

Enhancing Prosthetic Control with Ultrasound Images: A Convolutional Neural Network Approach for Hand Gesture Recognition^{*}

Yun Chen^{*} Xuefeng Bao^{**} Hongsheng He^{***} Qiang Zhang^{*}

^{*} Department of Mechanical Engineering, the University of Alabama, Tuscaloosa, AL 35487, USA (e-mail: qiang.zhang@ua.edu).

^{**} Department of Biomedical Engineering, the University of Wisconsin-Milwaukee, Milwaukee, WI 53045, USA.

^{***} Department of Computer Science, the University of Alabama, Tuscaloosa, AL 35487, USA.

Abstract: Human hand gesture recognition using biological signals from the forearm is an increasingly significant area of research, with implications across various fields such as prosthetic development, rehabilitation, and human-machine interaction. However, traditional hand gesture recognition with surface electromyography (sEMG) technique has some challenges, including cross-talk from neighboring muscles, low signal-to-noise ratio, and inability to measure deep muscles. In the current study, we proposed to use brightness mode (B-mode) ultrasound images from the muscles of the forearm anterior side as an alternative neuromuscular interface to recognize hand gestures. We designed a convolutional neural network (CNN) classifier to build the personalized mapping from static ultrasound images to eight different hand gestures. To evaluate the performance of the proposed CNN classifier, an ultrasound images dataset and labeled gestures from four young healthy participants were collected and analyzed. Results from offline intra-subject (personalization) validation, quasi-real-time validation, and real-time validation showed high classification accuracy of 99.65%, 97.47%, and 90.83%, respectively. In addition, real-time hand gesture recognition could be executed within 50 ms per image frame by using the proposed CNN classifier. Our findings demonstrated promising real-time hand gesture recognition with high accuracy by using B-mode ultrasound images and the proposed CNN classifier for prosthetic hand control.

Copyright © 2024 The Authors. This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Biomedical and medical image processing and systems, Rehabilitation engineering and healthcare delivery, Bio-signals analysis and interpretation, Man-machine interfaces, Bioinformatics, Estimation and control in biological systems.

1. INTRODUCTION

Human hand gestures are critical to daily functioning and overall quality of life. It refers to the ability to make coordinated, precise movements with the hands and fingers, enabling a vast range of activities from the mundane to the complex. For individuals with upper-limb disabilities, particularly those who may have lost the use of their hands or arms, hand gesture recognition can facilitate interaction with various devices and can be integrated into prosthetics to allow users to control these devices with intuitive movements, mimicking natural hand gestures. Globally, the number of people who suffer from the inconvenience of living with hand or forearm amputations is increasing, which poses great requirements for effective and efficient hand gesture recognition technologies. In recent years, the rapid development in biological signals-based gesture recognition methodologies has provided new possibilities

for function improvements and operational freedom of prosthetic hands (Parajuli et al. (2019)).

Gesture recognition plays a crucial role in prosthetic hand control. The prosthetic hand enables more precise and natural movement by accurately classifying and interpreting an amputee's gestures. Recently, researchers have adopted various sensors and methods to achieve gesture recognition. The clinical application of surface electromyography (sEMG) signals has been extensively studied (Drost et al. (2006)). It has been used to classify wrist movements (Young et al. (2012)), and independent finger movements (Taghizadeh et al. (2021)), and can be applied to the control of prostheses (Cimolato et al. (2022)). Lucas et al. have achieved a misclassification rate of $4.7 \pm 3.7\%$ for six hand movements with 8 channels of sEMG (Diogo et al. (2012)). However, sEMG inherently suffers from a low signal-to-noise ratio and poor spatial resolution, which hampers its accuracy in decoding precise finger movements (Fang et al. (2017)). This shortcoming makes the accuracy of using sEMG for classification decrease as the number of different

^{*} This work was supported by the Research Overhead and Startup funding at the University of Alabama with the FOAP of 13009-214271-200.

gestures increases. For example, Ding et al. used NinaPro's DB2 database (Atzori et al. (2012)), which contains eight-channel sEMG data of 50 kinds of gestures from 40 healthy subjects, and applied deep learning methods to train a classifier, achieving an accuracy of 83.79% (Ding et al. (2018)).

In contrast, ultrasonic signals have a higher signal-to-noise ratio and detection accuracy, so it is considered a non-invasive alternative technology for controlling prostheses. Usually, ultrasound signals are categorized into amplitude mode (A-mode) ultrasound and brightness mode (B-mode) ultrasound. A-mode provides one-dimensional information, while B-mode can provide two-dimensional images. Xia et al. used 4-channel A-mode ultrasound data and applied the SVM algorithm for classification (Xia et al. (2019)), reaching 84.25% on 20 gestures, which was 15.66% higher than the accuracy obtained by classifying sEMG data using the same method. This illustrates the signal advantages in recognition accuracy. B-mode ultrasound could visualize target tissues and demonstrate deep-muscle movements. Akhlaghi et al. used B-mode ultrasound images of forearm muscles, extracted features through image subtraction and aggregation, and achieved an average classification accuracy of 91% using the nearest neighbor classifier (Akhlaghi et al. (2015)).

Deep learning has been widely used in image classification (Wang et al. (2021)), human activity recognition (Chen et al. (2021)), medical image processing (Affonso et al. (2017)), and other fields, and has recently been introduced into gesture recognition research based on muscle information. Jia et al. compared the performance of various machine learning techniques for gesture classification using sEMG signal and found that the accuracy of the CAE(Convolutional Auto-Encoder)+CNN method was 99.38%, which was better than all other machine learning methods used for comparison, including SVM, CNN, and random forest. etc. (Jia et al. (2020)). McIntosh et al. used the optical flow method to process B-mode ultrasound image sequences and trained classifiers using SVM and MLP methods, respectively. The results showed that the MLP method was slightly better than the SVM method at four probe placement angles (McIntosh et al. (2017)).

In this study, we investigated a deep-learning approach to improve hand gesture recognition accuracy by using B-mode ultrasound images from the forearm. The contributions in this work are summarized as (1) a simple CNN design to improve the training efficiency and reduce real-time classification delay; (2) the inclusion of image augmentation to improve the robustness of the CNN model; and (3) hand gesture recognition performance validations in both offline manner and real-time manner.

2. MATERIALS AND METHODS

2.1 Experimental protocol and data Pre-processing

B-mode ultrasound is a technology that can display two-dimensional images of internal tissues' structural and morphological features of the human body and is widely used in medical diagnosis. Unlike amplitude mode (A-mode) ultrasound, which mainly provides one-dimensional informa-

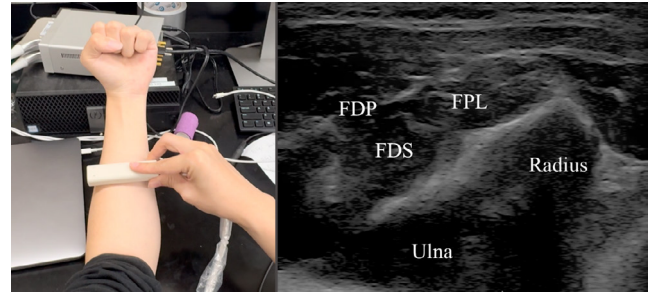


Fig. 1. Cross section of the forearm muscles and corresponding ultrasound image.

tion, B-mode ultrasound can provide a more intuitive two-dimensional image, which can clearly visualize muscles in the forearm that are specialized for finger motions (Diogo et al. (2012)). When the fingers move, the muscles in the forearm contract and pull on the ligaments, which then pull on the corresponding muscles to complete the hand movement. In this study, we mainly focused on the changes in ultrasound images of the flexor digitorum profundus (FDP), flexor digitorum superficialis (FDS), and flexor pollicis longus (FPL), as shown in Fig. 1. The FDP and FDS are respectively responsible for the four fingers except the thumb.

To evaluate the performance of B-mode ultrasound images in gesture recognition, we recruited 4 young healthy subjects (mass: 79.75 ± 9.32 kg, height: 178.25 ± 2.75 cm, and age: 26.5 ± 3.87 years old), and all study procedures were approved by the institutional review board (IRB) at the University of Alabama (Protocol ID: 24-01-7231). We used the ultrasound diagnostic system (ArtUs EXT-2H from Telemed, Lithuania) for data collection, as shown in Fig. 2. The maximum detection depth in the current study was set as 50 mm, and it can detect the movement of deep forearm muscles. The image refresh frequency is 60 fps, which means 60 images can be collected in one second. During the experiment, we placed the ultrasound probe at an angle of 30 degrees with the radius and fixed it on each subject's left forearm with a strap. The 30-degree oblique placement allows us to simultaneously obtain the deformation of the muscle cross-section and the muscle stretching and contraction along the radius from ultrasound images.

For each subject, we collected ultrasound image data and noted ground truth (hand gesture sequence from a screen) from five trials. Each trial comprised eight specific gestures, as shown in Fig. 2, performed according to the cues presented in timed slides on a screen. As detailed in Fig. 2, subjects were required to repeat each gesture three times. Each gesture was held for five seconds, followed by a rest period of equal duration. Throughout the experiment, we captured video recordings of the B-mode ultrasound images using an open-source screen recording software (OBS Studio). These videos were then segmented by using a custom Matlab code to extract frames from the 2.5 to 3.5-second mark of each gesture's stable maintenance phase, which were deemed the effective dataset. The images were subsequently converted to grayscale and resized to a resolution of 224×224 pixels. To mitigate the bad effect of speckle noise, we applied a thresholding technique, setting all pixel values in grayscale below 10 to zero.

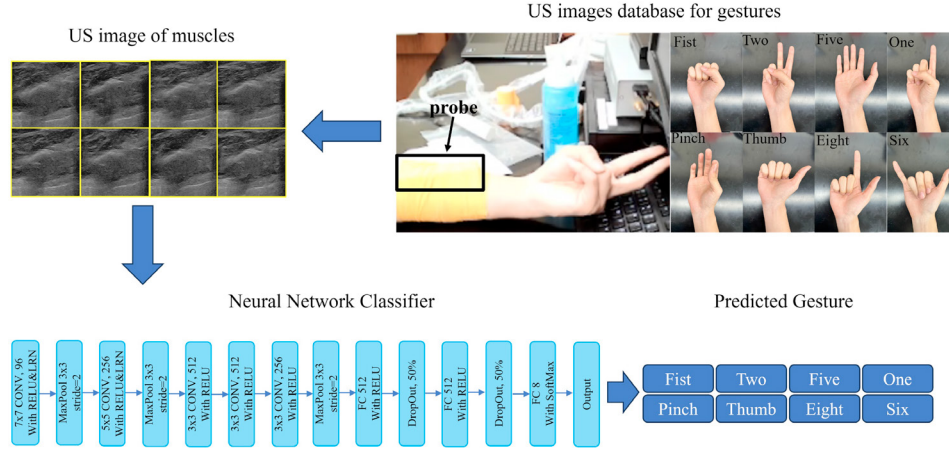


Fig. 2. Data collection and network design of CNN gesture classifier.

2.2 Hand gesture classifier development

The classifier is intrinsically a supervised learning problem that convolutional neural networks (CNNs) are designed to solve. Consider a scenario where we focus on classification, but limit the label (dependent variable) to a single dimension, specifically to one of eight possible hand gestures. Define $X \in \mathcal{X} \subset \mathbb{R}^{m \times n}$ and $Y \in \mathcal{Y} \subset \mathbb{R}$ to be two random variables. Typically, we have $Y = g(X)$ for some unknown function g . Considering a sample $(x_i, y_i)_{i=1,2,\dots,N}$ randomly drawn from the distribution of subsets of X and Y , the goal of supervised learning is to establish a mapping $\hat{g} : \mathcal{X} \rightarrow \mathcal{Y}$. This mapping aims to minimize a cost, as defined by a pre-designed cost function $L : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$. Since minimizing across the entire set of all possible functions from $\mathcal{X}$ and $\mathcal{Y}$ is impractical, we limit our search to a predefined subset $\mathcal{G}$, and define

$$\hat{g} = \arg \min_{g \in \mathcal{G}} E[L(Y, g(X))]. \quad (1)$$

As a class of biologically inspired neural networks that was first introduced by LeCun et al. (LeCun et al. (1989)), CNNs have been extensively applied to both classification and regression problems in various research fields (LeCun et al. (2015)). To solve equation (1), CNNs pass X through a series of convolutional filters and simple nonlinearities. A typical CNN architecture is shown in Fig. 2, where it has a hierarchical architecture starting from the input signal x , and each subsequent layer x_i is calculated as

$$x_i = \gamma(W_i x_{i-1}), \quad (2)$$

where W_i is a linear calculator and $\gamma(\cdot)$ is a nonlinearity. In general, W_i is a convolution, and γ is a rectifier $\max(x, 0)$ or a sigmoid function $\frac{1}{1+e^{-x}}$. In this context, W_i applies a series of convolutional filters, resulting in filter maps at each layer, which are computed as the sum of convolutions from the preceding layer below,

$$x_i(u, k_i) = \gamma\left(\sum_k (x_{i-1}(\cdot, k) * W_{i,k_i}(\cdot, k))(u)\right), \quad (3)$$

where $u \in \mathbb{R}^p$ ($p = 2$ for images) represents the input variables and k_i represents the k^{th} filter in the current layer. The symbol $*$ is the discrete convolution operator and can be defined as

$$(f * h)(u) = \sum_{v=-\infty}^{\infty} f(v)h(u-v). \quad (4)$$

The optimization problem defined by a CNN is inherently non-convex, which complicates the process of finding the global minimum. Typically, the weights W_i are learned using stochastic gradient descent, which relies on the back-propagation algorithm to compute the gradient efficiently, enabling iterative updates that seek to minimize the cost function.

2.3 CNN classifier training and validation

For each subject, four repeated trials out of five were selected following an index-increasing order for the CNN classifier training and intra-validation, where the image number ratio between training and intra-validation was set as 7 : 3. Then we applied data augmentation on the training set. Data augmentation is a method of randomly rotating, translating, cropping, and inverting images in a dataset to improve the robustness of the trained model. The remaining trial was regarded as a set of inter-validation sets (also defined as quasi-real-time validation) to expand the potential performance for real-time implementation, where no data from this trial was included in the classifier training or intra-validation.

In order to evaluate the effectiveness of the CNN classifier in the training, intra-validation, and quasi-real-time validation for a specific gesture category, we use the following formula to calculate the accuracy

$$\text{Acc}_j = \frac{TP_j + TN_j}{TP_j + TN_j + FP_j + FN_j}, \quad (5)$$

where TP_j are the number of instances correctly predicted as class j . TN_j are the number of instances correctly predicted as not class j . FP_j are the number of instances incorrectly predicted as class j . FN_j are the number of instances incorrectly predicted as not class j .

3. RESULTS AND DISCUSSIONS

In this section, we will present and discuss results from the training process, intra-validation, quasi-real-time validation, and real-time performance tests. The data collection

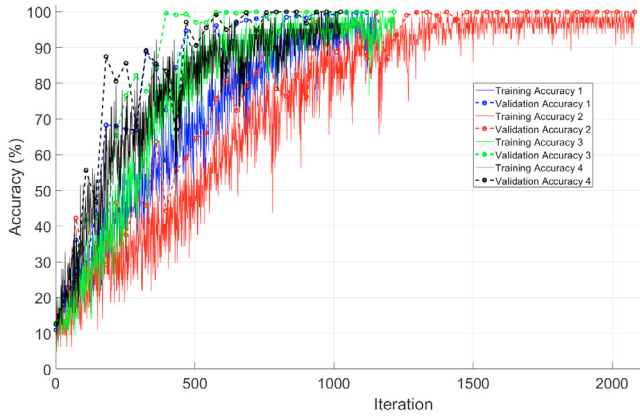


Fig. 3. Training and intra-validation accuracy over iterations.

between trials is independent, we can use any trial that has not been involved in the training process to test the classification performance of the real-time implementation.

3.1 Classifier training and intra-validation

For each subject, five training sessions (leave δ -trial out training, $\delta = 1, 2, \dots, 5$) were conducted to verify the effectiveness of the CNN classifier model design. The accuracy iterative changes during training and intra-validation are shown in Fig. 3, where four colored lines of blue, red, green, and black represent subjects 1 to 4 respectively. The convergence time of all classifier training sessions is between about 4 and 7 minutes, which facilitates the same-visit and same-session real-time implementation without removing and reattaching the ultrasound transducer. The intra-validation accuracy results of five training sessions from all participants are shown in Table 1. Due to the obvious and easy-to-extract features provided by B-mode ultrasound images, all models can achieve high accuracy in the intra-validation test, at least 99.65%.

Table 1. Hand gesture recognition accuracy of the intra-validation from four participants.

Leave δ -trial	Subject 1	Subject 2	Subject 3	Subject 4
Leave 1	100%	100%	100%	100%
Leave 2	100%	99.65%	99.59%	100%
Leave 3	99.65%	100%	99.83%	100%
Leave 4	99.83%	100%	99.88%	100%
Leave 5	100%	100%	99.88%	100%
Average	99.9%	99.93%	99.84%	100%

3.2 Quasi-real-time validation

While researchers have shown that offline training results did affect the real-time operations in terms of the sEMG-based human-machine interface that is used for wearable robotic devices control (Hinson et al. (2022)), the findings cannot be directly applied to the ultrasound imaging-based interfaces. To investigate the potential real-time implementations of the proposed hand gesture recognition approach, a close step of quasi-real-time validation is demonstrated in this work. The confusion matrix corresponding to each subject is shown in Fig. 4. The average accuracy values across 8 gestures from subject 1 to subject

4 are 92.4%, 99.74%, 98.55%, and 99.18%, respectively, and the overall quasi-real-time accuracy of our CNN classifier is 97.47%. In addition, we notice that for subject 1, in the quasi-real-time validations, several gestures have greater confusion. For example, 'pinch' and 'six' are incorrectly classified as 'fist', reaching an error rate of up to 19.2%. This may be due to the fact that the shape of the forearm muscles varies from person to person, and for some human participants, the muscle images between certain gestures are not sufficiently differentiable, resulting in the recognition mistake here.

3.3 Real-time Classification

To test the performance of our CNN classifier in real-time, we wrote custom code in Matlab to capture ultrasound images for the CNN classifier and predict the gesture. Detailed steps of the real-time classification procedure are as follows: 1) Collect a dataset of 5 trials and pre-process the images the same way as offline training 2) Use the dataset to train our CNN classifier 3) Start real-time testing without moving the ultrasound probe. The whole procedure took about 20 minutes. We demonstrated the real-time classification performance on one representative subject. Fig. 5 shows the experimental setup and classification results visualization interface on one computer screen. It took an average of around 50 ms to get the gesture recognition results frame to frame, which would be promising for prosthetic hand device control in real daily activities. For the real-time demonstration, we recorded 3 trials of 10 seconds per gesture to calculate the confusion matrix for real-time classification performance, which is shown in Fig. 4. The overall accuracy is 90.83% and the main misclassification appeared on the gesture of 'five'.

3.4 Discussions

The results of the quasi-real-time validations show that our CNN classifier demonstrates good extraction capabilities for the muscle deformation features presented in forearm B-mode ultrasound images and can classify them with a high degree of accuracy, although not all gestures are equally clear. We compared our method with some studies based on sEMG signals and A-mode signals, and our method was able to achieve the highest accuracy, which is shown in Table 2.

One existing study that used sEMG signals for real-time recognition of hand gestures achieved a delay of 180 ms to 220 ms Tepe and Demir (2022), and another existing study that used A-mode ultrasound signals got 10 Hz classification frequency Lu et al. (2022). Compared to these studies, our CNN classifier by using B-mode ultrasound images has achieved a lower delay (within 50 ms), enabling efficient and effective hand gesture control of prosthetic devices.

Through digging more into the current study, the accuracy of the CNN classifier in real-time is relatively lower than quasi-real-time validation. Especially, the gesture of 'five' was incorrectly classified as 'two' and 'pinch'. The possible reason lies in that in a real-time situation, the CNN classifier is sensitive to some environmental factors that influence the accuracy, for example, slight wrist turns and differences in finger position for the same gesture.

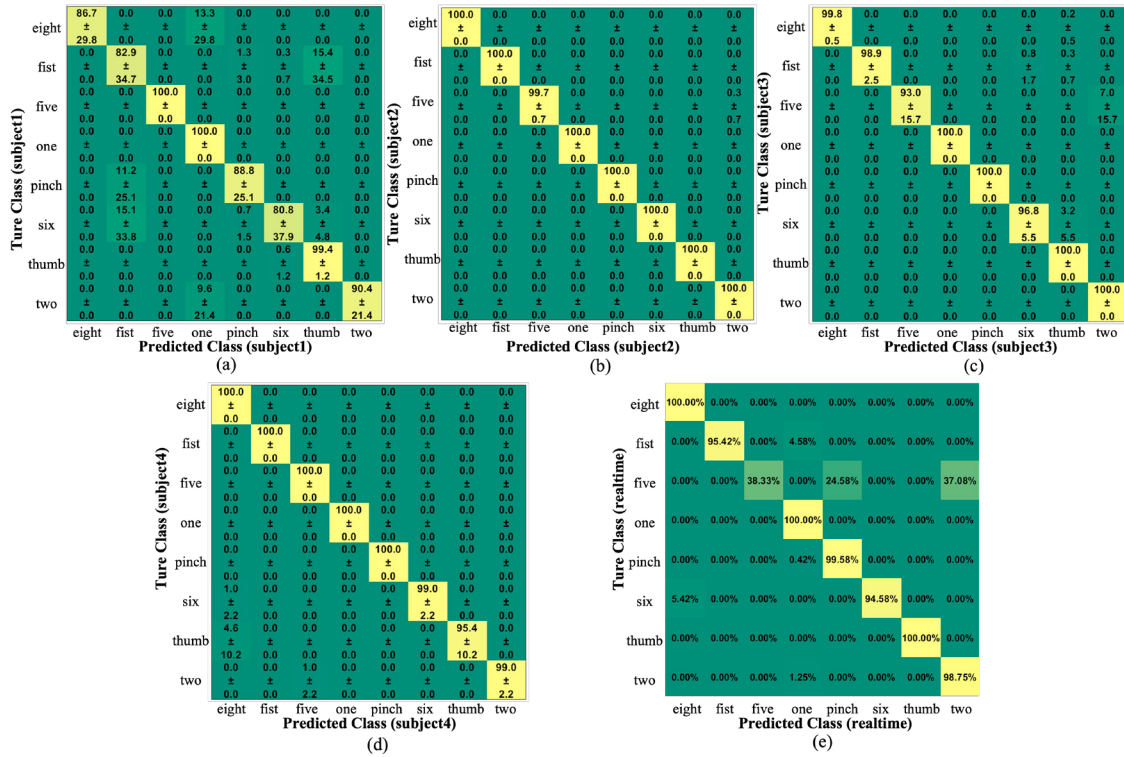


Fig. 4. Confusion matrix of the quasi-real-time validation from all human participants and real-time classification. Fig. (a)-(d) is the confusion matrix of quasi-real-time validation for subjects 1-4 respectively. Fig. (e) is the confusion matrix of real-time classification.

Table 2. Comparison of our study and recent studies

Research	Signal Type	Methods	Classification Accuracy	Number of Gestures
Fajardo et al. (2021)	sEMG	TL+LSTM LDA CNN	81.54%/88.54%/94.19%	8/6/5
Chen et al. (2020)	sEMG		97.34%	10
Lu et al. (2022)	A-mode US		96.92% $\pm$ 1.92%	10
our work	B-mode US		97.47 $\pm$ 3.41%	8

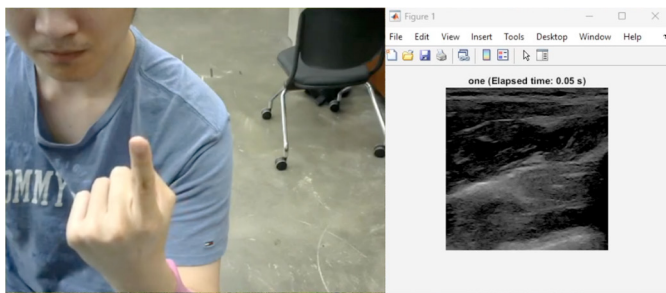


Fig. 5. Real-time hand gesture recognition experiments and results visualization interface by using Matlab.

Although promising results of real-time validation have been shown, there are still some limitations in the current work. Firstly, only 8 hand gestures were included and no wrist movements were allowed in the experimental protocol, and further investigations are required for more hand gestures and the involvement of wrist gestures. Secondly, the placement of the ultrasound transducer was set to 30° along with the mid-line of the forearm given the wide probe size and insufficient space on the forearm, which may caused missing useful muscle tissues during data collection.

Since the aim for developing hand gesture recognition method using US image is for prosthetic hand control, the portability of the devices included in this system should be considered. In this study we are using wired US imaging device which is not portable, but there are already some researches could be potential alternatives such as portable US imaging probe (Baribeau et al. (2020)) or smaller wearable device (Hu et al. (2023)).

For future work, using sensor fusion may be able to solve the shortcoming that, for some people, muscle image changes are not obvious enough in specific gestures. For example, sEMG signals are introduced to detect muscle movements that are not within the range of the ultrasound image, such as the inside of the arm near the elbow, to detect muscles such as extensor digitorum, extensor carpi ulnaris (Arteaga et al. (2020)), to compensate for muscle images between certain gestures that may have insufficient differentiation.

4. CONCLUSION

In this paper, we designed a CNN classifier for hand gesture recognition based on B-mode ultrasound images from the forearm anterior side. We trained the classifier model

using a dataset collected from four subjects comprising eight types of gestures. The results show that the classifier achieved a high accuracy of 99.65% in the intra-validation and performed well in quasi-real-time validation, reaching an overall accuracy of 97.47%. In real-time tests, it also shows a high accuracy of 90.83%. Therefore, it can be concluded that our classifier can be used for online classification and has the potential to be further used in controlling hand prostheses.

ACKNOWLEDGEMENTS

The authors would like to thank all participants for volunteering in this study.

REFERENCES

- Affonso, C., Rossi, A.L.D., Vieira, F.H.A., de Leon Ferreira, A.C.P., et al. (2017). Deep learning for biological image classification. *Expert Syst. Appl.*, 85, 114–122.
- Akhlaghi, N., Baker, C.A., Lahlou, M., Zafar, H., Murthy, K.G., Rangwala, H.S., Kosecka, J., Joiner, W.M., Pancrazio, J.J., and Sikdar, S. (2015). Real-time classification of hand motions using ultrasound imaging of forearm muscles. *IEEE Trans. Biomed. Eng.*, 63(8), 1687–1698.
- Arteaga, M.V., Castiblanco, J.C., Mondragon, I.F., Colorado, J.D., and Alvarado-Rojas, C. (2020). Emg-driven hand model based on the classification of individual finger movements. *Biomed. Signal Process Control*, 58, 101834.
- Atzori, M., Gijsberts, A., Heynen, S., Hager, A.G.M., Deriaz, O., Van Der Smagt, P., Castellini, C., Caputo, B., and Müller, H. (2012). Building the ninapro database: A resource for the biorobotics community. In *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 1258–1265. IEEE.
- Baribeau, Y., Sharkey, A., Chaudhary, O., Krumm, S., Fatima, H., Mahmood, F., and Matyal, R. (2020). Hand-held point-of-care ultrasound probes: the new generation of pocus. *Journal of cardiothoracic and vascular anesthesia*, 34(11), 3139–3145.
- Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z., and Liu, Y. (2021). Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Computing Surveys (CSUR)*, 54(4), 1–40.
- Chen, X., Li, Y., Hu, R., Zhang, X., and Chen, X. (2020). Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method. *IEEE J. Biomed. Health Inform.*, 25(4), 1292–1304.
- Cimolatto, A., Driessen, J.J., Mattos, L.S., De Momi, E., Laffranchi, M., and De Michieli, L. (2022). Emg-driven control in lower limb prostheses: A topic-based systematic review. *J. Neuroeng. Rehabil.*, 19(1), 43.
- Ding, Z., Yang, C., Tian, Z., Yi, C., Fu, Y., and Jiang, F. (2018). semg-based gesture recognition with convolution neural networks. *Sustainability*, 10(6), 1865.
- Diogo, R., Richmond, B.G., and Wood, B. (2012). Evolution and homologies of primate and modern human hand and forearm muscles, with notes on thumb movements and tool use. *J. Hum. Evol.*, 63(1), 64–78.
- Drost, G., Stegeman, D.F., van Engelen, B.G., and Zwarts, M.J. (2006). Clinical applications of high-density surface emg: a systematic review. *J. Electromyogr. Kinesiol.*, 16(6), 586–602.
- Fajardo, J.M., Gomez, O., and Prieto, F. (2021). Emg hand gesture classification using handcrafted and deep features. *Biomed. Signal Process Control*, 63, 102210.
- Fang, Y., Zhou, D., Li, K., and Liu, H. (2017). Interface prostheses with classifier-feedback-based user training. *IEEE Trans. Biomed. Eng.*, 64(11), 2575–2583.
- Hinson, R.M., Berman, J., Filer, W., Kamper, D., Hu, X., and Huang, H. (2022). Offline evaluation matters: Investigation of the influence of offline performance on real-time operation of electromyography-based neural-machine interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 31, 680–689.
- Hu, H., Huang, H., Li, M., Gao, X., Yin, L., Qi, R., Wu, R.S., Chen, X., Ma, Y., Shi, K., et al. (2023). A wearable cardiac ultrasound imager. *Nature*, 613(7945), 667–675.
- Jia, G., Lam, H.K., Liao, J., and Wang, R. (2020). Classification of electromyographic hand gesture signals using machine learning techniques. *Neurocomputing*, 401, 236–248.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436–444.
- LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., and Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Adv. Neural Inf. Process Syst.*, 2.
- Lu, Z., Cai, S., Chen, B., Liu, Z., Guo, L., and Yao, L. (2022). Wearable real-time gesture recognition scheme based on a-mode ultrasound. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 30, 2623–2629.
- McIntosh, J., Marzo, A., Fraser, M., and Phillips, C. (2017). Echoflex: Hand gesture recognition using ultrasound imaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1923–1934.
- Parajuli, N., Sreenivasan, N., Bifulco, P., Cesarelli, M., Savino, S., Niola, V., Esposito, D., Hamilton, T.J., Naik, G.R., Gunawardana, U., et al. (2019). Real-time emg based pattern recognition control for hand prostheses: A review on existing methods, challenges and future implementation. *Sensors*, 19(20), 4596.
- Taghizadeh, Z., Rashidi, S., and Shalbaf, A. (2021). Finger movements classification based on fractional fourier transform coefficients extracted from surface emg signals. *Biomed. Signal Process Control*, 68, 102573.
- Tepe, C. and Demir, M. (2022). Real-time classification of emg myo armband data using support vector machine. *IRBM*, 43(4), 300–308.
- Wang, P., Fan, E., and Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognit. Lett.*, 141, 61–67.
- Xia, W., Zhou, Y., Yang, X., He, K., and Liu, H. (2019). Toward portable hybrid surface electromyography/a-mode ultrasound sensing for human-machine interface. *IEEE Sensors Journal*, 19(13), 5219–5228.
- Young, A.J., Smith, L.H., Rouse, E.J., and Hargrove, L.J. (2012). Classification of simultaneous movements using surface emg pattern recognition. *IEEE Trans. Biomed. Eng.*, 60(5), 1250–1258.