A 3DCNN-LSTM Hybrid Framework for sEMG-Based Noises Recognition in Exercise

MIN-WEN LIN¹, SHANQ-JANG RUAN¹, (Senior Member, IEEE), AND YA-WEN TU²
¹Department of Electronics and Computer Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan
²Sijhih Cathay General Hospital, New Taipei City 221, Taiwan
Corresponding author: Ya-Wen Tu (yawentu@gmail.com)

ABSTRACT Recently, surface electromyography (sEMG) has been used to detect running-related works. sEMG provides a non-invasive and real-time method that allows quantification of muscle energy. However, noises in sEMG signals are a serious issue to be considered as these will interrupt the analysis of muscular activity. Hence, this work aims at distinguishing between sEMG valid signals and noises during running exercise by taking advantage of the combination of 3D-CNN and LSTM, which we called 3D-LCNN. Furthermore, according to the possible cases that happen in the sEMG data-collection procedure, we proposed two data-augmentation approaches to expend our sEMG dataset, which are the simulation of the surface electrodes displacement on the skin and the muscle fatigue. Experiment results show that the classification accuracy of the proposed 3D-LCNN can achieve 90.52%. Additionally, this work provides excellent service-oriented architecture (SOA). The recognition process can be done after the subject placed the sEMG sensors and performed a trial. Therefore, the process can help clinicians or therapists to distinguish between sEMG valid signals and noises more efficiently.

INDEX TERMS Data-augmentation, deep learning, long short-term memory, sEMG, sEMG noises, 3D convolution neural network.

I. INTRODUCTION
In recent years, the number of runners keeps increasing. Although running is a great way to achieve health, it is associated with a high risk of running-related injuries. sEMG signals are biomedical signals that measure electrical signals generated from muscles during muscle contraction [1]. Due to sEMG signals can be acquired non-invasive, real-time, and applicability, sEMG signals have been widely used to detect running-related injuries [2]–[4].

In previous studies, the majority of research in sEMG-based muscle evaluation have focused on the isokinetic and isometric contraction [3]. For isokinetic contraction, the subject is asked to hold a dumbbell, the subject uses biceps to curl the dumbbell, and then lower the dumbbell to his/her side repeatedly. For isometric contraction, the subject is asked to hold a dumbbell in a static position. Besides the aforementioned applications, there has been a dramatic increase in the number of publications on sEMG pattern recognition for hand gestures [5]–[7]. It should be noted, however, these researches have been documented especially for isometric contraction, isokinetic contraction, and static action conditions. In the research of sEMG-based signals analysis, previous studies have rarely focused on running exercise.

Compared with the aforementioned sEMG-based applications, running exercise produces more noises. The sEMG signals measured in dynamic conditions, such as running and intensive exercise will cause a lot of shaking in the sEMG measurement device. Moreover, the sEMG signals will be recorded by surface electrodes above the muscle on the skin during exercise. However, the surface electrodes are easily influenced by skin conditions, such as sweat accumulation that distorts the signals and degrades the skin contact conditions [8]. Furthermore, the effects of environmental interference are not negligible, like humidity changes [8]. Hence, it is difficult to maintain robust surface electrodes contacted the skin. Besides, sEMG signals are susceptible to various noises, which can be categorized into the following types: inherent noise in electronics equipment, ambient noise, cross talk, and inherent instability of signal [1]. In addition to the aforementioned problems, the noises also occur from multiple sensors receiving asynchronously. These noises in the sEMG
signals are serious issues to be considered, as these will cause distortion to the measured signals. In previous work, Malboubi et al. [9] proposed an adaptive IIR Laguerre filter for removing some certain noises. However, the filter only eliminates the power line noise of sEMG signals. Amrutha and Arul [1] discussed the basic noises removal techniques, which can be categorized into the following types:

- **Low pass differential filter**: The drawbacks of this method is that in low signal-to-noise ratio conditions, the high-frequency noises will be noticeable. Moreover, spectral leakage occurs when the wave and the sampling frequency are not the greatest common factor.
- **Adaptive noise cancellation**: As this technique relies on the noise model presented, the performance varies depending on the noise model. Additionally, the non-stationary nature of sEMG signals has a great impact on the performance of these methods.
- **Signal filtering based on wavelets**: The wavelets take a complex computational, therefore difficult to implement in real-time applications. The challenge in wavelets is selecting the most optimum mother wavelet for analyzing the signals, as various mother wavelets applied on the signal may produce different results.

All of the aforementioned techniques cannot completely eliminate all these noises. However, these noises and complicated patterns have effects on the result of feature extraction and diagnosis of the sEMG signals, especially in cases of running and intensive exercise.

In recent years, deep learning has become a powerful technique for various applications, it can extract the features from the mass of data automatically instead of adopting handcrafted features, which mainly depends on prior knowledge of designers. In [10]–[12] they have shown that deep learning have been applied to sEMG pattern recognition. Most importantly, they have shown deep learning methods are more suitable and effective than traditional classifiers. Therefore, deep learning approach is worth considering for solving noises identification problems. In this paper, we focus on deep learning-based classification methodologies for sEMG noises. Furthermore, according to the possible cases that happen in the sEMG data-collection procedure, we design two data-augmentation approaches to expend our sEMG dataset, which is very helpful for sEMG noises classification accuracy.

II. PROPOSED METHODOLOGY

sEMG signals are non-stationary and sensitive to many factors. sEMG signals will also change, even if he/she is the same person, which resets the surface electrodes. Therefore, the performance of the classifiers would degrade if we are not recalibrated. However, supervised recalibration of the classifiers by asking users to repeat a training protocol is difficult. To overcome these shortcomings, an efficient neural network architecture for sEMG-based recognition of valid signals is proposed in this paper. This work aims at distinguishing between sEMG valid signals and noises during running exercise to help users to get reliable signals. In sEMG pattern recognition applications, three main modules named pre-processing, feature extraction, and classification should be carefully considered. For this objective, this section is structured as follows. First, we describe the full details of the experimental setup that include the sEMG measurement device, experimental procedures, and participants. After which data pre-processing is presented. Finally, we describe
the implementation details of the proposed framework: 3D-LCNN.

Figure 1 shows that the proposed method consists of three processes: pre-processing, feature extraction, and classification.

A. EXPERIMENTAL SETUP

1) PARTICIPANTS
Thirty subjects (17 males and 13 females) volunteered to participate in the study. The subjects were all healthy and no known neuromuscular disorders or musculoskeletal injuries. Some instructions on how to perform running were given to subjects beforehand to ensure the consistency of the running actions and the rationality of the data analysis.

2) EXPERIMENTAL PROTOCOL
Each subject needs to at least one complete cycle of running. By definition, during one cycle of running, lower extremities undergo a phase of only one of the feet is in contact with the ground, one of swinging, and two phases of the feet are not in contact with the ground. Each participant ran on a standard running track and ran as long as they can.

3) sEMG DATA COLLECTION
During the running experiment, the sEMG signals were acquired by the MyoWare™ muscle sensor [13], [14]. Figure 2 shows the sEMG measurement module. The sensor captures the sEMG signals through the surface electrode on the skin over the muscles. A microcontroller is used to acquire sEMG signals and digitizes the input analog sEMG signals with an analog-to-digital converter (ADC).

The frequency of the sEMG data acquisition was set at 2000 Hz. Meanwhile, the sEMG signals are transmitted to the WiFi module which is responsible for sending the sEMG data to host. The block diagram of acquisition of EMG signal is shown in Fig. 3.

The main muscle groups used in running are the Vastus Intermedius, Vastus Lateralis, Vastus Medialis, and Rectus Femoris. Hence, we place sEMG sensors on the four muscles. The patterns of lower limb responses during dynamic muscle actions are with no differences [15]. To collect massive amounts of training data, the sEMG sensors were placed on both limbs in the data in the data collection process. However, the subject can place the sEMG sensors on one limb in the inference processing. Finally, we get a predicted probability for the valid sEMG signals and noises. Figure 4 shows the positions of four sensors attached on four muscles.

B. DATA PRE-PROCESSING

1) BAND-PASS FILTER
The energetic distribution of sEMG signals lies in the frequency range from 0 to 500 Hz and the dominant components lie in the frequency range from 20 to 500 Hz. However, outside the 0-500 Hz frequency range, signals with energy less than electrical noise level are unusable. Additionally, three main types of noise sources contribute to the process of sEMG signal acquisition, those are the inherent noise of electronic instrument, the ambient noise from the electromagnetic radiation in the environment, and motion artifact. These noise lies mostly in the frequency range from 0-20 Hz [16], [17]. Meanwhile, due to the quasi-random nature of the firing rate of the muscular motor units, the frequency components between 0 and 20 Hz are mostly unstable contains unstable components [16]. Therefore, a 20-500 Hz band-pass filter was used to eliminate the interfering signals and the artifacts. Figure 5 shows an example of a band-pass filtered sEMG signal.
overfitting in the network. The network uses dropout layers the 3D-LCNN, optimization methods are needed to prevent LSTM layer. Due to the small data sets used in the training of signals of the 3D convolutional layers are inputted into the layers have 16 filters with 3 filters with 3 kernel. The first and second convolutional layers have eight model. The 3D-LCNN consists of three 3D convolutional each feature map is a feature cube. Moreover, the 3D-CNN convolution kernel is three dimensions which process of a 2D and 3D convolution. Different from 2D-CNN, 3D-CNN convolution is con-ducted both spatially and temporally. Figure 9 illustrates the LSTM by combing them into one architecture, which we call at temporal modeling, we take advantage of the 3D-CNN and more outstanding in feature extraction and the LSTM is good in the training process. Table 1 clearly shows that the perfor-mance of the 3D-LCNN obtains higher accuracy. In this section, we presented the results of our experiments. The self-collected sEMG dataset is split to 80% training, 10% validation, and 10% testing chunks.

A. ACCURACY COMPARISON WITH THE DIFFERENT DNN MODELS
The experiment conducted in this section is mainly to com-pare the performance of CNN, 3D-CNN, CNN+LSTM and 3D-LCNN models based on the self-collected sEMG dataset. Moreover, these models are set the same filter size and layers in the training process. Table 1 clearly shows that the perfor-mance of the 3D-LCNN obtains higher accuracy.

B. EVALUATE THE EFFECTIVENESS OF DATA AUGMENTATION TECHNIQUES
To evaluate the model robustness after using the data augmen-tation method in the training process, we compare the mod-els trained with and without the augmentation techniques. As shown in Table 2, all the models trained with augmented

FIGURE 7. Creating the 3D-spectrogram using STFT.

2) SHORT TIME FOURIER TRANSFORM
In order to track the frequency content of a signal in time, a time-frequency transform must be utilized. Hence, the sEMG signals are translated to color maps consisting of frequency and time distribution using STFT. Figure 6 shows an example of the sEMG signal transformation using STFT.

There are four sEMG sensors placed on the lower limb. Each MyoWare™ will receive a sEMG signal set, which has 20000 samples. The sEMG signal set is divided into sequential intervals of time. Then, STFT is employed on the sEMG signal set to obtain the 2-dimension spectrogram (2D-spectrogram). After which all STFT results are concatenated together to form the final 3-dimension spectrogram (3D-spectrogram) as shown in Fig. 7.

C. PROPOSED 3D-LCNN ARCHITECTURE
For single image recognition, 2D-CNN has been very suc-cesful in feature extraction and classification. However, motion information is buried in between frames which would get lost when 2D-CNN is employed. To solve the afore-mentioned issues, we use the 3D-CNN to learn both the image-level information and motion information between consecutive frames. On the other hand, LSTM shows great effectiveness in capturing the temporal dependencies based on learning consecutive information. Due to the 3D-CNN is more outstanding in feature extraction and the LSTM is good at temporal modeling, we take advantage of the 3D-CNN and LSTM by combing them into one architecture, which we call 3D-LCNN.

Different from 2D-CNN, 3D-CNN convolution is con-ducted both spatially and temporally. Figure 9 illustrates the process of a 2D and 3D convolution. Different from 2D-CNN, the 3D-CNN convolution kernel is three dimensions which include two spatial dimensions and one temporal depth. Moreover, the 3D kernel is convolved with the video frames with stride 1 at all dimensions and the convolution result of each feature map is a feature cube.

Fig. 8 shows the network architecture of our 3D-LCNN model. The 3D-LCNN consists of three 3D convolutional layers. The first and second convolutional layers have eight filters with $3 \times 3 \times 3$ kernel and the third convolutional layers have 16 filters with $3 \times 3 \times 3$ kernel. Then, the output signals of the 3D convolutional layers are inputted into the LSTM layer. Due to the small data sets used in the training of the 3D-LCNN, optimization methods are needed to prevent overfitting in the network. The network uses dropout layers with dropout rate 0.5 which avoids overfitting by reducing the co-adaptation of hidden units. In addition, batch normal-ization layer normalizes each input channel across a mini-batch which aims to maintain the standard distribution of hidden layer activation and accelerate the convergence of the network. Moreover, the 3D-LCNN uses ReLU after each convolution layer for the optimization of the network. Finally, we use softmax as loss function in our design.

D. DATA AUGMENTATION
Data augmentation enhances the size and quality of training datasets, with the objective of achieving better generalization. Hence, we employ two data augmentation techniques for multichannel sEMG signals. The first data augmentation technique aims at simulating the surface electrodes displacement on the skin. Augmenting the training data by randomly shifting the channels of the sEMG signals can simulate the electrode shift that may occur in recording the sEMG signals. Hence, we shift the part of the spectrum from one channel to the next. This technique will be referred to as shifting channels augmentation. Another augmentation technique is to simulate muscle fatigue. The main effect of muscle fatigue on the sEMG signals is the spectrum gradually shifted to lower frequencies and higher frequencies are attenuated. Therefore, we reduce the median frequency of a channel by redistributing part of the power of a frequency to an adjacent lower frequency. Although the muscle fatigue could directly be evaluated by using AM-FM methods from the sEMG signal, our proposed model using the STFT spectrogram of the sEMG signal as input can reduce the median frequency of a channel by redistributing part of the power of a frequency to adjacent lower frequency [18]. Therefore, our proposed data-augmentation approach can be easily implemented in the frequency domain.

III. RESULTS
In this section, we presented the results of our experiments. The self-collected sEMG dataset is split to 80% training, 10% validation, and 10% testing chunks.

A. ACCURACY COMPARISON WITH THE DIFFERENT DNN MODELS
The experiment conducted in this section is mainly to com-pare the performance of CNN, 3D-CNN, CNN+LSTM and 3D-LCNN models based on the self-collected sEMG dataset. Moreover, these models are set the same filter size and layers in the training process. Table 1 clearly shows that the perfor-mance of the 3D-LCNN obtains higher accuracy.

B. EVALUATE THE EFFECTIVENESS OF DATA AUGMENTATION TECHNIQUES
To evaluate the model robustness after using the data augmen-tation method in the training process, we compare the mod-els trained with and without the augmentation techniques. As shown in Table 2, all the models trained with augmented
C. THE EFFECT OF sEMG NOISE ON MUSCLE FATIGUE DETECTION

Figure 10 shows the muscle fatigue results of raw sEMG signals with and without the proposed techniques. The experimental result shows the influence of sEMG noises on muscle fatigue detection. Additionally, the muscle fatigue plot decreases linearly and remains stable after removing the random noises from sEMG signals by using the proposed technique.

IV. DISCUSSION

CNN does not take into account the inter-frame motion information in the time dimension. However, 3D-CNN uses the 3D convolution kernel in the cube. Each feature map in the convolutional layer is connected to multiple adjacent consecutive frames in the previous layer, thus capturing muscle interaction information during running exercise. Hence, 3D-CNN achieves higher accuracy than traditional CNN. On the other hand, lower limb actions in running exercise have more periodic features. LSTM process inputs in a sequential way. The data from the previous input is considered when computing the output of the current step, which allows the neural network to carry information over different time steps rather than keeping all the inputs independent of each other. Compared
with using CNN or 3D-CNN alone, the recurrent layers can find out the time dependencies, which are crucial in time-sequence analysis, especially for the periodic features in running exercise. With the advantages of both 3D-CNN and LSTM, the 3D-LCNN presented a better performance than CNN, 3D-CNN, and CNN+LSTM.

Additionally, according to the possible cases that happen in the sEMG data-collection procedure, we designed two data-augmentation approaches to expand our sEMG dataset, those are the simulation of the surface electrodes displacement on the skin and the muscle fatigue. The experiment results showed that augmenting the training data by the simulation of the surface electrodes displacement can acquire higher prediction accuracy than the simulation of muscle fatigue. Since the dataset was recorded with the surface electrodes, the surface electrodes displacement is expected to occur. Moreover, compared with muscle fatigue, the probability of occurrence of the surface electrodes displacement is higher.

V. CONCLUSION
To the best of our knowledge, this paper is the first one to combine 3D-CNN and LSTM for sEMG-based noises recognition. Additionally, the 3D-LCNN does not need to re-learn when the user reseats the surface electrodes. Due to the motion information and time-frequency spectrograms are important for subsequent feature extraction, the 3D-LCNN architecture uses 3D-CNN to extract motion features and then passes to LSTM to learn temporal features. Moreover, this paper introduces two data-augmentation approaches, those are simulation of the surface electrodes displacement on the skin and muscle fatigue. The experiment results showed that these two-data-augmentation approaches have significantly enhanced the performance and robustness. Additionally, the classification accuracy of the 3D-LCNN achieves 90.52%. Especially, this work offered new research avenue for this field. In the future work, we will explore sEMG noises recognition by using incremental self-labeling GAN (ISL-GAN) [19] and neural embedding matching (NEM) [20] to solve the problem with small size samples at the first time, and other sEMG noise-related topics as well.

REFERENCES
[1] N. Amrutha and V. H. Arul, “A review on noises in EMG signal and its removal,” Int. J. Sci. Res. Publications, vol. 7, pp. 23–27, May 2017.
[2] M. Cifrek, V. Medved, S. Tonkovíc, and S. Ostojić, “Surface EMG based muscle fatigue evaluation in biomechanics,” Clin. Biomech., vol. 24, no. 4, pp. 327–340, May 2009.
[3] L. Wang, Y. Wang, A. Ma, G. Ma, Y. Ye, R. Li, and T. Lu, “A comparative study of EMG indices in muscle fatigue evaluation based on grey relational analysis during all-out cycling exercise,” BioMed Res. Int., vol. 2018, pp. 1–8, Apr. 2018.
[4] Y. Hotta, Y. Korakata, and K. Ito, “Verification of the muscle fatigue detection capability of a unipolar-leads system using a surface electromyogram model,” in Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Aug. 2014, pp. 110–113, doi: 10.1109/EMBC.2014.6943541.
[5] U. Cote-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, “Deep learning for electromyographic hand gesture signal classification using transfer learning,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 27, no. 4, pp. 760–771, Apr. 2019.
[6] Y. Yamanoi and R. Kato, “Control method for myoelectric hand using convolutional neural network to simplify learning of EMG signals,” in Proc. IEEE Int. Conf. Cyberg Bionic Syst. (CBS), Oct. 2017, pp. 114–118.
[7] F. Quivira, T. Koike-Akino, Y. Wang, and D. Erdogmus, “Translating sEMG signals to continuous hand poses using recurrent neural networks,” in Proc. IEEE EMBS Int. Conf. Biomed. Health Inform. (BHI), Mar. 2018, pp. 166–169.
[8] M. Kim, T. Kim, D. Kim, and W. Chung, “Curved microneedle array-based sEMG electrode for robust long-term measurements and high selectivity,” Sensors, vol. 15, no. 7, pp. 16265–16280, 2015.
[9] M. Malboubi, F. Razzaari, M. Aliyari Sh., and A. Davari, “Power line noise elimination from EMG signals using adaptive Laguerre filter with fuzzy step size,” in Proc. 17th Iranian Conf. Biomed. Eng. (ICRME), Nov. 2010, pp. 1–4, doi: 10.1109/ICRME.2010.5704932.
[10] P. Xu, Y. Liu, G. Gu, and X. Shen, “Research on the method of EMG pattern recognition based on neural networks,” in Proc. 2nd IEEE Adv. Inf. Manage.,Communicates, Electrom. Autom. Control Conf. (IMCCCE), May 2018, pp. 1401–1405, doi: 10.1109/IMCCE.2018.8469432.
[11] J. J. Bird, J. Kobylarz, D. R. Faria, A. Ektar, and E. P. Ribeiro, “Cross-domain MLP and CNN transfer learning for biological signal processing: EEG and EMG,” IEEE Access, vol. 8, pp. 54789–54801, 2020, doi: 10.1109/ACCESS.2020.2979074.
[12] M. G. R. Alam, S. F. Abedin, S. I. Moon, A. Talukder, and C. S. Hong, “Healthcare IoT-based affective state mining using a deep convolutional neural network,” IEEE Access, vol. 7, pp. 75189–75202, 2019, doi: 10.1109/ACCESS.2019.2919995.
[13] 3D Lead Muscle/ Electromyography Sensor for Microcontroller Applications. Advanced Technologies, USA. Accessed: Jan. 1, 2020. [Online]. Available: https://github.com/AdvanceTechnologies/MyoWare_MuscleSensor/raw/master/Documents/AT-04-001.pdf
[14] Y.-H. Yang, S.-J. Ruan, P.-C. Chen, Y.-T. Liu, and Y.-H. Hsueh, “Implementation of a low-cost wireless multi-channel surface EMG acquisition system,” IEEE Consum. Electron. Mag., vol. 9, no. 5, pp. 14–19, Sep. 2020, doi: 10.1109/MCE.2020.2986792.
[15] J. P. V. Anders, C. M. Smith, J. L. Keller, E. C. Hill, T. J. Housh, R. J. Schmidt, and G. O. Johnson, “Inter- and intra-individual differences in EMG and MMO during maximal, bilateral, dynamic leg extensions,” Sports, vol. 7, no. 7, p. 175, Jul. 2019.
[16] B. Rodriguez-Tapia, I. Soto, D. M. Martinez, and N. C. Arballo, “Myoelectric interfaces and related applications: Current state of EMG signal processing—A systematic review,” IEEE Access, vol. 8, pp. 7792–7805, 2020, doi: 10.1109/ACCESS.2019.2963881.
[17] M. Simao, N. Mendes, O. Gibaru, and P. Neto, “A review on electromyography decoding and pattern recognition for human-machine interaction,” IEEE Access, vol. 7, pp. 39564–39582, 2019, doi: 10.1109/ACCESS.2019.2906584.
[18] G. Biagetti, P. Crippa, A. Curzi, S. Orcioni, and C. Turchetti, “Analysis of the EMG signal during cyclic movements using multicomponent AM–FM decomposition,” IEEE J. Biomed. Health Inform., vol. 19, no. 5, pp. 1672–1681, Sep. 2015.
[19] X. Wei, X. Wei, W. Xing, S. Lu, and W. Lu, “An incremental self-labeling strategy for semi-supervised deep learning based on generative adversarial networks,” IEEE Access, vol. 8, pp. 8913–8921, 2020, doi: 10.1109/ACCESS.2020.2964315.
[20] Z. Wang, B. Du, and Y. Guo, “Domain adaptation with neural embedding matching,” IEEE Trans. Neural Netw. Learn. Syst., vol. 31, no. 7, pp. 2387–2397, Jul. 2020, doi: 10.1109/TNNLS.2019.2935608.

MIN-WEN LIN is currently pursuing the M.S. degree with the Department of Electronics and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan.
SHANQ-JANG RUAN (Senior Member, IEEE) is a Distinguished Professor with the Department of Electronics and Computer Engineering, National Taiwan University of Science and Technology. His research interests include embedded deep neural network processing, energy-efficient image processing, and embedded system design.

YA-WEN TU is the Chief Physician of physical medicine and rehabilitation with the Sijih Cathay General Hospital. She is currently in charge of two research proposals. One is related to oncology rehabilitation and the other is combining the latest AI’s convolutional neural network in the evaluation and treatment for speech disorders.

***