P3T: A Transformer Model for Enhancing Character Recognition Rates in P300 Speller Systems

Jiazhen Hong and Laleh Najafizadeh

Integrated Systems and NeuroImaging Laboratory, Department of Electrical and Computer Engineering, Rutgers University

Introduction

Background and Motivation

- Brain-computer interface (BCI) spellers offer a promising way to assist individuals who are unable to communicate through conventional means
- A type of BCI speller utilizes P300 event-related potential (ERP) in EEG signals, which occurs approximately 300 ms after the presentation of a target stimulus
- An ideal P300 BCI speller should detect the target character from a single repetition
- In current P300 BCI spellers, due to the small signal-to-noise ratio (SNR) of EEG, detecting a character requires multiple repetitions of the stimuli
- **Goal:** Enhance the character recognition rate with fewer number of repetitions

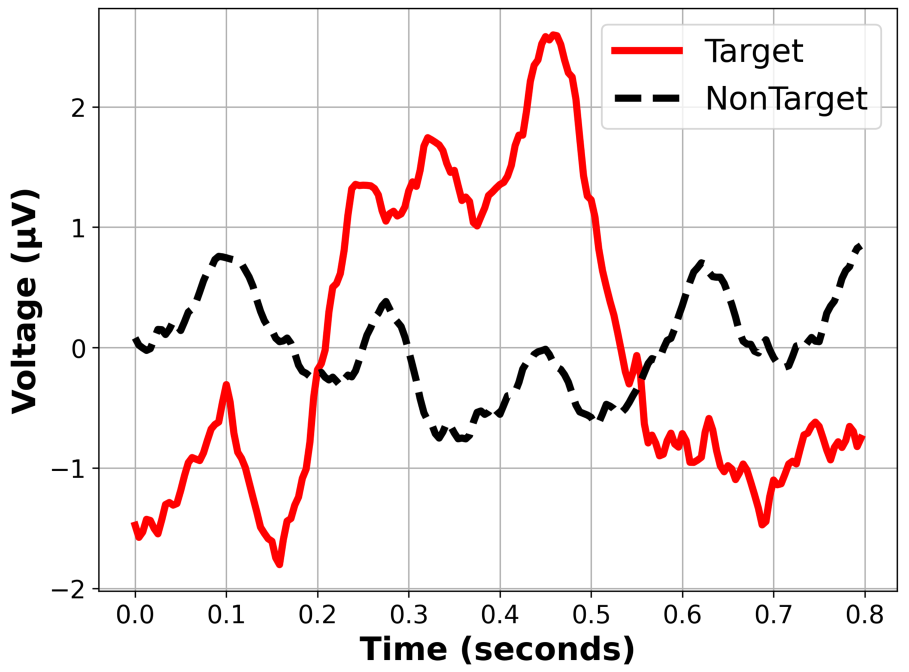

Fig. 1: Sample EEG in response to target (red) and non-target (black) stimuli. P300 is seen in response to target stimuli.

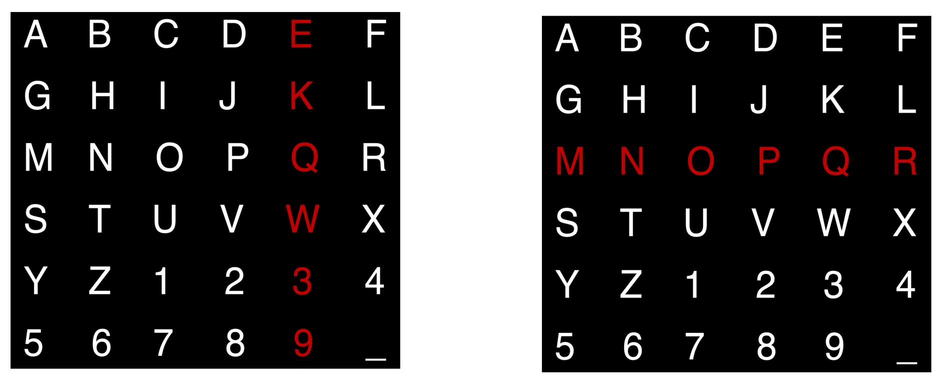

Fig. 2: BCI speller keyboard. The character "Q" will be selected as the the fifth column ($k = 5$) and the third row ($k = 9$) are the target stimuli (red).

Data & Processing

Table 1: Wadsworth BCI Dataset

Subject	Training (85 characters)		Testing (100 characters)	
	Target	Non-Target	Target	Non-Target
A	2550	12750	3000	15000
B	2550	12750	3000	15000

Preprocessing

- 64-channel cap @240Hz
- Bandpass filter (0.1-60 Hz)
- [0-650]ms post stimulus
- 156 time steps (vectors) × 64 channels (features)

Flashing Pattern:

- Each trial contains 12 flashes
- Each row/column lasts 175 ms (100 + 75)ms
- Each character repeats 15 times, pause 2.5 seconds
- $T_i = 2.5 + (0.175 \times 12) \times i$ (seconds)

Character Recognition

- $S_k = \sum_{i=1}^{n} S_k{}^{(i)}$
- $\mathrm{c} = \underset{k\in\{1,6\}}{\arg\max}\, S_k$
- $\mathrm{r} = \underset{k\in\{7,12\}}{\arg\max}\, S_k$

Character Recognition Rate (CRR)

- $\mathrm{CRR} = \frac{\text{Correct number of characters}}{\text{Total number of characters}} \times 100\%$

Information Transfer Rate (ITR)

- $\mathrm{ITR}_i = \frac{60(1-\mathrm{CRR}_i) + \log_2\frac{1-\mathrm{CRR}_i}{N-1} + \mathrm{CRR}_i \log_2 \mathrm{CRR}_i + \log_2 N}{T_i}$

Proposed Model

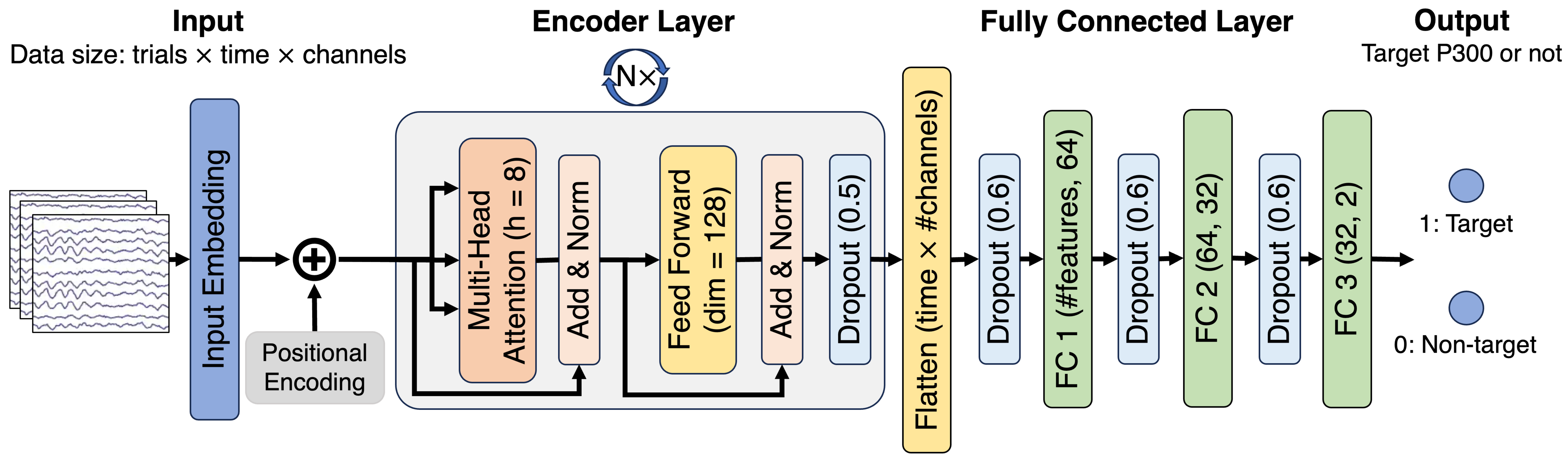

Fig. 3: Proposed P300-Transformer (P3T) for P300 ERP detection.

Results

Table 2: Character recognition rate (CRR) (%) by P3T and other state-of-the-art models averaged across subjects, for repetitions 1 to 15.

Model \ Repetitions	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CNN-1 [1]	25.5	42.5	53	60	70	73	79.5	83.5	88.5	88.5	90.5	90.5	91	92.5	94.5
MCNN-1 [1]	28.5	43	56	59	69	73.5	81	84	85	87	92	93.5	93	93.5	95.5
ERP-CapsNet [2]	30.5	48	59	65	74.5	80	81.5	86.5	86	90.5	**93**	94	94	95	97
ST-CapsNet [3]	29.5	46	59.5	**67**	**76.5**	**82.5**	**87**	**87.5**	**87.5**	**91.5**	92.5	94	93.5	**96**	97
MsCNN-TL-ESVM [4]	32	**48.5**	56.5	62	69.5	77	81	85.5	89	**91.5**	92.5	93.5	94.5	95.5	96
SWFP [3, 5]	28	44	58.5	64.5	73.5	76.5	82	85	86.5	89.5	90.5	92.5	91.5	94.5	95.5
Shrinkage LDA [2, 6]	26.5	39	55	61	66.5	72	81	81.5	83	88.5	89	92.5	91.5	92.5	95.5
P3T	**33**	47.5	**60**	65	75	80.5	84.5	85.5	88.5	**91.5**	92	**95**	**96**	**96**	**97.5**

Table 3: Comparison of character recognition rates (CRRs) (%) for different numbers of encoder layers.

Subject	# Encoder Layer	Accuracy	CRR (%) in # Repetitions		
			5	10	15
A	3	**74.73**	**69**	**91**	**99**
	2	73.56	67	89	**99**
	1	73.15	68	87	98
	0	72.95	21	42	46
B	3	77.64	**81**	92	**96**
	2	77.65	81	92	94
	1	**78.26**	80	**93**	95
	0	76.61	4	3	3

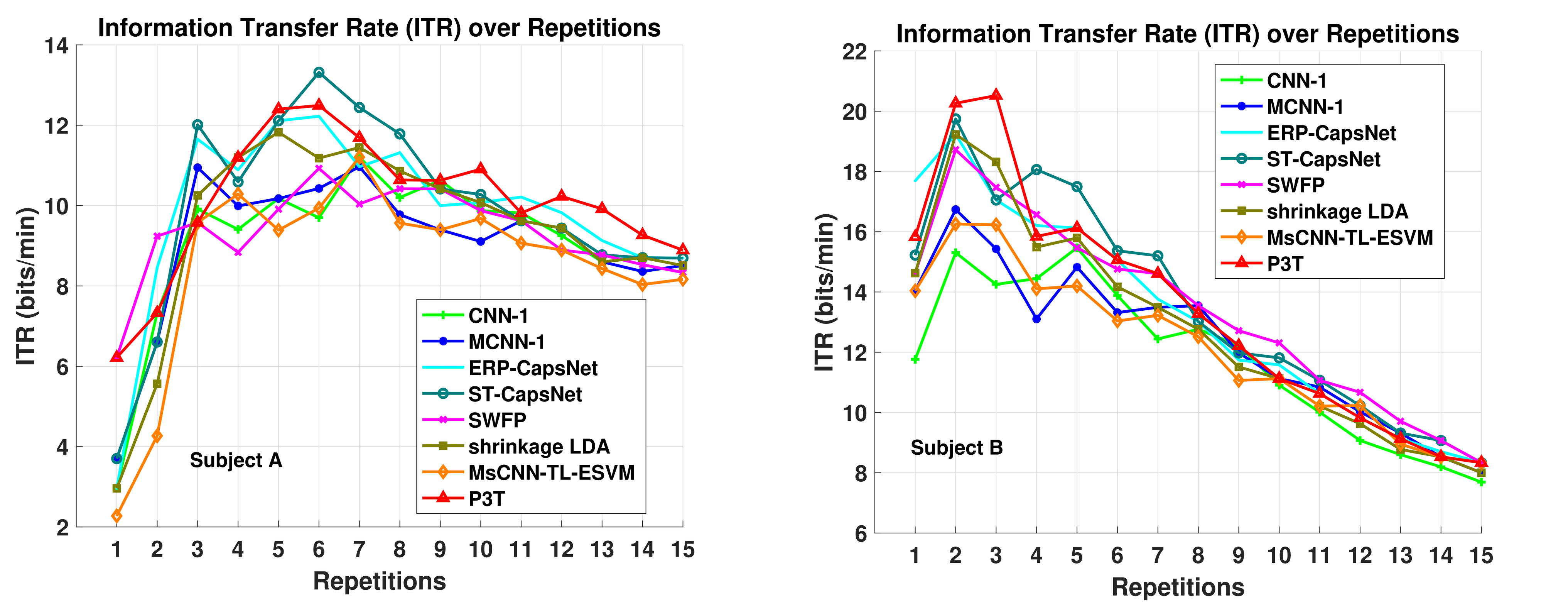

Fig. 4: Information Transfer Rate (ITR) across 15 repetitions for various models: (Left) Subject A, and (Right) Subject B.

- **P3T** achieved a **CRR** of **95%** after 12 repetitions, achieving the highest CRR compared to other models
- **P3T** achieved **ITR** of **20.52 bits/min** in 3rd repetition, the highest reported ITR for this dataset

Reference

[1] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain–computer interfaces," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 33, no. 3, pp. 433–445, 2010.
[4] S. Kundu and S. Ari, "MsCNN: A deep learning framework for P300- based brain–computer interface speller," IEEE Trans. on Medical Robotics and Bionics, vol. 2, no. 1, pp. 86–93, 2019.
[2] R. Ma, T. Yu, X. Zhong, Z. L. Yu, Y. Li, and Z. Gu, "Capsule network for ERP detection in brain–computer interface," IEEE Trans. on Neural Systems and Rehab. Eng., vol. 29, pp. 718–730, 2021.
[3] Z. Wang, C. Chen, J. Li, F. Wan, Y. Sun, and H. Wang, "ST-CapsNet: Linking spatial and temporal attention with capsule network for P300 detection improvement," IEEE Trans. on Neural Systems and Rehab. Eng., vol. 31, pp. 991–1000, 2023.
[5] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," IEEE Trans. on Neural Systems and Rehab. Eng., vol. 14, no. 2, pp. 174–179, 2006.
[6] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Muller, ¨ "Single-trial analysis and classification of ERP components—a tutorial," NeuroImage, vol. 56, no. 2, pp. 814–825, 2011.