Bistable Perception Analysis by Using Classification on EEG Signals

Amirreza Hatamipour, Arshak Rezvani, Mohammad Mahdi Mirrashid, and Ahmadreza Tavana

Student, Department, Sharif University of Technology

This manuscript was compiled on January 31, 2023

This paper conducts a novel classification method for a bistable visual stimuli or bistable perception. bistable perception is a perceptual phenomenon in which an observer experiences an unpredictable sequence of spontaneous subjective changes. bistable stimuli have the possibility of being perceived in two different ways. Due to their physical characteristics, these visual stimuli allow two different perceptions, associated with right-left and left-right modulating processes. This experimental article, a task is designed which the subject press a key whenever his perception changes. The identification of regions in the brain that are associated with a conscious perception of visual stimuli is one of the aims in this experiment. A CNN model is deployed on the data which is recorded by a 10-20 EGG recording, while the subject was doing the designed task.

Bistable Perception | EEG Signals Classification 2 | visual stimuli

The perceptual phenomenon by which an observer perceives the same stimuli in two different ways is known as bistable perception (1, 2). Without any variations in the stimuli (in the case of a static image), or with movement or fluctuations (in the case of bistable or multistable dynamic visual stimuli), the observer's perception often alternates between two possible interpretations, because the stimuli offer several interpretation possibilities that cannot be perceived simultaneously (3). Given that bistable visual stimuli admit two (or more, in the case of Multistable images) possible precept, they can also be called ambiguous figures(4). These changes between one precept and the other are what make bistability arise. It is essentially caused by alterations in patterns while observing the bistable stimulus and is also understood as a variation in gestaltic processes of organization (5).

The phenomenon of bistability in vision leads to a neural activity that implies operative of diverse neural substrates and several integrated perceptual processes (6). The use of invasive electrophysiological techniques has been considered the most suitable instrument to find explanations from the point of view of perceptual neural activity and its underlying factors. Nevertheless, and given that these techniques require the insertion of electrodes through the skull, they turn out to be dangerous methods, to the point that they can cause infections and injuries (7). Taking that into account, functional magnetic resonance imaging - fMRI has become the most common technique to establish neural correlates with bistable perception in human subjects (8). Besides this technique, electroencephalography (EEG), magnetoencephalography (MEG) and transcranial magnetic stimulation (TMS) are also recognized as useful and widely used. Through these methods, neural activity is measured (indirectly) in groups of neurons, which suggests extreme caution and methodic discipline to generate interpretations (9).

In this experiment we try to answer some questions about bistability effect on the subject brain and then figure out which part of brain would have more activity whenever the subject's perception changes from left to right and wise versa. The following four question are answered in this paper are:

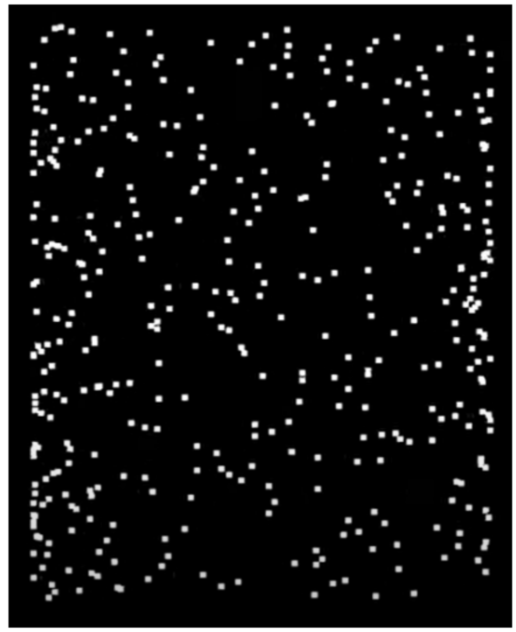

Fig. 1. The bistable stimuli task which is used in this study

Question 1: How can we detect bistable stimuli with EEG signals?
Question 2: Which part of brain and electrodes have the most response to the perception of this bistable stimuli?
Question 3: What is the accuracy of each classifier on the recorded EEG signal of this kind of bistable stimuli?
Question 4: How can tanscranial stimulations like tDCS or tACS effect on the EEG signal and results of the experiment?

Significance Statement

In this study, we try to show whether there is a significant change in our brain when we see a bistable stimulus or not. We try to analyze this task to address this question: which part of our brain has a more critical role than orders in causing this change in our perception? Answering this question helps scientists to improve their knowledge about visual bistable perception. Research in these fields helps people better recognize the brain's function.

This study aims to answer above questions using classification on recorded EEG signals while doing bistable perception task. The following figure is the bistable stimuli which is used in this study. A movie of this dots was played for the subject and he precept a cylinder (Figure. 1) that sometimes turn around from left to right and sometimes reverse of that. then the subject response to the stimuli by pressing a key on the keyboard whenever his perception of direction of the cylinder is changed and during that we recorded EEG signal of his brain.

Methods

In this study, we designed a task and record EEG signal of the subject and then we do a pre-processing by using Makato pipeline and then deploy a CNN model to the signals to figure out which part of the subject brain has the most response to the stimuli.

Task Design

We wrote a code with the Psycholab toolbox to show stimulus to our subject. Also, we recorded EEG signals from the subject during the stimulation. Our experiments contain fourteen trials. In each trial, we show a stimulus to our subject two times. Each time stimulus showed on the monitor for 85 seconds. Each stimulus contains some points that move around a cylinder(Figure. 2). During this time, our subject can select the orientation of his perception by clicking two keys. We saved time and directed the subject's thinking in order to analyze the EEG signal by using KbCheck. Our system records EEG signals from 30 channels with a frequency of 255 Hz.

Fig. 2. The Subject is doing the designed task and recording his EEG

Pre-Processing

For this section we made extensive use of the EEGLAB toolbox which is "an interactive Matlab toolbox for processing continuous and event-related EEG, MEG and other electrophysiological data" To start working with the recorded signals, the format of the raw EEG data was changed from TDMS to simple .set files to make them compatible with the standards of EEGLAB. 14 runs of the experiment were done on the subject, and each run consisted of a subject looking at a video twice, with a few seconds of rest in between. Each run was split into two separate datasets, resulting in 28 data files in total. Makoto's pipeline was used for preprocessing the EEG data. The steps of the pipeline are (from Makoto's preprocessing pipeline) 1. Change the option to use double precision 2. Import data 3. Downsample if necessary 4. High-pass filter the data at 1-Hz (for ICA, ASR, and CleanLine) 5. Import channel info 6. Remove line noise using CleanLine 7. Apply cleanrawdata to reject bad channels and correct continuous data using Artifact Subspace Reconstruction (ASR). 8. Interpolate all the removed channels 9. Re-reference the data to average 10. Run AMICA using calculated data rank with 'pcakeep' option (or runica using 'pca' option) 11. Estimate single equivalent current dipoles 12. Search for and estimate symmetrically constrained bilateral dipoles 13. Generate probabilistic IC labels using IClabel() for IC rejection 14. Epoch IC-rejected data to -1 to 2 sec to event onset (So far is the subject-level process) 15. Create final STUDY specifying full STUDY.design (This is the group-level process using STUDY)

Several minor changes were made to the steps above: The CleanLine method didn't work with our version of Matlab, and so a notch filter centered around frequencies 50 and 100 Hz was used instead. The results can be seen in (Figure. 3)

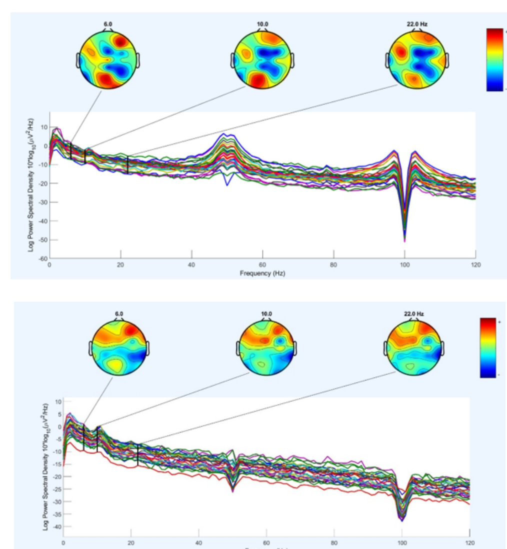

Fig. 3. EEG data of two example trials after applying the notch filter

ICA was applied to each dataset before extracting epochs. This is according to the guidelines of Makoto's pipeline, as the samples in each epoch would not be enough for an ICA decomposition.(Figure. 4)

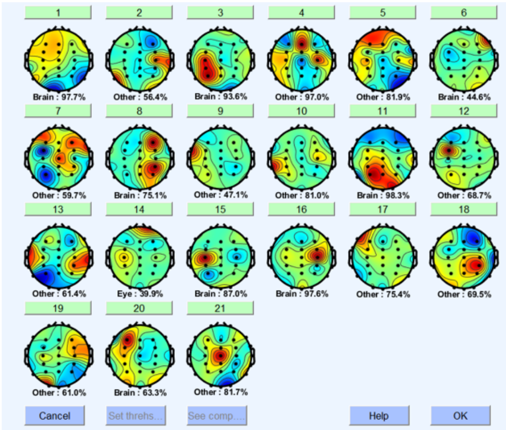

Fig. 4. An example of the resulting decomposition after using the ICLabel method

Also, epochs were extracted from -1 to 3 sec to event onset instead of the default -1 to 2 sec. This change was made to ensure that all the necessary information from a trial is kept in the epoch.(Figure. 5)

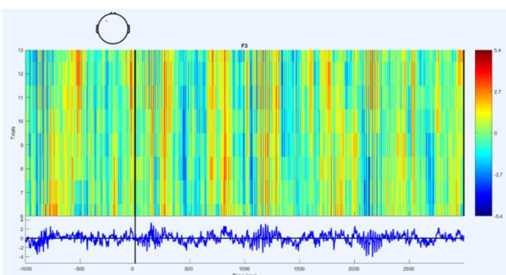

Fig. 5. ERP for the channel F3 of an example experiment run

EEG Signal Processing

After doing the pre-processing on the raw signal using the Makato pipeline, we want to find out which channels of EEG signal has the most amount of response when the subject switches, assuming a temporal-frequency coding happens in some particular channels, to find this channels the continuous wavelet transform of each trial for all channels is calculated, then averaged over each class, so we have 30 images for each class that contains the temporal-frequency content(channel wise), we find the absolute value of difference for each class then sum the differences of all values in the image and put them in a vector, this vector contains the amount of difference in temporal-frequency content for each class across the channels, then we plot the topographic map of this vector to see which regions of brain contribute the most the perception of subject.(Figure. 6)

The topographic map in (Figure. 6) shows that the channels CPz and FC4 contribute the most to the perception of the subject, from the atlas of brain the CPz channel is responsible for very detailed motor functions such as pressing buttons which is exactly the case here.

he wavelet transform of the best channel for two classes is shown in the (Figure. 7)

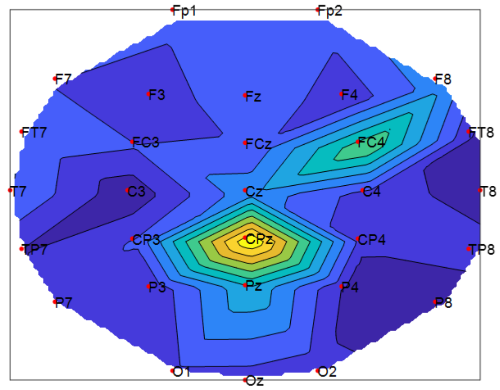

Fig. 6. The topographic map of the subject EEG and brain

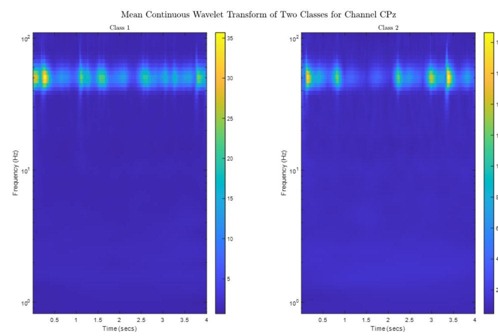

Fig. 7. Mean Continues Wavelet Transform of Two Classification for channel CPz

In the bandwidth of 35Hz to 70Hz there is an obvious coding happening and in the later stages an attempt is made to find this exact coding to classify the trials. Before starting the classification, the average time series for each class is shown in (Figure. 8) to find out if is there any possible ERDs in the EEG data.

For classification three methods are used: classification without feature extraction, classification using CNNs applied to the wavelet transform of the data, and classification using an ensemble method applied to features of the signal. The data was shuffled and 0.2 of the data (53 trials) was kept aside for testing the models, and the rest (208 trials) were used for training. The permutation was recorded in a Matlab file named 'indices' for reproducibility of the results.

Classification 1: No Feature Extracted

No feature extracted: for each trial the thousand time points for all channels is flattened into a 30*1000 vector, after normalization Principal Components Analysis is applied to reduce the number of features after that the data is classified using Support Vector Machine classifier. The first 20 PCs are used which will explain 78.5 percent variance of the data and the confusion matrix of trained model using five-fold cross validation is shown in (Figure. 9). As shown in the figure above this method performs poorly because it is taking the timepoints into features that are very noisy and are not coding anything particular that is detectable by a simple SVM, maybe an RNN on the times series of desired channels could detect the coding.

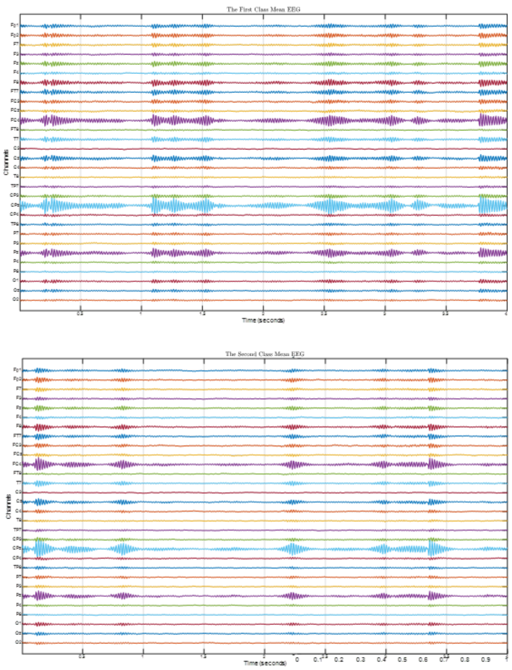

Fig. 8. The First Class Mean EEG

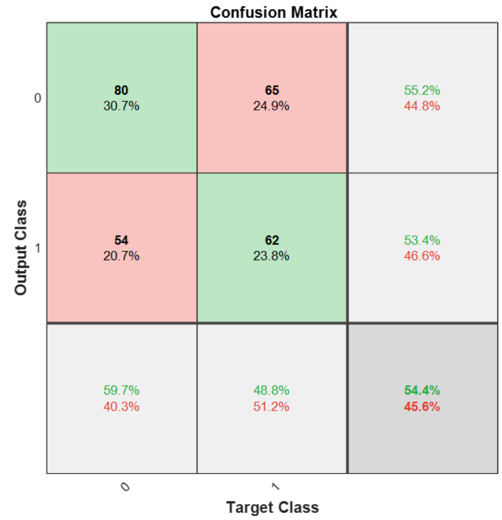

Fig. 9. Confusion Matrix of Classification 1

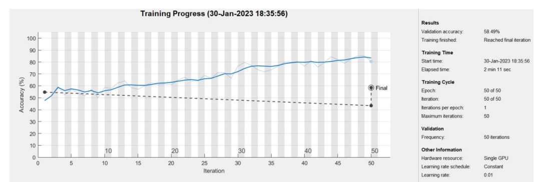

Fig. 10. CNN Result

Classification 2: Convolutional Neural Network

For the top 15 desired channels an image is constructed using wavelet transform(only the 35Hz to 71Hz frequency band is kept) , then these 15 images are concatenated to make a 37*1000*15 image, then these images are fed into a three-layer CNN with batch normalization and dropout layers. The network architecture is shown below. This technique tries to find the discriminative spatial patterns in the image constructed by the wavelet transform which means that tries to find the temporal-frequency coding that discussed earlier. At last, we want to find discriminative patterns for recognizing the perception of the subject, it comes natural to analyze and the accuracy of the CNN is shown in (Figure. 10)

The architecture of the CNN model is shown in (Figure. 11)

Classification 3: Ensemble Methods

This method performed better on the data because of the number of weak learners that boosted each other, for the features in this methos first the wavelet transform of the first 15 desired channels is computed then flattened into a vector, then two types of other features extracted for each channel times series are, Statistical Features: 1. Minimum Value 2. Maximum Value 3. Median Value 4. Variance 5. Arithmetic Mean 6. Standard Deviation 7. Normalized First Difference 8. Normalized Second Difference 9. Skewness 10. Log Energy Entropy 11. Shannon Entropy. Frequency Features: 1. Delta Band Power 2. Theta Band Power 3. Alpha Band Power 4. Beta Band Power 5. Gamma Band Power 6. Ratio of Band Power Alpha to Beta 7. Mean Energy.

Then PCA (because the number of features is enormous to keep the variance of the data 200 PCs are kept) is applied on the features to reduce the number of features. Then the default implementation of the ensemble model in Matlab was fitted on the data. The confusion matrix for this method is shown in (Figure. 12).

As shown in (Figure. 12) the accuracy on test data set is 60 percent which is very good compared to the other methods.

Tanscranial Stimulation

At first we try to find out what is tDCS and tACS. Transcranial direct current stimulation (tDCS) is a popular brain stimulation method that is used to modulate cortical excitability, producing facilitatory or inhibitory effects upon a variety of behaviors. There is, however, a current lack of consensus between studies, with many results suggesting that polarity-specific effects are difficult to obtain. Transcranial Direct-Current Stimulation (tDCS) is a portable, wearable brain stimulation technique that delivers a low electric current to the scalp. A fixed current between 1 and 2 mA is typically applied1. tDCS works by applying a positive (anodal) or negative (cathodal) current via electrodes to an area. tDCS is a neuromodulation technique that produces immediate and lasting changes in brain function. The position of the anode and cathode electrodes on the head is used to set how current flows to specific brain regions. The current delivered by tDCS is NOT strong enough to trigger an action potential in a neuron; instead its “sub-threshold” changes the pattern of already activity neurons.(10)

Transcranial alternating current stimulation (tACS) seems likely to open a new era of the field of noninvasive electrical stimulation of the human brain by directly interfering with cortical rhythms. It is expected to synchronize (by one single resonance frequency) or desynchronize (e.g., by the application of several frequencies) cortical oscillations. If applied long

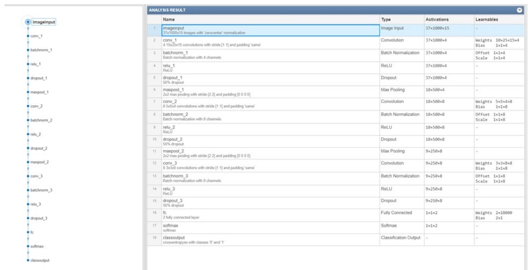

Fig. 11. CNN Architecture

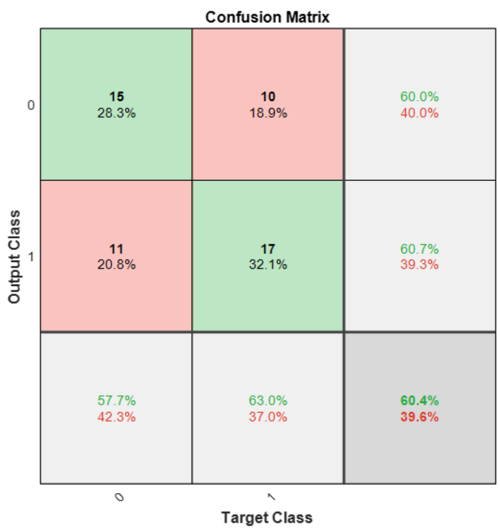

Fig. 12. Confusion Matrix of Classification 3

enough it may cause neuroplastic effects. In the theta range it may improve cognition when applied in phase. Alpha rhythms could improve motor performance, whereas beta intrusion may deteriorate them. TACS with both alpha and beta frequencies has a high likelihood to induce retinal phosphenes. Gamma intrusion can possibly interfere with attention. Stimulation in the “ripple” range induces intensity dependent inhibition or excitation in the motor cortex (M1) most likely by entrainment of neuronal networks, whereas stimulation in the low kHz range induces excitation by neuronal membrane interference. TACS in the 200 kHz range may have a potential in oncology.(11)

Transcranial Direct Current Stimulation, or tDCS, and transcranial Current Stimulation, or tACS, are close cousins when it comes to brain stimulation technology. Both tDCS and tACS are non-invasive (read “wearable”) devices that pass a little current to the brain to boost brain function and help treat brain diseases. Both are low-current, meaning the total voltage applied is about a 9V battery and the total current is a few mA. Both are considered well “tolerated” (read: does not hurt a lot) when a good device and technique is used. Though tACS is more likely to make you see imaginary white flashes, non-harmful “light shows” that scientists call phosphenes. Both require you to place electrodes on the head, which are then connected to a device. In fact, some devices can be switched between tDCS and tACS like this Soterix 1x1 tES device. Both tDCS and tACS can be upgraded to High-Definition: so HD-tDCS and HD-tACS respectively.

The key difference between tACS and tDCS is the waveform. tDCS uses a DC waveform. Think of the battery, the current coming out is just steady the whole time. In contract, tACS makes a sinusoidal waveform. Think of a wave that is constantly going up and down. Actually, with tACS, the direction of current flow switches as the wave swings one way and the other. Remember in tDCS people talk about anode and cathode (explained here) which is only possible because the current is always flowing one way. In tACS, people do not talk about anode and cathode because the direction of current flow is always switching back and forth. So when you set up tDCS you need to be very careful where you put the anode and cathode (see our montage guide). With tACS, there is no difference. With tACS, the wave can go up and down slowly (low frequency) or up and down fast (high-frequency). So with tACS, you need to specify the wave rate, as in 10 Hz tACS or 40 Hz tACS. Hz means how many times up and down per second. With DC there is no frequency to report (or you could say the frequency is zero). (12)

Results

Classification 1: No Feature Extracted
As shown in the (Figure. 9), this method performs poorly because it uses the timepoints as features that are very noisy and are not coding anything that is detectable by a simple SVM, maybe an RNN on the times series of desired channels could detect the coding.

Classification 2: Convolutional Neural Network
Various settings and layers were used for this method. The training results of the best performing architecture can be seen in (Figure. 10). We observe that the accuracy doesn't get far beyond the 50 percent mark, which seems to indicate that the wavelet transforms by itself might not have much predictive power for the two classes. But it's important to consider that the possible architectures for artificial neural networks are endless and other networks might have performed much better.

Classification 3: Ensemble Methods
This method outperformed the other methods used in this study. The features used in this method thus seem to have higher predictive power than features such as the wavelet transform of the signal. Many of them are simple statistical features, which shows that the most complex features might not work best in this task. The result of all of the classifications are shown in (Table 1).

Transcranial Stimulation
As it can be seen in (Figure. 10) CPz has higher response to the stimuli. We can test the result of a deploying tACS on this part for and see the result. The stimulation that we perhaps will have result is to deploy a tACS stimulation on CPz and see the functionality of this electrode just a 24 hours after that by doing the same experiment.

Discussion

As you can see in the previous sections, we reach a good result in dealing with detecting bistable perception from EEG signals. But, we have some limitations in this work. One of them is we need more data to extract features and train our models in order to reach higher precision. Another is we should employ other pre-processing steps to extract pure EEG

Table 1. Comparison accuracy of different methods of classification

Method	Best Validation Accuracy
1. Time series classification without feature extraction	55.2
2. CNN with the wavelet transform	58.5
3. Ensemble method with extra features	60

Best Accuracy of each model that we obtained in this study

signals. Solving these problems can lead to higher accuracy for this task.

References

1. Chik D. Kazanovich Y. Borisyuk, R. Visual perception of ambiguous figures: synchronization based neural models. *Biological Cybernetics*, 2009.
2. Eckardt J. Clément, G. Influence of the gravitational vertical on geometric visual illusions. *Acta Astronautica*, 2005.
3. Rinzel J. Rubin N. Moreno-Bote, R. Noiseinduced alternations in an attractor network model of perceptual bistability. *Journal of neurophysiology*, 2007.
4. van Ee R Brouwer, G. J. Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Research*, 2006.
5. Hupé J.-M. Pressnitzer, D. Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology*, 2006.
6. Ten Hove-M. W. Brammer M. Paré M. Munhall, K. G. Audiovisual integration of speech in a bistable illusion. *Current Biology*, 2009.
7. Rees G. Sterzer, P. Bistable perception and consciousness. *Encyclopedia of Consciousness*, 2009.
8. Bach M. Kornmeier, J. Bistable perception and consciousness. *Vision Research*, 2005.
9. Logothetis-N. K. Leopold, D. A. phenomena: changing views in perception. *Trends in cognitive sciences*, 1999.
10. Lugo CE Alexander ML, Alagapan S. Double-blind, randomized pilot clinical trial targeting alpha oscillations with transcranial alternating current stimulation (tacs) for the treatment of major depressive disorder (mdd). *Transl Psychiatry*, 2013.
11. Boros K. Poreisz C. Chaieb L. Terney D. Paulus W. Antal, A. Comparatively weak after-effects of transcranial alternating current stimulation (tacs) on cortical excitability in humans. *Brain stimulation*, 2008.
12. Paulus W. Ambrus, G. G. and A. Antal. Cutaneous perception thresholds of electrical stimulation methods: comparison of tdcs and trns. *Clin. Neurophysiol*, 2010.