Deep Learning Based Surface EMG Hand Gesture Classification for Low-Cost Myoelectric Prosthetic Hand

Nazmun Nahid, Arafat Rahman, M.A.R. Ahad
University of Dhaka
Dhaka, Bangladesh
Email: nazmunnahid@ieee.org, arafatrahman245@outlook.com, atiqahad@du.ac.bd

***Abstract*—In this paper, a comparative study of classifying different hand gestures of two well-known surface Electromyogram (sEMG) data sets, Rami Khusaba EMG repository, and UCI Machine Learning Repository, is shown. Applying transfer learning and CNN-LSTM neural network architectures, we find out a suitable control scheme for a myoelectric prosthetic hand (we mention it as DUFAB Hand). At first, the continuous wavelet transform (CWT) is exploited to create images from the sEMG signal, which serves as a powerful feature for the classification of different hand gestures. Then, we transferred the learning of various neural nets of image classification, e.g., AlexNet, and ResNet-18 to the sEMG image classification. Application of these deep neural networks outperformed general machine learning techniques with higher accuracy and performance. For example, the combination of CNN and LSTM has achieved the state of the art accuracies for these data sets, of 99.72% for UCI Machine Learning Repository and 99.83% for Rami Khusaba EMG repository respectively. The main contribution of this paper is, establishing an algorithmic pipeline using continuous wavelet transform (CWT) and CNN-LSTM deep neural networks to achieve high accuracy in two sEMG datasets.**

***Index Terms*—Deep Learning, Transfer Learning, sEMG, CNN, LSTM, Biomedical Signal Processing.**

I. Introduction

Approximately 2.4 million people around the world are suffering from upper-limb amputation [1]. Most of these people belong to developing or underdeveloped countries. To restore them to the workforce again, an upper-limb prosthesis is a must. Nowadays, upper-limb prostheses have a wide-range but efficient prosthetic hands may cost a fortune. Keeping in mind the economic condition of these people, Fab Lab DU designed a low-cost prosthetic hand, DUFAB Hand [2]. According to Biddiss et al. [3] two main key factors while designing myoelectric prosthetic hands to be considered are cost and functionality. So in this work, different hand gesture classification and control schemes are analyzed to get better functionality at a lower price for this hand.

Here two well known open source sEMG repositories; Rami Khusaba EMG repository [4] and UCI Machine Learning Repository [5], [6] have been used for developing the control scheme. Myoelectric control is chosen because, by far, it is mostly used as an interface to improve the dexterity of the prosthetic hands as it is relatively easy to use, comfortable, and promotes muscle tone. As the muscle activity required is relatively small, very low physical effort is needed for the operation [7]. Both of these repositories use 2-channel sEMG systems, which are highly cost-effective and a perfect match for the signal acquisition system used for DUFAB Hand.

Over the years, various control schemes are developed to translate the information in the sEMG signal. They are normally classified based on the character of control as sequential control and simultaneous control. Currently sequential control is mainly used for controlling. Here, the sEMG signals are translated by using different controlling methods : 1) on-off, 2) proportional, 3) direct, 4) finite state machine, 5) pattern recognition-based, 6) posture based, and 7) regression based.

A considerable increase in the popularity of deep learning in various applications in recent years has provided a new perspective to analyze sEMG signals for hand gesture recognition. For sEMG-based gesture recognition, the Convolutional Neural Network (CNN) has been exploited as it has shown excellent performance. Park and Lee [8] proposed a user-adaptive CNN model for improved performance in the multi-user sEMG dataset. Atzori et al. [9] proposed a simple CNN to classify 50 hand movements in three Ninapro datasets. Their average classification accuracy ($66.59 \pm 6.4\%$ on dataset 1, $60.27 \pm 7.7\%$ on dataset 2 and $38.09 \pm 14.29\%$ on dataset 3) is comparable with classical machine learning procedure. Geng et al. [10] developed a classification scheme using deep neural networks which takes high density sEMG images as input. Du et al. [11] introduced a high density sEMG dataset with deep learning based domain adaptation framework to increase inter session gesture recognition accuracy. For accurate predictions of deep learning models, a large amount of training data is needed. So, a particularly attractive research field in deep learning is Transfer Learning (TL), where a model can be pre-trained on numerous subjects before using it on a new participant. The ConvNets performance enhances when it learns from multiple users' data, and it reduces the size of the required training data set that typical deep learning algorithms need [12], [13]. Also, the hybrid CNN-RNN architecture has obtained good performance in this field [14] as both spatial and temporal information is provided by hybrid CNN-RNN architecture. Jia et al. [15] proposed a combination of convolutional auto-encoder and convolutional neural network (CAE+CNN) to classify 10 gestures of an EMG dataset. They

978-1-7281-9331-1/20/$31.00 ©2020 IEEE

used CAE to build a latent feature space which serves as the input of CNN. They achieved 99.38% test accuracy without adding noise and 98.13% accuracy with Gaussian noise of level 1e-5. Li et al. [16] used Principal component analysis (PCA) to extract features and deep neural regression model to predict force from sEMG signal. A fuzzy controller then takes the predicted force values to control a prosthetic hand. Ghazaei et al. [17] developed a deep learning-based computer vision system to enhance the grasping functionality of a commercial prosthesis. They trained a CNN with images of 500 graspable objects to augment the control functionality of prosthetic hands.

The main contributions of this paper are, we have analyzed both TL and CNN-RNN methods to compare accuracies and performances. To the best of our knowledge, this is the first time continuous wavelet transform (CWT) is used with CNN-LSTM for sEMG signal analysis. The control schemes developed here are not only suitable for controlling DUFAB hand but also any other prosthetic hand with closer configuration as the functionalities were tested in both real time cases and simulation.

The rest of the paper is organized as follows: we discuss the materials and methods containing the description of data sets, pre-processing steps, and different techniques applied for classification in Sec. II, discuss the results in Sec. III, and finally provide the conclusion and future work in Sec. IV.

II. Materials and Methods

A. Dataset Details

In this paper, we studied two open-source data sets; one is from Rami Khusaba EMG repository [4] and another one from UCI Machine Learning Repository [5], [6].

For the 1st data set, the sEMG data were collected from the arm surface of 8 healthy subjects using two channels of Delsys DE.2x sensors. The data were sampled at 4kHz using National Instruments, BNC-2090's 12-bit analog-to-digital converter. Then a Butterworth bandpass filter (20-450Hz) and a notch filter (50 Hz) were used to remove noise and line artifacts. Following ten classes of finger movements were performed – Hand Close (HC), Index (I), Little (L), Middle (M), Ring (R), Thumb (T), ThumbIndex (TI), ThumbMiddle (TM), ThumbRing (TR) and ThumbLittle (TL). Each subject performed each of these movements for about 5 seconds and repeated six times.

2nd data set has two parts; they are - UCI Data Set-1 and UCI Data Set-2. For the UCI Data Set-1, five healthy subjects performed six classes of grasps each for about 6 seconds with 30 trials. The six classes of grasps are - Spherical, Cylindrical, Hook, Lateral, Palmar, and Tip. The data were collected and sampled at 500 Hz using 2-channel Delsys Bagnolia, Handheld sEMG Systems. The sEMG data were collected from the surface of Flexor Capri Ulnaris and Extensor Capri Radialis muscles. For UCI Data Set-2 same acquisition system was used, and one healthy subject performed each of the six movements for about 5 seconds and repeated 100 times for three consecutive days.

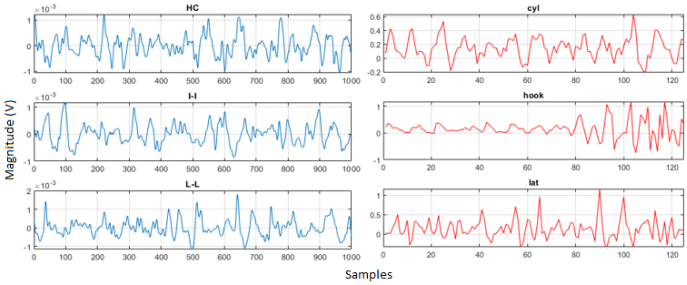

Fig. 1. Samples of data from 1st window for first 3 gestures in Rami (blue) and UCI (red) Data Set-1.

B. Data Pre-processing

For good accuracy and efficient classifier training, pre-processing is an important step. Signal pre-processing typically includes - amplification, filtering, windowing, etc. Two data sets used in this paper were amplified with a gain of 1000 and filtered as previously described. Input latency can significantly affect the performance of real-time control. A non-overlapping sliding window of 250 ms was chosen to keep the latency within 300 ms, as suggested by the study [18]. Although it is recommended to keep the delay between 100-250 ms [19], [20], priority was given to the accuracy over speed. Windows of 250 ms give 50 ms for pre-processing and classification. Fig. 1 shows the 1st window for the first three gestures in Rami and UCI data sets.

C. Continuous Wavelet Transform for sEMG Image Creation

The continuous wavelet transform (CWT) is a tool for representing a signal by varying the translation and scale parameter of the wavelets continuously. The CWT of a function $x(t)$ is defined as the following equation –

$$X_w(\tau, s) = \frac{1}{|s|^{\frac{1}{2}}} \int_{-\infty}^{\infty} x(t) \bar{\psi} \left(\frac{t-\tau}{s} \right) dt \qquad (1)$$

Where $s \in \mathbb{R}^+$ denotes the scale parameter and $\tau \in \mathbb{R}$ denotes the translation parameter [21]. The function $\psi(t)$ is called the mother wavelet, which is continuous in both the time and frequency domain. At first, this mother wavelet is generated, and then some daughter wavelets are created by simply translating and scaling this wavelet. So, for detecting low-frequency components, $\psi(t)$ should be stretched with larger values of s, and conversely, for high-frequency components, it should be compressed with smaller values of s.

sEMG is a non-stationary signal composed of smooth and abrupt oscillations, which are sometimes lost in the noise. So, to obtain useful information for sEMG classification, it is necessary to create joint time-frequency representation, which can be constructed by the CWT. Using various scale and shifting parameters of a wavelet, it outputs a 2-D time-frequency representation, which extracts time localized frequency components of the sEMG signal. The 2-D time-frequency representation is known as Scalogram, and it is generated for each window of the signal. Among various mother wavelets, the analytic Morse

wavelet was used here with 12 voices per octave, symmetry = 3, and time-bandwidth product = 60. For each sample of data, 85 coefficients are calculated, resulting in an 85x1000 matrix for one window of signal for 1st data set and 85x125 for 2nd data set. The values of the wavelet coefficients are treated as pixels of the image with various color intensities. Finally, each image is resized to 224-by-224 for ResNet-18 and 227-by-227 for AlexNet. From Fig. 2, we can see the Scalogram of Hand Close (HC) gesture for one window length of data of Rami Data Set, which shows the joint time-frequency representation of the signal. As magnitudes of the coefficients are color-coded in this graph, the intensity of the colors, for example - red indicates large magnitude, yellow indicates medium, and blue means small magnitude. The bright spots in the higher frequency portion of the graph indicate abrupt changes in the signal, whereas, in the lower frequency portion, they indicate smooth transitions or oscillations. So, observing Fig. 2, we can see the yellow spots occur in the frequency range of 86.65-115 Hz for the time range 0.05-0.07s and 108.5-162.6 Hz for the time range 0.18-0.20s. As these spots are located at a relatively lower frequency portion, they indicate smooth transition or oscillation. From the higher frequency portion, we can not find yellow spots that much, but the pattern shows that there are many small abrupt changes in the frequency range 300-500 Hz. The main reasons behind using CWT images as input images for developing the control scheme of DUFAB Hand are as follows -

1) **Joint time-frequency representation for improved classification:** Like other bio-electrical signals, sEMG is a non-stationary signal with varying frequency components with respect to time. So, only time or frequency domain features are not adequate to classify or differentiate between different hand gestures and control DUFAB Hand. For instance - Fast Fourier Transform is a well-known frequency-domain feature, but it will give almost the same output for index finger movement and thumb movement or other different gestures. This is true for other frequency or time-domain features. This difficulty is also evident from our previous work with DUFAB Hand [2] and various other literatures [22]–[25]. CWT can capture those time-varying frequency components and create an effective joint time-frequency representation. Therefore we used it as the image creator and used that image as input to the sEMG classifier of DUFAB Hand.
2) **Good time-frequency resolution:** As in CWT, the window size or scale parameter varies continuously, it gives good time-frequency resolution. It can detect and differentiate between abrupt changes and smooth oscillations of a signal effectively. Different hand gestures have different patterns of these sharp and steady changes. CWT identifies those sharp changes with small scale value and smooth oscillations with large scale value. Like CWT, the Short-Time Fourier Transform (STFT) is a method for creating time-frequency representation. But it can not resolve different events of a signal as

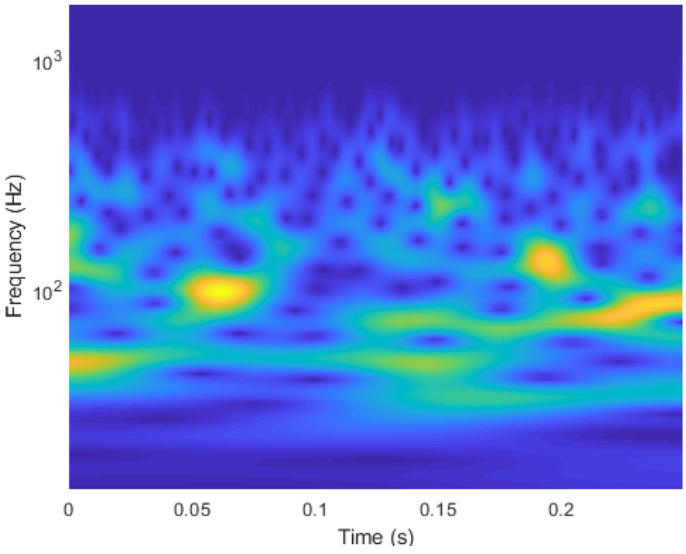

Fig. 2. Scalogram of HC-1 (Hand Close 1st trial) in Rami Dataset, which shows bright yellow spots at 86.65-115 Hz for the time range 0.05-0.07s and 108.5-162.6 Hz for the time range 0.18-0.20s. These spots indicate smooth transitions, whereas patterns visible in the frequency range 300-500 Hz indicate abrupt changes.

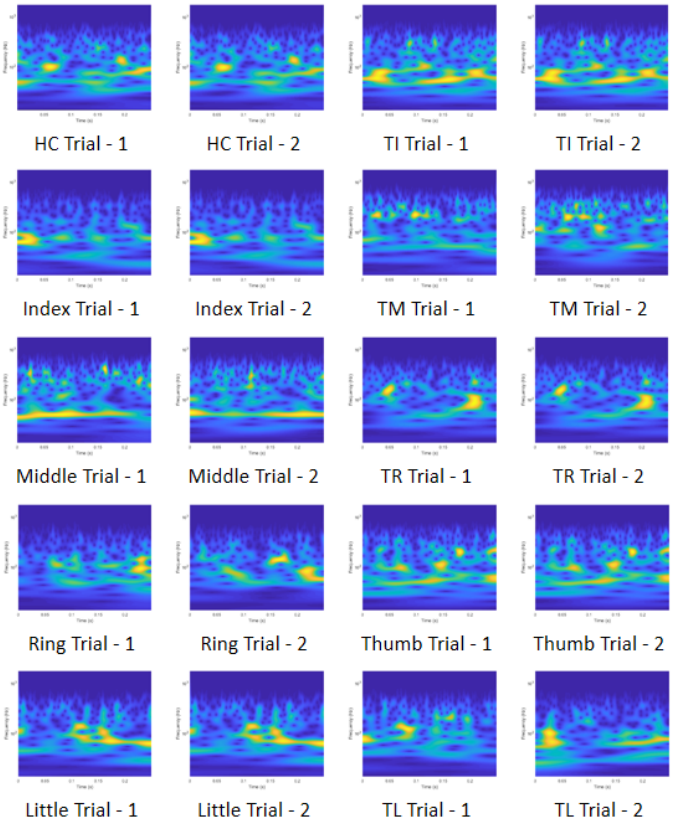

Fig. 3. Scalogram of all the gestures in Rami Dataset for 1st two trials, which shows visible similarity in patterns between multiple trials of same gestures and dissimilarity between different gestures.

good as CWT because STFT has a fixed window length. So, for classifying different gestures and better control of DUFAB Hand, CWT is an effective way of creating input images for the classifiers.

From Fig. 3, it is evident that CWT images for different hand gestures are different, while for the same gestures of different trials, it is quite similar.

D. Classification using Transfer Learning

1) AlexNet for sEMG Image Classification: AlexNet is a 8 layers deep pre-trained convolutional neural network [26], [27]. For sEMG image classification, we changed the last fully-connected layer with 10 class labels for Rami Data Set and 6 class labels for UCI Data Set. Images created using CWT are given as input, and the network extracts useful features from Scalogram images and classifies accordingly. For the 1st and 2nd data set, the training and validation partition was 70% and 30%, respectively. The network is trained with mini-batch size = 60, epoch = 3 and learning rate = 0.001. For 2nd data set, mini-batch size = 32, epoch = 3 and learning rate = 0.001.

2) Resnet-18 for sEMG Image Classification: Resnet-18 is a 18 layers deep pre-trained convolutional neural network [26], [28]. A similar procedure is followed as described before for AlexNet. For the 1st and 2nd data set, the training and validation partition was 75% and 25%, respectively. The network is trained with mini-batch size = 128, epoch = 3 and learning rate = 0.001.

E. Classification using CNN+LSTM

This is a hybrid network composed of CNN and LSTM. As LSTM takes sequences as input, we used feature extracted sEMG signals, and images created using CWT are given as input. This process can be called as sequence folding. Then we use convolutional layers to extract features from each image. The sequence structure is then restored by sequence unfolding layer and flatten layer that gives vector sequences as output. The feature extracted sequence is then given as input to the LSTM network and trained afterward. An LSTM network architecture is defined as follows - an input layer for sequences, a BiLSTM layer composed of 500 hidden units followed by a dropout layer, a fully connected layer with an output size corresponding to the number of classes, a softmax layer, and a classification loss layer. All the sequences are partitioned as 80% training data and 20% for testing. As Convolutional Neural Network modified, GoogLeNet and Resnet-18 are used. The scalogram images are converted to sequences of feature vectors by taking the output of the activation function of the last pooling layer named "pool5-7x7_s1" of GoogLeNet and "pool5" of Resnet-18.

III. Results and Discussion

We used accuracy, loss value, confusion matrix, prediction time, and training time as performance metrics. In the following subsections, these performances are discussed in detail.

A. Performance of Transfer Learning

1) Performance of Alexnet: For classifying sEMG scalogram images, Alexnet gave a satisfactory result with good accuracy. For Rami Data Set, it gave 99.64% accuracy in the validation set, classifying ten gestures. This accuracy has outperformed almost all normal machine learning algorithms applied previously. Fig. 4 shows the training progress graph of this network for the Rami Data Set. The difference between validation and training accuracy is 1.31%. This low value indicates that the network has achieved a good fit, and overfit or underfit has not occurred. The difference between validation and training loss is 0.0481, which is also small and gives the indication that the network can generalize well in test or validation data sets. Fig. 5 shows the confusion matrix for this network for Rami Data Set, which shows all the gestures are classified accurately with low error. From the confusion matrix, it can be observed that most misclassification occurs between R–R (ring finger) and M–M (middle finger). This might happen because of the signal similarity between these two gestures. For UCI Data Set-1, AlexNet gave 98.70% validation accuracy with 0.0408 training loss and 0.0385 validation loss.

2) Performance of ResNet-18: A similar procedure is followed for measuring the performance of ResNet-18, as described before. In this case, validation accuracy is observed as 99.55% in Rami Data Set with training and validation accuracy difference = 0.45%. This is again an indication of a good fit. For UCI Data Set-1, the accuracy is 98.61% with training and validation accuracy difference = 0.79%.

B. Performance of CNN+LSTM

Finally, we measured the performance of two combinations of CNN and LSTM networks. These combinations are - GoogLeNet+LSTM and ResNet-18+LSTM. Fig. 6 shows the training progress graph for ResNet-18+LSTM in UCI Data Set-1, which achieved 99.72% accuracy. It also performed well

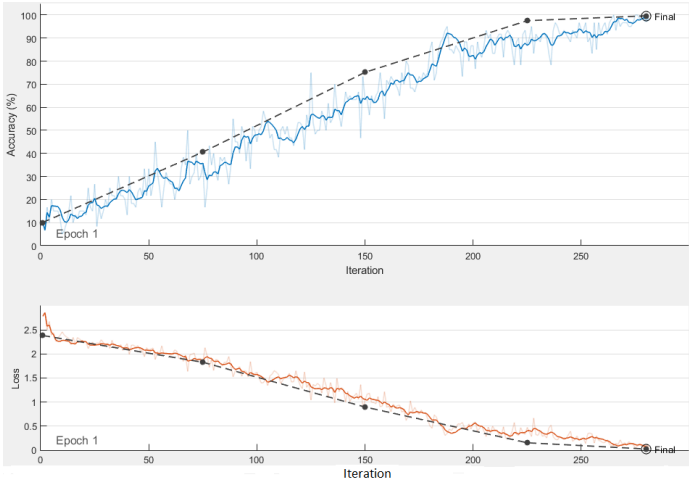

Fig. 4. Training progress graph for (Accuracy vs Iteration and Loss vs Iteration) AlexNet in Rami Data Set.

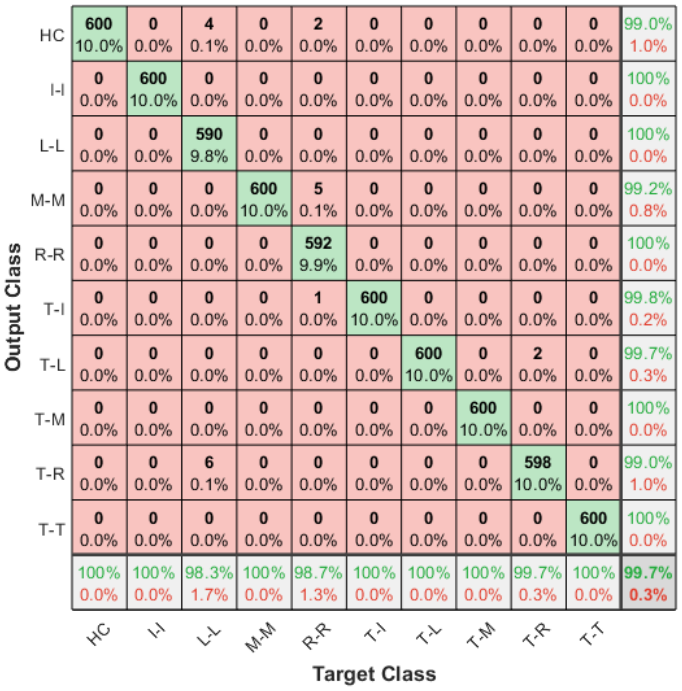

Fig. 5. Confusion Matrix for Alexnet in classifying 10 gestures in Rami Data Set.

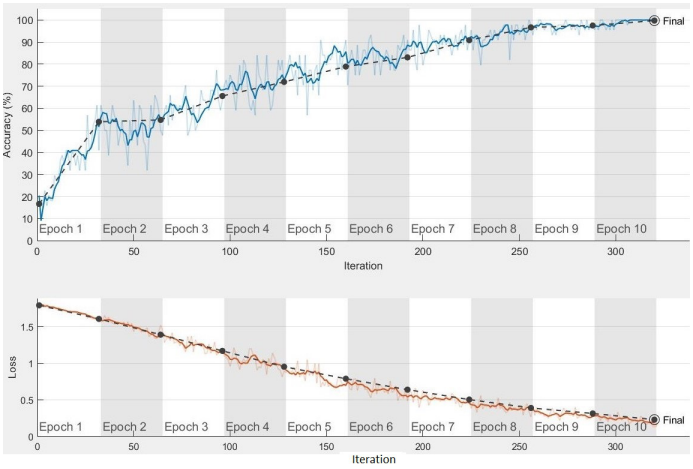

Fig. 6. Training progress graph for (Accuracy vs Iteration and Loss vs Iteration) ResNet-18+LSTM in UCI Data Set-1.

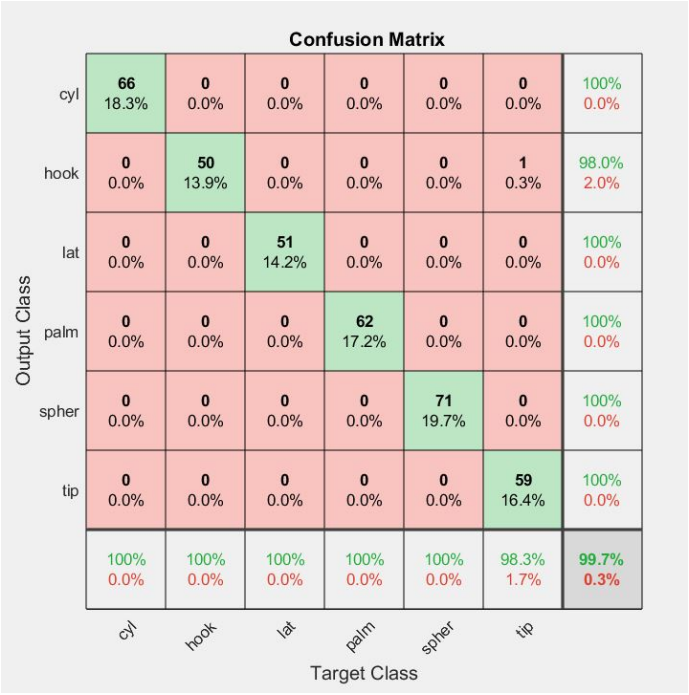

Fig. 7. Confusion Matrix for ResNet-18+LSTM in UCI Data Set-1.

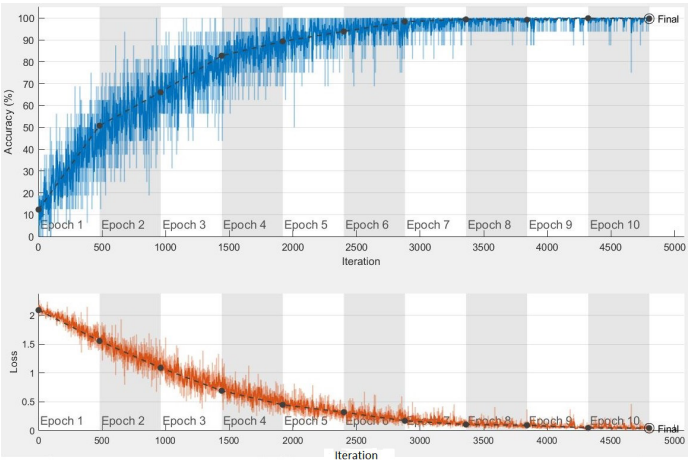

Fig. 8. Training progress graph for (Accuracy vs Iteration and Loss vs Iteration) ResNet-18+LSTM in Rami Data Set.

in UCI Data Set-2 with an accuracy of 99.23%. Similarly, the accuracy of GoogleNet+LSTM in UCI Data Set-1 is measured as 94.31% and for UCI Data Set-2, 98.35%. From Fig. 8 we can see that for Rami Data Set, ResNet-18+LSTM gave 99.83% accuracy. For GoogleNet+LSTM it gave 99.64% accuracy. From the last iteration, we measured validation and training accuracy difference of GoogleNet+LSTM is 0.36% for Rami Data Set, which is quite small, but for UCI Data Set-1, it is 5.69%. So, for UCI Data Set-1, GoogleNet+LSTM tends to overfit a little bit. From the last iteration, we measured validation and training accuracy difference of ResNet-18+LSTM is 0.17% for Rami Data Set, 0.28% for UCI Data Set-1, and 0.45% for UCI Data Set-2, all of which indicate a good fit.

C. Performance comparison of previous notable works and our works on the utilized data sets

From TABLE I, it can be observed that ResNet-18+LSTM combination has given the best accuracies for all data sets in this work. The classification accuracies of previous works can also be observed here. Most of them are based on general machine learning techniques. Only Alam et al. is based on deep learning. The table clearly shows that our proposed techniques have outperformed all of the previous works except Martinez et al. in UCI Data Set-1. sEMG signals are very stochastic in nature. The same sEMG signal from the same gesture can

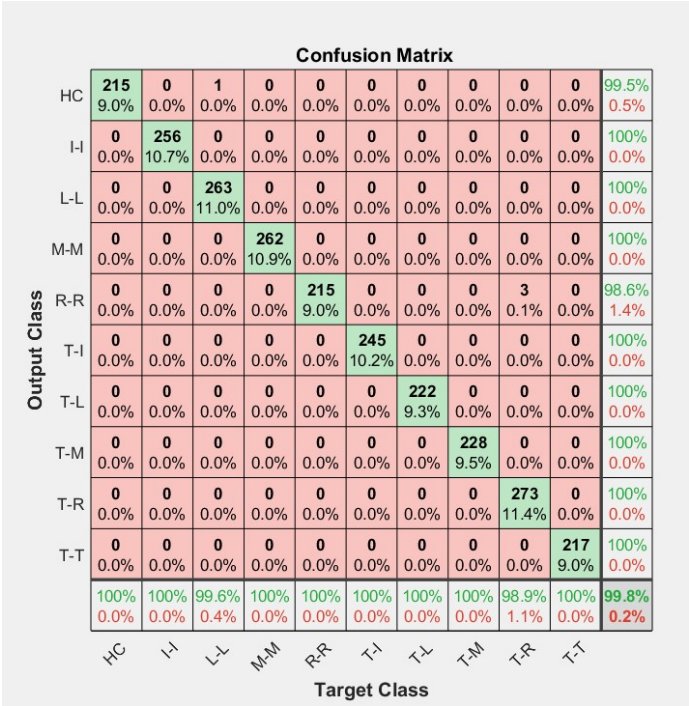

Fig. 9. Confusion Matrix for ResNet-18+LSTM in classifying 10 gestures in Rami Data Set.

TABLE I
ACCURACY COMPARISON OF PROPOSED NEURAL NETWORKS AND OTHER PREVIOUS WORKS ON SAME DATA SETS

Work	Data Set		
	Rami (%)	UCI Data Set-1 (%)	UCI Data Set- 2 (%)
Rami et al. [4]	90		
Sapsanis et al. [5]		89.21	
Sapansis et al. [6]			75
Anam et al. [29]	97.46		
Alam et al. [30]	98.88		
Martinez et al. [31]		99.88	
Kim et al. [32]			84.66
Tabatabaei et al. [33]		92.53	
Libal [34]			96.90
Yavuz et al. [35]		98.81	
AlexNet	99.64	98.70	
ResNet-18	99.55	98.61	
proposed GoogLeNet + LSTM	99.64	94.31	98.35
proposed ReseNet-18 + LSTM	99.83	99.72	99.23

TABLE II
COMPARISON OF PREDICTION TIME FOR ALL NEURAL NETWORKS IN ALL DATASET IN TESTING ONE SAMPLE IMAGE

Dataset	AlexNet (ms)	ResNet-18 (ms)	GoogleNet + LSTM (ms)	ResNet-18 + LSTM (ms)
Rami	34.51	98.54	211.46	173.20
UCI	41.72	99.90	207.47	171.31

vary greatly due to limb position, orientation, day to day activity, muscle fatigue, etc. So, reproducing the same result for the same classes of gestures can be difficult. Analyzing sEMG signals, we found that it has many slowly and abruptly changing portions. Its frequency changes with time, and specific gestures have specific change patterns. To capture this pattern, we used the Continuous Wavelet Transform (CWT). CWT is an excellent technique for capturing frequency and time information simultaneously. Using this technique, we created a full time-frequency image representation of the sEMG signal. CWT captures the slow and abrupt changes of a signal by shifting and scaling a wavelet continuously over the signal. By examining closely, we observed that each gesture has a distinct pattern of slow and abrupt changing features in their image. So, instead of engineering commonly used time or frequency domain features, (which are generally used in machine learning and known as feature engineering or extraction), we used CWT images of sEMG signal as feature image. Pre-trained ConvNets learned useful features themselves from those images. We could use raw sEMG for image creation. But then it would have been difficult for CNNs to extract useful information from those signals as they are highly stochastic and noisy. So, instead of raw sEMG, CWT offers an excellent feature representation, which reduces the work and computational load of CNNs, helping them to achieve very high accuracy.

Although the combination of CNN and LSTM is used heavily in many research fields, it is rarely used in EMG. To the best of our knowledge we used the combination of CWT, CNN, and LSTM in sEMG classification for the first time here. As sEMG is a sequential signal, LSTM suits best for this classification task. Usually, pre-trained networks are very large, and it can take a long time for a network to predict.

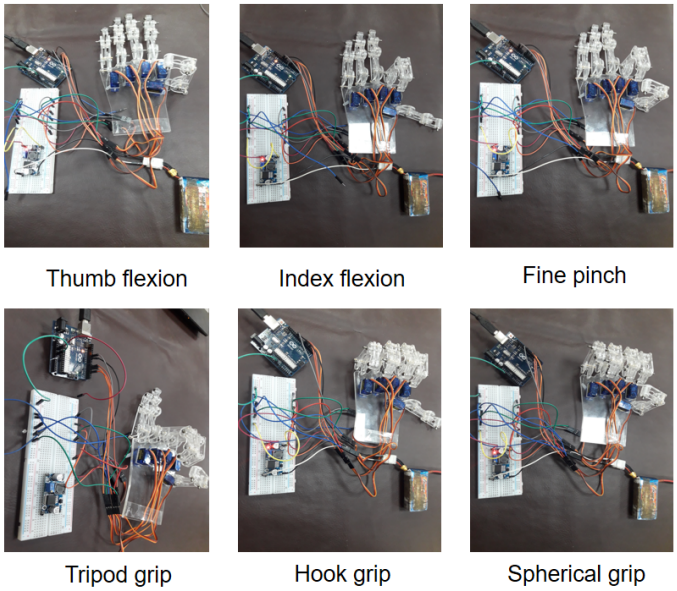

Fig. 10. Experimentation of different hand gestures of DUFAB Hand using the discussed control methods.

TABLE III
COMPARISON OF GRASP TIME OF SOME WELL KNOWN PROSTHETIC HANDS AND DUFAB HAND

	Vanderbilt Hand [36] (s)	Keio Hand [37] (s)	Bebionic Hand [38] (s)	DUFAB Hand (s)
Grasp Time	0.4	0.8	1.9	2

So, instead of using the whole network, only a small portion is used for extracting useful features, and these feature extracted sequences are fed as input to the LSTM. This is a useful method for sequence classification. It showed very good and consistent results, as previously discussed.

D. Performance comparison based on prediction time and grasp time

For real-time control of prosthetic hand or robot control, prediction time is an important factor to consider. As previously stated, a window size of 250 ms was chosen and the approximate time required for filtering this window was identified as 10 ms. From various experiments, the time required for generating one CWT image was found as 35 ms. TABLE II shows the comparison of average prediction time for all the neural networks studied here. From this table, it can be observed that AlexNet takes the lowest time. After considering the time required for windowing, pre-processing, CWT coefficient generation, and classification, the total time was found as 329.50 ms for AlexNet. So, with reasonable accuracy and lowest prediction time, AlexNet is the best possible choice for classifying sEMG signals for DUFAB Hand. From the comparison from TABLE III performance of DUFAB Hand after using deep network control scheme can be esteemed quite satisfactory.

IV. CONCLUSION AND FUTURE WORK

The main focus of this paper was to design a control scheme that may give the best performance for DUFAB Hand. For that, various pre-trained models were compared. Experiments were carried out several times to validate the results. sEMG signal is extremely noisy, stochastic, and prone to error. So, pre-processing is fundamental to achieve good performance. To get good performance from CNN, noiseless images are necessary with useful time-frequency features that were successfully developed with careful filter selection and sophisticated techniques like CWT. For both Rami and UCI Data sets, we have achieved state of the art results with accuracy 99.83% and 99.23% for LSTM+ResNet18. Though the performances of all the control schemes were highly accurate while tested in DUFAB Hand, we are considering the cost issues. So, among all the controlling methods, transfer learning using AlexNet is the best possible choice for controlling DUFAB Hand. Our main contributions in this paper are achieving state of the art result for both data sets, finding a better-suited control scheme for DUFAB Hand and developing two modified LSTM+CNN (GoogleNet+LSTM and ResNet-18+LSTM) architectures for EMG classification in MATLAB. For future development of this work, our next goals are to test benchmark data set like NinaPro or MeganePro to analyze the performance of these control schemes, designing our own CNN network with good accuracy and low processing time, and to create our very own EMG repository.

ACKNOWLEDGMENT

The authors would like to thank Fab Lab DU for funding the project and providing logistic supports and experimental facilities.

REFERENCES

[1] M. LeBlanc, "Estimates of amputee population." [Online]. Available: https://web.stanford.edu/class/engr110/2011/LeBlanc-03a.pdf (accessed Feb 24, 2020).

[2] N. Nahid, A. Rahman, T. K. Das, K. M. Khabir, A. Islam, and M. S. Alam, "Design and implementation of DUFAB Hand, a low-cost myoelectric prosthetic hand," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, Spokane, WA, USA, 2019, pp. 206–211.

[3] E. Biddiss, D. Beaton, and T. Chau, "Consumer design priorities for upper limb prosthetics," *Disability and Rehabilitation: Assistive Technology*, vol. 2, pp. 346–357, 2007.

[4] R. N. Khushaba, M. Takruri, S. Kodagoda, and G. Dissanayake, "Toward improved control of prosthetic fingers using surface Electromyogram (EMG) signals," *Expert Systems with Applications*, vol. 39, pp. 10 731–10 738, 2012.

[5] C. Sapsanis, G. Georgoulas, A. Tzes, and D. Lymberopoulos, "Improving emg based classification of basic hand movements using EMD," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013, pp. 5754–5757.

[6] C. Sapsanis, "Recognition of basic hand movements using Electromyography," Diploma Thesis, University of Patras, 2013.

[7] R. Merletti and P. Parke, "Control of powered upper limb prostheses," in *Electromyography: Physiology, Engineering, and Non-Invasive Applications*. Wiley-IEEE, 2004, pp. 453–475.

[8] K. Park and S. Lee, "Movement intention decoding based on deep learning for multiuser myoelectric interfaces," in *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, Yongpyong, 2016, pp. 1–2.

[9] M. Atzori, M. Cognolato, and H. Mller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in neurorobotics*, vol. 10, 2016.

[10] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Scientific Reports*, vol. 6, 2016.

[11] Y. Du, W. Jin, W. Wei, Y. Hu, and W. Geng, "Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation," *Sensors*, vol. 17, pp. 458–480, 2017.

[12] U. Ct-Allard, C. L. Fall, A. Campeau-Lecours, C. Gosselin, F. Laviolette, and B. Gosselin, "Transfer learning for sEMG hand gestures recognition using convolutional neural networks," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, AB, 2017, pp. 1633–1688.

[13] U. Ct-Allard *et al.*, "Deep learning for Electromyographic hand gesture signal classification using transfer learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, pp. 760–771, 2019.

[14] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid cnn-rnn architecture for semg-based gesture recognition," *PLOS ONE*, vol. 13, pp. 10–28, 2018.

[15] G. Jia, H.-K. Lam, J. Liao, and R. Wang, "Classification of electromyographic hand gesture signals using machine learning techniques," *Neurocomputing*, 2020.

[16] C. Li, J. Ren, H. Huang, B. Wang, Y. Zhu, and H. Hu, "PCA and deep learning based myoelectric grasping control of a prosthetic hand," *BioMedical Engineering Online*, 2018.

[17] G. Ghazaei, A. Alameer1, P. Degenaar, G. Morgan, and K. Nazarpour, "Deep learning-based artificial vision for grasp classification in myoelectric hands," *Journal of Neural Engineering*, vol. 14, 2017.

[18] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, pp. 82–94, 1993.

[19] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 15, pp. 111–118, 2007.

[20] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: Balancing the competing effects of classification error and controller delay," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, pp. 186–192, 2011.

[21] S. Yunhui and R. Qiuqi, "Continuous wavelet transforms," in *Proceedings 7th International Conference on Signal Processing*, Beijing, China, 2004, pp. 207–210.

[22] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable EMG sensors," *Sensors*, vol. 18, 2018.

[23] D. Tkach, H. Huang, and T. Kuiken, "Study of stability of time-domain features for Electromyographic pattern recognition," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, 2010.

[24] A. Phinyomark and E. Scheme, "A feature extraction issue for myoelectric control based on wearable EMG sensors," in *Proceedings of the IEEE Sensors Applications Symposium (SAS)*, Seoul, Korea, 2018, pp. 1–6.

[25] E. Scheme and K. Englehart, "On the robustness of EMG features for pattern recognition based myoelectric control; a multi-dataset comparison," in *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Chicago, IL, USA, 2014, pp. 650–653.

[26] "Imagenet." [Online]. Available: http://www.image-net.org (accessed Feb 25, 2020).

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, 2012.

[28] K. He, Z. Xiangyu, R. Shaoqing, and S. Jian, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2016, pp. 770–778.

[29] K. Anam and A. Al-Jumaily, "Real-time classification of finger movements using two-channel surface Electromyography," in *NEUROTECHNIX 2013 - Proceedings of the International Congress on Neurotechnology, Electronics and Informatics*, 2013, pp. 218–223.

[30] R. Alam, S. R. Rhivu, and M. A. Haque, "Improved gesture recognition using deep neural networks on sEMG," in *2018 International Conference on Engineering, Applied Sciences, and Technology (ICEAST)*, Phuket, Thailand, 2018, pp. 1–4.

[31] D. Martinez, M. Ponce, O. Pogrebnyak, and A. C. M. Prez, "Hand movement classification using burg reflection coefficients," *Sensors*, vol. 19, 2019.

[32] J. Kim and S. B. Pan, "A study on EMG-based biometrics," *Journal of Internet Services and Information Security (JISIS)*, vol. 7, pp. 19–31, 2017.

[33] S. M. Tabatabaei and A. Chalechale, "Local binary patterns for noise-tolerant sEMG classification," *Signal, Image and Video Processing*, vol. 13, pp. 491–498, 2018.

[34] U. Libal, "Pattern recognition using adaptive schur-like parametrization of signals with forgetting factor," in *2018 International Conference on Signals and Electronic Systems (ICSES)*, Krakow, Poland, 2018, pp. 151–156.

[35] E. Yavuz and C. Eyupoglu, "A cepstrum analysis-based classification method for hand movement surface EMG signal," *Medical and Biological Engineering & Computing*, pp. 1–23, 2018.

[36] S. A. Dalley, T. E. Wiste, T. J. Withrow, and M. Goldfarb, "Design of a multifunctional anthropomorphic prosthetic hand with extrinsic actuation," *IEEE/ASME Transactions on Mechatronics*, vol. 14, pp. 699–706, 2009.

[37] T. M. Y. Kamikawa, "Underactuated five-finger prosthetic hand inspired by grasping force distribution of humans," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Nice, France, 2008, pp. 717–722.

[38] "Rsl steeper." [Online]. Available: http://www.bebionic.com/wp-content/uploads/bebionic-ProductBrochure-Final.pdf (accessed Feb 25, 2020).