Wearable Real-Time Gesture Recognition Scheme Based on A-Mode Ultrasound

Zongxing Lu, *Member, IEEE*, Shaoxiong Cai, Bingxing Chen, Zhoujie Liu, Lin Guo, and Ligang Yao

***Abstract*— A-mode ultrasound has the advantages of high resolution, easy calculation and low cost in predicting dexterous gestures. In order to accelerate the popularization of A-mode ultrasound gesture recognition technology, we designed a human-machine interface that can interact with the user in real-time. Data processing includes Gaussian filtering, feature extraction and PCA dimensionality reduction. The NB, LDA and SVM algorithms were selected to train machine learning models. The whole process was written in C + + to classify gestures in real-time. This paper conducts offline and real-time experiments based on HMI-A (Human-machine interface based on A-mode ultrasound), including ten subjects and ten common gestures. To demonstrate the effectiveness of HMI-A and avoid accidental interference, the offline experiment collected ten rounds of gestures for each subject for ten-fold cross-validation. The results show that the offline recognition accuracy is 96.92% ± 1.92%. The real-time experiment was evaluated by four online performance metrics: action selection time, action completion time, action completion rate and real-time recognition accuracy. The results show that the action completion rate is 96.0% ± 3.6%, and the real-time recognition accuracy is 83.8% ± 6.9%. This study verifies the great potential of wearable A-mode ultrasound technology, and provides a wider range of application scenarios for gesture recognition.**

***Index Terms*— A-mode ultrasound, real-time, gesture recognition, human–machine interface (HMI).**

I. Introduction

UPPER limb defects seriously affect the ability of daily living (ADL) of transradial amputees. Surface electromyography (sEMG) is the mainstream method for prosthetic control in the human-machine interface (HMI). A great deal of effort has been devoted to the study of sEMG, see [1], [2], [3], [4], [5], [6], [7] and references therein. Despite sEMG having been progressively expanded in clinical applications, it still has inherent limitations [8], [9], [10]. First, sEMG has inherent noise, which will cause crosstalk. Secondly, lack of robust hierarchical signal, which makes the spatial resolution low. Moreover, the typical finger-relate muscles are far away from the surface tissue. The solution based on sEMG can not distinguish the depth and spacing of the muscles, so it is difficult to predict the dexterous gesture.

Recently, the ultrasound signal has been proposed as an alternative non-invasive technology to control prostheses due to its high signal-to-noise ratio, direct visualization of targeted tissue, and ability to access deep-seated muscles. Due to the extremely high commercial value of B-mode ultrasound in the clinical field, the research on B-mode ultrasound is earlier than on A-mode ultrasound [11], [12]. Shi *et al.* placed the B-mode ultrasound probe on the biceps brachii and found that there was a linear relationship between the thickness of the muscle and the torque generated by the elbow [13], and there was also a certain relationship between the muscle thickness and fatigue [14]. Moreover, by obtaining the deformation field between different gestures in the B-mode ultrasound image, the recognition accuracy of the five gestures reached 94.05% ± 4.10% [15]. Zhang *et al.* [16] proposed a dual-mode method combining B-mode ultrasound and sEMG to predict the dynamic dorsiflexion of the ankle joint, which can achieve better results than the single-mode method. Akhlaghi *et al.* [17] classified four gestures in real-time, and controlled the virtual hand using the HMI based on B-mode ultrasound.

Although the above studies have illustrated that B-mode ultrasound is a reliable substitute for sEMG in HMI for prosthetic control, B-mode ultrasound devices are bulky and difficult to apply to wearable prosthetic devices. Compared with B-mode ultrasound, the A-mode ultrasound transducer is cheaper, smaller, and more wear-resistant. Finger-related muscles are distributed around the forearm, and with the distribution of multi-channel transducers, A-mode ultrasound can detect the movement of these muscles. In addition, the echo signal of A-mode ultrasound is one-dimensional signal, which helps to shorten the calculation delay of HMI-A in real-time recognition.

Due to the advantages of A-mode transducer in prosthetic limb control, new transducers have emerged in recent years. Sun *et al.* [18] designed a dual-frequency ultrasound transducer with a ring array arrangement. The low frequency is responsible for deep detection and the high frequency is responsible for shallow detection. Compared with the single frequency transducer, the accuracy is improved by 7.3%. Yan *et al.* [19] developed a lightweight transducer made of polyvinylidene fluoride (PVDF) material, which can ensure

Manuscript received 14 June 2022; revised 26 July 2022; accepted 5 September 2022. Date of publication 8 September 2022; date of current version 19 September 2022. This work was supported by the National Natural Science Foundation of China under Grant 61801122. *(Zongxing Lu and Shaoxiong Cai contributed equally to this work.) (Corresponding authors: Bingxing Chen; Zhoujie Liu.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the School of Mechanical Engineering and Automation of Fuzhou University.

Zongxing Lu, Shaoxiong Cai, Bingxing Chen, Lin Guo, and Ligang Yao are with the School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, China (e-mail: luzongxing@fzu.edu.cn; qq442962182@gmail.com; bingxingchen@fzu.edu.cn; guolin_fzu@sina.com; ylgyao@fzu.edu.cn).

Zhoujie Liu is with the Department of Pharmacy, The First Affiliated Hospital of Fujian Medical University, Fuzhou 350005, China (e-mail: liuzhoujie_08@163.com).

Digital Object Identifier 10.1109/TNSRE.2022.3205026

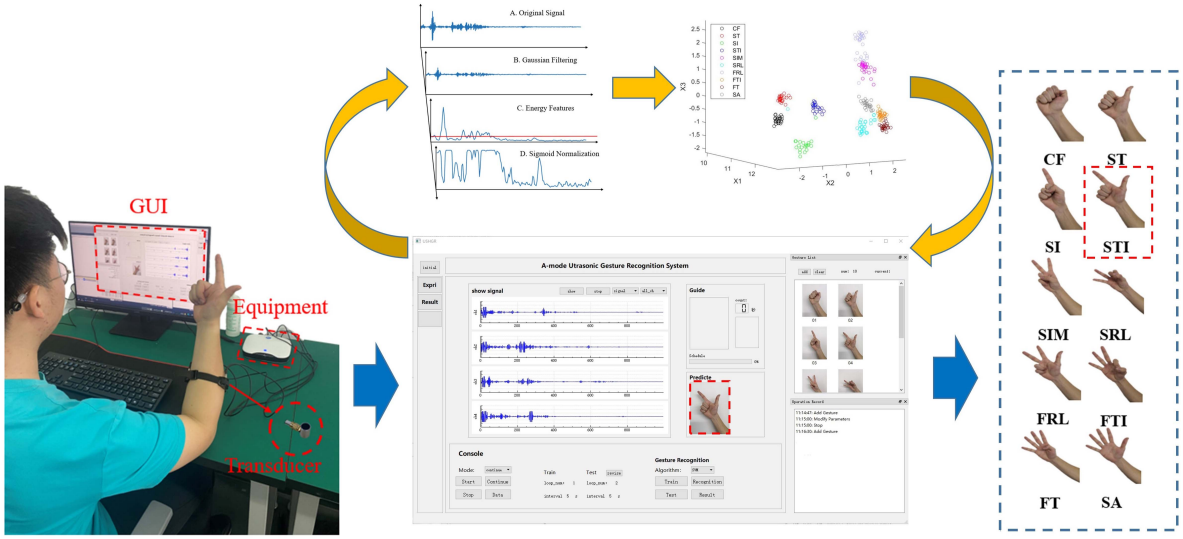

Fig. 1. The framework of real-time gesture recognition scheme based on wearable A-mode ultrasound.

lightweight and obtained a classification accuracy of 97.6%. Xia *et al.* [10] used bushing to integrate the A-mode ultrasound and the sEMG. Compared with the single-mode method, the accuracy was improved by 4.9%.

In addition, there are many practical studies on A-mode ultrasound [20], [21]. Guo *et al.* [22] used the A-mode ultrasound to obtain muscle deformation, and deduced the linear relationship between muscle deformation and wrist extension angle. Hettiarachchi *et al.* [23] increased the number of transducers to eight to recognize the finger movement of transradial amputees. Li *et al.* [24] proved that the average recognition accuracy using the A-mode ultrasound can reach 96% for six movements, including five fingers in bending and static state. In addition, there were studies on force recognition [25], wrist recognition [26], arm position [27] based on A-mode ultrasound.

Real-time performance is the premise for the practical application of wearable A-mode ultrasound. It is noted that most of the previous work related to A-mode ultrasound was analyzed offline, which can verify the feasibility of the algorithm. Real-time gesture recognition can be used in scenarios such as prosthetic control, presentations, and smart home appliances. These scenarios have higher requirements for real-time performance. Excessive response time will affect the interaction between the user and the HMI-A, resulting in delay or misjudgment.

In this paper, a gesture recognition scheme based on wearable A-mode ultrasound is proposed as shown in Fig. 1. The excellent recognition ability of this scheme is verified in the offline experiment. Then its real-time performance is further verified by a customized graphical user interface (GUI). The main contributions of this paper are as follows:

1) Realized real-time gesture recognition based on wearable A-mode ultrasound, providing a basis for practical use.

2) Real-time experiment was evaluated by four online performance metrics and achieved high performance.

3) Compared different features in three-dimensional map and introduced relative offset rate for quantitative analysis.

The rest of this paper is organized as follows: firstly, Section 2 introduces the A-mode ultrasound, experiment setup, data processing and statistical analysis. Section 3 describes the offline experiment and real-time experiment. Experimental results and extended discussion are presented in Section 4. Finally, Section 5 concludes the paper.

II. Method

A. A-Mode Ultrasound

A-mode ultrasound is a technology that uses the amplitude of pulse-echo to display. It is the earliest and longest-developed ultrasound technology. When the ultrasound signal propagates through various tissues of the human body, if it encounters a tissue mutation interface, the echo signal will be reflected accordingly due to the difference in acoustic impedance. Therefore, the echo signal reflects the structure of human tissue to a certain extent. When the fingers move, the ligaments are first pulled by muscle contractions, and then the ligaments pull the corresponding muscles to complete the movement. As an one-dimensional acoustic signal, the A-mode ultrasound can reflect the position and amplitude intensity from echo signal. Therefore, the A-mode ultrasound echo signal on the human forearm muscle can reflect the interface of the muscle-bone and muscle deformation. Muscle deformation under different gestures will lead to different ultrasound signals. By analyzing the echo signal of muscles, different gestures can be decoded, which can be used to build HMI.

As shown in Fig. 1, a 4-channel A-mode ultrasound signal acquisition instrument designed and manufactured by ELONXI Company was chosen, which was used to drive 4 A-mode ultrasound transducers and receive the echo signal. The transducers have a detection depth of 39mm, which can be used to detect the movement of deep muscles and tendons in the forearm. The repetition frequency was equal to the real-time recognition frequency and set as 10 Hz, which means it grabs 10 frames of data per second from the acquisition instrument. The number of sampling dots of each channel is

1000, and there are 4 channels in total. Therefore, each frame contains 4000 sampling dots, forming a matrix of 4 × 1000.

B. Experiment Setup

Ten common era subjects (eight males and two females; 23 ± 2 years, denoted as S1 to S10) without a history of neuromuscular and joint diseases participated in the experiment. Before participating, all subjects were informed about the experiment and provided informed consent. The testing procedure followed the Declaration of Helsinki and was approved by the Institutional Review Board of Fuzhou University.

As shown in Fig. 1, the subjects sat naturally with their elbows resting on the table and kept relaxed. Four A-mode transducers were attached to the forearm to get muscle information by a customized arm strap. Then, standard ultrasound couplant was applied between the skin and the transducers. To evaluate the performance of the A-mode ultrasound for gesture recognition, we designed 10 different gestures. As shown in Fig. 1, including clenching fist(CF), stretching thumb (ST), stretching index finger (SI), stretching thumb and index finger(STI), stretching index finger and middle finger(SIM), stretching ring finger and little finger (SRL), flexing ring finger and little finger(FRL), flexing thumb and index finger(FTI), flexing thumb (FT) and stretching all fingers(SA). Most of these gestures are commonly used in ADL [28].

Signal acquisition was carried out with a customized GUI as shown in Fig. 1. Each subject completed ten gestures in turn, which was regarded as one round of the experiment, and each gesture lasted 5 seconds. Since the subjects need a certain reaction time, the initial and final seconds were deducted, while the data of the middle 3 seconds were retained. Each time the subjects completed one round of data collection, they were given a 60-second break to reduce the weariness and discomfort associated with the experiment.

C. Data Processing

The first 20 sampling dots and the last 20 sampling dots for each echo signal were removed before filtering due to they were usually unstable. As the signal processing flow shown in Fig. 1, the first line is the raw signal, the ordinate is the echo amplitude, and the abscissa is the echo time. The second line is the Gaussian filtered signal. The third line is the feature value extracted after segmentation, the ordinate represents the value of the feature, and the abscissa represents the dimension of the feature. The fourth line is the normalized features. The processing method of the ultrasound signal is similar to our previous work [29]. It consisted of Gaussian filtering, feature extraction, and feature dimensionality reduction.

1) Gaussian Filtering: Considering the inherent noise and interference signal of ultrasound equipment. In the preprocessing stage of the original data, Gaussian filtering is chosen. As a zero phase shift filtering method with the smallest time-frequency window area, Gaussian filtering is widely used in the field of ultrasound signals. The ultrasound signal was filtered by the one-dimensional Gaussian function, as shown below:

$$G(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(f-\mu)^2}{2\sigma^2}} \tag{1}$$

where f denotes frequency, μ is the expectation, and σ is the standard deviation of the Gaussian function.

2) Feature Extraction: We extracted the ultrasound signal features with fixed window length and step size to obtain representative information. The window length was set as 10 sampling dots and step size as 5, then calculated the features Mean, Var, and ES of each window respectively. Suppose N is the length of window size, i is the ith sampling dot, $X_1\bar{x}$ is the average value of the window.

Denote Mean as the centralized trend of data, as follows:

$$Mean = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

The Var is used to measure the dispersion of data and is defined as follows:

$$Var = \frac{1}{N}\sum_{i=1}^{N} (x_i - \bar{x}) \tag{3}$$

The ES feature is expressed as the Energy feature normalized by sigmoid function, and the Energy feature is defined as:

$$Energy = \left(\sum_{i=1}^{N} x_i^2\right)^{\frac{1}{2}} \tag{4}$$

The sigmoid function is defined as:

$$S(x) = \left(1 + e^{-(x_i - b)}\right)^{-1} \tag{5}$$

where, i represents the ith feature, according to experience, we set threshold $b = 3$.

3) PCA Dimensionality Reduction: As a commonly used data dimensionality reduction algorithm, principal component analysis (PCA) transforms closely related variables into as few new variables as possible, so that these new variables are uncorrelated in pairs. Fewer comprehensive indicators represent various types of information respectively to achieve the purpose of data dimensionality reduction. Considering the computational cost, we used PCA to reduce the feature dimension to 40, and the cumulative contribution rate of principal components after dimensionality reduction is greater than 99%. Additionally, we used the top three principal components to draw a three-dimensional map of the features to compare different methods of feature reduction more intuitively.

D. Statistical Analysis

Average offline recognition accuracy and real-time performance metrics are reported with mean and standard deviation (SD). One-way repeated measure analysis of variance (ANOVA) and Mann-Whitney U test were used to compare different features (i.e., Mean, Var and ES) and different classifiers (i.e., LDA, SVM and NB). If the analysis data follows a normal distribution, one-way ANOVA was used. Otherwise, the Mann-Whitney U test was used. Tukey's honestly significant difference tests were used to assess whether there are significant differences between different groups. The statistical significance level of all comparisons was set to $p = 0.05$.

III. Experiment

A. Offline Experiment

1) Offline Test: For the offline experiment, we collected a total of 10 rounds of data, each containing 10 gestures.

Nine rounds of data are used as the training set, and the remaining round is used as the test set. In order to demonstrate the effectiveness of the offline experiment and avoid accidental interference, the experiment carried out ten-fold cross-validation.

The Linear Discriminant Analysis (LDA) classifier, Support Vector Machine (SVM) classifier and Naive Bayes (NB) classifier are effective in small sample experiment [29]. Therefore, we selected the LDA, SVM and NB classifiers as machine learning algorithms. The recognition accuracy (RA) was used to evaluate, which is defined as follows:

$$RA = \frac{Number\ of\ correctly\ recognized\ gestures}{Total\ number\ of\ testing\ gestures} \times 100\% \tag{6}$$

2) Relative Offset Rate: To compare the influence of different features on gesture recognition, we studied the feature space of different feature methods, and introduced relative offset rate (OR) to quantitatively analyzed the recognition effect. Suppose that the class center coordinates of two groups of data are $X_1\{x_1^1, x_1^2, \ldots, x_1^k\}$ and $X_2\{x_2^1, x_2^2, \ldots, x_2^k\}$ respectively, and the two groups of data come from the same gesture. Where k represent the dimension of the feature. Then, the offset distance (OD) of the two class center coordinates can be defined as:

$$OD = Euclidean(X_1, X_1) = \sqrt{\sum_{i=1}^{k} (x_1^i - x_2^i)^2} \tag{7}$$

Since different features differ in the order of magnitude, we used relative OR for comparison, which can be defined as:

$$OR = \frac{OD}{\sum_{i=1}^{n} \sum_{j=1}^{n} Euclidean(X^i, X^j)/C_n^2}, i \neq j \tag{8}$$

where n represents the number of gestures.

B. Real-Time Experiment

1) Real-Time Test: For the real-time experiment, we used the same data processing procedure for the offline test. In the offline experiment, the effect of the NB algorithm is slightly worse than that of the LDA and SVM. Therefore, only the LDA and SVM algorithms are used in the real-time experiment. To demonstrate the utility of the HMI-A, we tried to use very little training. For LDA and SVM algorithms, a round of gesture training is conducted respectively, and two rounds are repeated to evaluate the experimental results.

2) Performance Metrics: To evaluate the online performance of the HMI-A for gesture recognition, four metrics were chosen, including motion selection time (ST), motion completion time (CT), motion completion rate (CR), and real-time recognition accuracy (RA) [30].

Fig. 2 shows a detailed explanation of ST and CT. ST is the time from the start of the test to the first successful gesture recognition, which reflects the reaction speed of the recognition system. If a certain gesture is not recognized successfully within 5 seconds, it is considered that the recognition fails, and its ST is not included in the total statistics. CT is the time consumed to predict the current gesture rightly 10 times.

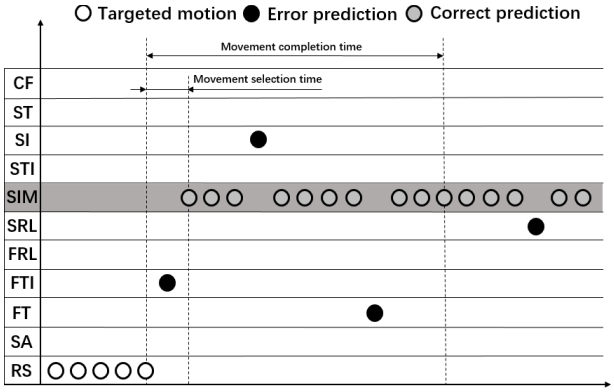

Fig. 2. Illustration of real-time performance metrics.

If the gesture is not successfully recognized 10 times within 5 seconds, the gesture is regarded as incomplete. Since the real-time recognition frequency is 10Hz, CT is greater than one second, and the closer it is to one second, the better the recognition effect.

CR is the percentage of actions completed within 5 seconds, which reflects the usability of the HMI to the user. If a gesture is successfully recognized 10 times within 5 seconds, the action is considered complete. RA is the recognition accuracy from the first correct prediction to the end of the 5 seconds prediction time, which can represent the prediction stability. Only when the current motion is marked as recognition completion is the ST, CT, and RA counted.

IV. Results and Discussion

A. Offline Analysis

After PCA dimensionality reduction, the top three principal components are retained to draw gesture scatter plots obtained by different feature extraction methods. As shown in Fig. 3, X1, X2 and X3 represent the values of the 1st, 2nd, and 3rd dimension principal components, respectively. Ten different colored circles correspond to ten different gestures. Circles of the same color are more concentrated in the ES method, and clusters of circles of different colors are more spread out. In contrast, the Var and Mean methods have lower aggregation degree of circles of the same color, and the cluster distance of circles of different colors is smaller, resulting in lower discrimination of the SRL, FTI and FT gestures. The ES method can make the same gesture more focused, while different gesture clusters have greater distance between them. Therefore, in the analysis of the feature space, the ES method is the best.

The relative OR can quantitatively analyze the recognition effect of different feature methods. Table I shows that the relative OR of the ES method is significantly lower than that of the Mean and Var methods. The lower the relative OR, the higher the degree of intra class aggregation, which is consistent with what we observe in the feature space. Table II lists the recognition accuracy obtained by the three feature methods using the three classification algorithms. Regardless of the algorithm, the recognition accuracy obtained by the ES method is higher than that of the Mean and Var methods, which further verifies the above statement.

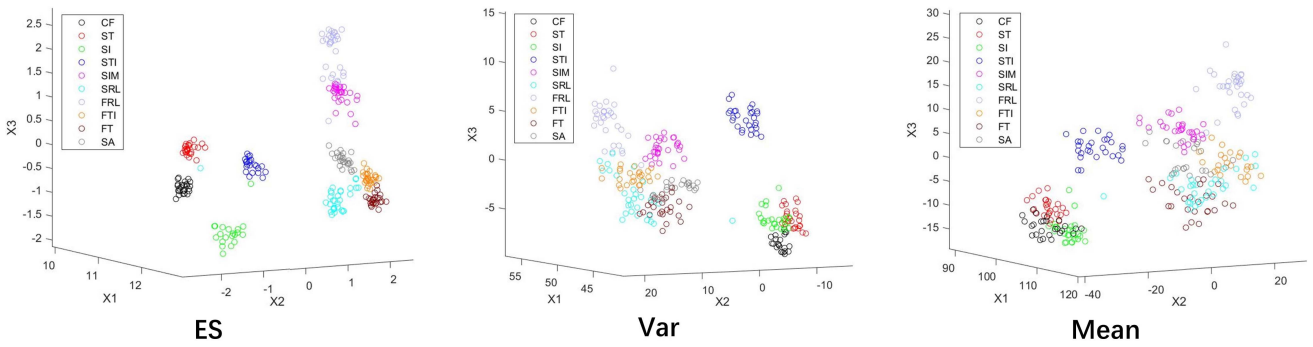

Fig. 3. Scatter plots of the top three principal components after PCA with different feature methods.

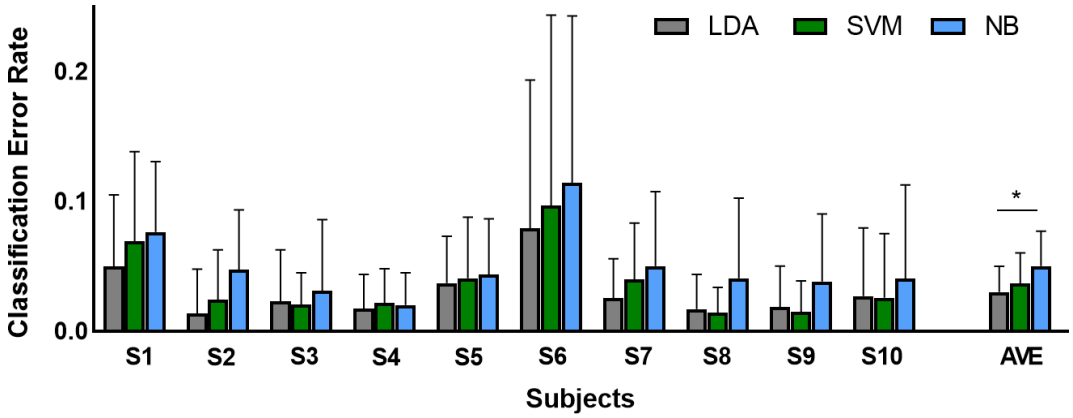

Fig. 4. Offline gesture recognition error rate of ten different subjects.

TABLE I
RELATIVE OFFSET RATE OF DIFFERENT FEATURE METHODS

Feature	OR	SD
ES	0.6396	0.156
Mean	0.8564	0.218
Var	0.8372	0.228

TABLE II
GESTURE RECOGNITION ACCURACY BASED ON COMBINATIONS OF DIFFERENT FEATURE METHODS AND ALGORITHMS

Feature	Algorithm	RA	SD
ES	LDA	0.9692	0.0192
	SVM	0.9632	0.0223
	NB	0.9498	0.0253
Mean	LDA	0.8971	0.0443
	SVM	0.8843	0.0490
	NB	0.8427	0.0629
Var	LDA	0.8980	0.0441
	SVM	0.8887	0.0439
	NB	0.8551	0.0565

Fig. 4 shows the gesture recognition error rate for ten different subjects using the ES method and three classification algorithms in offline experiment, respectively. Compared with the average value of recognition error rate, the NB algorithm has the highest error rate for gesture recognition, and ANOVA shows that the LDA algorithm is significantly better than the NB algorithm ($p < 0.05$). Fig.5 compares the ability of three classification algorithms to classify ten different gestures. The recognition performance of the LDA algorithm is the best, but the recognition accuracy of CF and SRL gestures is lower than that of the SVM algorithm. The recognition accuracy of almost all gestures, the NB algorithm is lower than that of the LDA and SVM algorithms. Therefore, we only used the LDA and SVM algorithms in the real-time experiment.

B. Online Performance

The online performance metrics of the HMI-A for gesture recognition are summarized in Table III and Table IV. Since it is a real-time experiment, the data of the LDA and SVM come from two experiments. The ST includes the subject's reaction time and the time for the software to recognize the correct gesture. The former depends on the response of the subjects, the ST of S1 and S9 are both within 0.18 seconds, while the slowest subject is 0.81 seconds. The CT can be considered as ST plus the time of ten successful predictions. Since the real-time recognition frequency is 10 Hz, the CT should be one second longer than the ST. The results show that the mean difference between CT and ST is within 1.1 seconds, indicating that the entire recognition result is basically correct during the completion of the gesture.

In practical applications, the CR is the most important performance metric, which reflects whether the user's expected

TABLE III
PERFORMANCE METRICS FOR TEN SUBJECTS BASED ON LDA

Metrics	Subjects										Mean±SD
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	
ST (s)	0.173	0.454	0.429	0.786	0.706	0.528	0.816	0.471	0.141	0.541	0.505±0.217
CT (s)	1.369	1.559	1.518	2.027	1.849	1.515	1.747	1.459	1.108	1.529	1.568±0.243
CR	1.000	0.900	1.000	0.950	1.000	0.950	0.950	0.850	0.950	1.000	0.955±0.047
RA	0.880	0.726	0.883	0.763	0.808	0.838	0.818	0.758	0.888	0.883	0.825±0.057

TABLE IV
PERFORMANCE METRICS FOR TEN SUBJECTS BASED ON SVM

Metrics	Subjects										Mean±SD
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	
ST (s)	0.166	0.672	0.495	0.549	0.502	0.648	0.610	0.700	0.122	0.207	0.467±0.199
CT (s)	1.282	2.169	1.541	1.624	1.488	1.702	1.603	1.711	1.110	1.388	1.562±0.258
CR	1.000	0.950	0.950	0.900	1.000	0.900	1.000	0.950	1.000	0.950	0.960±0.036
RA	0.918	0.719	0.807	0.786	0.910	0.745	0.884	0.814	0.943	0.849	0.838±0.069

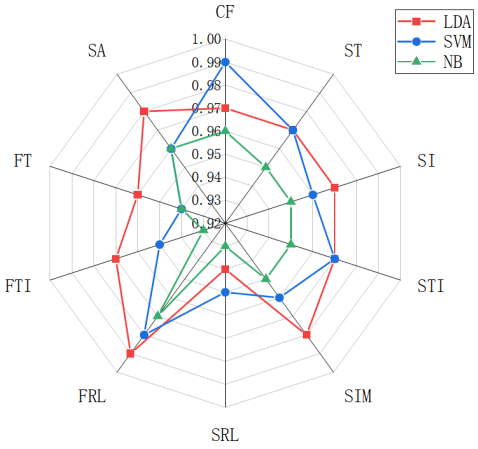

Fig. 5. Average recognition accuracy for ten different gestures.

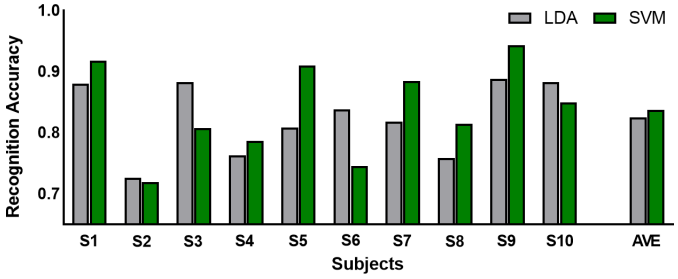

Fig. 6. Real-time gesture recognition accuracy of ten different subjects.

movement can be achieved. The average CR of the LDA and SVM algorithms is more than 95%, which means that the HMI-A can basically complete the instructions issued by the user. Fig. 6 lists the gesture recognition accuracy of ten different subjects in the real-time experiment, which

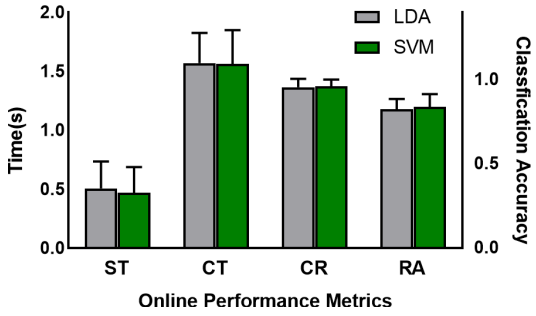

Fig. 7. Comparison of online performance metrics for LDA and SVM.

shows that both the LDA and SVM algorithms can get good recognition accuracy. The average recognition accuracy of the LDA algorithm is 82.5% ± 5.7%, and the SVM algorithm is 83.8% ± 6.9%. In fact, since the recognition results of the entire experimental process are considered, many wrong gestures are included in the process of switching gestures, so the actual recognition accuracy is higher than the experimental results. Fig. 7 compares the online performance metrics of the LDA and SVM algorithms. ANOVA shows that there is no significant difference between the performance of the LDA and SVM algorithms, which means that our gesture recognition scheme can achieve favorable results with simple machine learning algorithms.

V. CONCLUSION

In this paper, we proposed a real-time gesture recognition scheme for HMI based on wearable A-mode ultrasound. The experiments were performed on our customized GUI, and the signal was acquired by four A-mode ultrasound transducers. Data processing includes filtering, feature extraction and

dimensionality reduction. Based on this scheme, we have done offline and real-time experiments respectively. In the offline experiment, the feature three-dimensional map and relative OR were used to filter the features, and the ES feature with the best effect is selected. Then different classification algorithms are compared according to the recognition error rate. The effect of the NB algorithm is general, so the LDA and SVM algorithms are retained and applied to the next real-time experiment.

The real-time experiment was evaluated by four online performance metrics. ST is mainly determined by the user's reaction speed. CT and CR indicate that after the action selection is completed, the recognition of gestures is basically correct, and the completion rate is very high. RA counts the recognition accuracy throughout the experiment. Our scheme shows extremely high real-time performance and can be used in practical scenarios. This means the feasibility of HMI-A to replace traditional sEMG-based HMI, especially for the recognition of dexterous finger movements.

However, HMI-A also has some limitations due to its own signal characteristics. The A-mode ultrasound can only detect depth information in one dimension, so its anti-interference ability is weaker than that of the B-mode ultrasound. The offset of the transducers and different arm postures lead to a decrease in the recognition accuracy. In the future, we hope to work on HMI-A robustness research. We will obtain a recognition model containing various situations by deep learning, or propose algorithms for transducers offset or arm posture changes.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Z. Qing, Z. Lu, Y. Cai, and J. Wang, "Elements influencing sEMG-based gesture decoding: Muscle fatigue, forearm angle and acquisition time," *Sensors*, vol. 21, no. 22, p. 7713, Nov. 2021.

[2] M. Simáo, N. Mendes, O. Gibaru, and P. Neto, "A review on electromyography decoding and pattern recognition for human-machine interaction," *IEEE Access*, vol. 7, pp. 39564–39582, 2019.

[3] K. Koirala, M. Dasog, P. Liu, and E. A. Clancy, "Using the electromyogram to anticipate torques about the elbow," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 3, pp. 396–402, May 2015.

[4] L. Liu, P. Liu, E. A. Clancy, E. Scheme, and K. B. Englehart, "Electromyogram whitening for improved classification accuracy in upper limb prosthesis control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 5, pp. 767–774, Sep. 2013.

[5] L. Guo, Z. Lu, and L. Yao, "Human-machine interaction sensing technology based on hand gesture recognition: A review," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 4, pp. 300–309, Aug. 2021.

[6] A. Dwivedi, Y. Kwon, A. McDaid, and M. Liarokapis, "A learning scheme for EMG based decoding of dexterous, in-hand manipulation motions," *IEEE Trans. Neural Syst. Rehabil. Eng., Eng. Med. Biol. Soc.*, vol. 27, no. 10, pp. 2205–2215, Aug. 2019.

[7] L. Pan, K. Liu, and J. Li, "Effect of subcutaneous muscle displacement of flexor carpi radialis on surface electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1244–1251, 2022.

[8] X. Yang, J. Yan, and H. Liu, "Comparative analysis of wearable A-mode ultrasound and sEMG for muscle-computer interface," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 9, pp. 2434–2442, Sep. 2020.

[9] J. He, H. Luo, J. Jia, J. T. W. Yeow, and N. Jiang, "Wrist and finger gesture recognition with single-element ultrasound signals: A comparison with single-channel surface electromyogram," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1277–1284, May 2019.

[10] W. Xia, Y. Zhou, X. Yang, K. He, and H. Liu, "Toward portable hybrid surface electromyography/A-mode ultrasound sensing for human–machine interface," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5219–5228, Jul. 2019.

[11] S. Sikdar *et al.*, "Novel method for predicting dexterous individual finger movements by imaging muscle activity using a wearable ultrasonic system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 69–76, Jan. 2014.

[12] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 6, pp. 788–797, Nov. 2012.

[13] J. Shi, Y. P. Zheng, Q. H. Huang, and X. Chen, "Continuous monitoring of sonomyography, electromyography and torque generated by normal upper arm muscles during isometric contraction: Sonomyography assessment for arm muscles," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1191–1198, Mar. 2008.

[14] J. Shi, Y.-P. Zheng, X. Chen, and Q.-H. Huang, "Assessment of muscle fatigue using sonomyography: Muscle thickness change detected from ultrasound images," *Med. Eng. Phys.*, vol. 29, no. 4, pp. 472–479, 2007.

[15] J. Shi, J. Y. Guo, S. X. Hu, and Y. P. Zheng, "Recognition of finger flexion motion from ultrasound image: A feasibility study," *Ultrasound Med. Biol.*, vol. 38, no. 10, pp. 1695–1704, 2012.

[16] Q. Zhang, A. Iyer, Z. Sun, K. Kim, and N. Sharma, "A dual-modal approach using electromyography and sonomyography improves prediction of dynamic ankle movement: A case study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1944–1954, 2021.

[17] C. A. Baker, N. Akhlaghi, H. Rangwala, J. Kosecka, and S. Sikdar, "Real-time, ultrasound-based control of a virtual hand by a trans-radial amputee," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 3219–3222.

[18] X. Sun, X. Yang, X. Zhu, and H. Liu, "Dual-frequency ultrasound transducers for the detection of morphological changes of deep-layered muscles," *IEEE Sensors J.*, vol. 18, no. 4, pp. 1373–1383, Feb. 2018.

[19] J. Yan, X. Yang, X. Sun, Z. Chen, and H. Liu, "A lightweight ultrasound probe for wearable human–machine interfaces," *IEEE Sensors J.*, vol. 19, no. 14, pp. 5895–5903, Jul. 2019.

[20] X. Yang, Z. Chen, N. Hettiarachchi, J. Yan, and H. Liu, "A wearable ultrasound system for sensing muscular morphological deformations," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 6, pp. 3370–3379, Jun. 2021.

[21] X. Yang, J. Yan, Z. Chen, H. Ding, and H. Liu, "A proportional pattern recognition control scheme for wearable A-mode ultrasound sensing," *IEEE Trans. Ind. Electron.*, vol. 67, no. 1, pp. 800–808, Jan. 2020.

[22] J.-Y. Guo, "Dynamic monitoring of forearm muscles using one-dimensional sonomyography system," *J. Rehabil. Res. Develop.*, vol. 45, no. 1, pp. 187–196, Dec. 2008.

[23] N. Hettiarachchi, Z. Ju, and H. Liu, "A new wearable ultrasound muscle activity sensing system for dexterous prosthetic control," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 1415–1420.

[24] Y. Li, K. He, X. Sun, and H. Liu, "Human-machine interface based on multi-channel single-element ultrasound transducers: A preliminary study," in *Proc. IEEE 18th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Sep. 2016, pp. 1–6.

[25] X. Yang, Y. Li, Y. Fang, and H. Liu, "A preliminary study on the relationship between grip force and muscle thickness," in *Proc. 8th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, May 2017, pp. 118–121.

[26] X. Yang, J. Yan, Y. Fang, D. Zhou, and H. Liu, "Simultaneous prediction of wrist/hand motion via wearable ultrasound sensing," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 4, pp. 970–977, Apr. 2020.

[27] S. Cai, Z. Lu, L. Guo, Z. Qing, and L. Yao, "The LET procedure for gesture recognition with multiple forearm angles," *IEEE Sensors J.*, vol. 22, no. 13, pp. 13226–13233, Jul. 2022.

[28] B. Peerdeman *et al.*, "Myoelectric forearm prostheses: State of the art from a user-centered perspective," *J. Rehabil. Res. Develop.*, vol. 48, no. 6, p. 719, 2011.

[29] L. Guo, Z. Lu, L. Yao, and S. Cai, "A gesture recognition strategy based on A-mode ultrasound for identifying known and unknown gestures," *IEEE Sensors J.*, vol. 22, no. 11, pp. 10730–10739, Jun. 2022.

[30] W. C. Guo, X. Sheng, H. Liu, and X. Zhu, "Toward an enhanced human-machine interface for upper-limb prosthesis control with combined EMG and NIRS signals," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 4, pp. 564–575, Aug. 2017.