Proceedings of the 2nd International Conference on Mathematical Physics and Computational Simulation
DOI: 10.54254/2753-8818/39/20240651

Gesture recognition and classification based on Newton Raphson optimised XGBoost algorithm

Jiaming Guo

The College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, 200233, China

gym0360@163.com

Abstract. In this study, the XGBoost algorithm is improved using the state-of-the-art Newton-Raphson Optimisation Algorithm (NRBO) for the detection of EMG signals and for gesture recognition and classification. On the training set, we observed four gesture types: stationary hand, clenched fist hand, wrist flexion, and wrist extension, all of which achieved 100% accuracy in prediction. While a total of 210 gestures were correctly predicted on the test set, only three gestures were incorrectly predicted. Specifically, a hand that was supposed to be stationary was mispredicted as clenched fist, a gesture that was supposed to be wrist flexion was mispredicted as stationary, and a hand that was supposed to be stationary was mispredicted as wrist extension. Overall, our proposed XGBoost model based on NRBO optimisation exhibits 98.59% accuracy and performs well in gesture prediction and classification. This study is significant, not only improving the accuracy and efficiency of EMG signal processing techniques, but also providing useful insights for future research in related fields.

Keywords: Newton-Raphson optimisation algorithm, XGBoost, gesture recognition.

1. Introduction

Gesture recognition is an important human-computer interaction method, which can be applied to virtual reality, smart home, medical assistance and other fields. Traditional gesture recognition technology is usually based on cameras to capture hand movements, but there are problems such as lighting and occlusion in complex environments. In contrast, gesture recognition based on EMG signals is achieved by capturing the bioelectric signals generated by muscle activity, which has the advantages of being free from environmental restrictions, high accuracy and good real-time performance. EMG signals are weak currents generated by neuron excitation of muscles, which can reflect the intention and strength of human muscle movement [1]. Therefore, the use of EMG signals for gesture recognition can obtain the user's intention more directly, thus improving the interaction experience and application effect. Figure 1 shows that gesture recognition can be performed by recognising EMG signals.

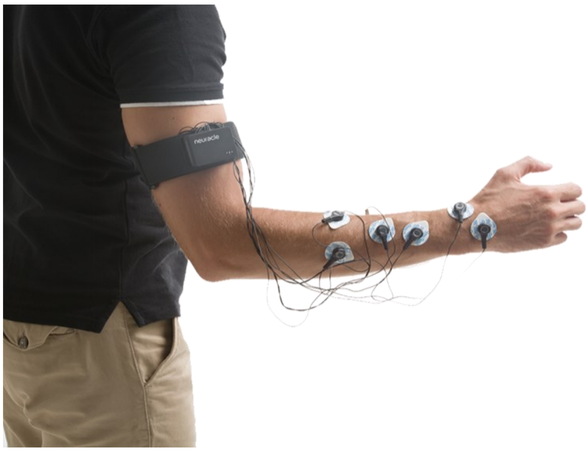


Figure 1. Myoelectric Signal Connection.

Machine learning algorithms play an important role in gesture recognition based on EMG signals. Firstly, machine learning algorithms can analyse and process a large amount of EMG signal data, extract effective features and build models, so as to achieve accurate classification and recognition of different gestures. Secondly, machine learning algorithms can also improve the adaptability of the system to individual differences, movement changes and other factors by training and optimising the data, thus improving the stability and robustness of the gesture recognition system [2]. Common machine learning algorithms include Support Vector Machines (SVM) [3], Random Forest [4], Deep Learning [5], and so on. These algorithms are able to classify EMG signals according to the relationship between different features, and they have strong generalisation ability, and they can also make accurate judgments when facing new samples.

In addition, with the development of deep learning technology, convolutional neural network (CNN) [6] has also been widely used in the field of EMG signal processing.CNN can automatically learn the feature representation and is suitable for dealing with data with temporal nature, so it has shown good results in EMG signal processing.

Machine learning algorithms play a crucial role in EMG signal-based gesture recognition, where they analyse and model large amounts of data to achieve accurate understanding and recognition of user intent. In order to take advantage of the great potential of current machine learning algorithms, we have improved the XGBoost algorithm using the current state-of-the-art Newton-Raphson Optimisation Algorithm (NRBO) for detecting EMG signals and performing gesture recognition and classification.

2. Related work

Gesture recognition and classification method based on EMG signals is a research area of great interest. Researchers have achieved recognition and classification of different gestures by capturing EMG signals, thus providing a new possibility for human-computer interaction. In related work, researchers typically capture EMG signal data generated when subjects perform various gestural actions, and analyse and process these data using machine learning algorithms.

Some research efforts have focused on exploring how to extract effective features to describe EMG signals in order to better distinguish between different gestures. These features can include time-domain features, frequency-domain features, amplitude features, etc. By extracting and selecting these features, it can help the algorithms to recognise gestures more accurately [7].

Other research works focus on selecting appropriate machine learning algorithms for gesture recognition and classification. Algorithms such as Support Vector Machines (SVM), Random Forest, and Deep Learning have been widely used in the field of EMG signal processing, and they can effectively model and classify EMG signal data to achieve accurate recognition of different gestures [8].

In addition, some studies have explored how to combine multimodal information (e.g., EMG signals, inertial sensor data, etc.) to improve the performance of gesture recognition systems. By fusing multiple sensor data, user behavioural characteristics can be captured more comprehensively and the system's understanding and accuracy of user intent can be improved [9].

3. Data set preparation

In this paper, experiments were conducted using an open source dataset, available at (https://www.kaggle.com/datasets/sojanprajapati/emg-signal-for-gesture-recognition/data), using a MYO Thalmic worn on the user's forearm bracelet and a PC with a Bluetooth receiver to receive EMG signals. The bracelet is equipped with eight equidistant sensors around the forearm, which can acquire the myogram signals simultaneously, and the signals are sent to the PC through the Bluetooth interface.A total of 712 data are used in this paper, and each data records the EMG signals received by the eight channels as well as the corresponding gestures, which are classified into four types, including the stationary hand, the hand with a clenched fist, the flexion of the wrist, and the extension of the wrist.Some of the datasets are shown in Table 1.

Table 1. Part of the dataset.

Channel1	Channel2	Channel3	Channel4	Channel5	Channel6	Label
-0.00128	-0.00041	-0.00014	-1.00E-05	-0.00012	0.00021	2
-0.00128	-0.00041	-0.00014	-1.00E-05	-0.00012	0.00021	2
-0.00128	-0.00041	-0.00014	-1.00E-05	-0.00012	0.00021	2
-0.00128	-0.00041	-0.00014	-1.00E-05	-0.00012	0.00021	2
-2.00E-04	0.00018	-6.00E-05	-4.00E-05	-5.00E-05	0.00113	3
-2.00E-04	0.00018	-6.00E-05	-4.00E-05	-5.00E-05	0.00113	3
-2.00E-04	0.00018	-6.00E-05	-4.00E-05	-5.00E-05	0.00113	3
-0.00012	-0.00025	-9.00E-05	-4.00E-05	-9.00E-05	-7.00E-05	1
-0.00012	-0.00025	-9.00E-05	-4.00E-05	-9.00E-05	-7.00E-05	1

4. Method

4.1. Newton-Raphson optimisation algorithm

The Newton-Raphson optimisation algorithm is an iterative numerical optimisation method for solving multivariate functions at their extreme points. The Newton-Raphson optimisation algorithm is based on the idea of Newton's method and Raphson's interpolation to approximate the extreme points of a function by continuous iteration. The algorithm contains four steps:

(a) Initialisation: an initial point is selected as the starting point.

(b) Iterative update: according to the function value and derivative information at the current point, it is iteratively updated by Newton's method. Specifically, the first-order derivatives (gradient) and second-order derivatives (Hessian matrix) at the current point are first calculated. Then, the inverse matrix of the Hessian matrix is used to estimate the local quadratic approximation of the function near the current point and to find the point of minimal value of this quadratic function. This minima point is used as the next iteration point and proceeds to the next round of iterations.

(c) Convergence judgement: repeat the iterative update until the preset convergence condition is satisfied.

(d) Output result: when the convergence condition is satisfied, output the final obtained extreme value point as the optimisation result.

The Newton-Raphson optimisation algorithm, by combining Newton's method and Raphson's interpolation method, uses the information of function derivatives and second-order derivatives to update the current point at each iteration step and approximates the extremum point of the multivariate function through the iterative process. This gives the algorithm better results and convergence performance in solving nonlinear optimisation problems.

DOI: 10.54254/2753-8818/39/20240651

4.2. XGBoost

XGBoost is an integrated decision tree based machine learning algorithm that combines gradient boosting and regularisation techniques with optimised algorithmic implementations to achieve significant success in various data mining and prediction tasks. The schematic diagram of XGBoost is shown in Fig. 2.

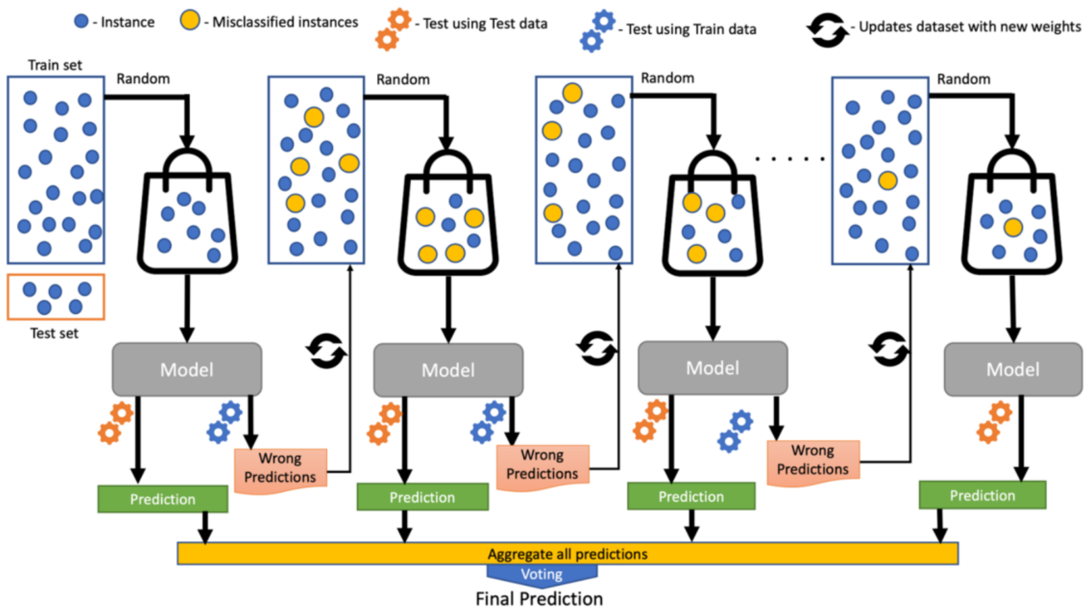


Figure 2. Schematic diagram of XGBoost.

When training the XGBoost model, it gradually builds multiple decision trees and minimises the loss function through continuous iteration. Specifically, it first initialises a simple decision tree model as the base model and then calculates the residuals of the current model on the training data. Then, in each iteration, XGBoost adds a new decision tree to fit these residuals. The newly added decision tree is designed to minimise the residuals of the current model on the training data, thus gradually improving the overall model performance. Such an iterative process will continue until a preset number of iterations is reached or some stopping condition is met.

In addition to the gradient boosting idea, XGBoost also uses regularisation techniques to prevent overfitting problems. This includes methods such as L1 and L2 regularisation, subsampling and column sampling.L1 and L2 regularisation helps to control the size of the model parameters and reduces the correlation between features, while subsampling and column sampling randomly selects a portion of the samples and features to be used in the training of each tree, increasing the diversity of the model and reducing the risk of overfitting.

In XGBoost, parameter updating using the Newton-Raphson method calculates the first-order derivatives (gradient) and second-order derivatives (Hessian matrix) of the current loss function, and then updates the model parameters based on this information. This second-order optimisation method is more efficient compared to the traditional gradient descent method because it is able to find the locally optimal solution faster.

In addition, regular terms are introduced in XGBoost to control the complexity of the model and avoid overfitting. By adding the regular term in the loss function and combining it with the Newton-Raphson optimisation algorithm for parameter tuning, the generalisation ability and stability of the model can be effectively improved.

In conclusion, the XGBoost model optimised based on the Newton-Raphson optimisation algorithm combines the gradient boosting algorithm, the regularisation technique and the efficient second-order optimisation method, and it performs well when dealing with large-scale datasets and complex features.

DOI: 10.54254/2753-8818/39/20240651

It not only can effectively improve the prediction accuracy, but also has strong robustness and generalisation ability, so it is widely used in practical applications.

5. Result

In the experimental part, this paper uses Matlab R2018b to carry out experiments with 32GB of memory, the data set is divided into training set and test set according to the ratio of 7:3, the training set is used for model training, the test set is used to test the trained model, and the accuracy is used as an evaluation index of the model's prediction effect. The adaptation curve of NRBO during training is shown in Figure 3, and the confusion matrices of the training set and the test set are outputted, and the results are shown in Fig. 4.

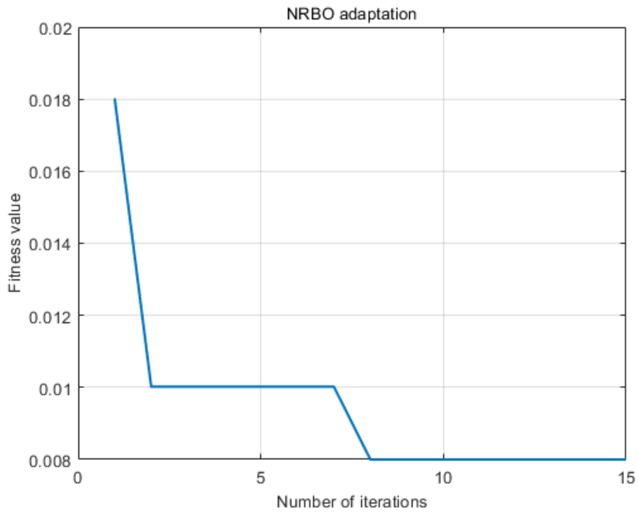
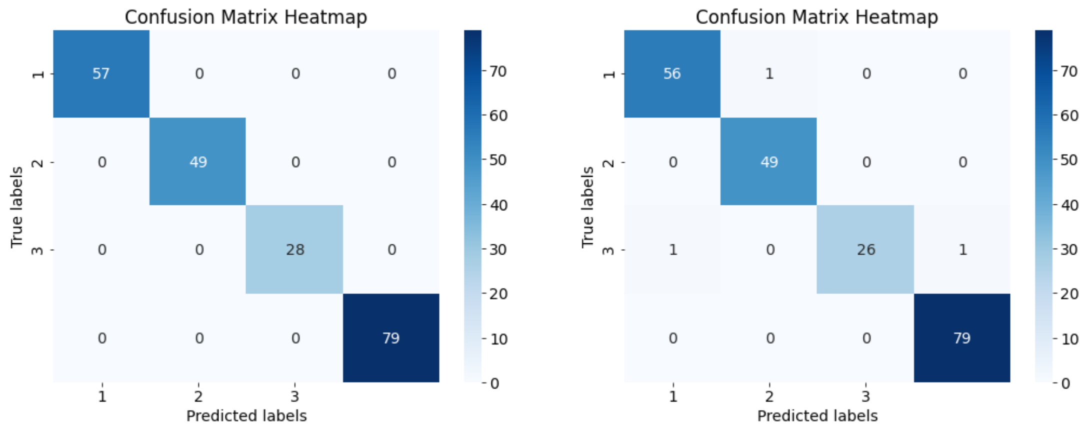


Figure 3. Fitness curve.



(a) (b)

Figure 4. (a) Confusion matrix for the training set. (b) Confusion matrix for the test set.

From the confusion matrix of the training set, it can be seen that the four types of prediction, namely, stationary hand, hand with clenched fist, wrist flexion and wrist extension, are all correct, and the prediction accuracy is 100%; in the test set, a total of 210 gestures are correctly predicted and 3 gestures are incorrectly predicted, in which one gesture that should be predicted as stationary hand is predicted as hand with clenched fist, and the other gesture that should be predicted as wrist flexion is predicted as

DOI: 10.54254/2753-8818/39/20240651

a stationary hand, and another gesture that should have been predicted as a stationary hand was predicted as a wrist extension gesture. Overall, the accuracy of gesture prediction is 98.59%, which proves that the XGBoost model based on the Newton-Raphson optimisation algorithm proposed in this paper can predict and classify gestures well.

6. Conclusion

In this paper, the Newton-Raphson optimisation algorithm (NRBO) is used to improve the XGBoost algorithm for the problem of gesture recognition and classification of EMG signals. The confusion matrix on the training set shows that all four gesture types (stationary hand, hand with clenched fist, wrist flexion and wrist extension) are correctly predicted with 100% accuracy. A total of 210 gestures were correctly predicted on the test set, and only three gestures were incorrectly predicted. Specifically, a hand that was supposed to be stationary was mispredicted as a hand with a clenched fist, a hand that was supposed to be wrist flexion was mispredicted as a stationary hand, and a hand that was supposed to be stationary was mispredicted as wrist extension.

Considering the overall performance, the XGBoost model based on NRBO optimisation proposed in this paper performs well in gesture prediction and classification with an accuracy of 98.59%. This means that our proposed method can effectively help identify different types of EMG signals and achieve accurate classification. Through this study, we provide a useful reference for the future development of smarter and more reliable EMG signal detection systems, and actively contribute to the research and applications in related fields.

References

[1] Qi, **g, et al. "Computer vision-based hand gesture recognition for human-robot interaction: a review." Complex & Intelligent Systems 10.1 (2024): 1581-1606.

[2] Wu, Meng. "Gesture Recognition Based on Deep Learning: A Review." EAI Endorsed Transactions on e-Learning 10 (2024).

[3] Kapitanov, Alexander, et al. "HaGRID--HAnd Gesture Recognition Image Dataset." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2024.

[4] Li, Deng, Bohao **ng, and **n Liu. "Enhancing micro gesture recognition for emotion understanding via context-aware visual-text contrastive learning." IEEE Signal Processing Letters (2024).

[5] Rastgoo, Razieh, et al. "Multi-modal zero-shot dynamic hand gesture recognition." Expert Systems with Applications 247 (2024): 123349.

[6] Liang, Hao, et al. "Mask-guided multiscale feature aggregation network for hand gesture recognition." Pattern Recognition 145 (2024): 109901.

[7] Karsh, Bhumika, R. H. Laskar, and R. K. Karsh. "mIV3Net: modified inception V3 network for hand gesture recognition." Multimedia Tools and Applications 83.4 (2024): 10587-10613.

[8] Balaji, Pranav, and Manas Ranjan Prusty. "Multimodal fusion hierarchical self-attention network for dynamic hand gesture recognition." Journal of Visual Communication and Image Representation 98 (2024): 104019.

[9] Khattak, Abid Saeed, et al. "Hand gesture recognition with deep residual network using Semg signal." Biomedical Engineering/Biomedizinische Technik 69.3 (2024): 275-291.

[10] Strobel, Maximilian, Stephan Schoenfeldt, and Jonas Daugalas. "Gesture recognition for FMCW radar on the edge." 2024 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT). IEEE, 2024.