PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Use of artificial intelligence for study of the visual perception

Anastasiya Runnova, Vladimir Maksimenko, Maksim Zhuravlev

Anastasiya Runnova, Vladimir Maksimenko, Maksim Zhuravlev, "Use of artificial intelligence for study of the visual perception," Proc. SPIE 11067, Saratov Fall Meeting 2018: Computations and Data Analysis: from Nanoscale Tools to Brain Functions, 1106704 (3 June 2019); doi: 10.1117/12.2527705

Event: International Symposium on Optics and Biophotonics VI: Saratov Fall Meeting 2018, 2018, Saratov, Russian Federation

Use of artificial intelligence for study of the visual perception

Anastasiya Runnova[a,b], Vladimir Maksimenko[a], Maksim Zhuravlev[a,c]

[a]Innopolis University, Innopolis, Russia
[b]Yuri Gagarin State Technical University of Saratov, Saratov, Russia
[c]Saratov State University, Saratov, Russia

ABSTRACT

In this report we propose an approach based on artificial neural networks for the classification and recognition of various states of the human brain associated with the spatial perception of ambiguous images. Based on the developed numerical methodology and analysis of the experimental multi-channel EEG data, we create and optimize an artificial neural network to ensure the accuracy of the classification of EEG states of the brain in visual perception close to 100%. Different interpretations of ambiguous images produce different oscillatory patterns in the EEG of a person with similar characteristics for each interpretation.

Keywords: electroencephalogram, artificial neuronal network, ambiguous image

1. INTRODUCTION

Currently, the dynamics of neural networks of the human brain attracts the attention of researchers in the natural and human sciences.[1,2] An interdisciplinary approach will come closer to understanding brain riddles and a better understanding of the neural mechanisms underlying its dynamics, open prospects in the field of medicine, neurophysiology and neurology in the near future.[3,4] The study of various aspects of the functioning of the human brain is usually based on objective data acquired in the course of psycho-physiological and cognitive experimental work.[5]The most convenient and inexpensive method for recording brain signals in cognitive studies today is EEG. A study of nonlinear processes in the brain neural network during perception of bitistable images is very important for the understanding of both the visual recognition of objects and the decision-making process in human brain.[5–8] Despite the considerable efforts of many researchers, the basic mechanisms underlying the interpretation of such images are not yet sufficiently clear. Currently, it is only known that perception is the result of non-linear processes occurring in the distributed neural network of the occipital, parietal and frontal cortical regions of the brain.[9,10] However, the question remains how the interpretation of the image affects the human EEG. Earlier, we demonstrated the effectiveness of the artificial neural network method in classifying EEG oscillations. It is well-known that using the different frequency bands of EEG signals leads to find additional information on brain dynamics.[11–13] In the paper we propose an approach based on artificial neural networks for classification and recognition of different human brain states related to spatial perception of ambiguous images.

2. EXPERIMENTAL SETUP AND SUBJECTS

Subjects were facing a display screen on which ambiguous images were displayed as visual stimulus (see Fig. 1). As an ambiguous image, we used the Necker cube,[5,14] a simple cube with transparent faces and visible ribs. A person with normal perception treats the Necker cube as a 3D-object thanks to a specific position of the cube's ribs. Visual bistability consists in the fact that this 3D-object can be treated as oriented in two different ways, especially if different ribs of the Necker cube are drawn with different intensity. Specifically, the contrast of the three middle lines centered in the left middle corner, $g \in [0, 1]$, was used as a control parameter of displayed images. The values $g = 1$ and $g = 0$ correspond, respectively, to 0 (black) and 254 (white) pixels luminance of the middle lines, using the 8-bit grayscale palette for visual stimulus presentation. Therefore, we can define a contrast parameter as $g = b/254$, where b is the brightness level of the middle lines in the used 8-bit grayscale

Further author information: (Send correspondence to Maksim Zhuravlev)
Maksim Zhuravlev: E-mail: zhuravlevmo@gmail.com, Telephone: +7 8452 99 88 31

Saratov Fall Meeting 2018: Computations and Data Analysis: from Nanoscale Tools to Brain Functions,
edited by Dmitry E. Postnov, Proc. of SPIE Vol. 11067, 1106704 · © 2019 SPIE
CCC code: 0277-786X/19/$18 · doi: 10.1117/12.2527705

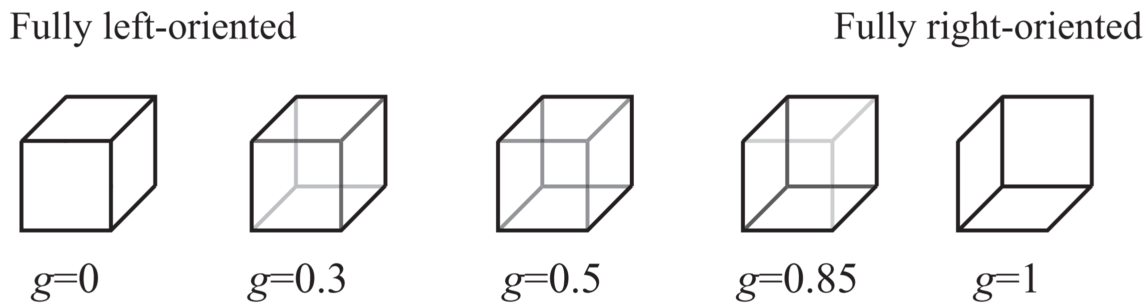

Figure 1. Examples of distinct Necker cube images with different wireframe contrasts characterized by control parameter g. The left-hand image with $g = 0$ corresponds to the fully left-oriented cube, while the right-hand image with $g = 1$ to the fully right-oriented cube

palette. The contrast of the three middle lines centered in the right middle corner was set to $(1 - g)$, and the normalized contrast of the six visible outer cube edges was fixed to 1.

The multi-channel EEG was recorded at 250-Hz sampling rate from $P = 19$ electrodes with two reference electrodes placed at standard positions of the 10–20 international system.[15] The EEG signals were filtered by a band-pass filter with cut-off points at 1 (HP) and 100 (LP) and a 50-Hz Notch filter. The electroencephalograph-recorder "Encephalan–EEGR–19/26" (Taganrog, Russian Federation) with multiple EEG channels and two-button input device (joystick) was used for amplification and analog-to-digital conversion of the EEG signals. Preliminary signal processing was provided by the original software for EEG registration artifact suppression. Machine learning algorithms were implemented with MATLAB. To demonstrate a grayscale stimulus we used a 24" BenQ LCD monitor with spatial resolution of 1080 pixels and refresh rate of 60 Hz. Each Necker cube drawn by black and gray lines was placed in the middle of a computer screen on a white background. The subjects were located at a distance of 70–80 cm from the monitor with visual angle approximately equal to 0.25 rad.

The experimental studies were performed in accordance to ethical standards[16] and approved by the local research ethics committee of the Yuri Gagarin State Technical University of Saratov. Twenty healthy subjects from a group of unpaid volunteers, male and female, between the ages of 20 and 45 with normal or corrected-to-normal visual acuity participated in the experiments. All subjects have provided informed consent before participating in the experiment.

3. ARCHITECTURE OF ANN AND ALGORITHM DESCRIPTION

The artificial neural network consists of a number of artificial neurons interconnected with each other by synaptic weights to form a network. There are many possible ANN architectures used for pattern recognition. In this study we use such class of ANN as a multilayer perceptron (MLP), which is characterized by signals traveling only in a forward direction (feedforward network) from left to right on a layer-by-layer basis.[17] MLP is the universal and popular class of ANN widely used for a broad range of applications including the classification problem.[18–23] In our case the classification problem is the recognition of two different multistable brain states corresponding to the perception of the ambiguous Necker cube image as left- or right-oriented.

Figure 2 shows the used ANN architecture of MLP with two hidden layers for EEG signal classification. First layer (*Input Layer – IL*) contains $P = 19$ inputs corresponding to 19 EEG channels. For each p-th input we used the functional EEG signal $s_p(t)$ with 1-s duration (250-sample time series) from p-th channel registered for the case of left- or right-oriented cube image perception. The signals from each input are fed to all computational nodes (artificial neurons) of the first hidden layer (*Hidden Layer 1 – HL1*) with the number of neurons equal to H_1. The resulting outputs of the first hidden layers are, in turn, applied to the second hidden layer (*Hidden Layer 2 – HL2*) with the same type of the computational nodes. The number of nodes in the second hidden layer is equal to H_2. Finally, the output signals of the second hidden layer neurons emerge at the single neuron of the output layer (OL). Since our classification problem is the recognition of two brain states by means of the analysis of the 19-channel EEG data set, the ANN contains only one output neuron, the output value of which classifies the current brain state corresponding to either left- or right-oriented cube interpretation.

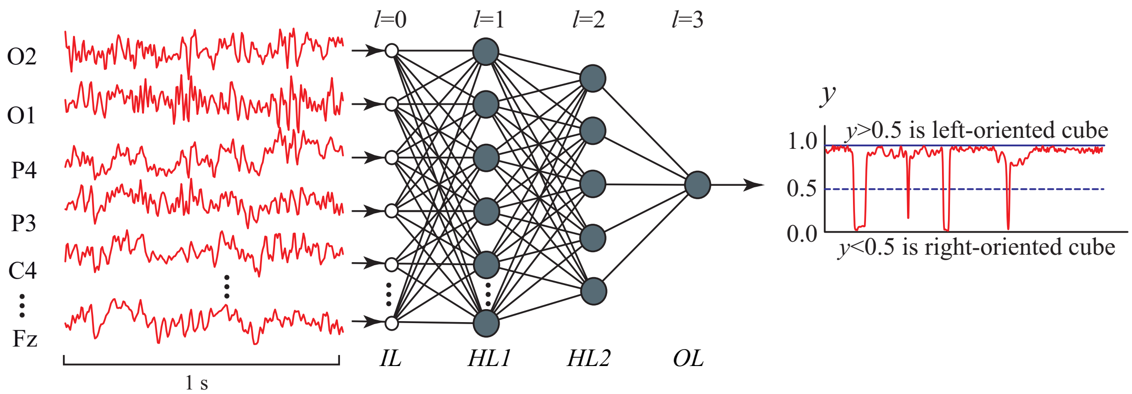

Figure 2. MLP architecture with two hidden layers in EEG signal classification. IL ($l = 0$) is the input layer, HL1 and HL2 are the first ($l = 1$) and second $l = 2$) hidden layers, respectively, whose nodes (artificial neurons) are characterized by nonlinear activation function (2), and OL is the output layer ($l = 3$) consisting of one artificial neuron with linear activation function (3). The number of inputs is $H^0 = P = 19$, the numbers of nodes within the hidden layers are $H^1 = P$ and $H^2 = 5$, respectively, and the number of output nodes is $H^3 = 1$

The evolution of an artificial neuron network is described by the following mathematical model[24]

$$u_i^l(t) = F^l \left(\sum_{p=1}^{H^{l-1}} w_{pi}^l u_p^{l-1}(t) - \theta_i^l \right), \tag{1}$$

where H^l is the number of neurons in the l-th layer (a layer with $l = 0$ is supposed to be the input layer), $u_i^l(t)$ is the output signal of the i-th neuron belonging to the l-th layer ($u_i^0(t)$ are signals from analyzed EEG channels), $\mathbf{W}^l = \{w_{pi}^l\}$ is the weight matrix of the l-th layer which dimension is $(H^{l-1} \times H^l)$, and w_{pi}^l ($p = 1, \ldots, H^{l-1}$, $i = 1, \ldots, H^l$) are the synaptic weights of input signals for the i-th neuron of the l-th layer, $\mathbf{\Theta}^l = \{\theta_i^l\}$ is the threshold vector for neurons of the l-th layer,

$$F^l(\eta) = f(\eta) = \frac{1}{1 + \exp(-\eta)} \tag{2}$$

is the nonlinear logistic activation function for neurons of hidden layers $l = 1, 2$, and

$$F^3(\eta) = \eta \tag{3}$$

is the linear activation function for the output layer ($l = 3$).

A class of recognized objects can be characterized by the mean squared value of output signal $u(t) = u_1^3(t)$:

$$y = \sqrt{\frac{1}{T} \int_0^T (u(t))^2 \, dt}. \tag{4}$$

Since the input signals $u_p^0(t)$ are trials $s_p(t_i)$ ($p = 1, \ldots P$, $t_i = i\Delta t$, $i = 1, \ldots N$) with the length T consisting of $N = 250$ samples ($T = 1$ s, $\Delta t = T/N$), Eq. (4) can be rewritten in the form

$$y = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (u(t_i))^2}. \tag{5}$$

For the left-oriented Necker cube perception the mean squared value of the output signal is supposed to be $y \geq 0.5$ and for the right-oriented cube $y < 0.5$.

The unknown matrices $\mathbf{W}^l$ and vectors $\mathbf{\Theta}^l$ can be obtained during the learning process by minimizing the classification error criterion:

$$\mu = \sqrt{\frac{1}{K}\sum_{k=1}^{K}(d_k - y_k)^2}, \qquad (6)$$

where K is the total number of objects in the training set, y_k is the mean squared value of the output signal calculated for the k-th object using Eq. (5), d_k is a desired output value of y_k which we wish the MLP learns ($d_k = 1$ corresponds to the left-oriented cube perception and $d_k = 0$ to the right-oriented one). To find unknown parameters of ANN we used the Levenberg-Marquardt algorithm (LMA).[25] By differentiating the error criterion (6) with respect to the unknown parameters, the LMA method gives better results in comparison with other optimization methods, but requires more computational time to determine the unknown parameters. For the learning process we created a data set consisting of 70 single trials with 1-s duration (250 samples) which were randomly selected from EEG records obtained from one volunteer. This data set consisted of 35 trials for each orientation of the Necker cube images with different contract parameters g. For a more reliable assessment of the result of ANN learning, we repeated the training procedure many times (1000 learning cycles in total). As a consequence, we obtained 1000 ANNs with different parameters and different values of classification error μ.

4. RESULTS

The recognition accuracy of the brain states classification during visual perception of ambiguous images (left-/right-oriented perception) for each of 12 subjects used for training ANN are shown in Fig. 3. To analyze the classification accuracy we took the remaining part of the EEG that was not used for training, i.e. about 280 EEG trials of the registered brain states after image demonstrations.

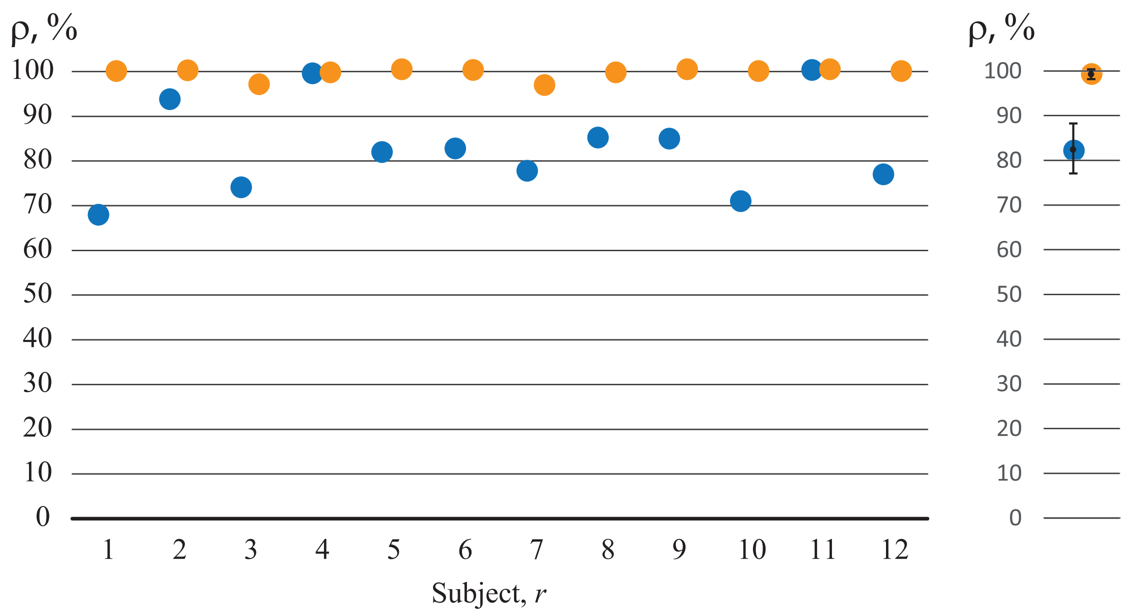

Figure 3. Recognition accuracy for all 12 subjects. Left blue columns represent accuracy for each subject using ANN trained on his/her own EEG ($h = r$), while right orange columns show accuracy using ANN trained for subject 4 ($h = 4$). The right panel represents the data averaged over all subjects under study.

We started the analysis of our classification algorithm from the training ANN for each subject under study. The training data set was formed individually for each subject and the optimal set of ANN parameters $\Gamma_r = (\mathbf{W}_r^1, \mathbf{W}_r^2, \mathbf{W}_r^3, \mathbf{\Theta}_r^1, \mathbf{\Theta}_r^2, \theta_r^3)$ ($r = 1, \ldots 12$ being the subject number) was obtained for classification of the brain states of subject r. In this case the mean accuracy for all 12 subjects was close to $83 \pm 5\%$ (mean $\pm$ S.D.) (left blue column in the right panel in Fig. 3). The recognition accuracy for every subject, shown in the blue left columns in the left panel of Fig. 3, vary between 68 and 100%.

It is remarkable that two subjects ($r = 4$ and $r = 11$) demonstrated recognition accuracy of classification of image perception close to 100%. When we used ANNs trained on these subjects for other subjects, we obtained much higher accuracy than if using their own ANNs. These results are shown in the right orange columns in Fig. 3. Using ANN with parameters Γ_4 evaluated for subject $r = 4$ the accuracy of classification was close to 100% for almost all subjects, only subjects with $r = 3$ and $r = 7$ demonstrate $\rho = 97\%$. Thus, we can conclude that features of EEG oscillatory patterns corresponding to perception of the left- or right-oriented cubes are typical for all subjects and a single ANN trained on the EEG data set of one person can classify with high accuracy the corresponding brain states of a large group of people.

5. CONCLUSION

In this paper we have proposed to use an artificial neuronal network for the classification and automatic recognition of human brain states associated with perception of ambiguous images. We have optimized the ANN architecture using obtained experimental data and have shown that it is possible to achieve close to 100% accuracy in the classification of the EEG patterns during perception of ambiguous images. We have found particular features of the EEG oscillatory patterns corresponding to different interpretations of the Necker cube, typical for all subjects, so that a single ANN trained on the EEG data set of one person can classify with high quality the corresponding brain states of a large group of people.

We firmly believe that the significance of our results is not limited to visual perception of the Necker cube images. We are sure that the proposed experimental approach and developed computational technique based on the ANN will be useful for studying and classifying different brain states analyzed by means of EEG and MEG recordings and can stimulate future research in the field of cognitive and pathological brain activity. The developed approach provides a solid experimentally approved basis for further understanding brain functionality. The rather simple way to quantitatively characterize brain activity related to perception of ambiguous images seems to be a powerful tool, which may be used in neurotechnology, e.g., for the brain-computer interface task,[26,27] and in medicine for diagnostic and prognostic purposes.[28] We expect that our work will be interesting and useful for scientists carrying out interdisciplinary research at the cutting edge of physics, mathematics, neurophysiology and medicine.

6. ACKNOWLEDGMENTS

This work has been supported by the Ministry of Education and Science of Russia (project 3.861.2017/4.6). V.A.M. acknowledges individual support from the Russian Presidents Council grant for the state support of young Russian PhD scientists (project MK-992.2018.2).

REFERENCES

[1] Hramov, A. E., Koronovskii, A. A., Makarov, V. A., Pavlov, A. N., and Sitnikova, E., [*Wavelets in Neuroscience*], Springer Series in Synergetics, Springer, Heidelberg, New York, Dordrecht, London (2015).

[2] Bear, M. F., Connors, B. W., and Paradiso, M. A., [*Neuroscience. Exploring the brain*], Woters Kluwer (2015).

[3] Maksimenko, V. A., Runnova, A. E., Zhuravlev, M. O., Protasov, P., Kulanin, R., Khramova, M. V., Pisarchik, A. N., and Hramov, A. E., "Human personality reflects spatio-temporal and time-frequency eeg structure," *PloS one* **13**(9), e0197642 (2018).

[4] Maksimenko, V. A., Heukelum, S., Makarov, V. V., Kelderhuis, J., Lüttjohann, A., Koronovskii, A. A., Hramov, A. E., and Luijtelaar, G., "Absence seizure control by a brain computer interface," *Scientific Reports* **7**(1), 2487 (2017).

[5] Pisarchik, A. N., Jaimes-Reategui, R., Magallón-Garcia, C. D. A., and Castillo-Morales, C. O., "Critical slowing down and noise-induced intermittency in bistable perception: bifurcation analysis," *Biological Cybernetics* **108**(4), 397–404 (2014).

[6] Leopold, D. A. and Logothetis, N. K., "Multistable phenomena: changing views in perception," *Trends in Cognitive Sciences* **3**(7), 254–264 (1999).

[7] Blake, R. and Logothetis, N. K., "Visual competition," *Nature Reviews. Neuroscience* **3**, 13–21 (2002).

[8] Hramov, A. E., Frolov, N. S., Maksimenko, V. A., Makarov, V. V., Koronovskii, A. A., Garcia-Prieto, J., Antón-Toro, L. F., Maestú, F., and Pisarchik, A. N., "Artificial neural network detects human uncertainty," *Chaos: An Interdisciplinary Journal of Nonlinear Science* **28**(3), 033607 (2018).

[9] Tong, F., Meng, M., and Blake, R., "Neural bases of binocular rivalry," *Trends in Cognitive Sciences* **10**(11), 502–511 (2006).

[10] Sterzer, P., Kleinschmidt, A., and Rees, G., "The neural bases of multistable perception," *Trends in Cognitive Sciences* **13**(7), 310–318 (2009).

[11] Sitnikova, E., Hramov, A. E., Grubov, V., and Koronovsky, A. A., "Time-frequency characteristics and dynamics of sleep spindles in wag/rij rats with absence epilepsy," *Brain Research* **1543**, 290–299 (2014).

[12] van Luijtelaar, E. L. M., Hramov, A. E., Sitnikova, E., and Koronovskii, A. A., "Spike - wave discharges in WAG/Rij rats are preceded by delta and theta precursor activity in cortex and thalamus," *Clinical Neurophysiology* **122**, 687–695 (2011).

[13] Pavlov, A. N., Hramov, A. E., Koronovskii, A. A., Sitnikova, Y. E., Makarov, V. A., and Ovchinnikov, A. A., "Wavelet analysis in neurodynamics," *Physics-Uspekhi* **55**(9), 845–875 (2012).

[14] Necker, L. A., "Observations on some remarkable phenomena seen in switzerland; and an optical phenomenon which occurs on viewing of a crystal or geometrical solid," *Philos. Mag.* **3**, 329–343 (1832).

[15] Niedermeyer, E. and da Silva, F. L., [*Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*], Lippincot Williams & Wilkins (2004).

[16] "World medical association (2000) declaration of helsinki: ethical principles for medical research involving human subjects," *The Journal of the American Medical Association* **284**(23), 3043–3045 (2000).

[17] Haykin, S., [*Neural Networks: A Comprehensive Foundation*], Pearson; 3d edition (2008).

[18] Haselsteiner, E. and Pfutscheller, G., "Using time-dependent neural networks for eeg classification," *IEEE Transactions on Rehabilitation Engineering* **8**, 457–463 (2000).

[19] Fontoura da, C. L. and Cesar, J. R. M., [*Shape analysis and classification: theory and practice*], Boca Raton, CRC Press (2001).

[20] Garrett, D., Peterson, D. A., Anderson, C. W., and Thaut, M. H., "Comparison of linear, nonlinear, and feature selection methods for eeg signal classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **11**(2), 141–144 (2003).

[21] Dias, N. S., Kamrunnahar, M., Mendes, P. M., Schiff, S. J., and Correia, J. H., "Comparison of eeg pattern classification methods for brain-computer interfaces," *Conf Proc IEEE Eng Med Biol Soc* , 2540 (2007).

[22] Hasan, M. R., Ibrahimy, M. I., Motakabber, S. M., and Shahid, S., [*Classification of Multichannel EEG Signal by Linear Discriminant Analysis*], ch. 0, 279–282, Springer International Publishing (2015).

[23] Maksimenko, V. A., Kurkin, S. A., Pitsik, E. N., Musatov, V. Y., Runnova, A. E., Efremova, T. Y., Hramov, A. E., and Pisarchik, A. N., "Artificial neural network classification of motor-related eeg: An increase in classification accuracy by reducing signal complexity," *Complexity* **2018** (2018).

[24] Yao, X., "Evolving artificial neural networks," *Proceedings of the IEEE* **87**(9), 1423–1447 (1999).

[25] Strutz, T., [*Data Fitting and Uncertainty (A practical introduction to weighted least squares and beyond)*], Springer, 2nd edition ed. (2016).

[26] Bell, C. J., Shenoy, P., Chalodhorn, R., and Rao, R., "Control of a humanoid robot by a noninvasive brain computer interface in humans," *Journal of Neural Engineering* **16**(5), 432–441 (2008).

[27] McFarland, D. J., Parvaz, M. A., Sarnacki, W. A., Goldstein, R. Z., and Wolpaw, J. R., "Prediction of subjective ratings of emotional pictures by eeg features," *Journal of Neural Engineering* **14**(1), 016009 (2017).

[28] Ovchinnikov, A. A., Luttjohann, A., Hramov, A. E., and Luijtelaar van, G., "An algorithm for real-time detection of spike-wave discharges in rodents," *Journal of Neuroscience Methods* **194**, 172–178 (2010).