Poster Abstract: Toothbrushing Recognition using Neural Networks

Hua Huang
Stony Brook University
New York 11794
hua.huang@stonybrook.edu

Shan Lin
Stony Brook University
New York 11794
shan.x.lin@stonybrook.edu

ABSTRACT

Daily toothbrushing is essential for maintaining oral health. Recently, a wrist-watch based system was designed to monitor the effectiveness of toothbrushing at home [3]. Since toothbrushing involves complex motions of the hand and the toothbrush, advanced machine learning techniques are needed for accurate recognition. In this paper, we design a neural network classifier to recognize toothbrushing surfaces based on wrist watch sensor data, including accelerometer, gyroscope and magnetometer data. Experiments shows that our algorithm successfully recognized toothbrushing surfaces with a recognition precision of 91.2%, outperforming existing solutions.

CCS CONCEPTS

•Human-centered computing →Ubiquitous and mobile computing systems and tools; •Computing methodologies → Neural networks;

KEYWORDS

Toothbrushing Monitoring, Neural Network, Activity Recognition

ACM Reference format:
Hua Huang and Shan Lin. 2017. Poster Abstract: Toothbrushing Recognition using Neural Networks. In *Proceedings of The 2nd ACM/IEEE International Conference on Internet-of-Things Design and Implementation, Pittsburgh, PA USA, April 2017 (IoTDI 2017),* 2 pages.
DOI: http://dx.doi.org/10.1145/3054977.3057287

1 INTRODUCTION

Systematic reviews have shown that the Bass technique for manual brushing has consistently demonstrated good dental hygiene results [4]. The general descriptions of the Bass technique are: a) to brush front outer, back outer and back inner surfaces, place the toothbrush at a 45-degree angle to the gums. Brush these areas with a vertical movement of the toothbrush bristles. Each brushing stroke begins from the gum line to the tip of the teeth. b) To brush chewing surfaces, move the bristles of the toothbrush along the chewing surface of the back teeth in a back and forth motion. c) To brush front inner surfaces, tilt the brush vertically and make up-and-down strokes. To monitor the compliance of Bass technique at home and recognize the surfaces of teeth brushed, an automatic toothbrushing monitoring system has been recently designed [3].

In the toothbrushing monitoring system, a very challenging issue is to recognize the toothbrushing surfaces with the Bass technique because there are similarities among the brushing gestures for different tooth surfaces and variations in brushing gestures for a same surface. Therefore, it is necessary to utilize advanced machine learning techniques to capture the essential characteristics of brushing gestures for different surfaces. Recently, neural networks have achieved promising results in many pattern recognition applications, including computer vision [2], speech recognition [1], and time series analysis [5]. Neural networks can autonomously learn the complex and non-linear mapping between the input features and output labels. Depending on the application scenarios, the neural network models can be learned in either supervised or unsupervised manners.

In this paper, we tackle the toothbrushing surface recognition problem by adopting a neural network model. The proposed approach can recognize complicated toothbrushing gestures based on the data collected from the accelerometer, gyroscope, and magnetic sensors on the wrist watch. A set of motion features from the raw sensor data [3] are extracted. Then we design a two layer neural network to conduct toothbrushing surface classification. Finally, we compare the performances of the neural network classifier and the Bayes classifier in terms of toothbrushing recognition accuracy. We show that when we train the neural network classifier with more than five sets of training data, it can achieve 91.2% of recognition precision, which outperforms the Bayes classifier.

2 ALGORITHM DESIGN

The sensor data are collected using the accelerometer, gyroscope and magnetic sensor in a smart watch. We firstly extract the statistical features from these sensor data. For the accelerometer and gyroscope, gravity sensor and magnetic sensor data, we extract their means and standard deviations. We denote them as $\overline{a}$,$\sigma(a)$, $\overline{g}$, $\sigma(g)$, $\overline{gr}$, $\sigma(gr)$, and $\overline{m}$, $\sigma(m)$, respectively.

We also extract the PCA-based motion direction feature from the accelerometer data. For each pair of dimensions i and j, $i, j \in \{x, y, z\}, i \neq j$, we feed the accelerometer data to the Principal Component Analysis (PCA) algorithm and compute the axes of principal component. We record the slope d_{ij} to represent the motion direction in the ij plane. The PCA-based feature is effective in recognizing brushing motion direction, and is essential in classifying the gestures of brushing left and right teeth.

Finally, we extract specialized features from the sensor data. We compute the power spectrum of the accelerometer data and extract the peak value of the power spectrum $max(P_{acc})$ and the percentage of energy concentrated around the peak frequency ($perc_Y$).

IoTDI 2017, Pittsburgh, PA USA

DOI: http://dx.doi.org/10.1145/3054977.3057287

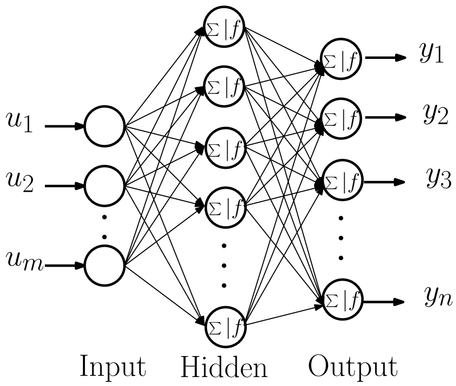

Figure 1: The Structure of the Neural Network

We also compute the correlation $C_{y,z}$ between accelerometer data along the y and z axes.

2.1 Neural Network Design

After the activity features are extracted, we use a neural network to classify the toothbrushing gestures. The structure of the neural classifier is shown in Figure 1. The network consists of an input layer, a hidden layer, and an output layer. The input to this network, represented by $\mathbf{u} = [u_1, u_2, ..., u_m]$, is the set of features extracted from the sensor data. m represents the dimension of the input feature. The network output vector is $\mathbf{y} = [y_1, y_2, ..., y_n]$. In this vector, $y_i = 1$ and $y_j = 0, \forall j \neq i$ if the prediction result is gesture class i.

To connect different layers of this network, we need to define activation functions. There are many different activation functions, such as linear, hyperbolic, step-wise, and sigmoid. Sigmoid functions are often used in neural networks because the sigmoid function has a simple relation with its own derivative, which makes it computationally easy to handle. Specifically, a sigmoid function $f(t)$ and its derivative $f'(t)$ are defined in the following equations:

$$\begin{aligned} f(t) &= \frac{1}{1+e^{-t}} \\ f'(t) &= f(t)(1 - f(t)). \end{aligned} \tag{1}$$

In the network, each neuron computes a linear combination of its input signals and applies a sigmoid function to the result, which becomes the input signal to the next layer. Finally, to train this network, we apply the classical back-propagation algorithm.

3 EVALUATION

In the evaluation, the sensor data were collected from 12 volunteers who conducted daily toothbrushing for three weeks. The ground truth labels were collected either by manual input or by a second observer. In this experiment, we test the performance of our system when we train the recognition algorithm using a different number of training data sets. By one set of training data, we mean the amount of sensor data collected in one toothbrushing session. To evaluate the system performance, we use the precision, which is defined as (True Positive/(True Positive + False Positive)).

As for the baseline algorithm, we implement the Bayes Classifier. Specifically, given the training data set $(\mathbf{u}_1, \mathbf{y}_1), (\mathbf{u}_2, \mathbf{y}_2)$

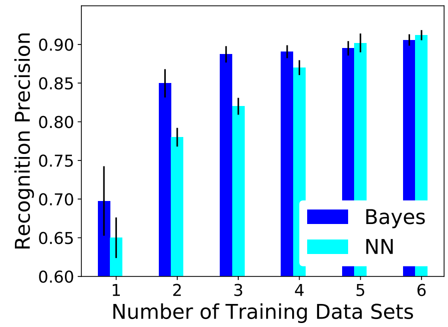

Figure 2: Toothbrushing Surface Recognition Precision

, ..., $(\mathbf{u}_N, \mathbf{y}_N)$, the algorithm computes the following three quantities: the conditional probability $p(u_i|y_i)$, the class prior probability $p(y_i)$ and the feature probability distribution $p(u_i)$. Then we use the Bayes rule $p(y|u') = \frac{p(u'|y_i)*p(y_i)}{p(u')}$ to compute the label y' that maximizes the probability for any input feature u'

From Figure 2, we can see that when only one set of data is used, the surface recognition precision of the Neural Network (NN) is 65%, which is lower than the 69% achieved by the Bayer classifier. The recognition accuracy improves gradually as the number of training data sets increases. When the number of training data sets is five, the recognition precision of NN is 90.2%, which surpass the performance of the Bayes classifier. The recognition precision of NN is further improved to 91.2% when six sets of training data sets are used. Therefore, we can see that the neural network based approach has lower recognition accuracy than the Bayes classifier when the training data set is limited, but eventually achieves better performance when the training data set is large enough (more than 5 sets).

4 FUTURE WORK

To improve recognition accuracy, we plan to introduce deep learning techniques such as the Convolution Neural Network (CNN) that can capture subtle patterns from the sensor data. We will also work on extending our recognition algorithm to other toothbrushing techniques such as the Charters and the Stillman techniques.

ACKNOWLEDGMENT

This research was supported in part by the DE-EE0007682, NSF CNS 1553273, CNS 1536086, CNS 1463722, and IIS 1460370.

REFERENCES

[1] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, and others. 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine* 29, 6 (2012), 82–97.

[2] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. 2006. A fast learning algorithm for deep belief nets. *Neural computation* 18, 7 (2006), 1527–1554.

[3] Hua Huang and Shan Lin. 2016. Toothbrushing Monitoring using Wrist Watch. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. ACM, 202–215.

[4] M Poyato-Ferrera, JJ Segura-Egea, and P Bullón-Fernández. 2003. Comparison of modified Bass technique with normal toothbrushing practices for efficacy in supragingival plaque removal. *International journal of dental hygiene* 1, 2 (2003), 110–114.

[5] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiaoli Li, and Shonali Krishnaswamy. 2015. Deep Convolutional Neural Networks on Multichannel Time Series for Human Activity Recognition.. In *IJCAI*. 3995–4001.