EMG SIGNALS FOR FINGER MOVEMENT CLASSIFICATION BASED ON SHORT-TERM FOURIER TRANSFORM AND DEEP LEARNING

Ivana Kralikova[1], Branko Babusiak[1], Lubomir Kralik[2]

[1]Department of Electromagnetic and Biomedical Engineering, Faculty of Electrical Engineering and Information Technology, University of Zilina, Zilina, Slovakia

[2]Faculty of Management Science and Informatics, University of Zilina, Zilina, Slovakia

Abstract

An interface based on electromyographic (EMG) signals is considered one of the central fields in human-machine interface (HCI) research with broad practical use. This paper presents the recognition of 13 individual finger movements based on the time-frequency representation of EMG signals via spectrograms. A deep learning algorithm, namely a convolutional neural network (CNN), is used to extract features and classify them. Two approaches to EMG data representations are investigated: different window segmentation lengths and reduction of the measured channels. The overall highest accuracy of the classification reaches 95.5% for a segment length of 300 ms. The average accuracy attains more than 90% by reducing channels from four to three.

Keywords

classification, EMG signals, finger movements, HCI

Introduction

The Human-Computer Interface (HCI) can be defined as a communication medium between a human and a computer system or terminal device. The HCI is considered a significant part of current research in medicine, industry, education, and entertainment, and thanks to rapidly evolving computed technology permeates people's daily lives. One of the most challenging approaches in this area of research is connecting bioelectric signals from the human body to a computer system. Bioelectric signals in the human body directly correlate with its state. For example, when performing movements, changes occur in the measured electromyographic (EMG) signals from the corresponded area, the electrooculographic (EOG) signals are linked with eye movement, and the electroencephalographic (EEG) signals correspond to mental state and external stimuli. These signals can be processed quantitatively or qualitatively to be related to human intent, making them applicable to the HCI systems.

The human movement or gesture can be considered the most natural, intuitive, and non-verbal interaction medium. One way to detect it is surface electromyography (sEMG). The EMG signal is a bioelectric signal generated during skeletal muscle contraction controlled by the nervous system. The EMG signals are acquired from many active motor units of examined muscle. The active motor unit's action potentials are electrically superposed and detected by surface electrodes placed on the skin above the investigated muscle. The typical value of the non-stationary sEMG signal ranges between 1 mV and 10 mV. A significant tissue volume between electrodes and muscle fibers and the electrode-skin interface limits the upper limit usable frequency band at 500 Hz. The dominant signal energy power interval corresponds to 50–150 Hz [1, 2].

As muscle moves, an EMG signal is generated that contains unique patterns. Recognition of these patterns can be used in a wide range of HCl systems. The output of the classification process then represents a control command to perform specific activities of the respective terminal device.

The EMG signals have considerable potential in HCI applications such as a bionic limb or robotic arm control. Shi et al. in [3] designed a prototype system for the recognition of five hand gestures to control the bionic hand based on real-time analysis of two bipolar EMG channels. The classification accuracy using the k-Nearest Neighbor (k-NN) algorithm achieved 94% for online classification.

Other studies deal with electric wheelchair control. Kumar et al. in [4] designed an electric wheelchair assistive device for patients with lower cervical spinal cord disorders. The wheelchair can move left, right,

Lekar a technika – Clinician and Technology 2021, vol. 51(1–4), pp. 15–20, DOI: 10.14311/CTJ.2021.1.02
ISSN 0301-5491 (Print), ISSN 2336-5552 (Online)

straight, and stop. The rest command, is related to muscle relaxation, based on a set threshold corresponding to muscle activation. The recognition accuracy rate using the Support Vector Machine (SVM) was 93.50% for an online case.

Several studies have also been conducted on the conversion of EMG signals into sign language. Tateno et al. in [5] proposed a high-precision, real-time hand motion recognition system based on American Sign Language that helps hearing-impaired people communicate with healthy people. The achieved accuracy was 97.7% using the Long Short-Term Memory (LSTM) algorithm for the recognition of 20 different word expressions.

Surface EMG signals can also be used to control a virtual keyboard, computer mouse, or virtual reality. Rahim et al. in [6] proposed a procurement system character using a virtual keyboard based on EMG signal analysis while performing hand movements. They used five hand movements to enter 20 characters of the alphabet from A to Z on the virtual keyboard. Achieved accuracy was 96.75% by the SVM algorithm. David et al. in [7] presented the computer mouse system controlled by real-time surface EMG signals, which correspond to four hand movements, including a relaxed hand condition with an accuracy of 87.13% by the Multilayer Perceptron (MLP). Li et al. in [8] dealt with the design of a real-time virtual reality management system based on wireless sensing of EMG signals with a classification accuracy of 83.33%. The virtual reality scene is the kitchen, in which the entity can perform four actions. The system can be used for muscle training rehabilitation, with the virtual kitchen scene having a positive psychological impact on patients.

This research investigates 13 finger movement recognition based on EMG signals represented by spectrograms in the time-frequency domain. Features are extracted from the obtained spectrograms and classified using a deep learning algorithm.

Materials and Methods

Hardware for Data Acquisition

A biopotential amplifier (Fig. 1) was used to obtain EMG signals whose dominant component is the integrated circuit ADS1298 (Texas Instruments, USA). This integrated circuit allows simultaneous measurement of analog signals via eight differential inputs in 24-bit resolution. The ADS1298 circuit allows setting the sampling frequency from 250 Hz to 32 kHz, setting programmable gain, lead-off detection, or configuring the driven right leg (DRL) circuit for individual inputs.

The Atmega328PB (Microchip Technology, USA) MCU is the primary control unit for the entire device through SPI. A USB-UART converter is used for communication between the device and the computer, specifically the FT232RL circuit (FTDI chip, GB).

Fig. 1: Biopotential amplifier.

Software for Data Acquisition

The whole acquisition chain is configurable via application software with a graphical user interface developed in MATLAB R2021b (MathWorks, USA). After starting the application, the user will see a clear separate window (Fig. 2) containing the main menu bar at the top; below it is a bar with tools, and in the central part located space for simultaneous rendering of received data.

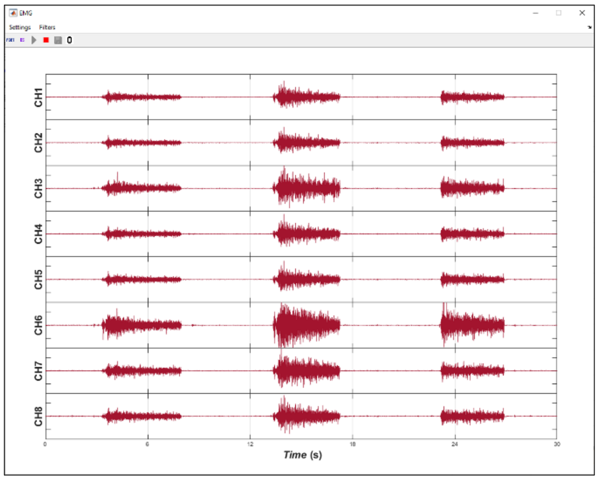

Fig. 2: Full window - simultaneous EMG measurement from eight channels in a unipolar configuration.

The *main menu* allows setting the conditions of data recording, i.e., selection of the acquisition mode (test signal or biosignal), change of the sampling frequency, setting of the programmable gain, and enabling DRL for required inputs. Another attribute is digital filtering, which provides the ability to extract the proper frequency range of the EMG signal using a bandpass filter with cut-off frequencies of 10 Hz and 350 Hz, and power line-specific frequency suppression using a notch bandstop filter set to 50 Hz.

The *toolbar* contains six icons that perform specific functions:

- the serial port selection,
- changing the range of the *x*-axis and *y*-axis,
- starting and stopping EMG signal recording,
- saving the recorded data,
- isoelectric line offset regulation.

Lekar a technika – Clinician and Technology 2021, vol. 51(1–4), pp. 15–20, DOI: 10.14311/CTJ.2021.1.02
ISSN 0301-5491 (Print), ISSN 2336-5552 (Online)

In the central part of the window the *space for displaying EMG waveforms* is situated. There are eight separate graphs for plotting the data read from the device, which is used to visually check the measurement's accuracy. Rendering is performed for newly loaded samples in sections bounded by the updated *x*-axis limit.

Dataset of EMG Data

The EMG signals were acquired simultaneously using four channels of the biopotential amplifier in a bipolar configuration for subsequent classification. Measurements were made on five healthy volunteers aged 21 to 53 years, separately for three days. The electrodes were placed on the forearm muscles of the right hand in the arrangement shown in Fig. 3. The goal was to create an EMG data set for the 13 finger movements illustrated in Fig. 4. The subject performed the individual movements for approximately five seconds, followed by a relaxation period of approximately five seconds. Each movement was repeated five times.

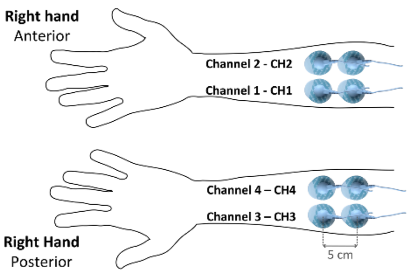

Fig. 3: Electrode placement with channel labels.

Fig. 4: 13 Investigated Movements; Ex. = Extension, Fl. = Flexion.

EMG Data Pre-processing

The application described in the previous section was used to record the data. Both types of digital filters provided by the application were applied to the initial preprocessing of the measured EMG data.

Each respective EMG record was segmented using a sliding window with a fixed length of 100 ms, 200 ms, 250 ms, and 300 ms in a step of 50 ms to form a dataset inserting the following classification process.

Subsequently, the transformation of EMG signals from the time domain to the time-frequency domain was performed using the Short-Term Fourier Transform (STFT), which can be defined as follows [9]:

$$X(m,k) = \sum_{n=-\infty}^{\infty} x(n)w(n-m)e^{-j\frac{2\pi}{N}kn}, \qquad (1)$$

where $X(m,k)$ is the signal representation in the time-frequency domain, $x(n)$ represents the sequence of input samples, w is a window function, n is sample order, k is the order of the spectral component, m is the window interval centered around zero, and N is the number of discrete frequencies.

This transformation made it possible to obtain signal power information for individual frequencies in the form of spectrograms, which can be defined as [10]:

$$spectrogram = |X(m,k)|^2. \qquad (2)$$

The spectrograms were received from individual signal segments of four investigated lengths using a Hamming window with a length of 50 samples and an overlap of 34 samples.

Finger movement recognition

A convolutional neural network (CNN) was designed to classify spectrograms representing 13 finger movements. Therefore, the spectrograms represented input samples, measuring $64\times L\times CH$, where L = 4, 10, 14, 16 for 100 ms, 200 ms, 250 ms, and 300 ms segment lengths, respectively, and CH denotes the number of channels. The *z-score* method was used to normalize data based on training samples, where the standardized value x' is defined as follows:

$$\boldsymbol{x}' = \frac{\boldsymbol{x}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}, \qquad (3)$$

where x denotes the value of the original sample, μ is the mean value and σ is the standard deviation.

The utilized CNN architecture (Fig. 5) for the feature extraction stage consists of four *convolution* layers along with an equal number of *max-pooling* layers. For convolutional layers, the number of filters is 32, 64, 128, and 384, with a kernel size of 3×3. The *rectified linear unit* (*ReLU*) activation function is applied next. The dimensions of the feature map produced by convolution operations are reduced through pooling operations based on maximum value with a pooling size of 2×2 and stride 2. After the second and fourth *max-pooling* layer, a *dropout* layer with a probability of 0.1 is included. The classification stage consists of the first fully connected layer with 32 neurons and the second fully connected layer with a *softmax* activation function. Therefore, the output was created as a vector describing the probability distribution of the input spectrograms corresponding to each of the 13 defined classes.

The *Adam* optimizer was used as an optimization method. The training hyperparameters are set as follows; *mini-batch size* to 512, a number of *epochs* to 50, and *learning rate* to 0.001. Training and validation samples were shuffled in every epoch. Output model

Lekar a technika – Clinician and Technology 2021, vol. 51(1–4), pp. 15–20, DOI: 10.14311/CTJ.2021.1.02
ISSN 0301-5491 (Print), ISSN 2336-5552 (Online)

corresponding to the training iteration with the lowest validation cross-entropy loss.

A 5-fold cross-validation (CV) method was applied to evaluate the predictive performance of the classification models in terms of different properties of the input data. The original dataset was divided into 80% training data, 20% test data (for an overall evaluation of the most accurate model), and 20% of the training data, hereinafter referred to as validation data, were used for cross-validation assessment. The training and validation processes were performed via Deep Learning Toolbox in MATLAB R2021b using the GPU NVIDIA GeForce GTX 1660 (Nvidia Corporation, USA).

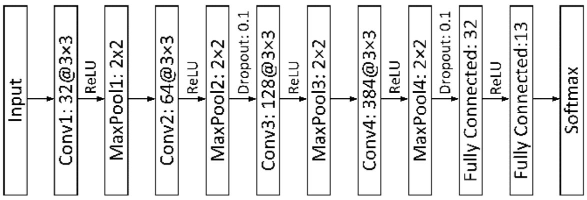

Fig. 5: Architecture of utilized CNN.

Results and Discussion

The classification results corresponding to the validation subtest of the spectrograms representing the 13 finger movements for the different segment lengths are depicted in Fig. 6. The Average accuracies across the 5-fold CV with standard deviation in the form of boxplots are reproduced. It can be observed that with increasing segment length, the average accuracy rate increases, from (81.1 ± 0.5)% for 100 ms to (95.4 ± 0.3)% for 300 ms. Thus, it can be stated that the more extended segment contains more representative information about the movement of the individual finger movement.

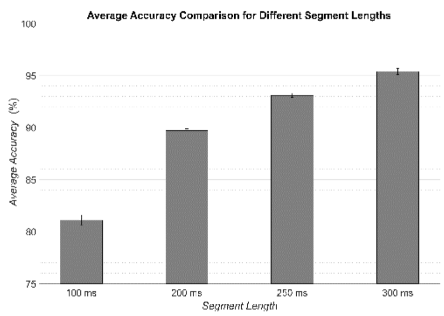

Fig. 6: Comparison of average accuracy (5-fold CV) comparison for different segment lengths.

Subsequently, the influence of individual channels and groups of these channels was investigated for the information corresponding to the most successful segment length of 300 ms. The box plots describe the results of this experiment in Fig. 7. A total of 15 combinations of channels were used, which can be divided into the following groups: one channel, two channels, three channels, and all channels. When using only one and two channels, satisfactory results were not achieved. The average accuracy is less than 55% for single-channel and using two channels is lower than 85% for all cases. The classification accuracy for the three-channel configuration ranges from (90.4 ± 0.6)% (channels CH2, CH3, and CH4) to (91.6 ± 0.6)% (channels CH1, CH2, and CH3). The case for all four electrodes has been evaluated previously.

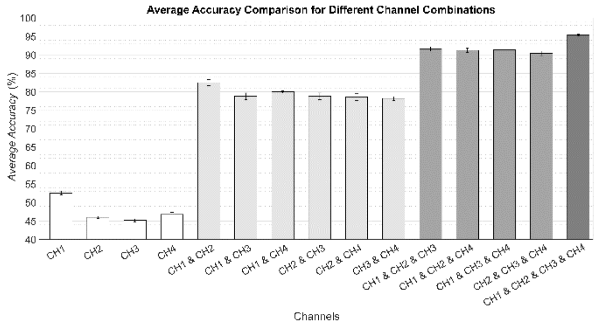

Fig. 7: Comparison of average accuracy (5-fold CV) for different channel combinations.

Finally, the CNN model was trained for the input conditions achieving the highest results, i.e., 300 ms segment length and all channels. This model was tested using a test subset (20% of original data). Fig. 8 depicts the graphical representation of these results in the form of a confusion matrix. The rows reflect the output (predicted) classes, while the columns provide information about the target (true) classes based on the labeling of the input data. The diagonal cells represent the number of correctly classified input data for individual finger movements. Cells outside the diagonal reflect the observations that have been improperly allocated to the given class. The last column denotes classification precision for individual classes, and the last row represents classification recall (sensitivity). The cell in the lower right corner contains acquaintance about the overall classification accuracy.

The overall classification accuracy reaches 95.5% on the test subset. It can be stated that the classification of the little finger flexion movements reached the highest precision of 98.1%. Contrariwise, the lowest precision value is achieved for the extension of this finger, namely 92.1%. The highest recall value of 99.9% belongs to the relaxed hand class, and the lowest recall of 92.7% is achieved for the ring finger flexion. Most samples of little finger extension were incorrectly classified as ring finger flexion. This fact may find a correlation in the difficulty of performing given movements across participants.

Table 1 reflects certain selected previous studies dealing exclusively with the classification of gestures based on the measurement of EMG signals using machine learning algorithms. Direct comparison of the results of different research is demanding as a consequence of methodological reasons. Therefore, studies examining only finger movements with a comparable number of subjects are mentioned.

Lekar a technika – Clinician and Technology 2021, vol. 51(1–4), pp. 15–20, DOI: 10.14311/CTJ.2021.1.02
ISSN 0301-5491 (Print), ISSN 2336-5552 (Online)

The number of finger movements investigated and classified in this research is more compared to previous studies. Another fact is that compared to others, only four channels were used to classify up to 13 comparable classification accuracy results were still achieved. The utilization of a low number of measured movements. Even the reduction of channels to three channels has a significant effect on the complexity reduction of the measuring system, increasing the comfort and also decreasing the intrusiveness of the measured subject. Another advantage is the reduction of computing power requirements. The presented approach can be applied in a wide range of real applications, such as rehabilitation, solving conditions after spinal cord injury, prosthetic limbs control, robotic arm control, work in dangerous conditions, computer mouse control, keyboard control, etc.

Fig. 8: Confusion matrix for the test subset. 1—Closed Hand, 2—Opened Hand, 3—Relaxed Hand, 4—Thumb Extension, 5—Index Finger Extension, 6—Middle Finger Extension, 7—Ring Finger Extension, 8—Little Finger Extension, 9—Thumb Flexion, 10—Index Finger Flexion, 11—Middle Finger Flexion, 12—Ring Finger Flexion, 13—Little Finger Flexion.

Table 1: Comparison of different studies that investigated finger movement classification based on EMG signals.

Reference	Subjects	Channels	Movements	Window Length	Sampling rate	Classifier	Result
[11]	5	7	11	200 ms	1500 Hz	ANN	90.52%
[12]	10	8	6	N/A	200 Hz	DNN	95% OFF 92% ON
[13]	4	8	12	250 ms	200 Hz	CNN	94.9 %
[14]	5	6	8	132 ms	2000 Hz	CNN	97.5%
[15]	5	6	8	125 ms	2000 Hz	PNN	92.2%
[16]	11	1	4	300 ms	1024 Hz	TSVM	93% H (11) 81% A (1)
Proposed	5	4	13	300 ms	1000 Hz	CNN	95.5%

ANN—Artificial Neural Network, DNN—Deep Neural Network, PNN—Probabilistic Neural Network, TSVM—Twin Support Vector Machine, OFF—Offline, ON—Online, H—Health, A—Amputee.

Lekar a technika – Clinician and Technology 2021, vol. 51(1–4), pp. 15–20, DOI: 10.14311/CTJ.2021.1.02
ISSN 0301-5491 (Print), ISSN 2336-5552 (Online)

Conclusion

This research presents the classification of 13 finger movements, including a relaxed hand condition, based on EMG signals acquired by four channels. Five healthy volunteers participated in the study, wherein the EMG data were measured over three different days from the upper limb. The obtained EMG records were segmented into segments of four lengths, transformed into the time-frequency domain by spectrograms, and subsequently classified using the CNN algorithm. The best validation result was obtained for data corresponding to a segment length of 300 ms using the 5-fold CV method, while the final model achieved 95.5% accuracy. Therefore, the accuracy of the classification of such spectrograms was further investigated with combinations of reduced channel numbers. Based on the results, it can be concluded that the use of one or two channels was not satisfactory using the presented procedures. However, input data classification accuracy of over 90% was achieved using only three channels.

Nevertheless, there is still oddments space for global results improvements by selecting more optimal hyperparameters of the classification model and including a larger dataset with more exact EMG information about finger movements. With some adjustments, the presented approach can be used for real-time finger movements classification and accordingly utilized in a specific application.

Acknowledgment

This publication has been produced with the support of the Integrated Infrastructure Operational Program for the project: Creation of a Digital Biobank to support the systemic public research infrastructure, ITMS: 313011AFG4, co-financed by the European Regional Development Fund.

A preliminary version of the results published in this article was presented at the Trends in Biomedical Engineering 2021 conference.

References

[1] Webster JG. Medical Instrumentation: Application and Design. 4th edition. USA: Wiley; 2010. 736 p. ISBN: 0471676004.

[2] De Luca CJ. Electromyography. USA: Wiley; 2006. p. 98–109. (Webster JG, editor. Encyclopedia of Medical Devices and Instrumentation, vol. 3. 2nd ed. 598 p.). ISBN: 0471263583.

[3] Shi WT, Lyu ZJ, Tang ST, Chia TL, Yang CY. A bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study. Biocybern Biomed Eng. 2018;38(1):126–35. DOI: 10.1016/j.bbe.2017.11.001

[4] Kumar B, Paul Y, Jaswal RA. Development of EMG Controlled Electric Wheelchair Using SVM and kNN Classifier for SCI Patients. Commun Comput Inf Sci. 2019 Sep;1076:75–83. DOI: 10.1007/978-981-15-0111-1_8

[5] Tateno S, Liu H, Ou J. Development of sign language motion recognition system for hearing-impaired people using electromyography signal. Sensors. 2020 Oct 14;20(20):5807. DOI: 10.3390/s20205807

[6] Rahim MA, Shin J. Hand movement activity-based character input system on a virtual keyboard. Electronics. 2020 May 8;9(5):774. DOI: 10.3390/electronics9050774

[7] David RL, Cristian CL, Humberto LC. Design of an electromyographic mouse. Proceedings of the 20th Symposium on Signal Processing, Images and Computer Vision; 2015 Sep 2–4; Bogota, Colombia. IEEE; 2015 Nov 19. pp. 1–8. DOI: 10.1109/STSIVA.2015.7330450

[8] Li X, Zhou Z, Liu W, Ji M. Wireless sEMG-based identification in a virtual reality environment. Microelectronics Reliability. 2019 Jul;98:78–85. DOI: 10.1016/j.microrel.2019.04.007

[9] Yildiz A, Zan H, Said S. Classification and analysis of epileptic EEG recordings using convolutional neural network and class activation mapping. Biomed Signal Process Control. 2021 Jul;68:102720. DOI: 10.1016/j.bspc.2021.102720

[10] Mandhouj B, Cherni MA, Sayadi M. An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis. Analog Integr Circuits Signal Process. 2021 Feb 15;108:101–10. DOI: 10.1007/s10470-021-01805-2

[11] Zhang Z, Yu X, Qian J. Classification of finger movements for prosthesis control with surface electromyography. Sensors and Materials. 2020 Apr 30;32:1523–32. DOI: 10.18494/SAM.2020.2652

[12] Naseer N, Ali F, Ahmed S, Iftikhar S, Khan RA, Nazeer H. EMG Based Control of Individual Fingers of Robotic Hand. Proceedings of the 2018 International Conference on Sustainable Information Engineering and Technology; 2018 Nov 10–12; Malang, Indonesia. IEEE; 2019 Apr 18. pp. 6–9. DOI: 10.1109/SIET.2018.8693177

[13] Stephenson RM, Chai R, Eager D. Isometric Finger Pose Recognition with Sparse Channel SpatioTemporal EMG Imaging. Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2018 Jul 18–21; Honolulu, HI, USA. IEEE; 2018 Oct 29. pp. 5232–5. DOI: 10.1109/EMBC.2018.8513445

[14] Fu J, Cao S, Cai L, Yang L. Finger Gesture Recognition Using Sensing and Classification of Surface Electromyography Signals With High-Precision Wireless Surface Electromyography Sensors. Front Comput Neurosci. 2021 Nov 11;15:770692. DOI: 10.3389/fncom.2021.770692

[15] Fu J, Xiong L, Song X, Yan Z, Xie Y. Identification of finger movements from forearm surface EMG using an augmented probabilistic neural network. Proceedings of the IEEE/SICE International Symposium on System Integration; 2017 Dec 11–14; Taipei, Taiwan. IEEE; 2018 Feb 5. pp. 547–52. DOI: 10.1109/SII.2017.8279278

[16] Kumar DK, Arjunan SP, Singh VP. Towards identification of finger flexions using single channel surface electromyography – able bodied and amputee subjects. J Neuroeng Rehabil. 2013 Jun 7;10:50. DOI: 10.1186/1743-0003-10-50

Ing. Ivana Kralikova
Department of Electromagnetic and Biomedical Engineering
Faculty of Electrical Engineering and Information Technology
University of Zilina
Univerzitna 8215/1, SR-010 26 Zilina

E-mail: ivana.kralikova@feit.uniza.sk
Phone: +421 415 135 097