Human upper limb motions recognition for stroke rehabilitation with smartphone sensors

Chen Xiaochen, Han Ping, Xu Wenchao, Yang Yanqin and Zhu Xuanming

April 23, 2019

Human upper limb motions recognition for stroke rehabilitation with smartphone sensors

Abstract

In order to improve the effective of rehabilitation training for stoke elders and provide accurate rehabilitation guidance for therapist, we construct a neural network based on Multi-Layer Perceptron(MLP) to recognize five upper limb motions basing on mobile phone sensors. Five rehabilitation movements of upper limb are chosen to be recognized include hand horizontal, hand turn left and right, hand scroll down, elbow flexion. In this experiment, the raw data are collected from the three-dimensional data of accelerometer of smartphone. After preprocessing and feature extraction of the data, the neural network can identify the five motions and some combined motions. Through the experiment, this system has a nicely performance with 98.2% accuracy.

Keywords: Rehabilitation, Smartphone, Neural network, Upper Limb Motion Recognition

1 Introduction

Stroke is a significant cause of adult disability that affects both cognitive and physical functions. It occurs when blood flow to an area of brain is cut off, and brain cells die from oxygen deprivation. In order to improve the motor abilities, such as physiological recovery, body function recovery, and activity recover, it always needs a long therapy process and much research efforts are devoted to finding the most effective forms of therapy.

Studies have shown functional improvements and induced neural plasticity in people who are a year or more post-stroke when appropriate therapy, based on motor learning principles, is administered. According to the existing rehabilitation methods for stroke people, it is mainly divided in two parts, traditional rehabilitation equipment and new rehabilitation equipment. The old rehabilitation equipment mainly includes two aspects. One is the traditional medical rehabilitation equipment, which is more professional and is generally used in medical institutions or specialized rehabilitation centers. It has a systematic program of rehabilitation and doctors can develop different training programs for different patients. But the equipment is expensive and the time available to patients is too limited. Another is specialized rehabilitation training doctors. On the other hand, the new rehabilitation equipment mainly includes five aspects. One is remote rehabilitation technology. Remote rehabilitation (TR) technology, as another way to provide rehabilitation services, enables medical staff to establish contact with patients in remote areas through information and communication technology and enables patients to receive rehabilitation training at home without the direct involvement of rehabilitation therapists. And the virtual reality (VR) technology refers to the interactive processing created by computer hardware and software to present users with a virtual environment similar to the real world, it can achieve the training that cannot be completed in reality, such as training attention in the situation of traffic intersection. The research also mentions the rehabilitation robot for stroke as well as the fabric technology electronic, which refers to the electronic characteristics of textiles or integrated into the components in the textiles. And this paper discusses a type of training method is base the sensors of mobile phone, which has not been widely applied in the rehabilitation training. It has the advantages of convenience, safety, low cost, easy recording and high accuracy.

Arm rehabilitation is an important process after stroke to regain movement skill lost. The main goal of stoke rehabilitation must begin as early as possible after a stroke attack. The common rehabilitation movement include the movements of shoulders, upper limb and the elbow. The research introduced five movements for upper limb, which is hand horizontal, hand turn left and right, hand scroll down, elbow flexion. These five movements can help stroke to recover their body function as well as activity function.

The rest of the paper is organized as follows. Section 2 introduces the related work and compare four papers. Section 3, we introduce Multi-layer perception and the new type of neural network based on MLP. In section 4, we show data processing. And in section 5, it introduces the six combined motions types. We discuss the errors and validation in section 6. In section 7, we present the result and future work. Finally, we summarize this system and propose some improvement direction for future work.

2 Related work

Recently, more and more attentions are focused on rehabilitation training. Most of researchers use wearable equipment to recognition human activity and help patients or elders to train their motion. With the same time, the development of smartphone and AI help researchers much more convenient and accuracy to record the training records.

Liquan Guo [1] use support vector machine to recognition the upper limb motion and develop an automated system by wireless inertia sensors attached to patients' arms to help stroke people do rehabilitation. This paper gets 100% accuracy from their model, but inconvenient to stroke people to use and only can recognize two typical motion with unreal time.

Yong-Joong Kim [2] recognize people's activity by smartphone thought hidden Marko model ensemble with 83.51% accuracy. It has two simple features, mean and standard deviation. After all, this experiment has nicely performance than support vector machine and multilayer perceptron. But its accuracy is not to high and need much better features to extract.

FALEH Rabeb [3] in his paper use multi-Layer Perceptron to classify the electronic nose with nice performance. The network is designed in this work for classification of seven classes using gas sensor array.

Xing Su [4] presents a human activity recognition application on smartphone. It is a real-time app and can recognize eight types of different activities with 98.7% accuracy with MLP at most. This paper has its own application work flow chart and will be released soon in Apple's App Store.

Compare with [1], we get much convenient and cheap for stroke people to use in daily life. And patients can record their rehabilitation in real-time. And we have the accuracy 99.5% compare with [2,3,4], while use the new neural network. We use wavelet transform denoising method to reduce the noise and get high accuracy. Our neural network has 2 hidden layer, back propagation network, it performs nicely in recognize single activities and combined activities.

3 Method

In this section, we describe the architecture of our model, which is show in figure 1. We make a comparison with traditional MLP network. Our network is based on MLP network and that of David Weiss and Chris Alberti (2015) which use for dependency parsing. We discuss the algorithm and regularization at the end of this section.

3.1 Multi-Layer Perceptron

MLP is the most widely used architecture for neural networks. It organized in a regular structure with several layers: an input layer, some hidden layers and an output layer. For classification problems, only one winning node of the output layer is active for each input pattern.

In the figure 1, there is a traditional MLP network show left. It has one input layer, with three input, which collected from smartphone's sensors. For the first layer of the MLP it was activation parameters and bias variables which associated with each hidden node. Each line between the input layer and hidden layer has a weight parameter. So we can have:

$$a_j^1 = \sum_{i=1}^{6} w_{ji}^{(1)} x_i + b_j^{(1)} \tag{1}$$

Where i is the number of input units and the j is the number of hidden units. The activation parameters are then fed by the non-linear activation functions of the hidden layer. There are many types of activation functions can be chosen in this part for hidden layer. We can get the output units (y_j) as the follow function:

$$y_{j=}\tanh\,(a_j^{(1)}) \tag{2}$$

3.2 Neural Network Model

In order to improve the performance of MLP, a new type MLP network has been reconstructed. It shows in the left of Figure 1.

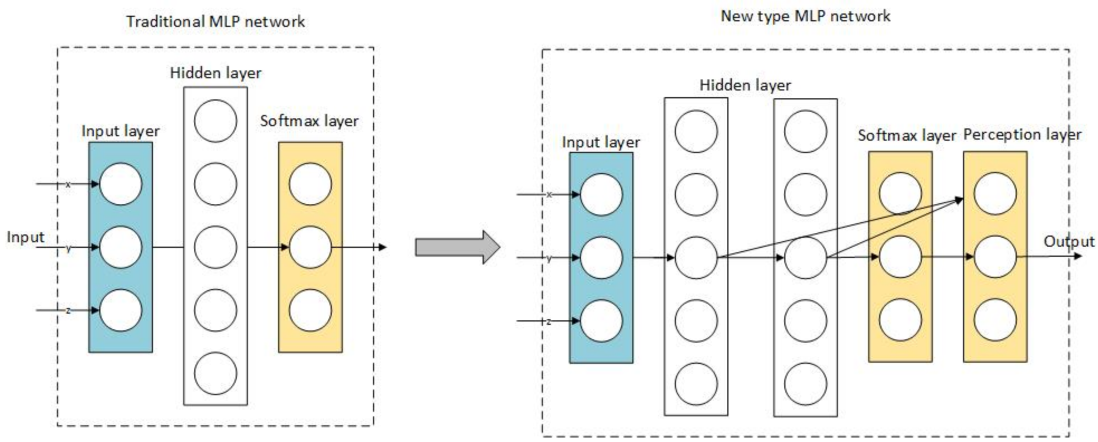

Figure 1 The structure of new type of neural network

In this figure, it adds one hidden layer and one perception layer. The data flows from the first hidden layer, and then makes the same operations in the second hidden layer. The activation functions are not same as in the traditional function, it uses third power activation function as show in table 1:

$$\mathrm{h} = (\mathrm{W}_1^{\mathrm{x}}\mathrm{x} + \mathrm{W}_1^{\mathrm{y}}\mathrm{y} + \mathrm{W}_1^{\mathrm{z}}\mathrm{z} + \mathrm{b})^3 \tag{3}$$

Those data then pass from Softmax layer using Softmax function:

$$\sigma(h_2) = \frac{e_{h_2}}{\sum_{k=1}^{k} e_{h_2}^{k}} \tag{4}$$

After the Softmax layer, the data connect to one new layer: perception layer (David Weiss and Chris Alberti 2015). The data would combine the data from the first and second hidden layer, make the result:

$$\underset{y \in GEN(x_i)}{\mathrm{argmax}} \sum_{j=1}^{m} v(y) \cdot \emptyset(\mathrm{x}, c_i) \tag{5}$$

Where $\emptyset(\mathrm{x}, c_i) = [h_1 h_2 P(y)]$. This neural network use backpropagation to learn the $\emptyset(\mathrm{x}, c_i)$ during perceptron training.

3.3 Relationship to David Weiss and Chris Alberti (2015)

Our model is clearly based on the work of David Weiss and Chris Alberti (2015). There are a few structural differences: 1) we use for three-dimensional data gather from mobile phone 2) we reduce the embedding layer 3) we use for recognize single activity and continues activity which combine two or more actions.

3.4 Algorithm and regularization

$$\mathrm{L(w)} = -\frac{1}{N}\sum_{n=1}^{N}[y_n log\hat{y}_n + (1-y_n)\log(1-\hat{y}_n)] \tag{6}$$

Our error measure is calculated from the cross-entropy error via Eq. (1). The term w denotes the weight vector, and N denotes the sample number. The terms y_n and $\hat{y}_n$ represent the real and expected outputs, respectively. The learning objective is to essentially minimize the learning error. The stochastic gradient descent method is adopted to find the optimal parameters.

4 Data preprocessing and settings

4.1 Dataset

Three volunteers are gather together to record the five activities, i.e., hand horizontal, hand turn left and right, hand scroll down, elbow flexion. They use three different type of smartphone which afford by our research team. One is galaxy Note3

Lite 4G made by Samsung, one is MX made by MEIZU, and last is Mate10 by Huawei. A debugged application in android are installed in the three smartphones and gather accelerometer data via the accelerometer sensor embedded in the smartphones. The values gather from smartphone sensor are three-dimensional.

Each activity sample of the dataset is made by fixed-width sliding window of 2.56sec and 50% overlap (128 readings/window). More detail of the dataset is summarized in table 1.

Table 1
The dataset from three types of smartphone

Motion		**Description**		
Device	**Galaxy Note3**	**MEIZU MX**	**Huawei Mate10**	**Percentage**
Hand Horizontal	2270	2310	2419	13.44%
Hand Turn Left	2220	3233	4122	18.39%
Hand Turn Right	2765	4322	3599	20.52%
Hand Scroll Down	4495	4111	2788	21.88%
Elbow Flexion	4527	3998	4888	25.76%
Total	16277	17974	17816	100%

And the five activities, which is hand horizontal, hand turn left and right, hand scroll down and elbow flexion, that volunteers use to rehabilitate, is show in figure 2.

Fig. 2. Five activities

4.2 Data preprocessing

4.2.1 Mean method

After the raw data collected from phone sensors, 3-axis tilt sensors can be got. Before transform the data into neural network, the raw data need to preprocessing. In this part, we use mean method that be defined as:

$$x = (x - mean)/standard\ deviation \tag{7}$$

4.2.2 Wavelet threshold denoising method

Since the complex of raw data, the three-dimensional data always full of noise. It may affect the result of deep learning. In order to reduce the impress of noise, we use the wavelet threshold denoising method to get rid of noise. After dealing with the wavelet transform denoising method, the data can be suppressed nicely.

The method can be defined as below:

$$X_w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t-b}{a}\right) dt \tag{8}$$

The steps of wavelet threshold denoising method is as below:

· Do wavelet transform

· Remove the coefficients that fall outside one standard deviation

· Only new coefficients are used to generate new denoising sensor data

To compare the result of the performance, we choose one activity, which is hand scroll down (show in the Figure 2).

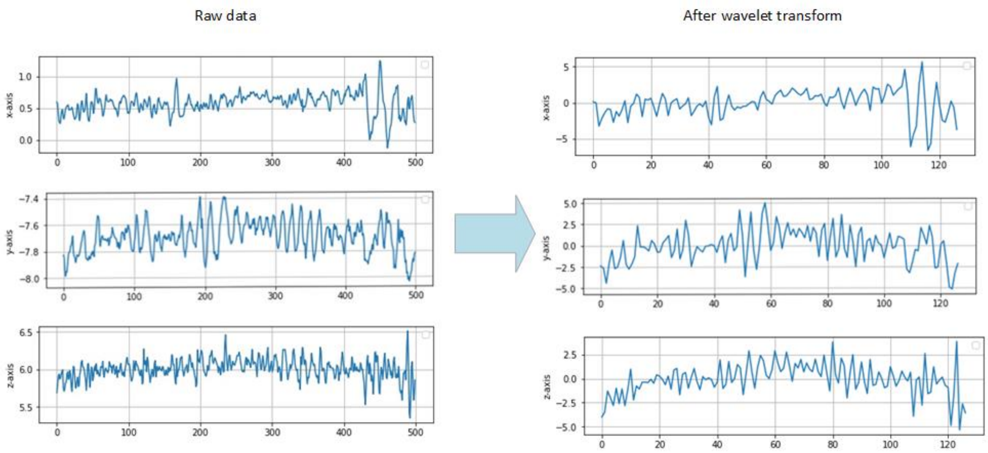

Fig.2 The data after wavelet threshold denoising

4.3 Model Initialization and Hyperparameters

The neural network method for recognize upper limb motion has two hidden layers composed of M rectified linear (Relu) units. Each unit in the hidden layers is fully connected to the previous layer:

$$h_i = \max\{0, W_i h_{i-1} + b_i\} \tag{9}$$

where W_1 is a $M_1 \times$ E weight matrix for the first hidden layer and W_i are $M_i \times M_{i-1}$ matrices for all subsequent layers. The weights b_i are bias terms.

In the start of case, we initialized W_i and β randomly using a Gaussian distribution with variance 10^{-4}. We used fixed initialization with b_i = 0.2, to ensure that most Relu units are activated during the initial rounds of training.

This research uses the following hyper-parameter values in this experiment: hidden layer sizes=28, activation=Relu, learning rate=invscaling, solver=lbfgs, max iteration=100. And split the data with 20% to test and 80% to train.

For the perceptron layer, we used $\emptyset(\mathrm{x}, c_i) = [h_1 h_2 P(y)]$.When not tri-training, we used hyperparameters of γ = 0.2, η = 0.05, μ = 0.9, early stopping after roughly 40 minutes of training time. With the tri-training data, we de- creased η= 0.05, increased γ

= 0.5, and decreased the size of the network to M1 = 1024, M2 = 256 for run-time efficiency, and trained the network for approximately 80 minutes.

5 Combined motion types

In this paper, the method of new neural network can recognize the five upper limb motions, but also can recognize the combined motions. We define the five upper limb motions as a symbol below:

- a represents hand horizontal
- b represents hand turn left
- c represents hand turn right
- d represents hand scroll down
- e represents elbow flexion

It contains as below:

Table 2 The five combined motions type for limb upper

Combined motions	**Include actions**
Type I	a+c+a
Type II	a+b+a
Type III	a+d+a
Type IV	a+e+a
Type V	a+c+a+b+a
Type VI	a+d+a+e+a

6 Errors metrics and Validation

6.1 Error metrics

We use confusion matrices to evaluate the effectiveness of our classifiers, it has false positives, false negatives, true positives and true negatives. As shown in Table3:

Table 3. Confusion matrices

	Positive	**Negative**
TRUE	TP	FN
FALSE	FP	TN

Table 3. Confusion matrices

Table 3 includes four kinds of values. True positives (*TP*) represent the number of pulsars that were correctly classified as pulsars; True negatives (*TN*) represent the number of non-pulsars that were incorrectly classified as non-pulsars; False negatives (*FN*) represent the number of pulsars that were classified as non-pulsars; False positives (*FP*) represent the number of non-pulsars that were not classified as pulsars. But in this paper, we have five classifiers, so we have the confusion matrices shown below:

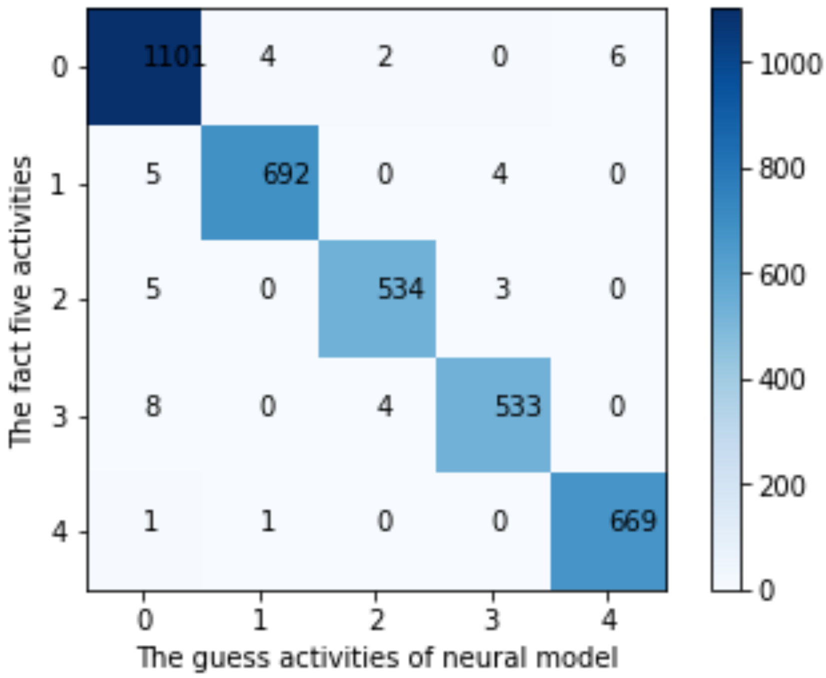

Fig.3 The Confusion matrices for five activities

In the figure 3, we take 20% part of our raw data to test the result, the 0,1,2,3,4 are separated represent hand horizontal, hand turn left and right, hand scroll down, elbow flexion.

6.2 Validation

We used a ten-fold cross-validation to get 10 results with 5 different dropout values and 5 different sample ratios to evaluate our model. We randomly divided all labeled data into ten subsets. We then used nine of the subsets to train our model. We used the remaining subset to validate our model. After ten cycles of validation, we got the average value. This procedure used to make sure the validation data set is clean and connect validation error with the error out of the sample. Furthermore, we repeated the validation procedure 5 times to ensure the performance is reliable. This can make sure our model can work well on new data. Therefore, we can choose the best model by using best validation error.

7 Result and Discussion

7.1 *Different Smartphone*

We use galaxy Note3 Lite 4G made by Samsung、MX made by MEIZU and Huawei, the difference accuracy we can see in table 4:

Table.4 The accuracy of different smartphone

Upper Limb Motion	**Galaxy Note3 Lite 4G**	**MX**	**Huawei**
Hand horizontal	98.9%	98.2%	**99.1%**
Hand turn left	97.1%	96.9%	**97.8%**

Hand turn right	96.8%	95.6%	**97.7%**
Hand scroll down	**99.7%**	98.1%	98.1%
Elbow flexion	99.3%	**99.6%**	99.5%

Form table 4, we can see different type of smartphone has fewer effect on the result of recognition.

7.2 Single activity comparison with SVM and MLP

We perform a ten-fold cross-validation on the data set. The result is summarized in Table 5. Compared with the Support Vector Machine and MLP, in five different upper limb motion, the neural network has a better performance with 98.2% average accuracy.

Table 5 The Comparison with SVM and MLP

	SVM	**MLP**	**Neural network**
Hand horizontal	91.1%	93.5%	**98.8%**
Hand turn left	89.1%	94.2%	**97.3%**
Hand turn right	87.5%	92.5%	**96.8%**
Hand scroll down	92.8%	91.9%	**98.6%**
Elbow flexion	90.2%	96.7%	**99.4%**

7.3 Combined action comparison with MLP

There are five combined actions define in this paper. Because of the back propagation, the neural network has a much better performance than MLP.

Table 6 The combined motion accuracy

	MLP	**Neural network**
Type I	89.1%	92.1%
Type II	87.2%	91.2%
Type III	86.7%%	96.8%
Type IV	83.1%	97.6%
Type V	91.2%	94.4%
Type VI	92.5%	93.2%

8 Conclusion

Nowadays, rehabilitation is playing an increasingly important role for stoke patients. It is also an urgent issue for stoke to interact with the rehabilitation easily and efficiently. The new neural network in this paper can effectively solve this problem. The system based on smartphone sensors, which is much convenience and easy to use for rehabilitation in real time. This neural network can recognize not only single upper limb motions but also the combination upper limb motions. The performance of this model is reliable and have a big future for much more complex activity than usual network like MLP or SVM.

On the other hand, this method also has some deficiencies. For instance, it cannot completely eliminate in real time. Although, it can only predict five activities. Therefore, more research should be done for the improvement of this method. In the coming future, the real-time control will be install in android smartphone and much more different upper limb motions can be recognized in the last work.

References

[1] Liquan Guo Lei Yu and Qiang Fang, Upper Limb Motion Recognition for Unsupervised Stroke Rehabilitation Based on Support Vector Machine

[2] Feigin VL,Barker- Collo S,McNaughton H,Brown P,Kerse N. Long- term neuropsychological and functional outcomes in stroke survivors:current evidence and perspectives for new research. Int J Stroke,2008,3:33- 40.

[3] Brainin M,Norrving B,Sunnerhagen KS,Goldstein LB,Cramer SC,Donnan GA,Duncan PW,Francisco G,Good D,Graham G, Kissela BM,Olver J,Ward A,Wissel J,Zorowitz R.Poststroke chronic disease management:towards improved identification and interventions for poststroke spasticity- related complications. Int J Stroke,2011,6:42- 46.

[4] Madureira S,Guerreiro M,Ferro JM.Dementia and cognitiveimpairment three months after stroke.Eur J Neurol,2001,8:621- 627.

[5] Yong-Joong Kim, Bong-Nam Kang, and Daijin Kim, Hidden Markov Model Ensemble for Activity Recognition using Tri-axis Accelerometer.

[6] FALEH Rabeb, BEDOUI Souhir, KACHOURI Abdennaceur, an electronic nose for detection pollutant odorant and olfaction classification using neural network.

[7] Xing SU, Ping Ji, Activity Recognition with Smartphone Sensors.

[8] Wavelet, https://pywavelets.readthedocs.io/en/latest/regression/index.html, (retrieved: Apr 24, 2018)

[9]StandardScaler,https://blog.csdn.net/sinat_33761963/article/details/54911020, (retrieved: Feb 07,2017)

[10] Michael M. Pires, Tshilidzi Marwala, American Option Pricing Using Multi-Layer Perceptron and Support Vector Machine

[11] Suhaimi, Rashidah; Analysis of EMG-based muscles activity for stroke rehabilitation.

[12] Duff, Margaret, An adaptive mixed reality training system for stroke rehabilitation.

[13] Alex,Marylyn, A review of sensor devices in stroke rehabilitation

[14] Ellsworth, Christopher, an innovative system to enhance upper-extremity stroke rehabilitation.

[15] David Weiss, Chris Alberti, Structured Training for Neural Network Transition-Based Parsing