validation. Convolutional, LSTM, and fully connected layers were activated with ReLU functions, and output layers used softmax activation for classification.

In order to compare to existing literature, we tested each dataset on an LDA (see [29]) and a k-NN ($k = 5$) (see [16, 30]). Since these classifiers cannot autonomously select informative features from data, we extracted 57 linear acceleration features, 54 angular velocity features, and 112 sEMG-based features as input, as recommended by Koskimaki et al. [29]. In contrast, our network algorithms were only fed raw data. This allows us to compare the efficacy of expert-recommended features against those generated autonomously by NNs for classifying SCE tasks.

C. Evaluation

All classifiers were trained separately on both datasets until convergence. We evaluated performance metrics based on leave-n-trials-out cross-validation. In order to ensure sufficient training data was available, we used $n = 1$ for MIT-UCSD, and $n = 3$ for MyoGym. We did not perform resampling or class-balancing on the test data to simulate a robot perceiving human actions in real-time.

We report micro-F_1 score as our evaluation metric, as it more faithfully represents classification performance across unbalanced classes compared to accuracy and macro-F_1 score. To analyze the variation in outcome measures, we performed a three-way repeated-measures analysis of variance (ANOVA) across classifier, data modality, and dataset.

IV. RESULTS

All effects are reported significant at $p < 0.05$. Mauchlys tests indicated that that the assumption of sphericity was violated for the main effect of classifier, as well as the interaction effects of classifier and data modality, and that of classifier and dataset. We corrected for this using Greenhouse-Geisser estimates of sphericity. Each of our measures had a significant main effect on F1 score (classifier: $F(2.06, 267.36)$=472.1, modality: $F(1, 130) = 5.9$, dataset: $F(1, 130) = 642.3$). There were also significant interaction effects between modality and classifier, $F(2.79, 362.76) = 10.9$, between dataset and classifier, $F(1.952, 253.76) = 359.7$, and between modality and dataset $F(1, 130) = 35.1$. This suggests that the type of sensing capabilities as well as dataset have different effects on classification accuracy depending on the classifier used.

Contrasts reveal that the CNN-LSTM performed significantly better than the other classifiers overall on the MyoGym dataset. This architecture performed consistently better than the LSTM and k-NN across all evaluations. The CNN-LSTM also performed significantly better than the LDA on MyoGym when sEMG was present, but saw no significant improvement over LDA in the other scenarios. On MIT-UCSD, there was no significant advantage shown using CNN-LSTM instead of LDA or CNN. There was no significant difference between the performance of the CNN-LSTM and LDA or CNN on the MIT-UCSD dataset. The classifiers performance decreased slightly but significantly across both datasets when no sEMG

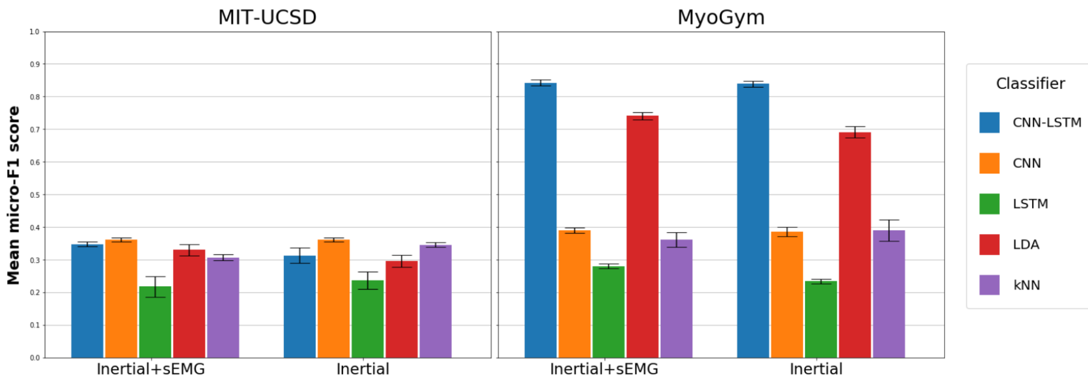


Fig. 4. Average micro-F1 scores (across all trials) for the classifier type and sensor data channels, separated by dataset.

signal was available, suggesting our system performs adequately even with less information available.

The average F1 score for the MIT-UCSD dataset across all classifiers and modalities was 31.2 $\pm$ 5.0%, significantly lower than the performance on the MyoGym dataset (51.6 $\pm$ 23.6%.). All classifiers performed significantly better on the MyoGym dataset or had no significant change. Including sEMG in training and classification had significant overall positive effect on F1 score across classifier. The inclusion of sEMG significantly assisted every classifier except the overall performance of the LSTM, and the performance of the CNN on the MIT-UCSD dataset.

V. Discussion

Our evaluation suggests that a hybrid CNN-LSTM architecture offers superior identification of SCE tasks in a realistic environment compared to prominent rival techniques. We found that CNN-LSTM architecture excels in environments that exhibit strenuous pushing, pulling, and lifting tasks, attaining 84% accuracy across 30 different actions on the MyoGym dataset. Additionally, the hybrid architecture is on par with other state-of-the-art classifiers over the MIT-UCSD dataset. This suggests that the combination of convolutional and recurrent layers with forearm sEMG and inertial signals is a promising approach for supporting robot understanding of complex human activities in real-world environments.

Although popular in recent literature, our evaluation suggests that *k*-NN is unsuited to NVHAR in SCEs, even when aided by expert feature selection. This is interesting, as it has been widely validated as a suitable means for identifying ADLs [16, 30, 31]. This implies that more specific, alternative approaches, such as a hybrid CNN-LSTM, may be necessary to support safe and robust NVHAR in SCEs and other complex environments. Tasks in SCEs are complex and stochastic, and take even humans substantial time to learn when newly introduced [37]. In order to ensure safe and accurate NVHAR, it is important not to take a previously successful classifier's effectiveness for granted. Instead, one must evaluate all robot systems on realistic data for the target environment.

Beyond classifiers, our results also suggest benefits of supplementing inertial data with sEMG. We found that sEMG signals were informative for pushing and pulling tasks, and assisted most classifiers in broadly categorizing tasks that involved targeted hand movements, such as assembling blocks. They also helped in discerning between tasks with similar movements, such as reaching forward to receive an automobile part versus doing so to install it in the dashboard.

However, one limitation of this work is that due to the small size of the MIT-UCSD dataset (approximately 4000 1-second sequences), the relative performance of each classifier is difficult to gauge. In particular, the CNN-LSTM still has significant room for improvement, as it is widely known that NNs require substantial training data to learn informative features. This premise is supported by the poor performance of the other classifiers when trained on this dataset. Nevertheless, given our system requires only an unobtrusive wearable sensor to gather data, a real-world implementation should have little issue collecting ample training data for robust performance.

While we found that sEMG signals are beneficial in some cases, sEMG had a detrimental effect when classifying tasks that involved raising the arms and lifting. Furthermore, several deep learning approaches performed worse when sEMG was included, suggesting that the additional modalities may confound classification on smaller or more intricate datasets. Caution and careful testing should be used when exploring whether sEMG sensing benefits future NVHAR applications. Future work will explore several avenues for expanding this NVHAR system. As the purpose of this work was to identify the most effective technique for SCEs, we performed no hyperparameter optimization. Moving forward, we will fine-tune hyperparameters and explore other NN architectures. In addition, in order to continue developing systems that perform in real SCEs, we intend to to gather a larger dataset of real-world manufacturing and ED tasks. We will make these datasets publicly available to empower the robotics community to investigate NVHAR in SCEs.

In this paper, we propose a novel NVHAR system that aims to to support robot integration into SCEs by enabling robust identification of human activity. To our knowledge, this system is the first NVHAR approach that is able to robustly identify complex full-body tasks using a single, unobtrusive sensor feasible for real-world use in SCEs. With the ability to more accurately distinguish between complex activities, robots will no longer be excluded from human-dense, safety-critical environments. Through our work, researchers will be able to develop advanced robots that can improve health and quality of life of the millions of workers worldwide.

References

[1] I. Bragard, G. Dupuis, & R. Fleet, "Quality of work life, burnout, and stress in emergency department physicians: a qualitative review," *Eur. J. Emerg. Med.*, 2015.

[2] J. L. Vandergrift *et al.*, "Physical and psychosocial ergonomic risk factors for low back pain in automobile manufacturing workers," *Occup Environ Med*, 2012.

[3] K. S. Welfare *et al.*, "Considering the human work experience when integrating robotics in the workplace," in *HRI*. IEEE, 2019.

[4] T. D. Shanafelt *et al.*, "Burnout and satisfaction with work-life balance among us physicians relative to the general us population," *Arch Intern Med.*, 2012.

[5] L. D. Riek, "Healthcare robotics," *Commun. ACM*, 2017.

[6] J. T. James, "A new, evidence-based estimate of patient harms associated with hospital care," *J Patient Saf*, 2013.

[7] T. Iqbal, S. Rack, & L. D. Riek, "Movement coordination in human–robot teams: a dynamical systems approach," *T-RO*, 2016.

[8] T. Iqbal & L. D. Riek, "Coordination dynamics in multihuman multirobot teams," *RA-L*, 2017.

[9] M. Cornacchia *et al.*, "A survey on activity detection and classification using wearable sensors," *IEEE Sensors J.*, 2017.

[10] O. D. Lara, M. A. Labrador *et al.*, "A survey on human activity recognition using wearable sensors." *Commun. Surveys Tuts.*, 2013.

[11] A. Taylor & L. D. Riek, "Robot-centric human group detection," *HRI Social Robots in the Wild*, 2018.

[12] L. Riek & D. Howard, "A code of ethics for the human-robot interaction profession," *Proc. of We Robot*, 2014.

[13] I. Batzianoulis *et al.*, "Emg-based decoding of grasp gestures in reaching-to-grasping motions," *RAS*, 2017.

[14] J. Wu, L. Sun, & R. Jafari, "A wearable system for recognizing american sign language in real-time using imu and surface emg sensors," *JBHI*, 2016.

[15] A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Applied Soft Computing*, 2018.

[16] A. Kubota *et al.*, "Activity recognition in manufacturing: The roles of motion capture and sEMG+inertial wearables in detecting fine vs. gross motion," in *ICRA*. IEEE, 2019.

[17] S. Kumar, L. Moro, & Y. Narayan, "Perceived physical stress at work and musculoskeletal discomfort in x-ray technologists," *Ergonomics*, 2004.

[18] S. Inoue *et al.*, "Mobile activity recognition for a whole day: recognizing real nursing activities with big dataset," in *UbiComp*. ACM, 2015.

[19] T. Stiefmeier *et al.*, "Wearable activity tracking in car manufacturing," *IEEE Pervasive Comput*, 2008.

[20] P. Lukowicz *et al.*, "Recognizing workshop activity using body worn microphones and accelerometers," in *PERVASIVE*. Springer, 2004.

[21] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE Sensors J.*, 2015.

[22] Y. Vaizman, N. Weibel, & G. Lanckriet, "Context recognition in-the-wild: Unified model for multi-modal sensors and multi-label classification," *IMWUT*, 2018.

[23] R. Chereshnev & A. Kertész-Farkas, "Rapidhare: A computationally inexpensive method for real-time human activity recognition from wearable sensors," *J Amb Intel Smart En*, 2018.

[24] H. Ghasemzadeh, R. Jafari, & B. Prabhakaran, "A body sensor network with electromyogram and inertial sensors: Multimodal interpretation of muscular activities," *IEEE Trans. Inf. Technol. Biomed.*, 2010.

[25] Y. Sun, G.-Z. Yang, & B. Lo, "An artificial neural network framework for lower limb motion signal estimation with foot-mounted inertial sensors," in *BSN*. IEEE, '18.

[26] T. S. Saponas *et al.*, "Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces," in *CHI*. ACM, 2008.

[27] T. L. Inc, "Myo gesture control armband," https://www.myo.com/, 2016, accessed: 2018-01-08.

[28] J. G. Abreu *et al.*, "Evaluating sign language recognition using the myo armband," in *SVR*. IEEE, 2016.

[29] H. Koskimäki, P. Siirtola, & J. Röning, "Myogym: introducing an open gym data set for activity recognition collected using myo armband," in *Ubicomp & ISWC*. ACM, 2017.

[30] M. S. Totty & E. Wade, "Muscle activation and inertial motion data for noninvasive classification of activities of daily living," *BME*, 2017.

[31] K. Safi *et al.*, "Physical activity recognition using inertial wearable sensorsa review of supervised classification algorithms," in *ICABME*. IEEE, 2015.

[32] M. Hasan & A. K. Roy-Chowdhury, "A continuous learning framework for activity recognition using deep hybrid feature models," *MM*, 2015.

[33] N. Y. Hammerla, S. Halloran, & T. Ploetz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," *IJCAI*, 2016.

[34] S. Hochreiter & J. Schmidhuber, "Long short-term memory," *Neural computation*, 1997.

[35] R. Yao *et al.*, "Efficient dense labelling of human activity sequences from wearables using fully convolutional networks," *Pattern Recognition*, 2018.

[36] S. S. Khan & B. Taati, "Detecting unseen falls from wearable devices using channel-wise ensemble of autoencoders," *Expert Systems with Applications*, 2017.

[37] J. J. Van Merrienboer & J. Sweller, "Cognitive load theory and complex learning: Recent developments and future directions," *Educational psychology review*, 2005.