BCI NU

NAZARBAYEV UNIVERSITY

Non-oddball ERP paradigms with joint temporal-frequency learning in convolutional neural network

Madina Saparbayeva | Master Degree of Computer Science

Min-Ho Lee | Lead Supervisor | Assistant Professor of Computer Science

Berdakh Abibullaev | Co-Supervisor | Assistant Professor of Robotics

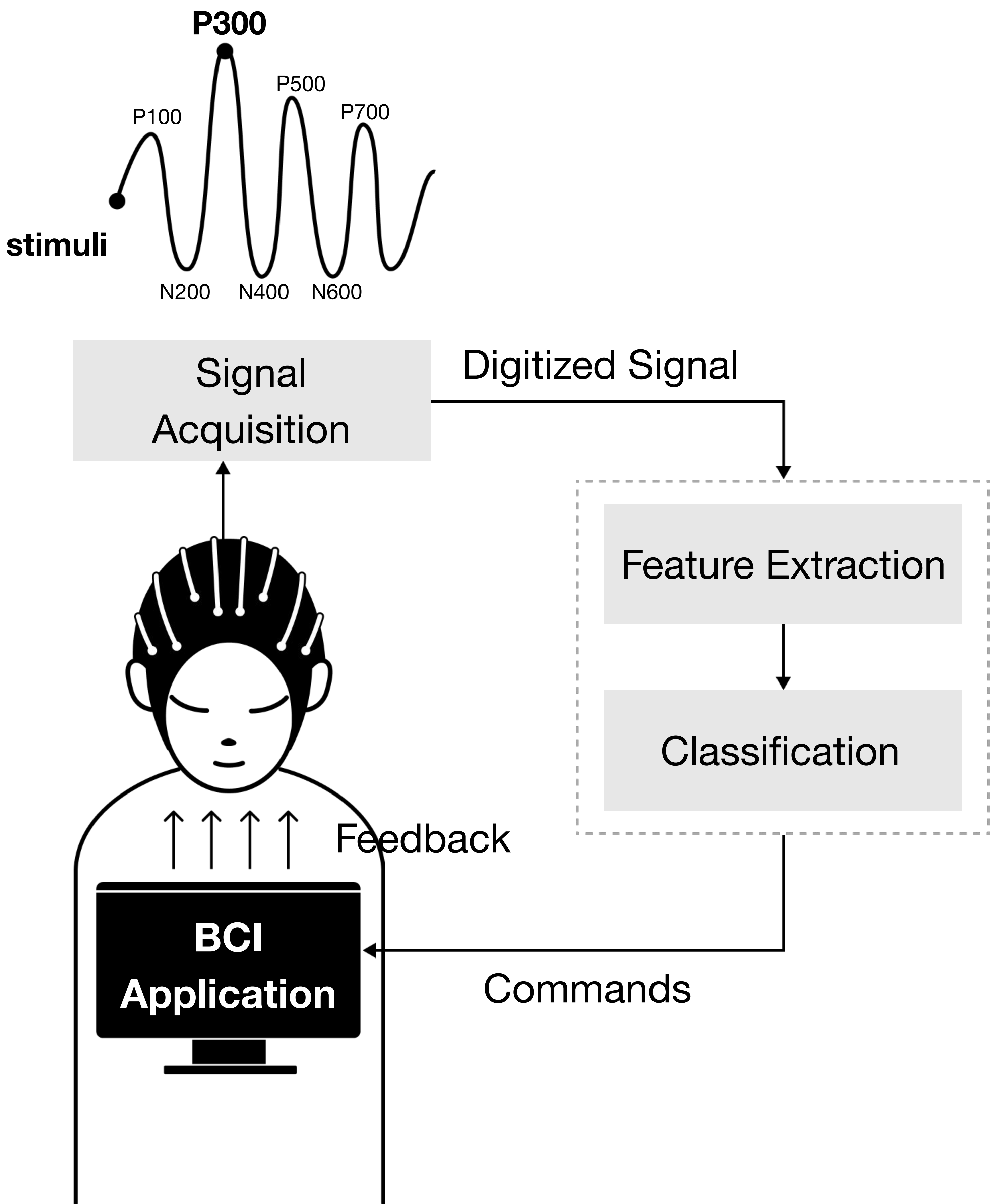
25-04-2022 | Nazarbayev University | School of Engineering and Digital Science

Introduction

A brain-computer interface (BCI) [1] allows the user to control an external device for the users with diagnoses such as locked-in-syndrome, paralysis, or spinal cord injury.

Electroencephalography (EEG) is widely used for BCI-purpose because of its non-invasive, low-risk, and easy-to-use method [2].

Event-Related Potential across the parieto-central area of the skull that usually occurs around 300 ms after stimuli presentation called P300 is larger after the target stimulus.



Oddball Paradigm

The oddball paradigm is an experiment where an “odd” event in a stream of typical events would elicit a distinct scalp-recorded potential pattern while the subject was concentrating on the stream of external auditory or visual stimuli.

P300 is evoked when the target stimuli is **unpredictable** while user focus on typical events.

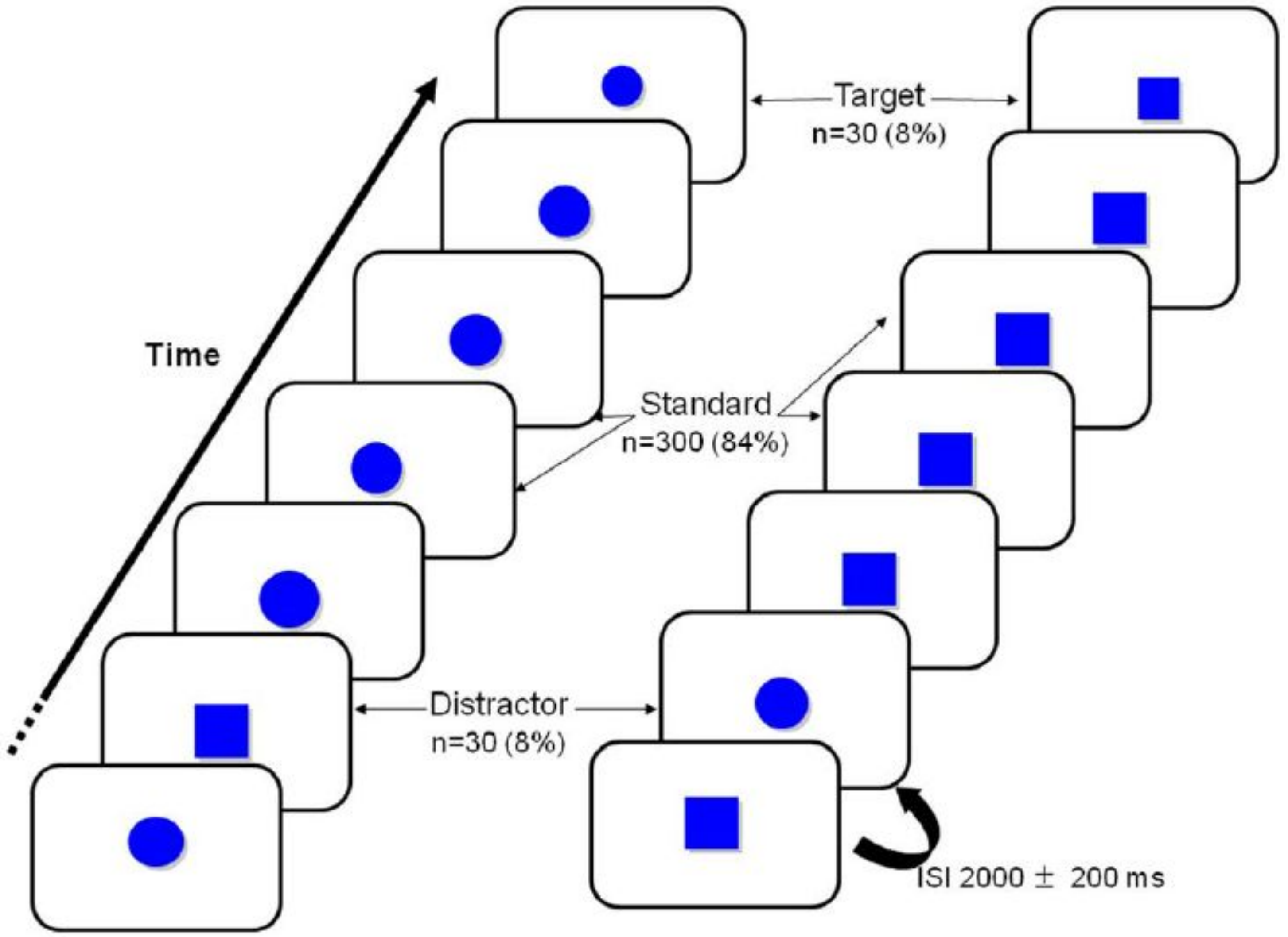


Figure 3. Typical oddball visual paradigm [3]

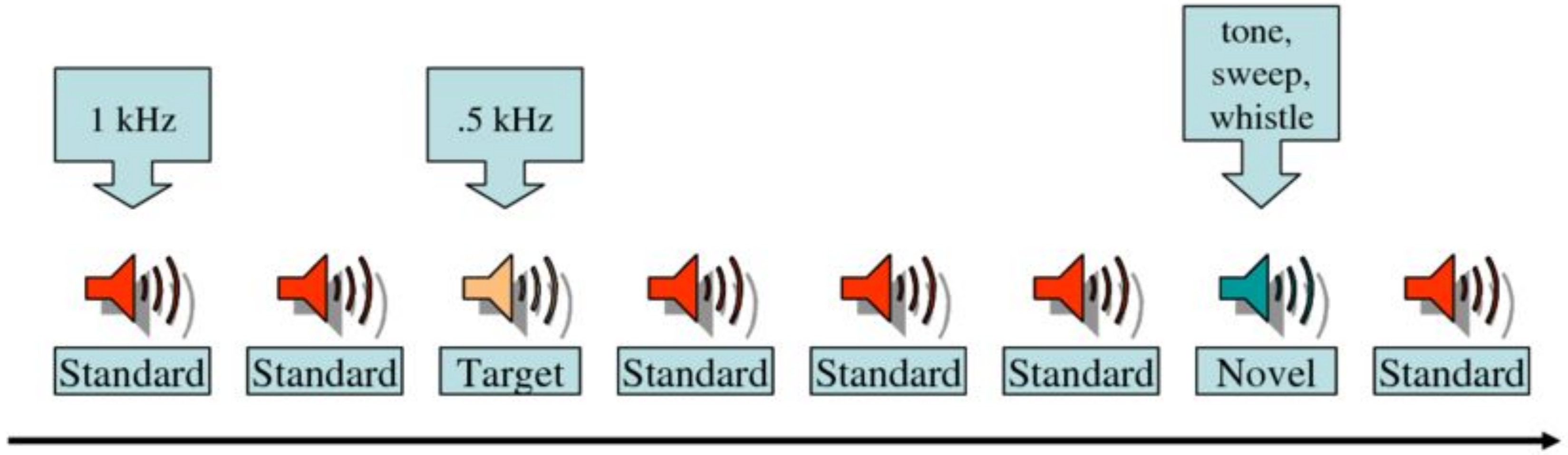


Figure 4. Typical oddball auditory paradigm [3]

Oddball Paradigm

Feature discrimination of ERP components in the oddball paradigm **highly depends on the external properties** of target stimulus.

Therefore, a common approach **to enhance the performance** of ERP-based BCI systems is primarily a development of manipulations of the stimuli parameters:

Auditory BCIs:

increasing volume

spatial arrangement (Fig. 1)

modulated frequency

Visual BCIs:

big shapes

light colour (Fig. 2)

increasing intensity

highly recognizable images (Fig. 3)

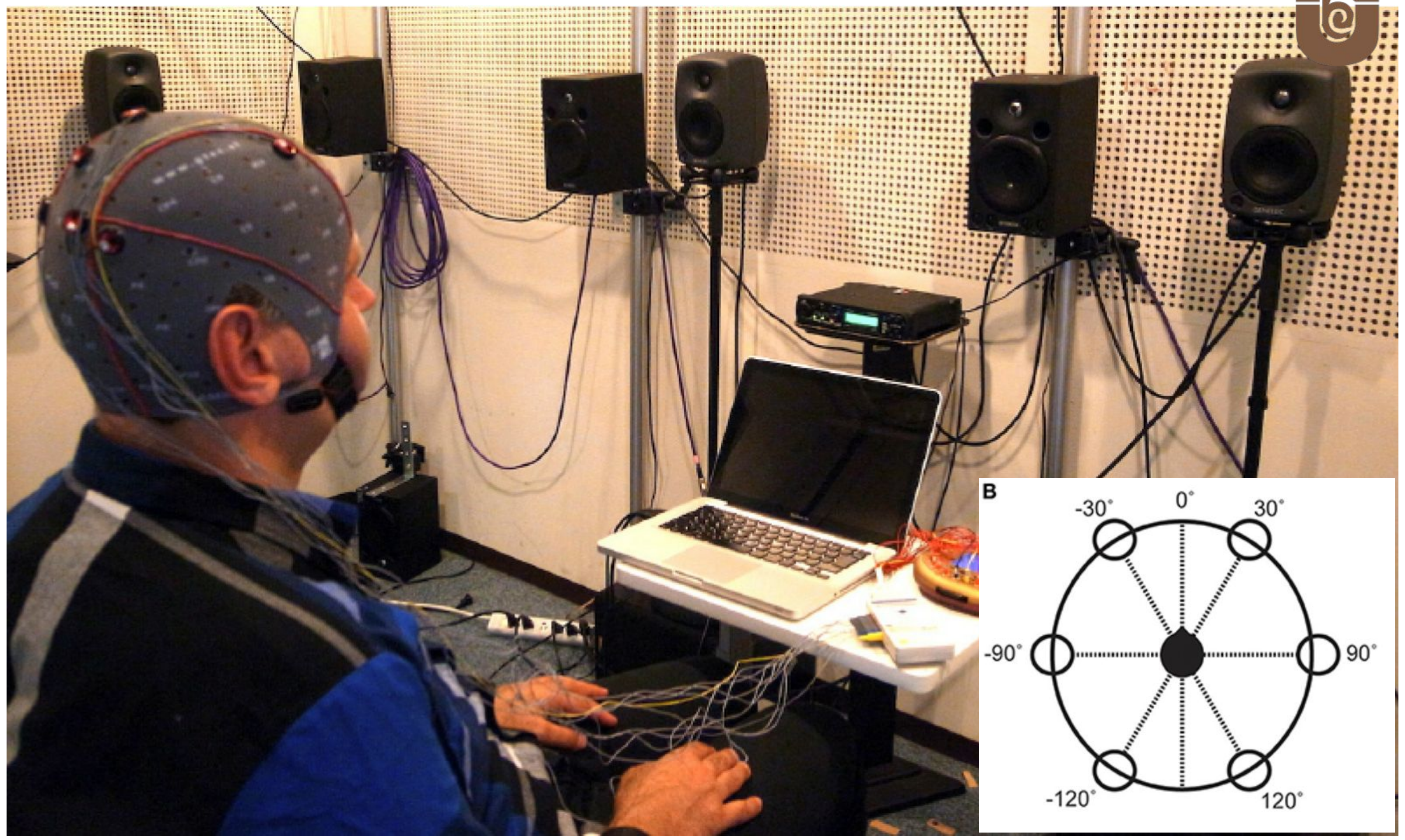


Figure 1. The spatial environment for auditory paradigm [3]

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	0	1	2	3
4	5	6	7	8	9

Figure 2. The spelling matrix with light colors [3]

Figure 3. The spelling matrix with colored smiling cartoon faces [5]

Literature Review

Reference	Year	Feature Extraction	Classification Method	Accuracy Level	BCI Task
3	2020		LR	95%	EEG signal categorization
4	2020		LSTM	97.13%	EEG signal categorization
5	2019	Riemannian geometry, CSP and PSO	CNN	80.44%	EEG signal categorization
6	2018	PCA	SVM	92.50%	BCI therapy stage classification
9	2018	WT	SVM	>90%	EEG signal categorization
10	2018	FFT	KNN SVM	100% 100%	EEG signal categorization
12	2016		SVM	92.50%	P300-based BCI operation
10	2012		SVM	90.55%	ERP signal categorization
11	2010		CNN	95.50%	ERP signal classification
15	2006	PCA	SVM	>95%	EEG signal categorization

The most of papers did classification on ERP-based BCI method. Therefore, non-oddball auditory paradigm tested.

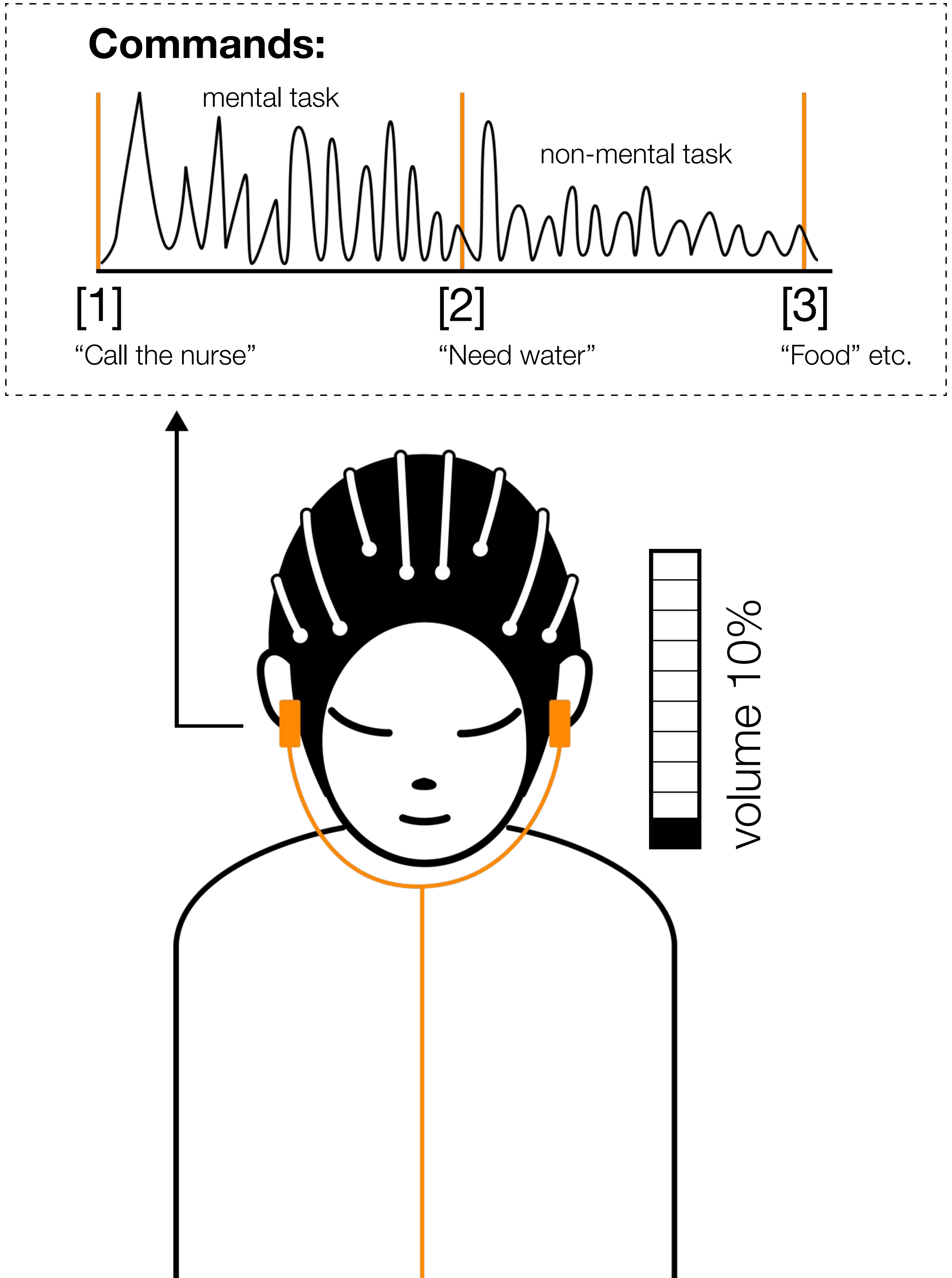
Thesis's Motivation

Problem statement:

The user's voluntary mental task **could generate strong endogenous potentials** that can be detected in neural network model.

Binary BCI system that **extremely minimizes the impact of external stimulus** (e.g., size or volume) where the user is comfortable to the given stimulus but maintaining comparable performance to the conventional approaches.

The intensity (i.e., size and volume) of proposed visual and auditory cues were extremely reduced as its role is priorly **letting the user know the timing of performing a mental task**.



Methodology

Participants

14 subjects participated in this experiment:

- 25-33 years old
- 4 women, 5 newbies in BCI field

BCI set

ActiCap EEG amplifier (Brain Products, Munich, Germany)

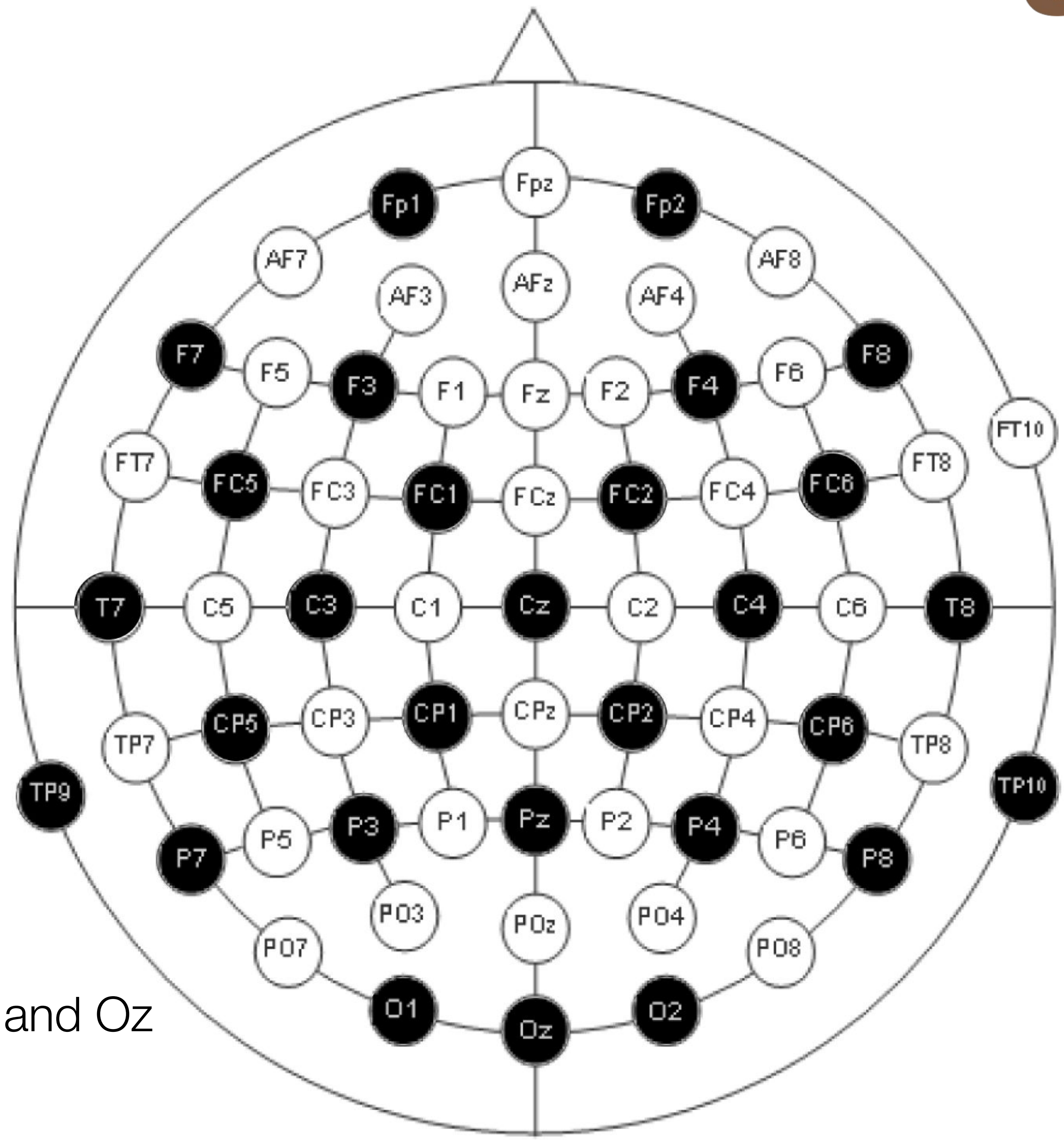
32 EEG channels:

Fp1-2, F3-4, Fz, F7-8, FC5-6, FC1-2, T7-8, C3-4, Cz,

CP1-2, CP5-6, TP9-10, P3-4, P7-8, Pz, PO9-10, O1-2, and Oz

EEG amplifier referenced on the nose with a forehead ground

Ag/AgCl electrodes (international 10-20 system)



Experimental Paradigms

Two experiments:

The binary systems in **non-oddball visual-cue** conditions.

The binary systems in **non-oddball auditory-cue** conditions.

The experiments were mainly designed to eliminate the oddball characteristics in order to fully derive the endogenous potentials.

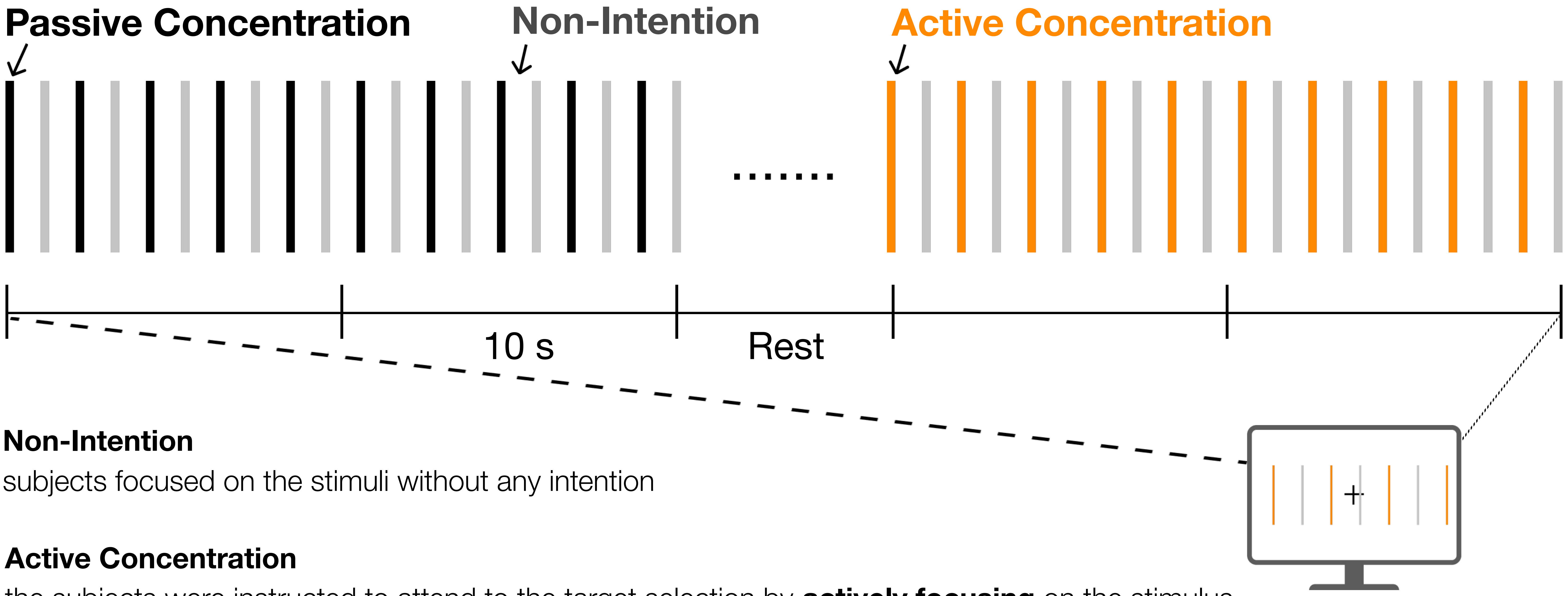
The experiments consisted of **two phases**:

Training phase: estimate the classifier parameters

Test phase: validate the decoding accuracy

All experimental paradigms were developed with the Psychophysics Toolbox(http://psychtoolbox.com) and OpenBMI [15] in Matlab (MathWorks; MA, USA). This study was reviewed and approved by the Institutional Review Board at Korea University [1040548-KUIRB-16-159-A-2], and written informed consent was obtained from all participants before the experiments.

Task Definitions



Non-Intention

subjects focused on the stimuli without any intention

Active Concentration

the subjects were instructed to attend to the target selection by **actively focusing** on the stimulus.

Passive Concentration

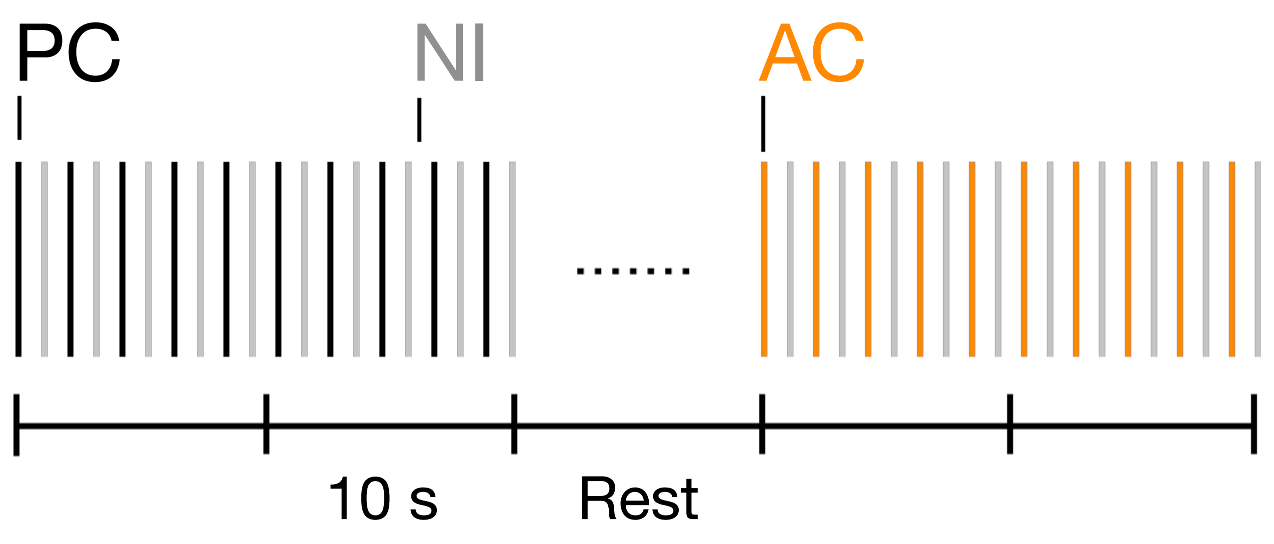
the subjects were instructed to attend to the target selection by **passively focusing** on the stimulus.

Experiment I: Non-oddball Visual Cue (NV)

Training phase

Participants fix their eyes on the cross-symbol and perform the designated tasks when the **bar-stimulus exactly overlapped the fixed cross-symbol**

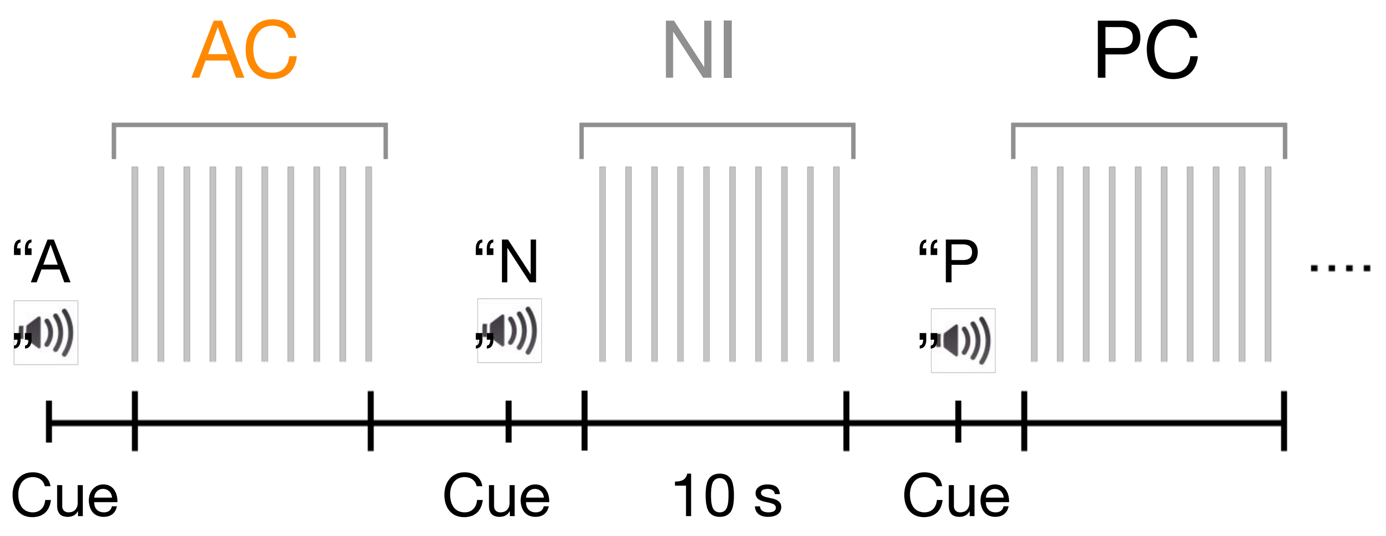
At the end of the training phase: a total of 1080 trials (360 trials for each NI, PC and AC)



Testing phase

Subjects perform a specific task ten times designated by **the given voice cue 5s before the visual stimuli**

At the end of the testing phase: 10 attempts in each class, 300 trials (10 attempts × 10 sequences × 3 classes)



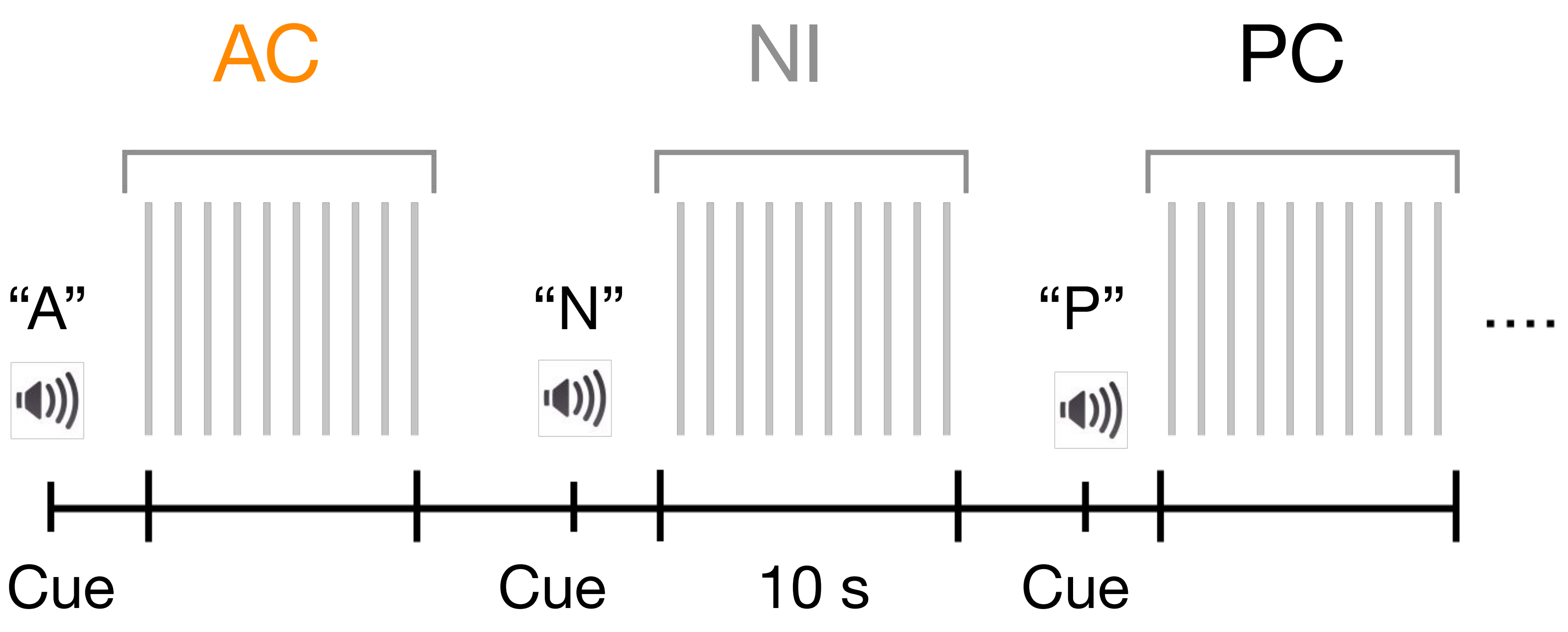
Experiment II: Non-oddball Auditory Cue (NA)

The auditory experiment was designed to explore the **ERP responses in visually blinded conditions**

Before the experiment:

1. **Beep-type auditory stimulus at a frequency of 8000 Hz** was presented to the participant

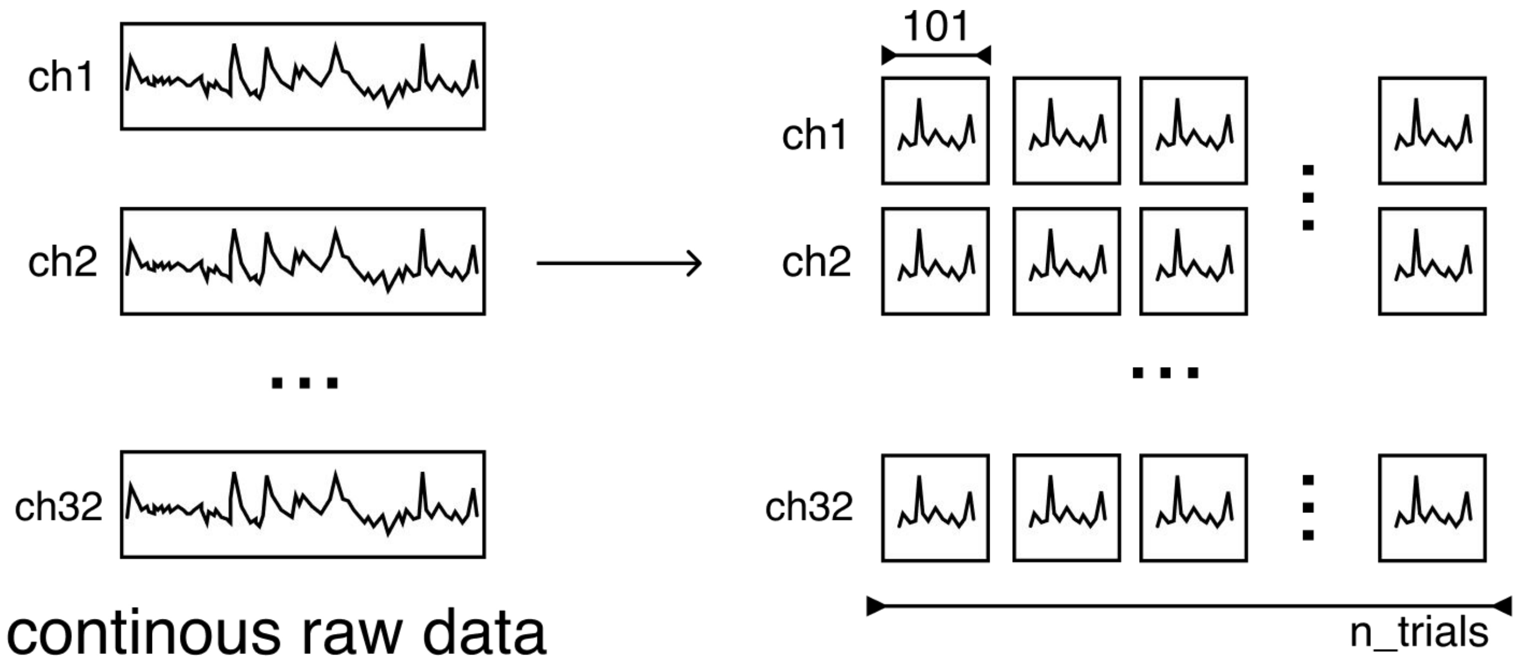
2. The **volume** of the auditory stimulus was adjusted **as low as possible**
   (subject was only able to recognize the moment of the given auditory stimulus)



Data Representation

Data Representation 1

n x 32 x 101

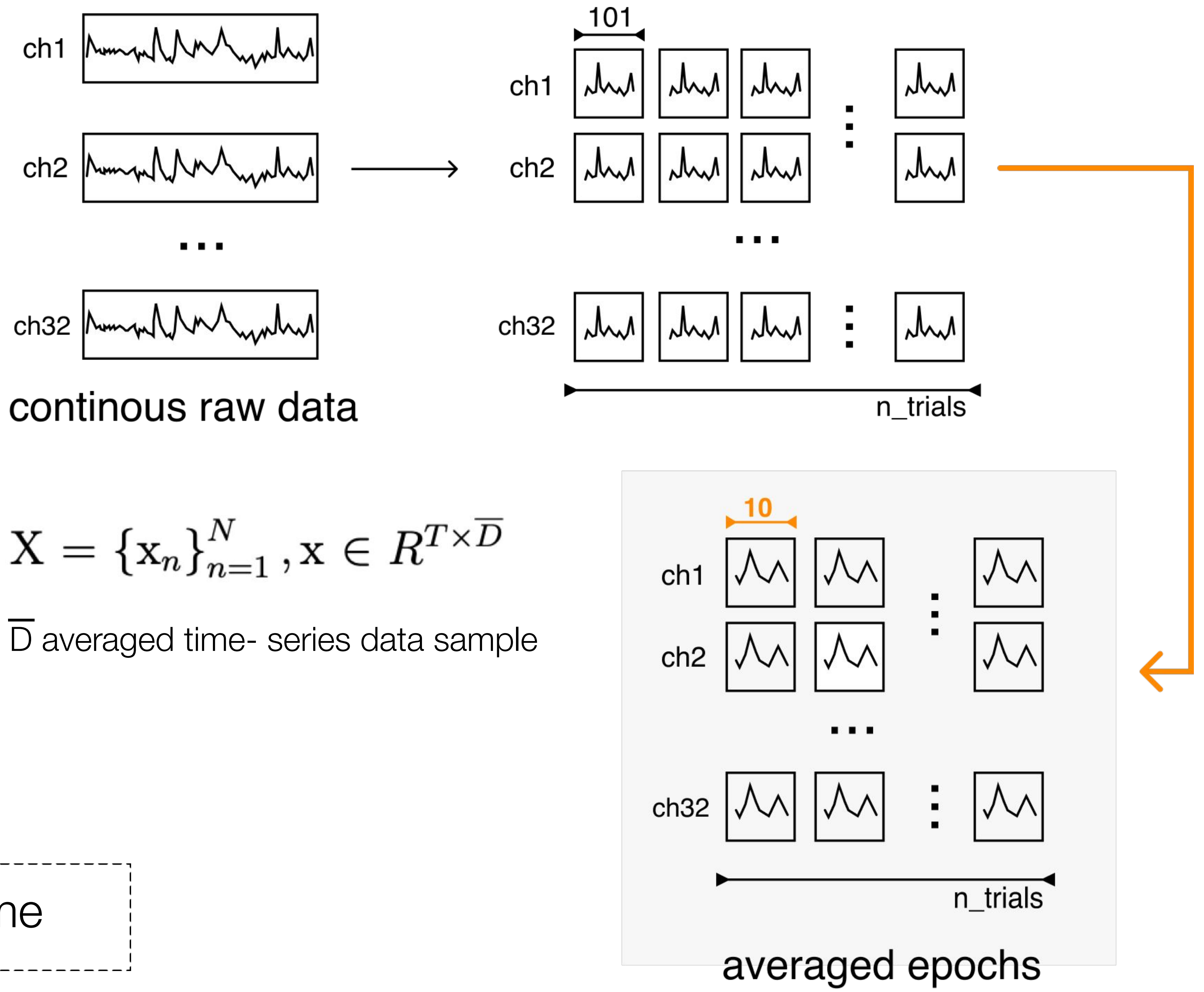


$$\mathrm{X} = \{\mathrm{x}_n\}_{n=1}^{N}, \mathrm{x} \in R^{T\times D}$$

N - the total trial number,
T - time-series data samples,
D - channel numbers

Data Representation 2

n x 32 x 10

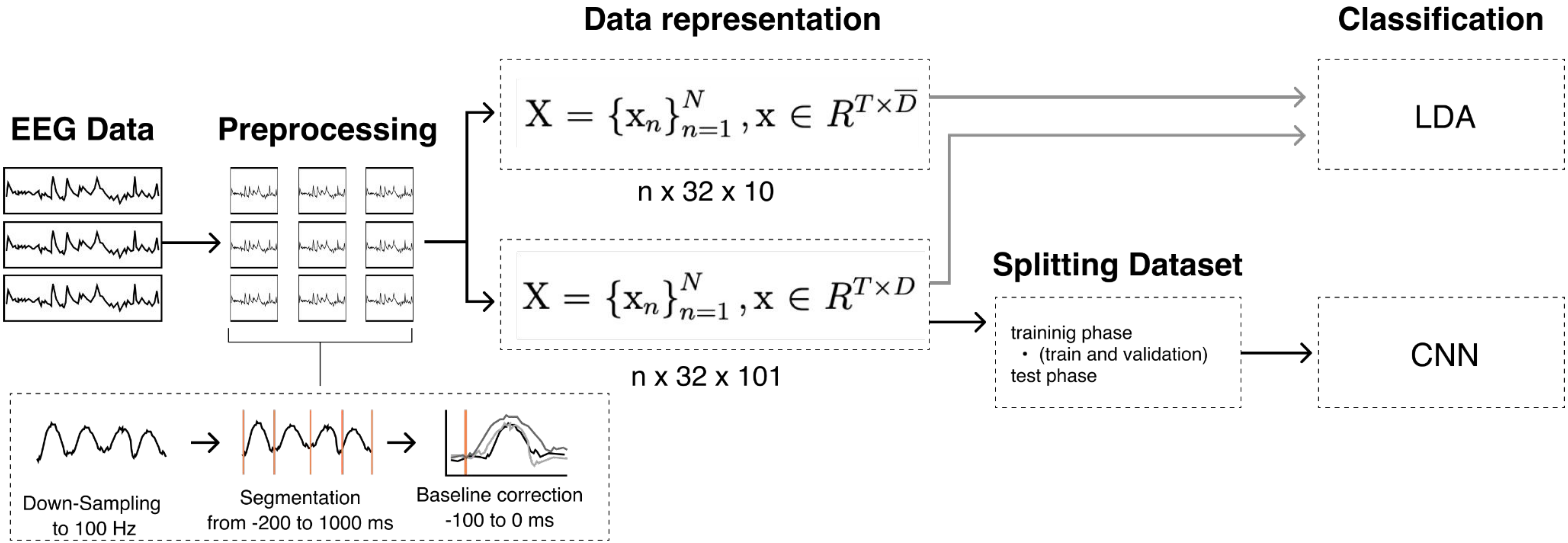


$$\mathrm{X} = \{\mathrm{x}_n\}_{n=1}^{N}, \mathrm{x} \in R^{T\times \overline{D}}$$

$\overline{D}$ averaged time- series data sample

Examine the frequency of robust detecting signal over time

Architecture



Total 14 subjects were held in experiment

Three classes will be classified: Active Concentration (AC), Passive Concentration(PC) and Non-Intention (NI)

CNN: **CNN, EEGNet**

Data Analysis: ERP Data Augmentation for CNN model

Unlimited number of synthetic ERP trials can be created by **adding Gaussian random noise** to the task-relevant component:

- We estimated the task-relevant signal $\tilde{p}(t)$ by averaging all training trials, i.e., $\tilde{p}(t) = \bar{X} = 1/N \sum_{n=1}^{N} \mathrm{x}_n$, where we assumed that the $r(t)$ is approximated to zero by sufficient averaging procedure.
- The certain number of trials was randomly selected from the X and then averaged (denoted by $\bar{\mathrm{Z}} = 1/K \sum_{k=1}^{K} \mathrm{z}_k$)
- We calculated $\bar{\mathrm{Z}} - \bar{\mathrm{X}} = \tilde{p}(t) - p(t) + r_k(t)$. The output is then $r_k(t) \sim \mathcal{N}(0, \sigma^2/K)$ by our previous assumption.
- The noise parameter $(\text{i.e., } \sigma^2)$ was then estimated by $K \cdot var(r(t))$

The random variable v were defined with the probability density function P as follows:

$P(v) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{(v-\mu)^2/2\sigma^2}$ where the μ is zero and the $\sigma^2 = K \cdot var(r(t))$

Finally, the new data samples were created by adding the Gaussian noise to $\bar{\mathrm{X}}_t$ and $\bar{\mathrm{Z}}_t$,

i.e., $\tilde{p}(t) + v(t),\ p(t) + r_k(t) - r_k(t) + v(t)$, respectively.

LDA Model

k time intervals were defined from the stimulus onset to 1000 ms with **a length of 100 ms and a step size of 50 ms** (i.e., {[0 – 100], [50 – 150], ..., [900 – 1000]}).

From the **EEG trials X**, mean amplitude features in the specific time intervals were calculated across all channels, and then concatenated.

$\mathrm{V} = \{\mathrm{v}\}_{n=1}^{N}$ defined the **feature vector set** and formed as $R^{(D \times k) \times N}$

N - the total trial number, k - time intervals, D - channel numbers

From the extracted feature set V, **a regularized linear discriminant analysis** (RLDA).

$f(v) = w^T \cdot v + b$ defined the **decision function**.

w - the hyperplane for separation of binary classes,
b - a bias term.

CNN Model 1

$$X = \{x_n\}_{n=1}^{N}, x \in R^{T \times D}$$

Batch size: 64

Epochs: 100

Optimizer: Adam

Learning Rate: 0.001

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 32, 101, 5)	85
conv2d_1 (Conv2D)	(None, 32, 101, 25)	650
average_pooling2d (AveragePooling2D)	(None, 32, 25, 25)	0
conv2d_2 (Conv2D)	(None, 32, 25, 100)	7600
flatten (Flatten)	(None, 80000)	0
batch_normalization (BatchNormalization)	(None, 80000)	320000
dense (Dense)	(None, 200)	16000200
dropout (Dropout)	(None, 200)	0
dense_1 (Dense)	(None, 100)	20100
dropout_1 (Dropout)	(None, 100)	0
dense_2 (Dense)	(None, 2)	202

Total params: 16,348,837
Trainable params: 16,188,837
Non-trainable params: 160,000

Scikit-learn: Machine Learning in Python,
Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

EGGNet

$$X = \{x_n\}_{n=1}^{N}, x \in R^{T \times D}$$

[0.5] dropoutRate : dropout fraction

[25] kernelLength : length of temporal convolution in first layer.

[32, 16] F1, F2 : number of temporal filters (F1) and number of pointwise filters (F2) to learn. Default: F1 = 4, F2 = F1 * D.

[8] D : number of spatial filters to learn within each temporal convolution. Default: D = 2

[2] dropoutType : Either SpatialDropout2D or Dropout, passed as a string.

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	[(None, 32, 101, 1)]	0
conv2d_5 (Conv2D)	(None, 32, 101, 8)	256
batch_normalization_5 (BatchNormalization)	(None, 32, 101, 8)	32
depthwise_conv2d_2 (DepthwiseConv2D)	(None, 1, 101, 16)	512
batch_normalization_6 (BatchNormalization)	(None, 1, 101, 16)	64
activation_2 (Activation)	(None, 1, 101, 16)	0
average_pooling2d_3 (AveragePooling2D)	(None, 1, 25, 16)	0
dropout_2 (Dropout)	(None, 1, 25, 16)	0
separable_conv2d (SeparableConv2D)	(None, 1, 25, 16)	384
batch_normalization_7 (BatchNormalization)	(None, 1, 25, 16)	64
activation_3 (Activation)	(None, 1, 25, 16)	0
average_pooling2d_4 (AveragePooling2D)	(None, 1, 6, 16)	0
dropout_3 (Dropout)	(None, 1, 6, 16)	0
flatten (Flatten)	(None, 96)	0
dense (Dense)	(None, 2)	194
softmax (Activation)	(None, 2)	0

Total params: 1,506
Trainable params: 1,426
Non-trainable params: 80

Scikit-learn: Machine Learning in Python,
Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Results: ERP Responses

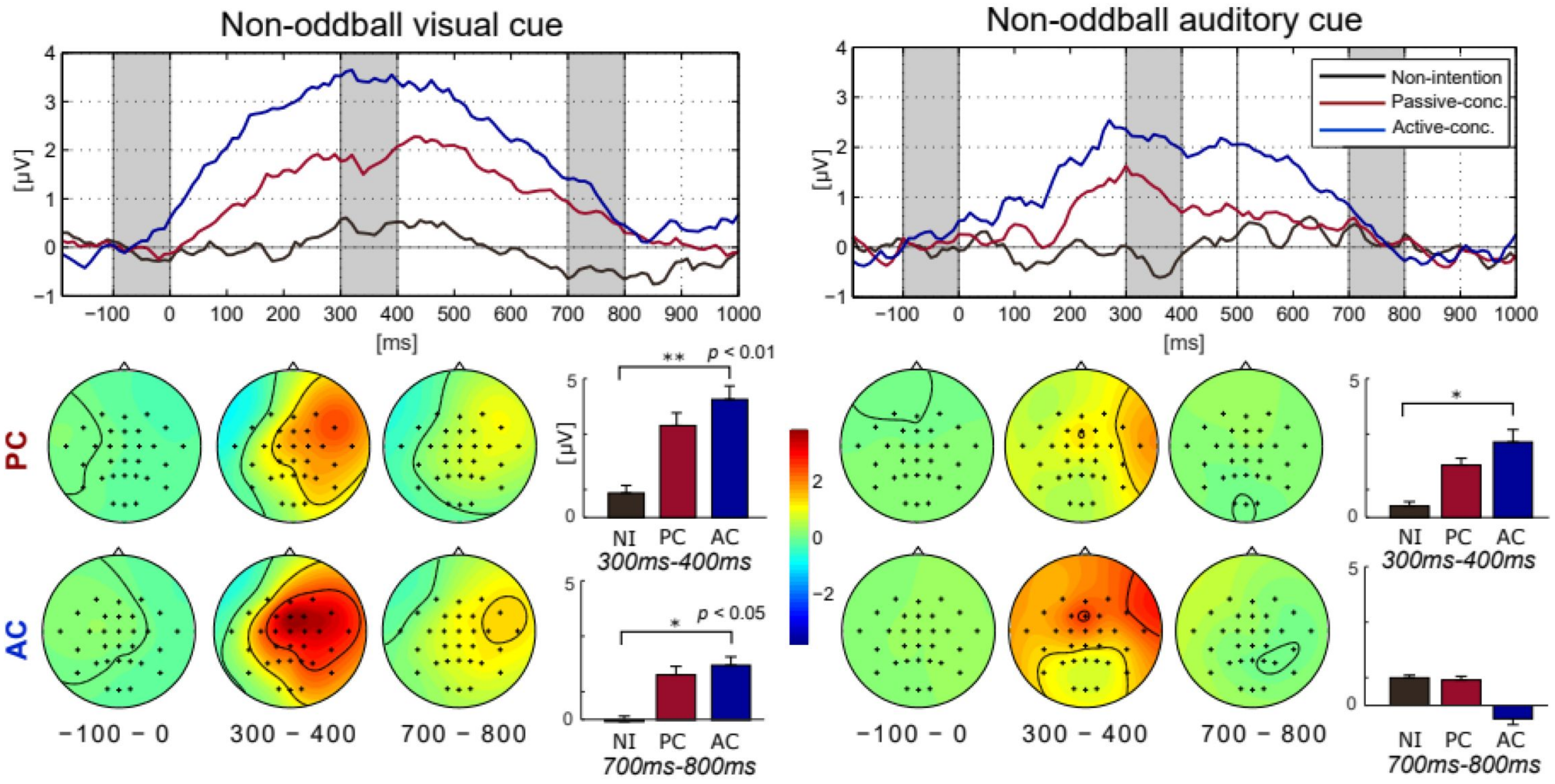


Figure 3. Averaged ERP responses at Cz electrode for three sessions, i.e.,non-oddball visual/auditory cue.

The means of peak amplitude in the interval of 300-400 ms:

Non-oddball Visual cue

NI - 0.558 (±2.083uV),

PC - 2.277 (±3.461)uV,

AC - 3.547 (±3.637)uV

Non-oddball Auditory cue

NI - 0.204 (±1.079)uV

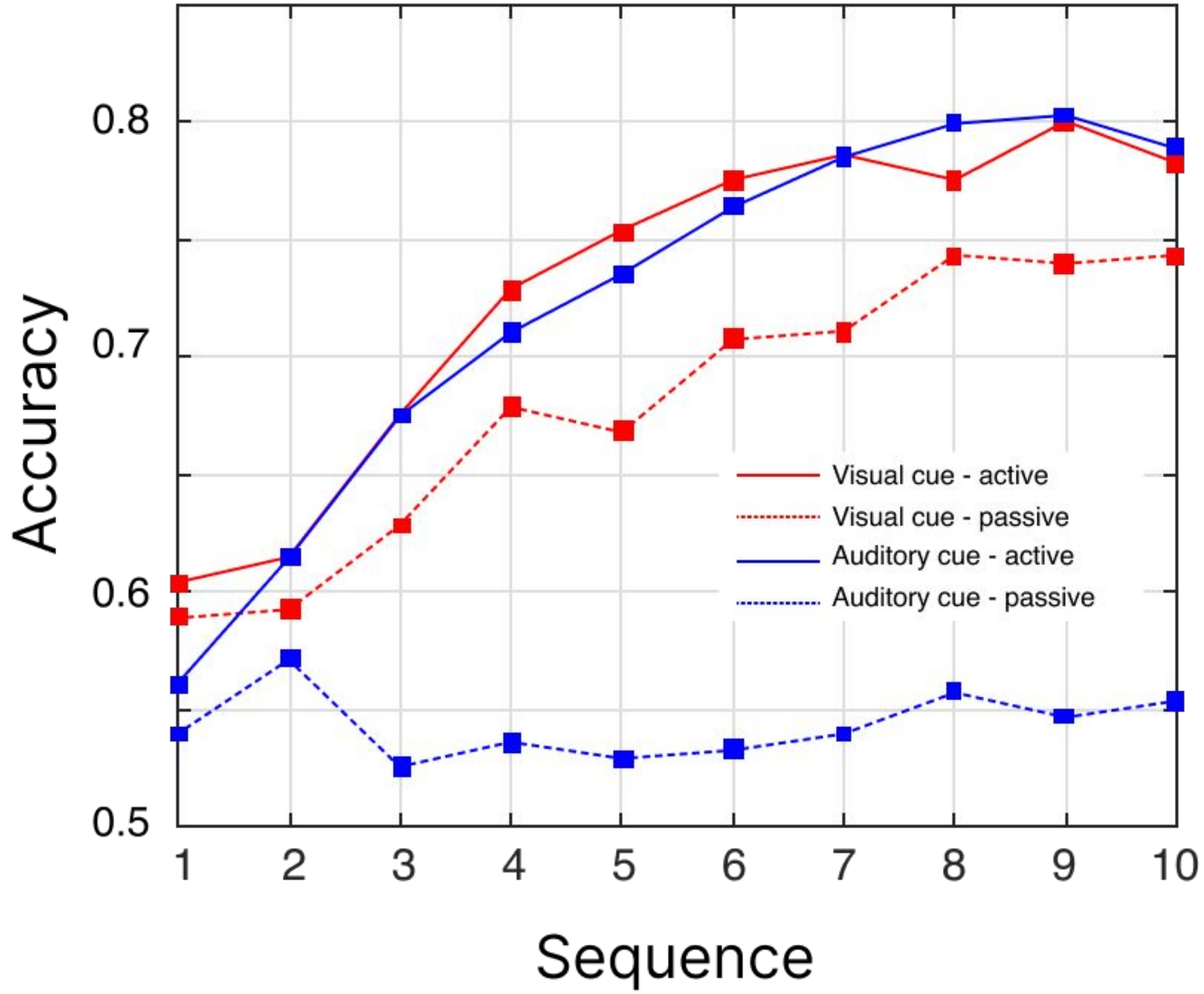
PC - 1.210 (±1.683)uV

AC - 2.257 (±3.090)uV

ref [8, 9, 10, 11]

Results: LDA Decoding Accuracy

Decoding accuracy of non-oddball visual and auditory paradigms



For the given number of sequences, the decoding accuracy of active and passive tasks for both non-oddball visual and non-oddball auditory paradigms.

Much **higher accuracy** is achieved for **active tasks** compared to passive tasks.

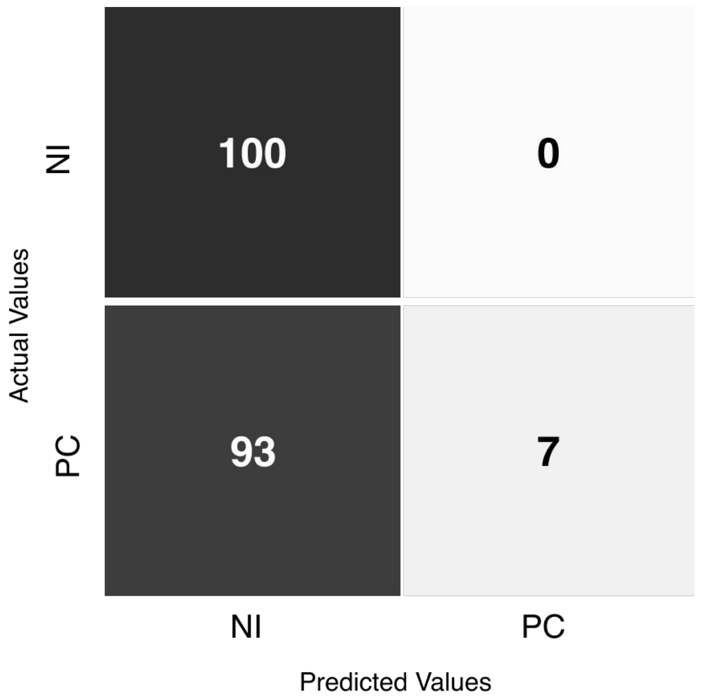
ref [8, 9, 10, 11]

Results: CNN Decoding Accuracy

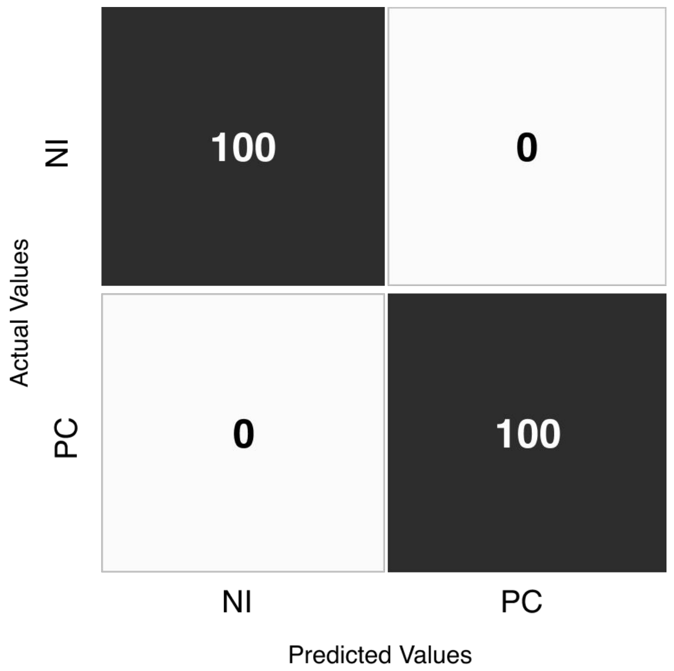
CNN 1 model showed minimum 53.5% for PC vs NI visual paradigm and maximum 100% for PC vs NI visual paradigm.

	Task	Subject 1	2	3	4	5	6	7	8	9	10	11	12	13	14	Avg	Min	Max
	AC vs NI	0.995	0.975	0.595	0.855	0.995	0.700	0.935	0.830	0.995	0.970	0.930	0.970	0.975	0.970	**0.906**	**0.595**	**0.995**
AU	**PC vs NI**	0.995	0.960	0.970	0.615	0.995	0.885	0.920	0.725	0.990	0.570	0.895	0.970	0.670	0.985	**0.868**	**0.570**	**0.995**
	AC vs NI	0.890	0.970	0.995	0.785	0.960	0.940	0.960	0.980	0.980	0.950	0.935	0.675	0.960	0.940	**0.923**	**0.675**	**0.995**
VI	**PC vs NI**	1.000	0.995	0.905	0.995	0.805	0.950	0.960	0.960	0.960	0.980	0.655	0.535	0.575	0.880	**0.868**	**0.535**	**1.000**

Confusion Matrix



PC vs NI visual paradigm
min acc: 53.5%
Subject 12



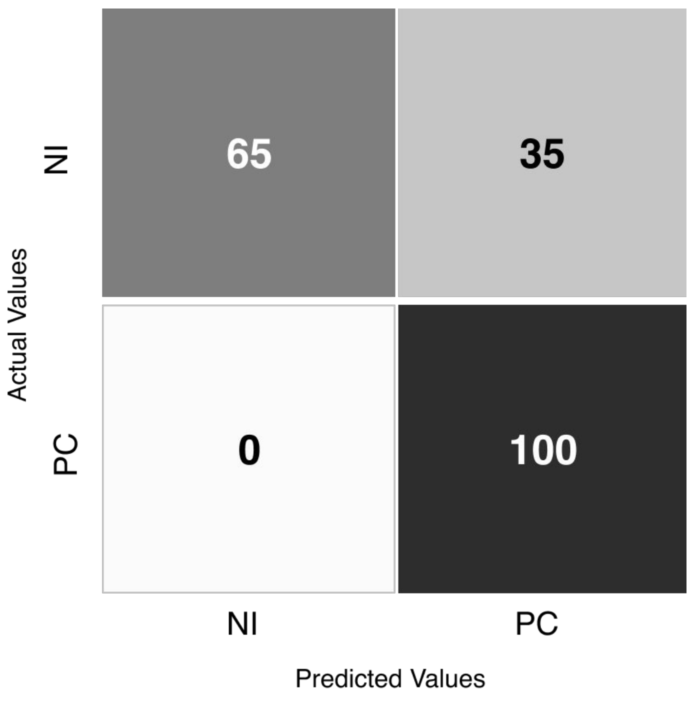
PC vs NI visual paradigm
max acc: 100%
Subject 1

Results: EEGNet Decoding Accuracy

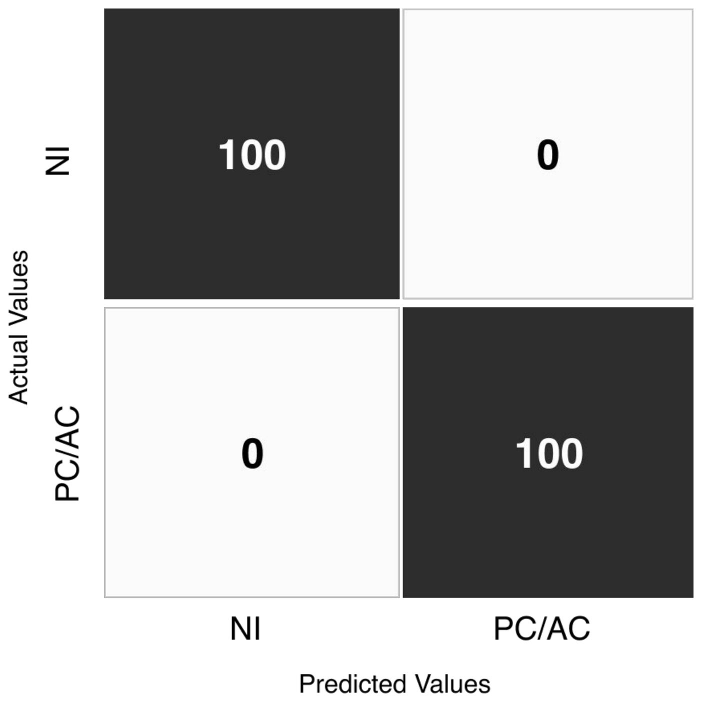
EEGNet model showed minimum 82.5% for PC vs NI visual paradigm and maximum 100% for PC vs NI visual paradigm and both tasks of auditory paradigm.

	Task	Subject 1	2	3	4	5	6	7	8	9	10	11	12	13	14	**Avg**	**Min**	**Max**
	AC vs NI	0.980	0.985	0.995	0.995	1.000	0.980	0.960	1.000	0.985	0.955	0.980	0.990	0.980	0.985	**0.984**	**0.955**	**1.000**
AU	**PC vs NI**	0.965	0.940	0.985	1.000	0.995	0.995	0.945	0.980	0.965	0.995	0.995	0.950	0.985	0.965	**0.976**	**0.940**	**1.000**
	AC vs NI	0.995	0.995	0.875	0.990	0.880	0.970	0.990	0.995	0.975	0.995	0.985	0.970	0.995	0.995	**0.972**	**0.875**	**0.995**
VI	**PC vs NI**	1.000	0.990	0.960	0.995	0.975	0.975	0.990	0.825	0.970	0.970	1.000	0.990	0.970	0.995	**0.972**	**0.825**	**1.000**

Confusion Matrix



PC vs NI visual paradigm
min acc: 82.5%
Subject 8



AC vs NI auditory paradigm
PC vs NI auditory paradigm
PC vs NI visual paradigm
max acc: 100%

Discussion

We proposed a **binary visual/auditory system with minuscule stimulus effects**.
Robust endogenous potentials were validated where the users elicited significant ERP components by themselves. Consequently, we pursue **a stimulus free ERP paradigm** to overcome the current limitation where system performance is excessively dependent on external factors.

We present **unobtrusive continuous visual and auditory stimuli** where the magnitude of the auditory and visual **stimuli is reduced** to the point that they are barely noticeable. This approach greatly reduces the intensity of external stimulation.

The **grand averaged ERP responses of AC and PC for the visual cue** were greater than those resulting from the auditory cue. This is due to the different experimental protocols of each paradigm.

Specifically, subjects were instructed to perform the mental task 10 times, once every second in the auditory paradigm, while they alternatively performed the mental tasks and ignoring state in the visual paradigm.

The most remarkable result is that CNN model showed **more than 75% of accuracy** for all cases.

Limitation and Future Work

Proposed systems are highly appropriate for the severely debilitated or patients at the later stages of a disease who might be sensitive to intense external stimulation. Experiment were done on healthy people, therefore experiment should be done on focus group.

Additionally, the number of channels should be decreased by taking account User Experience. The hospital condition cannot afford many electrodes on the head due to condition of patient. BCI patients usually are in hard conditions as lock-in-syndrome where patients cannot move and lay on bed. The electrodes located on the back of head are uncomfortable in this situation.

Conclusion

Our design sought to increase the **user's endogenous potential with less intense auditory stimuli**.

To achieve that goal we focused on the LPP response which is a result of the neural activity from an active cognitive process.

Result:

- Reduces interference from peripheral stimulation
- Minimizes the size of the stimuli
- High accuracy at decoding model
- Mobile and light-weighted algorithm
- Examine the frequency of robust detecting signal over time

References

[1] Nicolas-Alonso, L.F., Gomez-Gil, J.: Brain-computer interfaces, a review. Sensors 12(2), 12111279 (2012)

[2] Squires, N.K., Squires, K.C., Hillyard, S.A.: Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalography and Clinical Neurophysiology 38(4), 387401 (1975)

[3] Yeom, S.K., Fazli, S., Mu ̈ller, K.R., Lee, S.W.: An efficient ERP-based brain- computer interface using random set presentation and face familiarity. PloS one 9(11), e111157 (2014)

[4] Li,Q.,Liu,S.,Li,J.,Bai,O.:Use of a green familiar faces paradigm improves P300-speller brain-computer interface performance. PloS one 10(6), e0130325 (2015)

[5] Li, Q., Lu, Z., Gao, N., Yang, J.: Optimizing the performance of the visual P300-speller through active mental tasks based on color distinction and modulation of task difficulty. Frontiers in Human Neuroscience 13, 130 (2019)

[6] Lee, M.H., Williamson, J., Kee, Y.J., Fazli, S., Lee, S.W.: Robust detection of event-related potentials in a user-voluntary short-term imagery task. PloS one 4(12), e0226236 (2019)

[7] Lee, M.H., Williamson, J., Lee, Y.E., Lee, S.W.: Mental fatigue in central-field and peripheral-field steady-state visually evoked potential and its effects on event-related potential responses. NeuroReport 29(15), 1301 (2018)

[8] Hara, K., Kataoka, H., Satoh, Y.: Learning spatio-temporal features with 3D resid- ual networks for action recognition. In: Proceedings of the IEEE International Conference on Computer Vision Workshops. pp. 31543160 (2017)

[9] Kwon, O.Y., Lee, M.H., Guan, C., Lee, S.W.: Subject-independent brain-computer interfaces based on deep convolutional neural networks. IEEE Transactions on Neural Networks and Learning Systems 31(10), 3839-3852 (2019)

[10] Kang, Y., Hyndman, R.J., Li, F.: Gratis: Generating time series with diverse and controllable characteristics. Statistical Analysis and Data Mining: The ASA Data Science Journal 13(4), 354376 (2020)

[11] Ku ̈bler, A., Birbaumer, N.: Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients. Clinical Neurophysiology 119(11), 26582666 (2008)

NAZARBAYEV UNIVERSITY

Thank you for the attention

Madina Saparbayeva | madina.saparbayeva@nu.edu.kz
Min-Ho Lee | minho.lee@nu.edu.kz
Berdakh Abibullaev | berdakh.abibullaev@nu.edu.kz