Optimal Feature Selection for EMG-Based Finger Force Estimation Using LightGBM Model

Yuhang Ye, Chao Liu, Nabil Zemiti, Chenguang Yang

Optimal Feature Selection for EMG-Based Finger Force Estimation Using LightGBM Model

Yuhang Ye[1], Chao Liu[*2], Nabil Zemiti[2], Chenguang Yang[*3]

Abstract— **Electromyogram (EMG) signal has been long used in human-robot interface in literature, especially in the area of rehabilitation. Recent rapid development in artificial intelligence (AI) has provided powerful machine learning tools to better explore the rich information embedded in EMG signals. For our specific application task in this work, i.e. estimate human finger force based on EMG signal, a LightGBM (Gradient Boosting Machine) model has been used. The main contribution of this study is the development of an objective and automatic optimal feature selection algorithm that can minimize the number of features used in the LightGBM model in order to simplify implementation complexity, reduce computation burden and maintain comparable estimation performance to the one with full features. The performance of the LightGBM model with selected optimal features is compared with 4 other popular machine learning models based on a dataset including 45 subjects in order to show the effectiveness of the developed feature selection method.**

I. INTRODUCTION

Human-robot interaction (HRI) has been an active research topic during past few decades as the advances in robot technology and artificial intelligence (AI) have enabled the robotic systems to work side-by-side or in a collaborative fashion with humans. The applications have extended beyond traditional industrial domain to new areas such as service, health care, etc [1].

Efficiency and safety are of paramount importance in practical HRI applications. As indicated in [2], one principle of efficient HRI is to "Directly Manipulate the World", meaning interfaces should allow the task to be done without drawing attention to the robot and the interface per se. Conventional human-robot interfaces (such as cameras, force sensors) are not sufficient due to their limitations in working condition requirements, cost, flexibility, level of intelligence, etc. Myoelectric or Electromyogram (EMG) signal has been used as a complementary source of information for HRI since it originates from the eletrophysiological and mechanical activation of muscle fibers in vivo and hence can be considered as a natural pathway to detect human motion intention from the nervous system without constraining human motion. This signal has been especially and almost exclusively used in the

domain of robotic rehabilitation in order to control artificial limb prosthesis and exoskeleton system since early literature [3], [4]. In rehabilitation literature, the EMG signal has been mainly used for motion and posture control of artificial limbs prostheses and hands. Much less but recent research efforts have also been made for force control of artificial hands using EMG signal [5].

During the past few years, along with the booming development of artificial intelligence, new research results indicate that plain and traditional EMG signals can be further explored in a better way by improving the signal analysis methods. Except conventional motion and posture control of upper limb prostheses and hands, recent research works have been able to derive new biomechanical information such as arm stiffness and to study finer scale such as hand fingers with the aide of new machine learning methods [6]. And the applications of EMG is no longer limited to rehabilitation devices, and they also find wider potential applications in HRI such as teleoperation of robots, haptic devices, and so on. In these new HRI applications, human-robot interaction force is of increasing importance in successful completion of task, with the haptic teleoperation as a good example [7]. An increasing number of researchers have been attracted to work on improving the performance of surface EMG (sEMG)-based force estimation from various aspects. In literature, numerous models have been proposed to map the EMG-force relationship, including the Hill model [8], polynomial fitting model [9], fast orthogonal search (FOS) [10], and parallel cascade identification (PCI) [11]. As state-of-the-art, models of two main groups are widely used: the ones based on neural network (NN) [12] and the ones based on decision tree [13]. Both kinds of methods can be effective for different estimation tasks, and both have their advantages and limitations.

In this work, the specific application is to estimate in real time the finger force of intact subject based on forearm surface EMG signal measurement. Although not new, this serves as the first step towards our study on the effects of surgeons' arm and hand muscle activation on their skill performances in both robot-assisted surgical training and operation, where the force and motion generated by surgeon's hand finger and wrist are the decisive factors. To achieve accurate and efficient finger force estimation, the proper estimation model is to be first decided. Considering the requirement of generalisability due to large number of potential users, NN-based methods may not be the best choice as neural networks are usually hard to train and tune due to the subject-dependant cost for training and optimizing

[1]Yuhang Ye is with Key Laboratory of Autonomous Systems and Networked Control, College of Automation Science and Engineering, South China University of Technology, Guangzhou, 510640, China `yhang.au@gmail.com`

[2]Chao Liu and Nabil Zemiti are with the Department of Robotics, LIRMM, CNRS-University of Montpellier, UMR5506, 34095 Montpellier, France. `liu,zemiti@lirmm.fr`

[3]Chenguang Yang is with Bristol Robotics Laboratory, University of the West of England, Bristol, BS16 1QY, UK. `cyang@ieee.org`

*Co-corresponding authors

the models [14]. A LightGBM (Gradient Boosting Machine) model [15] under the general class of Gradient Boosting Decision Tree (GBDT) methods has been chosen for this estimation task with its wide generalisability, robustness against signal noises and less proneness to overfitting. However, it is noticed that the estimation/regression performance of LightGBM model depends on the features extracted from the sample datasets. The feature(s) selection is often empirical and hence cannot guarantee the best performance. Generally, the performance is better with more features used. On the other hand, more feature extraction implies more computation cost and therefore longer computation time, which is critical for real time application.

Therefore, in this study, we develop a method to automatically select the optimal features among various time-domain and frequency-domain features to be used in the LightGBM model for the specific aforementioned application scenario. The number of selected optimal features should be minimal in order to save the calculation burden but do not cause noticeable performance deterioration at the same time. The feature selection is performance-based and therefore is not affected by the designer's experience. The comparisons of the finger force estimation performance obtained from the LightGBM with optimal features are made to the one of LightGBM with full features and also to the estimation performances of 4 other popular models (linear regression (LR), support vector regression (SVR), convolutional neural network (CNN), and Multilayer Perceptron (MLP)) with all features available in order to confirm the effectiveness of the selected optimal features and the automatic selection method.

II. MATERIALS AND METHODS

A. Experimental Setup

In this work, we use the online open dataset "putEMG-Force" [1] [16] provided by the Biomedical Engineering and Biocybernetics Team of Poznan University of Technology, Poland, to serve as the basis for evaluation of the method we propose. The putEMG Dataset is a database of surface electromyographic activity recorded from forearm which allows the development and evaluation of algorithms for both gesture recognition and hand finger force estimation. The essential information on the experiment platform setup and data collection is provided in following, and more details can be found on the putEMG website and [16].

The experiment platform is illustrated in Fig. 1. The system is dedicated to sEMG signal acquisition of forearm muscle activity for a single subject. In this work, we use the EMG signals recorded from 24 electrodes fixed around subject right forearm and the finger forces measured by tensometer sensors as shown in Fig. 2. The dataset includes experiment data collected from 45 healthy, fully-abled subjects (8 females, 37 males) aged 19 to 37 years old. It should be noted that the value of measured force from the tensometer has no unit of physical meaning, since the

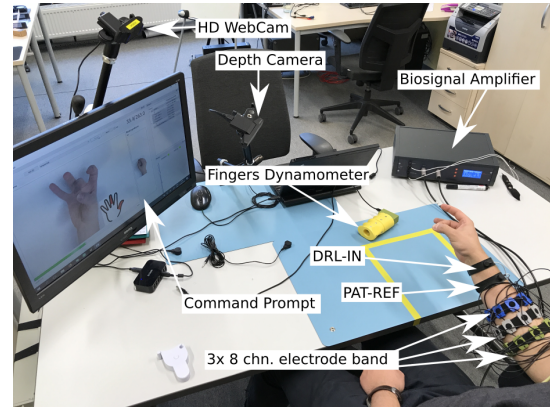conversion has not been done yet according to the dataset provider.
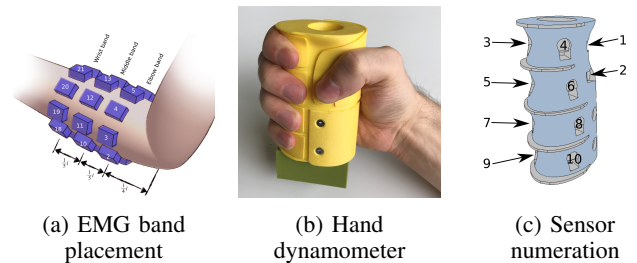

Fig. 1: putEMG experiment platform

| (a) EMG band placement | (b) Hand dynamometer | (c) Sensor numeration |

Fig. 2: EMG electrodes placement and finger force sensors

B. Background

EMG signal is a non-stationary signal containing a variety of noises and artifacts. These noises will deteriorate the models in performance if the raw EMG data is used as the input for classification or regression. Therefore, features extraction is very important in EMG signal processing. The widely used features can be summarized into three categories: time domain (TD) features, frequency domain (FD) features and time-frequency domain (TFD) features.

Time-domain statistical analysis is the most commonly used method for EMG signal processing because they are faster and smaller than other features, and this is very important for real-time systems. In 1993, Hudgins et al. proposed five TD feature of sEMG, including mean absolute value (MAV), mean absolute value slope (MAVSLP), slope sign changes (SSC), waveform lengths (WL) and zero crossings (ZC) [18], [19], [20]. Zardoshi et al. extracted the features of integral of absolute value (IVA), zero crossing (ZC), variance(VAR), Willison amplitude (WAMP) to control the artificial limb by EMG signal [17]. Also for the applications of prosthetic hand, Ahmad and Chappel adopted the skewness (Skew), Kurtosis (Kurt) and moving approximate entropy (mApEn) as the TD feature for the first time [21]. In order to investigate the stability of time-domain EMG features on a task of low and high forces classification by EMG, Dennis Tkach et al. [22] used 11 TD features in

their experiment and they found the classifier performance could be improved by the use of at least four combined EMG features and the multi-dimensional features, such as autoregression coefficients (AR) and cepstrum coefficients (CC), could get the best performance.

Although time-domain features are easy to extract, a lot of studies have shown that when the muscle contraction force changes slightly, the frequency-domain description of EMG signals is relatively more stable than the time-domain features [23]. The frequency domain analysis of EMG mainly applies power spectrum analysis methods. As early as the late 1980s, Christensen used Fourier transform to analyze the power spectrum of surface EMG signals [24]. The difference of power spectrum between normal individuals and patients with neuromuscular diseases was revealed by the amplitude ratio of high frequency to low frequency. The features of power spectrum extraction mainly include peak frequency (PKF), mean power frequency (MNP), Mean frequency, median frequency (MDF) and so on, which have been successfully used as muscle force, fatigue and geometry indices [25]. To further improved the EMG performance, Al-Timemy et al. [26] considered the time-frequency characteristics of EMG and proposed a set of TD-PSD feature for hand movement classification with three force levels.

Traditional Fourier transform can better depict the signal characteristics in global frequency but can hardly provide the frequency information in the time domain, which is insufficient for the analysis of the non-stationary EMG signal. The time-frequency analysis methods combine both the time and the frequency information for EMG signal analysis and thus received wide research attention. At present, the main time-frequency analysis methods used in EMG analysis are short-time Fourier transform (STFT), Wigner-Ville transform (WVD), Choi-Williams distribution (CWD) and wavelet transform (WT). Veer et al. used Short-time Fourier transform and wavelet transform for the recognition of arm movements and their result showed that wavelet transform performs better [27].

Benefiting from the improved computation power and the development of big data, deep learning algorithms have been studied and applied in the fields of image recognition, speech recognition, and intelligent robots. It has a strong capability to automatically extract features by end-to-end learning. In recent years, many researchers have tried to apply deep learning to the field of sEMG signal recognition. Atzori et al. [28] used the proposed CNN architecture for hand movement recognition based on the EMG signal. After the preprocessing of sub-sample and low-pass filter, EMG data are fed into the CNN model which is composed of convolutional, pooling, and softmax layers. They compared the CNN method with other classical classification methods and got the conclusion that the classification accuracy of CNN's is higher than the average results of classical methods but lower than the best reference methods using random forests algorithm with the features of discrete wavelet transform (DWT), histogram (HIST), waveform length (WL) and root mean square(RMS). Therefore, CNN is not always

superior in feature extraction, and how to extract the well-performed features and find optimal feature combination is very important in EMG recognition field, which is also the focus in this work.

C. Optimal Features Selection Strategy

1) Data preprocessing and feature extraction: In order to extract the effective signal of EMG and reduce noise, a bandpass filter of 10-350Hz is used on the 24-channels EMG data while a 50Hz low pass filter is adopted on the force data. Taking the accuracy and real-time requirements of force estimation into consideration synthetically, we use a sliding window with the length of 256 sample points (50 ms) and move across 32 sample points (6.25 ms) for each step to segment the EMG and the force signals. The segmenting process is illustrated in Fig. 3.
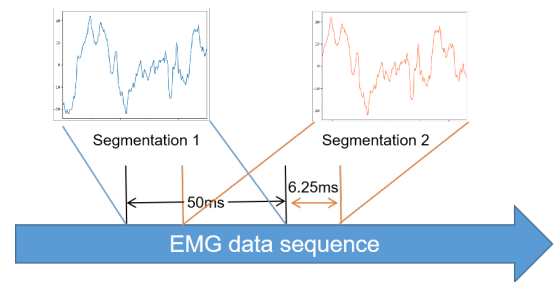

Fig. 3: EMG segmentation with 50ms sliding window

To evaluate the feature importance and find the best feature combination, we consider the most commonly used features both in time and frequency domain and most of them have been introduced in subsection II-B. We do not select the time-frequency domain feature because our features were extracted in a 50ms window so the Fourier transform can be seen as a Short-time Fourier transform which contains the time and frequency information. The features we consider is shown in Table I.

After data filtering and feature extraction, EMG raw data of one sample is transformed into a feature vector $\{x_1, x_2, ..., x_d\}$. Because the features are calculated by different methods as shown in Table I, their scale are different which will affect the convergence speed and precision of the algorithm. A min-max scaled normalization method was adopted here to convert the features to the same scale,

$$\hat{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where $\hat{x}_i$ is the normalized features. x_{min} and x_{max} are the minimum and maximum value of this feature. After that, all the features will be normalized to a range from 0 to 1.

2) Measure of the Feature importance: Gradient Boosting Decision Tree (GBDT) [37] is an ensemble learning algorithm that utilizes decision or the regression trees as weak classifiers. It continuously adds new regression trees by the method of gradient descent, so that the total model residuals are continuously reduced.

Supposed that we have a training set with n samples of EMG data after normalization represented as $\{x_1, x_2...x_n\}$,

TABLE I: List of Features

	Features	Abbreviation	References
1	Integral Absolute Value	IAV	[22], [29]
2	Average Amplitude Change	AAC	[30], [29]
3	Auto-Regressive Coefficients	AR	[22], [29]
4	Cepstral Coefficients	CC	[30], [29]
5	Kurtosis	Kurt	[31], [29]
6	LOG Detector	LOG	[30], [29]
7	Maximum Amplitude	MAX	[36], [29]
8	Mean Absolute Value	MAV	[22], [30], [29]
9	Mean Absolute Value Slope	MAVSLP	[30], [29]
10	Multiple Hamming Windows	MHW	[30], [29]
11	Myopulse Percentage Rate	MYOP	[30]
12	Skewness	Skew	[32], [29]
13	Slope Sign Change	SSC	[30], [29]
14	Absolute Temporal Moment	TM	[30]
15	Variance	VAR	[30], [29]
16	Willison Amplitude	WAMP	[30], [29]
17	Waveform Length	WL	[30], [22], [29]
18	Zero Crossing	ZC	[30], [22], [29]
19	Mean Frequency	MNF	[34], [30], [29]
20	Median Frequency	MDF	[34], [30], [29]
21	Mean Power frequency	MNP	[33], [29]
22	Frequency Ratio	FR	[30], [29]
23	Variance of Central Frequency	VCF	[30], [29]
24	Power Spectrum Ratio	PSR	[30], [29]
25	Signal-to-Noise Ratio	SNR	[34], [29]
26	Maximum-to-minimum Drop in Power Density Ratio	DPR	[35]

and each $x_i \in R^d$ is a feature vector with d dimensions extracted using the methods in Table I. $\{g_1, g_2, ..., g_n\}$ are the negative gradients with respect to the output of the GBDT model which are calculated to select the most informative features for the tree node split. The most informative features can split samples in the largest information gain so that the tree can converge faster. The information gain of GBDT is usually measured by the variance after splits. In each iteration, we calculate the information gain of each segmentation point (j, s) by traversing each feature j and the split value s in turn. The information gain is formulated in following form:

$$V(j, s) = \frac{1}{n} \left(\frac{\left(\sum_{x_{ij} \leq s} g_i\right)^2}{n_l^j(s)} + \frac{\left(\sum_{x_{ij} > s} g_i\right)^2}{n_r^j(s)} \right) \quad (2)$$

where n is the number of the data in training set. x_{ij} denotes the feature j of the ith sample. $n_l^j(s)$ and $n_r^j(s)$ are the numbers of the samples split into left and right child node, so $n_l^j(s) = \sum I[x_{ij} \leq s]$, $n_r^j(s) = \sum I[x_{ij} > s]$ and $I[\cdot]$ is the indicator function. For each feature j, the GBDT algorithm selects the best split value by $s_j^* = \arg\max_s V(j, s)$ and calculates the largest gain $V(j, s_j^*)$. Then the algorithm will select the best feature and split value for the data split.

Here, we make full use of the advantages of GBDT algorithm in feature selection and measure the importance of a feature by calculating the number of times this feature is used for data split. Obviously, those features that can achieve greater information gain often provide better discrimination because they always put similar samples together and sepa-

rating different samples in data split.

3) Features selection: From the above subsection, the importance of each feature can be easily caculated by counting the number of times the features used in data split for the tree growth. To select the optimal feature combination, the root mean square error (RMSE) is used here to measure the regression performance, which is defined as below:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - p_i)^2} \quad (3)$$

where n is the number of the EMG samples, y_i is the label of the force signal and p_i is the predict output of x_i. Firstly, the result of RMSE using all the features by GBDT model is set as a baseline performance. Then, an iterative algorithm is used and drop the last 10% or 20% features according to the feature importance in each time. Performance of the remaining features is tested in GBDT model and the feature importance is re-evaluated. If the performance loss (compared with the baseline performance) is less than certain threshold after deleting the feature, the algorithm will continue to drop the last features based on the new feature importance until the performance loss exceeds the threshold or the number of features is less than the preset minimum, and the last feature combination will be the optimal feature. It should be noted that for a specific application the feature selection is expected to be conducted in advance to accelerate the online EMG signal processing, thus the calculation time for the offline feature selection phase is not demanding. The algorithm is outlined in Algorithm 1.

III. EXPERIMENTAL EVALUATION

In robot-assisted surgical operation, surgeons usually use their thumb and index finger to control the joystick, so the muscle activation related to the index finger shows more potential in surgeons' skill performances analysis. Therefore, force 4 (channel 4 in the force sensor shown in Fig.2c) estimation of index finger is used for experimental evaluation in this work. For the putEMG database, there are five kinds of trajectories consisting of varying combinations of pressure magnitude, duration and shape, and the trajectories of "repeats_long". Each trial of "repeats_long" contains 5 action blocks (each block for 70s and 350 seconds in total) where thumb, index, middle, ring+small and all fingers take turns. For each block, press action repeated 11 times with varying parameters. The press action of the index finger is mainly applied between 70 and 140 seconds in a trial, which was selected for our experiment. For each trial, We divided the data into three parts: the last 20% data for testing, the randomly select 20% of the first 80% data for validation, and the rest for training. Force prediction for multiple channels and subject-independent estimation are our future works and thus not considered in this paper.

In the experiment, the proposed optimal features selection strategy is tested in six randomly selected subjects from the all 45 subjects. The GBDT model we choose is the LightGBM (LGB), which performs extremely well in both

Algorithm 1 Automatic Algorithm for Optimal Feature Selection

Input: Training data with high dimension features x_i; Label of the training data y_i; The percentage of features retained for each iteration $P\%$; Performance loss threshold θ. Minimum feature numbers k

Output: The optimal features combination $best$.

Initial: The number of iteration $t = 0$; The features selected are all the features $Feature = all$; The RMSE performance of the all the features in validation set R_0.

Repeat:

1: t = t + 1.
2: The best feature combination $best = Feature$.
3: Train the GBDT model by gradient boosting. Use the variance as shown in eq.2 for the informative gain measure and split the data by the best feature and best split value.
4: Caculate the feature importance by computing the number of time for each features using in data split. And sort the feature by feature importance.
5: Use the RMSE to evaluate the performance in validation set and get the result of R
6: Select the top $P\%$ feature tp according to the feature importance, $Feature = tp$. And $len(tp)$ is the number of the feature in top $P\%$. Update the training set.
7: **Until** : $R - R_0 > \theta$ **or** $len(tp) < k$

(a) Prediction with all features for train&valid sets

(b) Prediction with one final selected feature for train&valid set

(c) Prediction with all features for test set

(d) Prediction with one final selected feature for test set

(e) Features importance in iteration 1 (Top 3 features: TM, AAC, VAR)

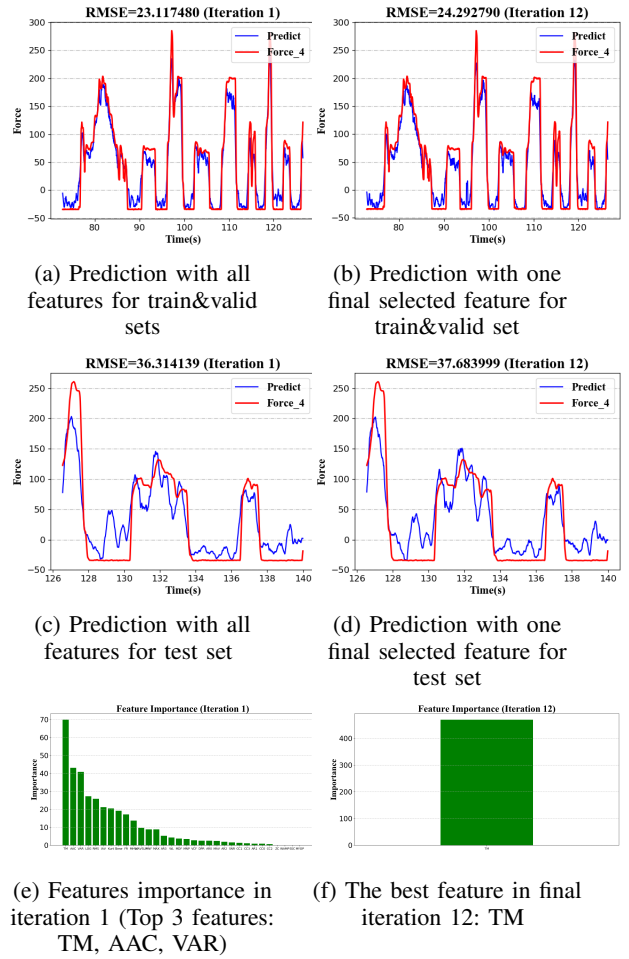(f) The best feature in final iteration 12: TM

Fig. 4: The force estimation performance with the selected features and all features for subject 35

TABLE II: The selected feature of the six subjects

Subject	s12	s14	s15	s35	s37	s39
Selected feature	AAC	AAC	AAC	TM	TM	AAC

performance and speed [15]. For the hyper parameters, the threshold is set to 10% performance loss in the validation set compared with the result using all the features. The Minimum feature number is 1 and the percentage of features retained for each iteration is 80%. These features will be evaluated by the LGB and the feature importance will be calculated again for the next iteration until the performance loss exceeds the threshold or the number of the retained features is less than or equal to 1. Our experiments show that the algorithm for all 6 subjects was stopped when the number of features reduced to 1. The feature selection results and performances of subject 35 and 39 are shown in Fig. 4 and Fig. 5 respectively. It should be noted again that the unit of "Force" has no physical meaning as the tensometer measurements were not yet converted to real force in Newton.

From the subgraph (e) of Fig. 4 and Fig. 5 we can see that the importance of different feature varies greatly but the top important features are similar in subject 35 and 39. Features that have the best importance at the beginning of the iterations don't stay at the top all the time. The importance of the features changes slightly because the combination of features keeps changing during the process of feature selection. So, in subject 37 we can see that the AAC feature is at the top in iteration 1 but the TM feature finally won. Similar conclusions can be drawn from other subjects and the best features selected by the six subjects were shown in Table II.

From Table II, it is seen that AAC and TM both appear as the best feature depending on different subject, and we can therefore conservatively draw a conclusion that AAC and TM features may be considered as the optimal feature combination in the total 26 features.

To further evaluate whether the selected features contain enough information for force estimation with small performance loss, the results of selected optimal features (AAC and TM) in lightGBM model are compared with 4 widely used models using all the features. The 4 models are linear regression (LR), support vector regression (SVR), convolutional neural network (CNN), and Multilayer Perceptron (MLP). The CNN model contains 6 convolution layers, 3 pooling layers, and 1 fully-connected layer. The activation function of scaled exponential linear units (SELU), batch normalization, and dropout are also used in CNN models. There are 3 hidden layers in MLP model and the activation function is rectified
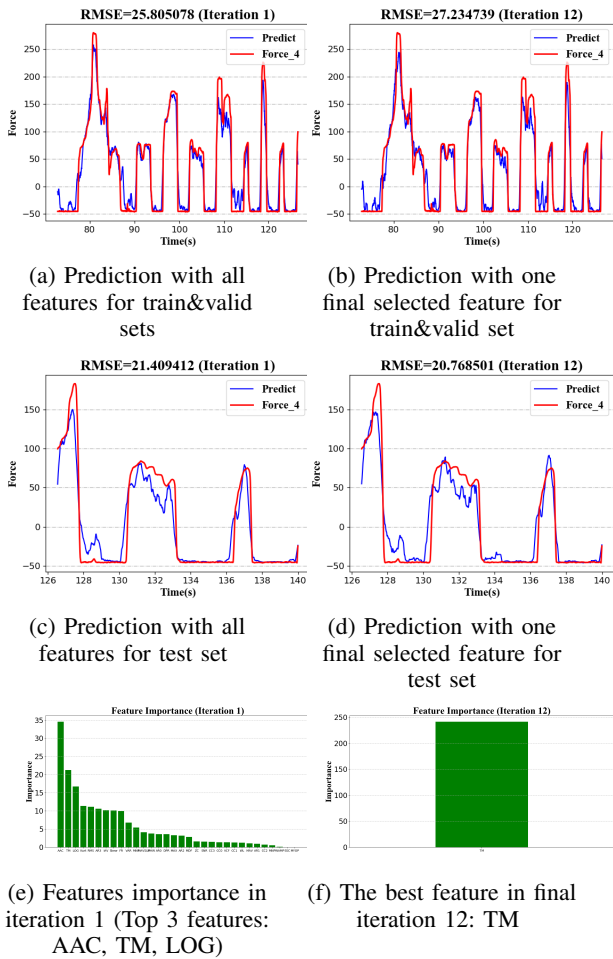
(a) Prediction with all features for train&valid sets

(b) Prediction with one final selected feature for train&valid set

(c) Prediction with all features for test set

(d) Prediction with one final selected feature for test set

(e) Features importance in iteration 1 (Top 3 features: AAC, TM, LOG)

(f) The best feature in final iteration 12: TM

Fig. 5: The force estimation performance with the selected features and all features for subject 37

TABLE III: Comparison result of the five models evaluated by RMSE. Models of LR, SVR, MLP, and CNN use all the features while LGB are tested on all and the selected two feature. In the column of LGB, on the left side of "/" are the all features and on the right side the selected features.

Subject	LR	SVR	MLP	CNN	LGB
s12	45.89	49.16	43.69	43.03	38.54/38.33
s14	53.83	44.25	43.82	29.00	37.91/40.34
s15	38.29	30.69	26.42	22.75	25.04/25.32
s35	45.77	43.68	44.85	37.88	36.31/36.64
s37	21.54	26.25	23.62	24.69	21.41/21.47
s39	46.79	43.69	38.20	40.23	39.15/38.49
Average	42.01	39.62	36.77	32.93	33.06/33.43

algorithms. Since the EMG signal is complex and noisy, various features can be used in the force estimation using LightGBM model and hence may lead to real time computation issue. The developed optimal feature selection method is demonstrated capable of minimizing the number of important features without causing noticeable performance degradation. This result provides an efficient tool towards more complex real time hand motion and force estimation using EMG signals, which is part of our future research on robot-assisted surgical training and operation augmented with surgeon's muscle activation state information.

ACKNOWLEDGMENT

linear units (RELU). The results of comparisons are shown in Table III. From Table III, we can see that the performances for models of LR, SVR, and MLP are much worse than that of the LGB in terms of RMSE even though the LGB only uses the two selected features. It is not surprising that CNN has comparable performance because it has been proven to be very effective in the field of recognition. The selected two-feature LGB is slightly worse than the full-feature LGB medel, but compared with the models of LR, SVR, MLP and even the CNN its performance is still very competitive and the small performance loss is acceptable. The comparison results fully demonstrated that the proposed features selection method has potential to accelerate the processing of real-time EMG signal by reducing the number of involved features with marginal estimation performance loss.

IV. CONCLUSIONS

In this paper, an optimal feature selection method is presented for EMG-based finger force estimation. The force estimation resorts to the LightGBM model which represents one of the most successful state-of-the-art machine learning

REFERENCES

[1] Goodrich M. A and. Schultz A. C. Human-Robot Interaction: A Survey. Foundations and Trends in Human-Computer Interaction, 1(3), 27, pp 23-275, 2007.

[2] Goodrich M.A. and Olsen D.R. Seven principles of efficient human robot interaction, Proceedings of IEEE International Conference on Systems, Man and Cybernetics, pp. 3943–3948, Oct. 2003, Washington, DC, USA.

[3] McLean L. and Scott R. N., The early history of myoelectric control of prosthetic limbs (1945–1970), Powered Upper Limb Prostheses, Berlin, Germany: Springer, 2004, pp. 1–15.

[4] Andreasen D. S.,. Alien S. K and Backus D. A., Exoskeleton with EMG based active assistance for rehabilitation, 9th International Conference on Rehabilitation Robotics, Chicago, IL, 2005, pp. 333-336.

[5] Castellini C., Van der Smagt P., Sandini G. and Hirzinger G., Surface EMG for Force Control of Mechanical Hands, IEEE Int. Conf. Robotics and Automation, pp. 725–730, Pasadena, CA, USA, May 2008.

[6] Yang C., Luo J., Liu C., Li M. and Dai S., Haptics Electromyogrphy Perception and Learning Enhanced Intelligence for Teleoperated Robot, IEEE Transactions on Automation Science and Engineering, OnlineFirst, 2018.

[7] Guo J., Liu C. and Poignet P., A Scaled Bilateral Teleoperation System for Robotic-Assisted Surgery with Time Delay, Journal of Intelligent and Robotic Systems, OnlineFirst, August 2018.

[8] Hill A.V., The Heat of Shortening and the Dynamic Constants of Muscle. Proc. R. Soc. Lond. Ser. B, 126, pp. 136–195, 1938.

[9] Clancy E.A. and Hogan N., Relating agonist-antagonist electromyograms to joint torque during isometric, quasi-isotonic, nonfatiguing contractions. IEEE Trans. Biomed. Eng. 44, pp. 1024–1028, 1997.

[10] Mobasser F., Eklund J. M., Hashtrudi-Zaad K., Estimation of Elbow-Induced Wrist Force with EMG Signals Using Fast Orthogonal Search. IEEE Trans. Biomed. Eng., 54, pp. 683–693, 2007.

[11] Hashemi J., Morin E., Mousavi P., Mountjoy K. and Hashtrudi-Zaad K., EMG-Force Modeling Using Parallel Cascade Identification. J. Electromyogr. Kinesiol., 22, pp. 469–477, 2012.

[12] Schmidhuber J., Deep Learning in Neural Networks: An Overview, Neural Networks, 61, pp. 85–117, 2015.

[13] Chen T. and Guestrin C., Xgboost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.

[14] Xu L., Chen X., Cao S., Zhang X. and Chen X., Feasibility Study of Advanced Neural Networks Applied to sEMG-Based Force Estimation, Sensors, 18(10):3226, Sep. 2018.

[15] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems. 3146-3154, 2017.

[16] Kaczmarek P., Mańkowski T. and Tomczyński J., "putEMG - a surface electromyography hand gesture recognition dataset", arXiv:1907.07428 and under review of MDPI Sensors, 2019

[17] Zardoshti-Kermani M, Wheeler B C, Badie K, et al. EMG feature evaluation for movement control of upper extremity prostheses. IEEE Transactions on Rehabilitation Engineering, 3(4): 324-333, 1995.

[18] Hudgins B, Parker P, Scott R N. A new strategy for multifunction myoelectric control. IEEE Transactions on Biomedical Engineering, 40(1): 82-94, 1993.

[19] Englehart K, Hugdins B, Parker P. Multifunction control of prostheses using the myoelectric signal. Intelligent systems and technologies in rehabilitation engineering, 153-208, 2000.

[20] Englehart K, Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. IEEE transactions on biomedical engineering, 50(7): 848-854, 2003

[21] Ahmad S A , Chappell P H . Surface EMG pattern analysis of the wrist muscles at different speeds of contraction. Journal of Medical Engineering & Technology, 33(5):376-385, 2009.

[22] Tkach D , Huang H , Kuiken T A . Study of stability of time-domain features for electromyographic pattern recognition. Journal of NeuroEngineering and Rehabilitation,7,1(2010-05-21), 7(1):21, 2010.

[23] Doerschuk P C, Gustafon D E, Willsky A S. Upper extremity limb function discrimination using EMG signal analysis. IEEE Transactions on Biomedical Engineering, (1): 18-29, 1983.

[24] Ronager J, Christensen H, Fuglsang-Frederiksen A. Power spectrum analysis of the EMG pattern in normal and diseased muscles. Journal of the Neurological Sciences, 94(1):283-294, 1989.

[25] Thongpanja S, Phinyomark A, Limsakul C, et al. Application of mean and median frequency methods for identification of human joint angles using EMG signal, Information Science and Applications. Springer, Berlin, Heidelberg, 689-696, 2015.

[26] Al-Timemy A H, Khushaba R N, Bugmann G, et al. Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(6): 650-661, 2016.

[27] Veer K, Agarwal R. Wavelet and short-time Fourier transform comparison-based analysis of myoelectric signals. Journal of Applied Statistics, 42(7): 1591-1601, 2015.

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

[29] Nazmi N, Abdul Rahman M, Yamamoto S I, et al. A review of classification techniques of EMG signals during isotonic and isometric contractions. Sensors, 16(8): 1304, 2016.

[30] Phinyomark A, Phukpattaranont P, Limsakul C. Feature reduction and selection for EMG signal classification. Expert systems with applications, 39(8): 7420-7431, 2012.

[31] Nazarpour K, Al-Timemy A H, Bugmann G, et al. A note on the probability distribution function of the surface electromyogram signal. Brain research bulletin, 90: 88-91, 2013.

[32] Khushaba R N, Al-Ani A, Al-Jumaily A. Orthogonal fuzzy neighborhood discriminant analysis for multifunction myoelectric hand control.

IEEE Transactions on Biomedical Engineering, 57(6): 1410-1419, 2010.

[33] Oskoei M A, Hu H. Support vector machine-based classification scheme for myoelectric control applied to upper limb. IEEE transactions on biomedical engineering, 55(8): 1956-1965, 2008.

[34] Kendell C, Lemaire E D, Losier Y, et al. A novel approach to surface electromyography: an exploratory study of electrode-pair selection based on signal characteristics. Journal of neuroengineering and rehabilitation, 9(1): 24, 2012.

[35] Altin C, Er O. Comparison of different time and frequency domain feature extraction methods on elbow gestures EMG. European journal of interdisciplinary studies, 2(3): 35-44, 2016.

[36] Daud W M B W, Yahya A B, Horng C S, et al. Features extraction of electromyography signals in time domain on biceps Brachii muscle. International Journal of Modeling and Optimization, 3(6): 515, 2013

[37] Jerome H Friedman. Greedy function approximation: a gradient boosting machine. Annals of statistics, pages 1189–1232, 2001.