

Bayesian Approach for Adaptive EMG Pattern Classification Via Semi-Supervised Sequential Learning

Seitaro Yoneda¹ and Akira Furui¹

Abstract—Intuitive human-machine interfaces may be developed using pattern classification to estimate executed human motions from electromyogram (EMG) signals generated during muscle contraction. The continual use of EMG-based interfaces gradually alters signal characteristics owing to electrode shift and muscle fatigue, leading to a gradual decline in classification accuracy. This paper proposes a Bayesian approach for adaptive EMG pattern classification using semi-supervised sequential learning. The proposed method uses a Bayesian classification model based on Gaussian distributions to predict the motion class and estimate its confidence. Pseudo-labels are subsequently assigned to data with high-prediction confidence, and the posterior distributions of the model are sequentially updated within the framework of Bayesian updating, thereby achieving adaptive motion recognition to alterations in signal characteristics over time. Experimental results on six healthy adults demonstrated that the proposed method can suppress the degradation of classification accuracy over time and outperforms conventional methods. These findings demonstrate the validity of the proposed approach and its applicability to practical EMG-based control systems.

I. INTRODUCTION

Surface electromyogram (EMG) signals are electrical activities that capture the action potentials of motor units in muscles from the skin surface, reflecting internal states of muscle contraction. A classification model based on machine learning can estimate the correspondence between measured EMG signals and human’s motion intention, aiding the construction of interfaces for operating devices such as myoelectric prostheses [1]–[4].

The continuous utilization of EMG-based interfaces leads to a gradual decrease in motion recognition performance owing to variations in signal characteristics over time [5], [6]. These variations are caused by distribution shifts between training and test data, which can be attributed to factors such as muscle fatigue [7], change in motor imagery [8], and electrode shifts [9], [10]. Existing methods have predominantly employed static pattern classification models, which cannot adapt to such variations in signal characteristics.

To prevent the decline in classification performance over time, it is necessary to update the classification model sequentially based on changes in signal characteristics. For instance, Vidovic et al. and Zhu et al. demonstrated that updating the parameters of linear discriminant analysis (LDA) using a small amount of calibration data can sustain good performance over a prolonged period [11], [12]. In addition,

Karrenbach et al. showed that a similar strategy is effective in transfer learning of deep neural networks [13]. However, these approaches require the periodic acquisition of labeled calibration data, a burden on users.

Various techniques have been proposed to update the classification models without using labeled calibration data within the framework of semi-supervised learning to achieve adaptive motion recognition without increasing the burden on users [14]–[18]. The typical approach is based on self-training, which predicts labels for test data using a classification model trained with labeled data, subsequently updating the model through retraining with the predicted labels as pseudo-labels [14], [15]. However, this method requires the constant retention of updating data for retraining, and the amount of stored data gradually increases unless appropriate data replacement is performed. In addition, since all pseudo-labels are used for updates, unreliable classification may lead to incorrect updating. Therefore, if we could develop a method that enables the selective updating of the model for highly reliable data without retaining data instances, practical sequential learning for EMG would be realized.

This paper outlines a Bayesian approach for adaptive EMG pattern classification using semi-supervised sequential learning. The proposed method employs a Gaussian classification model (GCM) to predict motion class labels and estimate their associated confidence. The posterior distributions of GCM are then recursively updated via semi-supervised Bayesian sequential learning (SS-BSL) for test data with high predictive confidence, enabling adaptive motion recognition without the need for labeled calibration data. Furthermore, information from previous trials is stored as the prior distribution parameters; thus, the model can be updated without retaining data instances.

II. PROPOSED CLASSIFICATION METHOD

Fig. 1 shows an overview of the proposed method. The measured EMG signals at trial t are first processed for feature extraction and then input into a trained GCM to compute class predictions and their associated confidence. The resulting output is then utilized to update the posterior distributions via SS-BSL. These procedures are repeated each trial, thereby adapting to changes in EMG signal characteristics over trial progress.

A. Gaussian Classification Model (GCM)

GCM describes the relationship between the EMG signal $\mathbf{x}_n \in \mathbb{R}^D$ captured at data point n from D electrodes and the corresponding motion class label $\mathbf{y}_n \in \{0, 1\}^C$ (C is

This work was partially supported by JSPS KAKENHI Grant Number JP23H0343800.

¹Seitaro Yoneda and Akira Furui are with Graduate School of Advanced Science and Engineering, Hiroshima University, Higashi-hiroshima, Japan (e-mail: seitaroyoneda@hiroshima-u.ac.jp; akirafurui@hiroshima-u.ac.jp).

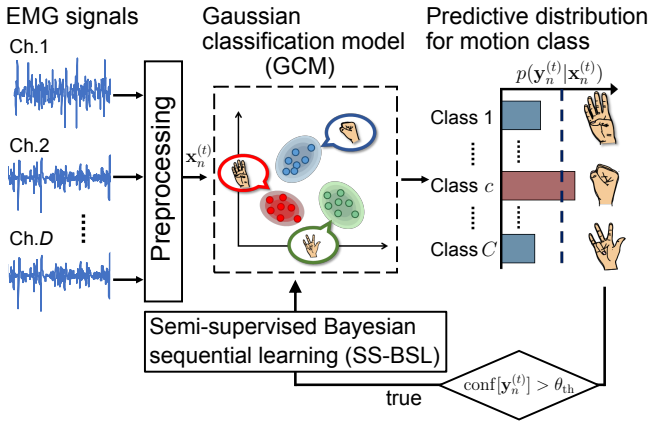


Fig. 1: Overview of the proposed GCM with SS-SBL

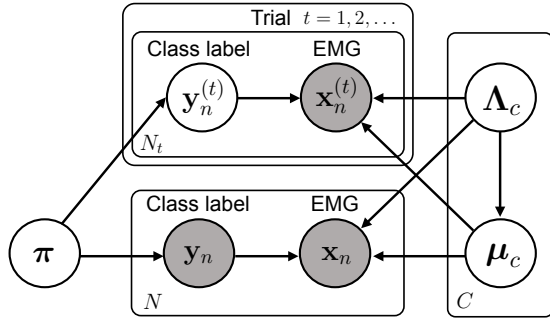


Fig. 2: Graphical model of GCM for EMG classification. The gray and white nodes represent the observed and unobserved variables, respectively.

the number of classes) using a Bayesian probabilistic model (Fig. 2). The observed model of the EMG signal \mathbf{x}_n for class $c \in \{1, \dots, C\}$ is expressed via the following Gaussian distribution:

$$p(\mathbf{x}_n | \mu_c, \Lambda_c) = \mathcal{N}(\mathbf{x}_n | \mu_c, \Lambda_c^{-1}), \quad (1)$$

where $\mu_c \in \mathbb{R}^D$ and $\Lambda_c \in \mathbb{R}^{D \times D}$ represent the mean vector and precision matrix (the inverse of the covariance matrix) of each class, respectively. To assign observations to each class, we introduce a one-hot representation for the motion class label $\mathbf{y}_n = \{y_{n,c}\}$. If $y_{n,c} = 1$ for a given c , it means that the c -th class is selected. Therefore, the distribution of \mathbf{y} can be expressed via the following categorical distribution:

$$p(\mathbf{y}_n | \boldsymbol{\pi}) = \text{Cat}(\mathbf{y}_n | \boldsymbol{\pi}) = \prod_{c=1}^C \pi_c^{y_{n,c}}, \quad (2)$$

where $\boldsymbol{\pi} = \{\pi_c\}$ is the mixing coefficients ($\pi_c \in [0, 1]$ and $\sum_{c=1}^C \pi_c = 1$).

To treat this model as a Bayesian model, a prior distribution is specified for each model parameter. For ease of computation, we assign Gaussian-Wishart and Dirichlet distributions, which are conjugate priors for $\{\mu_c, \Lambda_c\}$, and $\boldsymbol{\pi}$, as follows:

$$p(\mu_c, \Lambda_c) = \mathcal{NW}(\mu_c, \Lambda_c | \mathbf{m}, \beta, \nu, \mathbf{W}), \quad (3)$$

$$p(\boldsymbol{\pi}) = \text{Dir}(\boldsymbol{\pi} | \boldsymbol{\alpha}), \quad (4)$$

where $\mathbf{m} \in \mathbb{R}^D$, $\beta \in \mathbb{R}^+$, $\nu > D - 1$, $\mathbf{W} \in \mathbb{R}^{D \times D}$, and $\boldsymbol{\alpha} = \{\alpha_c\}$ are the hyperparameters and are common among classes.

B. Semi-Supervised Bayesian Sequential Learning (SS-BSL) with Pseudo-labels

In the proposed method, we introduce semi-supervised sequential learning to handle changes in the characteristics of EMG signals over time. First, the initial learning of the model is conducted using a labeled training dataset. Subsequently, the model is updated using sequential learning on newly acquired unlabeled test data. To achieve this, the predicted labels from the trained model are utilized as pseudo-labels. The entire process, comprising both initial and sequential learning, is performed within a unified Bayesian updating framework.

1) *Initial learning*: Let us consider the initial learning of a model with a labeled training dataset $\mathcal{D}_0 = \{(\mathbf{x}_n, \mathbf{y}_n)\}_{n=1}^N$. This can be achieved by computing the posterior distributions of the model parameters, μ , Λ , and $\boldsymbol{\pi}$, using \mathcal{D}_0 . The posterior distributions for $\{\mu_c, \Lambda_c\}$ and $\boldsymbol{\pi}$ for each class can be calculated as follows:

$$p(\mu_c, \Lambda_c | \mathcal{D}_0) \propto p(\mathcal{D}_0 | \mu_c, \Lambda_c) p(\mu_c, \Lambda_c), \quad (5)$$

$$p(\boldsymbol{\pi} | \mathcal{D}_0) \propto p(\mathcal{D}_0 | \boldsymbol{\pi}) p(\boldsymbol{\pi}). \quad (6)$$

From conjugacy, each posterior distribution is attributed to the following Gaussian-Wishart and Dirichlet distributions, which are of the same type as the priors:

$$p(\mu_c, \Lambda_c | \mathcal{D}_0) = \mathcal{N}(\mu_c | \hat{\mathbf{m}}_c, (\hat{\beta}_c \Lambda_c^{-1})) \mathcal{W}(\Lambda_c | \hat{\nu}_c, \hat{\mathbf{W}}_c), \quad (7)$$

$$p(\boldsymbol{\pi} | \mathcal{D}_0) = \text{Dir}(\boldsymbol{\pi} | \hat{\boldsymbol{\alpha}}), \quad (8)$$

where $\hat{\beta}_c$, $\hat{\mathbf{m}}_c$, $\hat{\nu}_c$, $\hat{\mathbf{W}}_c$, and $\hat{\boldsymbol{\alpha}}$ are the hyperparameters of the posterior distributions and are defined as follows:

$$\hat{\beta}_c = \sum_{n=1}^N y_{n,c} + \beta, \quad (9)$$

$$\hat{\mathbf{m}}_c = \frac{\sum_{n=1}^N y_{n,c} \mathbf{x}_n + \beta \mathbf{m}}{\hat{\beta}_c}, \quad (10)$$

$$\hat{\nu}_c = \sum_{n=1}^N y_{n,c} + \nu, \quad (11)$$

$$\hat{\mathbf{W}}_c^{-1} = \sum_{n=1}^N y_{n,c} \mathbf{x}_n \mathbf{x}_n^\top + \beta \mathbf{m} \mathbf{m}^\top - \hat{\beta}_c \hat{\mathbf{m}}_c \hat{\mathbf{m}}_c^\top + \mathbf{W}^{-1}, \quad (12)$$

$$\hat{\alpha}_c = \sum_{n=1}^N y_{n,c} + \alpha. \quad (13)$$

2) *Sequential learning*: Let $\mathcal{D}_t = \{(\mathbf{x}_n^{(t)}, \tilde{\mathbf{y}}_n^{(t)})\}_{n=1}^{N_t}$ denote the set of test data and pseudo-labels at trial $t \in \{1, 2, \dots\}$. The pseudo-label $\tilde{\mathbf{y}}_n^{(t)}$ represents the predicted class labels determined by the trained model at trial $t - 1$. The posterior distribution of each model parameter can be

obtained by repetitively applying Bayes' theorem as follows:

$$p(\boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_t) \propto p(\mathcal{D}_t | \boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c) p(\boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}), \quad (14)$$

$$p(\boldsymbol{\pi} | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_t) \propto p(\mathcal{D}_t | \boldsymbol{\pi}) p(\boldsymbol{\pi} | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}). \quad (15)$$

As in the initial learning, conjugacy results in the following form for each posterior distribution:

$$p(\boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_t) = \mathcal{NW}(\boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c | \hat{\mathbf{m}}_c^{(t)}, \hat{\beta}_c^{(t)}, \hat{\nu}_c^{(t)}, \hat{\mathbf{W}}_c^{(t)}), \quad (16)$$

$$p(\boldsymbol{\pi} | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_t) = \text{Dir}(\boldsymbol{\pi} | \hat{\boldsymbol{\alpha}}^{(t)}). \quad (17)$$

The hyperparameters of the posterior distributions at trial t can be updated by substituting $y_{n,c}$ with $\hat{y}_{n,c}^{(t)}$ and the hyperparameters of the prior distributions with those of the posterior distributions at trial $t-1$ in (9)–(13). Thus, the recursive substitution of the hyperparameters of the prior distributions can update the model with sequentially obtained test data for $t = 1, 2, \dots$, resulting in adaptive motion recognition.

3) *Motion class prediction and assignment of pseudo-labels*: To perform classification on the test data $\mathbf{x}_n^{(t)}$ obtained at trial t , we can compute the following predictive distribution:

$$p(\mathbf{y}_n^{(t)} | \mathbf{x}_n^{(t)}, \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}) = \int p(\mathbf{y}_n^{(t)} | \mathbf{x}_n^{(t)}, \boldsymbol{\theta}_c) p(\boldsymbol{\theta}_c | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}) d\boldsymbol{\theta}_c, \quad (18)$$

where $\boldsymbol{\theta}_c = \{\boldsymbol{\mu}_c, \boldsymbol{\Lambda}_c, \boldsymbol{\pi}_c\}$ represents the set of parameters, and $p(\boldsymbol{\theta}_c | \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1})$ denotes its joint posterior distribution. We determine the classified motion \hat{c}_n as the class for which the class posterior probability, defined by the above predictive distribution, is maximized:

$$\hat{c}_n = \arg \max_c p(y_{n,c}^{(t)} = 1 | \mathbf{x}_n^{(t)}, \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}). \quad (19)$$

In the proposed SS-BSL, the predicted label constructed based on \hat{c}_n is used as the pseudo-label \hat{y}_n . However, since the prediction results may contain errors, assigning all predicted labels as pseudo-labels may lead to inappropriate model updates. To avoid this issue, we compute the confidence for the predictive distribution on each data point obtained at trial t as follows:

$$\text{conf}[\mathbf{y}_n^{(t)}] = \max_c p(y_{n,c}^{(t)} = 1 | \mathbf{x}_n^{(t)}, \mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{t-1}). \quad (20)$$

We assign pseudo-labels for model updates only when $\text{conf}[\mathbf{y}_n^{(t)}] > \theta_{\text{th}}$, where θ_{th} is an arbitrary threshold. This approach excludes ambiguous data with a high likelihood of misclassification from the sequential learning.

III. SIMULATION EXPERIMENTS

A. Methods

To assess the effectiveness of the proposed sequential learning approach in handling data with generating distributions that shift gradually, simulation experiments were conducted on artificial data comprising two classes. Specifically, the data pertaining to classes 1 and 2 were randomly generated from two-dimensional Gaussian distributions denoted as $\mathcal{N}(\mathbf{x} | [-6 + 1.2t, 3]^\top, 3.0\mathbf{I})$ and $\mathcal{N}(\mathbf{x} | [6 - 1.2t, 3]^\top, 3.0\mathbf{I})$, respectively, where $\mathbf{I} \in \{0, 1\}^2$ represents the identity matrix. The mean vector of the generating distribution for each class varies as a function of the progression of trials, with $t = 1, 2, \dots, T$.

The experiments generated a total of 300 data points for each class at every trial t , thereby constructing the dataset \mathcal{D}_t . We set the upper limit of t to $T = 10$. First, we considered \mathcal{D}_1 as the training dataset and performed initial learning of GCM. Subsequently, each of the datasets $\mathcal{D}_2, \dots, \mathcal{D}_{10}$ was treated as an individual test dataset and provided as input to the proposed method to perform class prediction and sequential learning with pseudo-labels.

We evaluated the classification accuracy of the proposed method (GCM with SS-BSL) for each test dataset. In addition, we also evaluated the classification accuracy when GCM was fixed after initial learning with no sequential learning performed. The prior distribution parameters for GCM were set as $\mathbf{m} = \mathbf{0}$, $\beta = 1$, and $\nu = D + 1$. The covariance matrix based on the training dataset was set to \mathbf{W} , and $\boldsymbol{\alpha}$ was initialized with standard normal random numbers. The confidence threshold for sequential learning was set at $\theta_{\text{th}} = 0.9$.

B. Results and Discussion

Fig. 3 shows the scatter plot of the artificial data generated for each t , along with the predictive distributions of each class output by each method. The colored regions in the figure represent the predictive distributions, where the hues of red and blue indicate the probabilities of belonging to classes 1 and 2, respectively, being close to 1.0. In addition, Fig. 4 shows the changes in the classification accuracy for each trial. The proposed method with sequential learning (GCM with SS-BSL) and the method without sequential learning (GCM) are represented by the red and blue lines, respectively.

In Fig. 3, the class 1 data shifts toward the right, and the class 2 data shifts toward the left as the trial progresses. This indicates that the decision boundary derived from initial learning failed to accurately classify the data obtained from later trials. Consequently, the accuracy of GCM without sequential learning noticeably decreased as the trial progressed (Fig. 4). In contrast, the proposed GCM with SS-BSL gradually adapts the decision boundary to the shifting data, resulting in relatively high accuracy even in the later part of the trial. These results suggest that adaptive classification can be accomplished by sequentially updating the classification model using SS-BSL in response to the distribution shift of the test data.

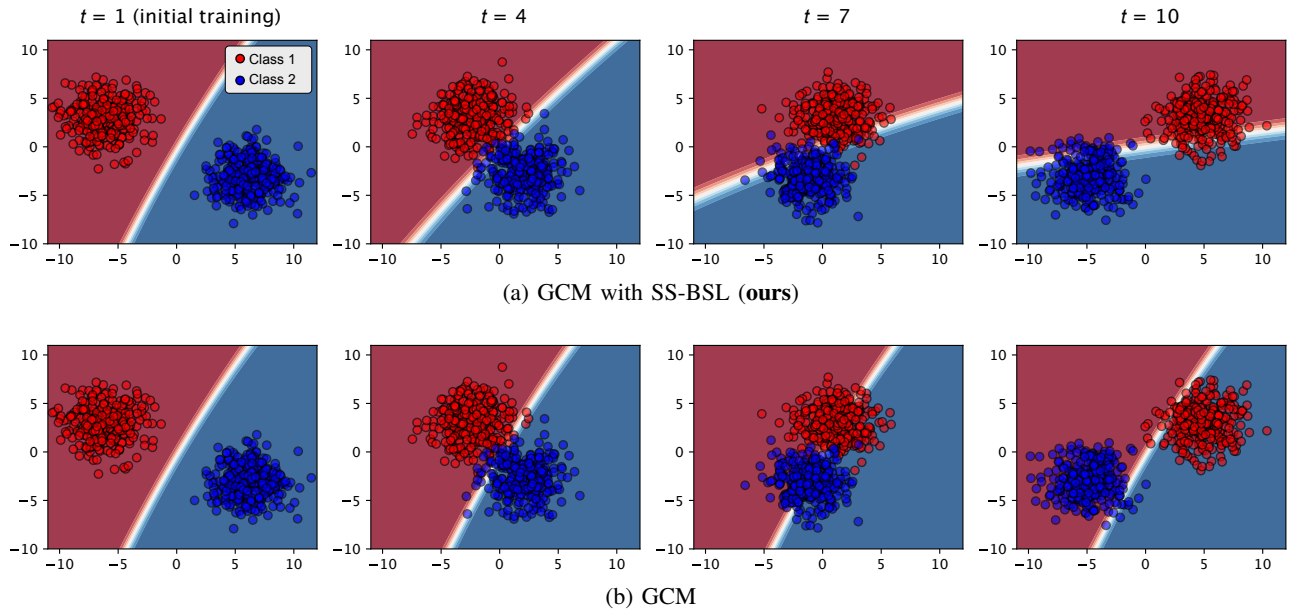


Fig. 3: Snapshots of the sequential learning process for each GCM. The red and blue areas indicate predictive distribution of class labels close to 1.0 for classes 1 and 2, respectively.

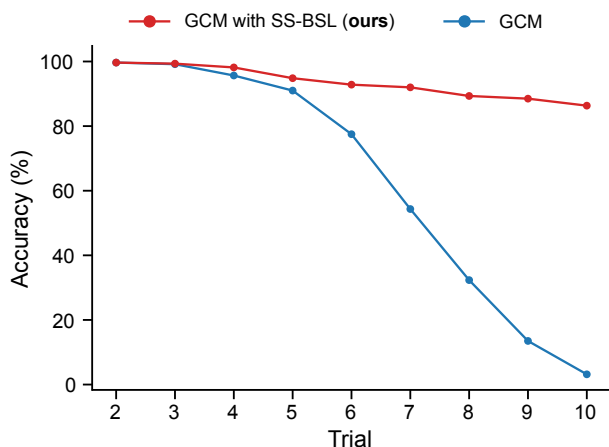


Fig. 4: Per-trial classification accuracy for each method on simulated data

IV. EMG PATTERN CLASSIFICATION EXPERIMENTS

A. Methods

To verify the effectiveness of the proposed method for actual EMG signals, EMG pattern classification experiments were conducted. Six healthy adult males (mean age: 22.6 ± 0.47 years) participated in the experiments. Four electrodes ($D = 4$) were attached at equal intervals around the arm near the elbow on the skin surface of the forearm. We used a wireless measurement system (Delsys, Trigno; sampling frequency: 2000 Hz; bandwidth: 20–450 Hz) to acquire and store EMG signals. During the experiment, the participants were instructed to maintain a seated posture and perform 20 trials of six different motions ($C = 6$): hand grasp, hand open, wrist flexion, wrist extension, supination, and pronation, with the right elbow resting on the desk. Each

motion was measured for 7 s, with rest periods of 5 s between motions.

For feature extraction, full-wave rectification processing and smoothing using a second-order Butterworth low-pass filter with a cutoff frequency of 1 Hz were performed. To exclude the transition state from the resting posture, we removed the first 10% of each data sample and used the remaining portion in the analysis. Of the 20 trials, the first and second trials were treated as training trials, and the remaining trials were treated as test trials. In the analysis, all data from the training trials were used as the training dataset for initial learning of GCM. Subsequently, the data from each test trial were input into the model, and sequential learning was conducted with the class label predictions and pseudo-labels. The parameters of the prior distribution were set the same as in the simulation experiment, and the confidence threshold for sequential learning was set to $\theta_{th} = 0.9$.

We compared the classification accuracy of the proposed method (GCM with SS-BSL) with the following two baselines and two conventional methods.

- **GCM** does not perform sequential learning, fixing the model after initial learning. We consider this baseline as the lower bound of performance for GCM-based methods.
- **GCM with fully supervised BSL (FS-BSL)** performs supervised sequential learning on the GCM using true class labels of test data instead of pseudo-labels. This method is unrealistic as the class labels of test data are unknown in practice. We consider this baseline as the upper bound of performance for GCM-based methods.
- **Adaptive LDA** [14] is a classification method that introduces semi-supervised sequential learning into LDA. Model updating employs a dataset combining training

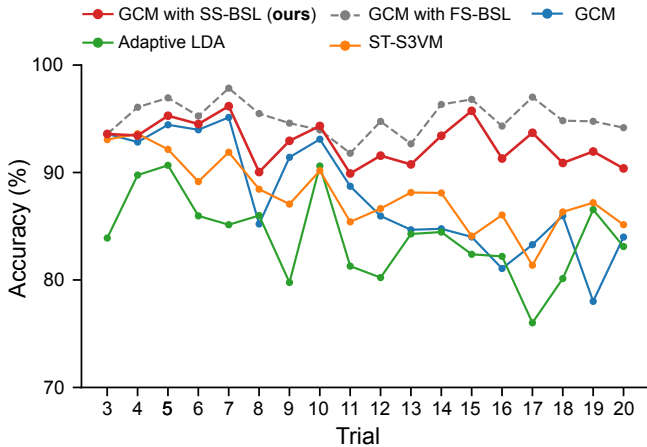


Fig. 5: Averaged per-trial classification accuracy of each method over all participants

and test data. To suppress unexpected changes in parameters, a part of the updating dataset is fixed, and the rest is replaced cyclically. The replacement rate of the updating dataset must be set in advance, and we set it to 50%, which is recommended in [14].

- **Self-training semi-supervised support vector machine (ST-S3VM)** [15] is a classification method that combines self-training (ST) with the semi-supervised SVM (S3VM). Similar to adaptive LDA, this method combines the training dataset with a test dataset for updating the model. To prevent the expansion of the updating dataset with increasing trials, a maximum threshold was established to limit its size, and subsampling was executed on the entire dataset if the threshold was surpassed. In this experiment, the maximum threshold was identical to the sample size of 10 trials. Three hyperparameters in S3VM were optimized by a $7 \times 7 \times 7$ grid search based on 2-fold cross-validation on the training dataset. The search range was set to $C_1 = 2^{-13}, \dots, 2^{-7}$ and $C_2 = 2^{-3}, \dots, 2^3$ for the cost parameters for labeled and unlabeled data, respectively, and $\gamma = 2^{-20}, \dots, 2^{-14}$ for the kernel coefficient for the RBF kernel.

B. Results

Fig. 5 shows the changes in averaged classification accuracy for trial t . The results of GCM with FS-BSL, which is the upper-bound baseline, represent the ideal case where the true labels for test data, which should be unknown, are known. The proposed GCM with SS-BSL outperforms the simple GCM without sequential learning and the conventional methods (i.e., adaptive LDA and ST-S3VM) in all trials.

Fig. 6 summarizes the classification accuracy of each method in a boxplot over trials. The median and mean values for each boxplot are represented by a horizontal line and a triangle mark, respectively. The averaged accuracy for each method is as follows: GCM with SS-BSL: 93.4%, GCM with FS-BSL: 95.5%, GCM: 88.9%, adaptive LDA: 85.9%,

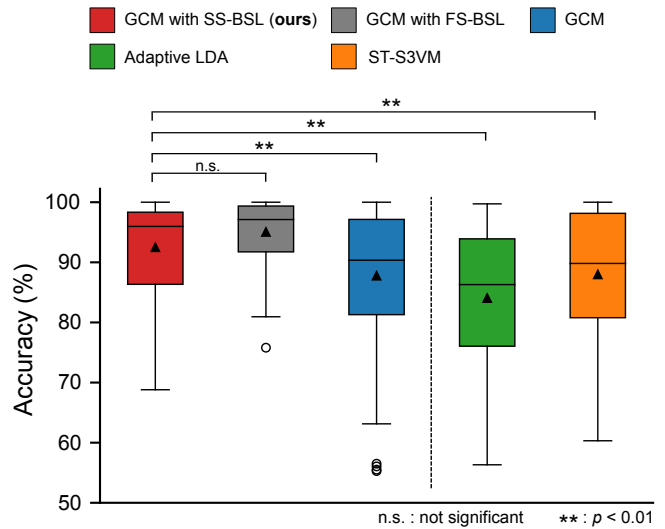


Fig. 6: Boxplot of classification accuracy for each method over trials and participants

and ST-S3VM: 87.9%. The figure also presents the results of pairwise comparison tests (significance level: 1%) with the proposed method as the control group. To account for individual differences, a linear mixed model was introduced where the classification method was set as a fixed effect, and the trials were treated as a random effect. The p -values between groups, calculated based on the estimated marginal means, were adjusted using the Holm method.

C. Discussion

In Fig. 5, the accuracy of GCM without sequential learning decreases as the number of trials increases. This may be attributed to variations in the characteristics of EMG signals caused by muscle fatigue or changes in motor imagery, resulting in a gradually increasing mismatch between the distributions of training and test data. In contrast, the proposed GCM with SS-BSL suppressed the decrease in accuracy and maintained an accuracy of approximately 90% in the later trials. These results suggest that the proposed BSL approach can adaptively recognize motions based on changes in EMG characteristics over time.

The classification accuracies of the conventional methods, adaptive LDA and ST-S3VM, were lower by 8% and 6%, respectively, compared to the proposed method. These methods use all prediction results for model updates, which may decrease accuracy by incorporating uncertain data into the model. In particular, ST-S3VM determines support vectors by considering unlabeled data, causing significant changes in decision boundaries with the test dataset. While this can be effective for long-term datasets with a large distribution shift of test data relative to training data [15], for short-term datasets, such as the one used in this experiment, excessive updating of the decision boundary may occur, causing a decrease in accuracy. Additionally, adaptive LDA sets the replacement rate for updating data to the value recommended in the previous study [14], which may have

been inappropriate for the dataset used in this experiment. In contrast to these conventional methods, the proposed GCM with SS-BSL can exclude uncertain data from the sequential learning process by determining the data used for model updates based on a confidence threshold. Furthermore, the BSL framework stores information from previous trials as prior distributions; thus, there is no need to additionally maintain updating data or adjust the replacement rate. These advantages enable the proposed method to achieve superior performance compared to the conventional methods in the sequential learning of EMG patterns.

Although the accuracy of the proposed SS-BSL is slightly lower than that of FS-BSL, which assumes the ideal scenario where the test labels are known, the discrepancy was not substantial (Fig. 5). Moreover, the statistical test results did not reveal a significant difference between the two methods (Fig 6). This indicates that the pseudo-labeling based on predictive confidence in the proposed method is effective. In practical applications, class labels for test data are unknown, and obtaining labeled calibration data separately from the test data can be challenging and increase the user's burden. Therefore, the proposed method, which does not require labeled data for updating, can maintain good performance over extended periods without causing any additional burden.

V. CONCLUSION

In this paper, we propose an adaptive EMG pattern classification method, GCM with SS-BSL, which utilizes Bayesian model-based sequential learning. The proposed method uses GCM to predict the motion class from EMG signals and estimate its confidence. Furthermore, pseudo-labels are assigned to data with high predictive confidence, and the posterior distributions of the model are sequentially updated by SS-BSL, thereby enabling adaptive motion recognition against changes in signal characteristics.

Simulation experiments demonstrated the applicability of the proposed method under conditions of gradual distribution shifts occurring in test data. EMG classification experiments show that the proposed method outperforms conventional adaptive classification methods while suppressing accuracy degradation in continuous motion trials

In future research, we intend to increase the number of participants and trials and verify the effectiveness of the proposed method under more diverse conditions, such as changing postures during the motion execution.

REFERENCES

- [1] A. Tsoli and O. C. Jenkins, "Robot grasping for prosthetic applications," in *Robotics Research*. Springer, 2010, pp. 1–12.
- [2] I. M. Rezazadeh, M. Firoozabadi, H. Hu, and S. M. R. H. Golpayegani, "Co-adaptive and affective human-machine interface for improving training performances of virtual myoelectric forearm prosthesis," *IEEE Trans. Affect. Comput.*, vol. 3, no. 3, pp. 285–297, 2012.
- [3] A. Fougner, Ø. Stavadahl, P. J. Kyberd, Y. G. Losier, and P. A. Parker, "Control of upper limb prostheses: Terminology and proportional myoelectric control—a review," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 20, no. 5, pp. 663–677, 2012.
- [4] N. Irastorza-Landa *et al.*, "EMG discrete classification towards a myoelectric control of a robotic exoskeleton in motor rehabilitation," in *Converging Clinical and Engineering Research on Neurorehabilitation II*. Springer, 2017, pp. 159–163.
- [5] J. W. Sensinger, B. A. Lock, and T. A. Kuiken, "Adaptive pattern recognition of myoelectric signals: exploration of conceptual framework and practical algorithms," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 17, no. 3, pp. 270–278, 2009.
- [6] H. Zhang, D. Yang, C. Shi, L. Jiang, and H. Liu, "Robust EMG pattern recognition with electrode donning/doffing and multiple confounding factors," in *Proc. Int. Conf. Intel. Robot. Appl.*, 2017, pp. 413–424.
- [7] P. K. Artemiadis and K. J. Kyriakopoulos, "An EMG-based robot control scheme robust to time-varying EMG signal features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 582–588, 2010.
- [8] O. W. Samuel *et al.*, "Resolving the adverse impact of mobility on myoelectric pattern recognition in upper-limb multifunctional prostheses," *Comput. Biol. Med.*, vol. 90, pp. 76–87, 2017.
- [9] J. L. Betthausen *et al.*, "Limb position tolerant pattern recognition for myoelectric prosthesis control with adaptive sparse representations from extreme learning," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 4, pp. 770–778, 2017.
- [10] A. Fougner, E. Scheme, A. D. Chan, K. Englehart, and Ø. Stavadahl, "Resolving the limb position effect in myoelectric pattern recognition," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 19, no. 6, pp. 644–651, 2011.
- [11] M. M.-C. Vidovic *et al.*, "Improving the robustness of myoelectric pattern recognition for upper limb prostheses by covariate shift adaptation," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 24, no. 9, pp. 961–970, 2015.
- [12] X. Zhu, J. Liu, D. Zhang, X. Sheng, and N. Jiang, "Cascaded adaptation framework for fast calibration of myoelectric control," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 25, no. 3, pp. 254–264, 2016.
- [13] M. Karrenbach, P. Preechayasomboon, P. Sauer, D. Boe, and E. Romboakas, "Deep learning and session-specific rapid recalibration for dynamic hand gesture recognition from EMG," *Front. Bioeng. Biotechnol.*, vol. 10, 2022.
- [14] H. Zhang *et al.*, "An adaptation strategy of using LDA classifier for EMG pattern recognition," in *Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 4267–4270.
- [15] Y. Okawa, S. Kanoga, T. Hoshino, and T. Nitta, "Sequential learning on sEMGs in short-and long-term situations via self-training semi-supervised support vector machine," in *Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2022, pp. 3183–3186.
- [16] Q. Ding, X. Zhao, J. Han, C. Bu, and C. Wu, "Adaptive hybrid classifier for myoelectric pattern recognition against the interferences of outlier motion, muscle fatigue, and electrode doffing," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 27, no. 5, pp. 1071–1080, 2019.
- [17] S. Amsüss *et al.*, "Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 4, pp. 1167–1176, 2013.
- [18] Q. Huang *et al.*, "A novel unsupervised adaptive learning method for long-term electromyography EMG pattern recognition," *Sens.*, vol. 17, no. 6, p. 1370, 2017.