BCNN: A Bantamweight Convolutional Neural Network for P300 Detection

M. Liu, S. Wang and F. R. Beyette Jr.

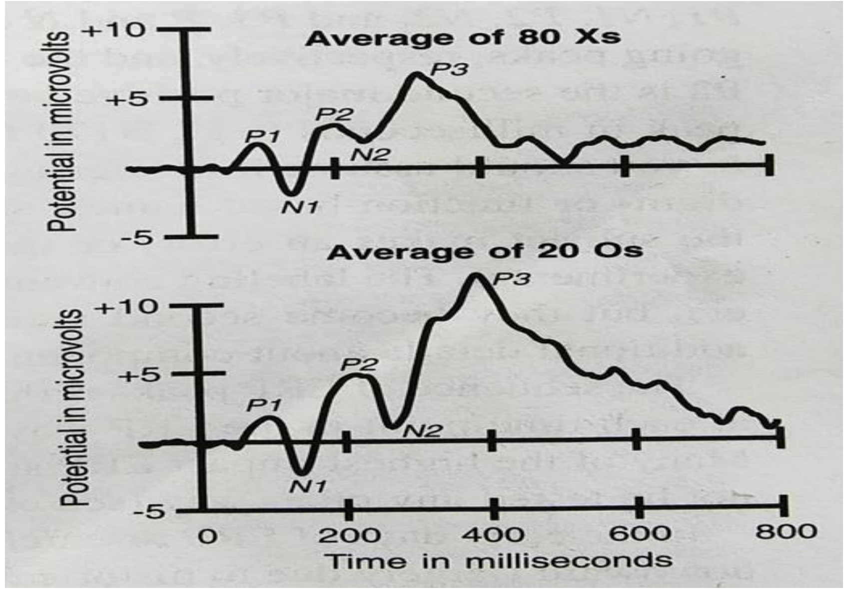
Presenter: Maohua Liu

The 2023 IEEE Signal Processing in Medicine and Biology Symposium
December 2nd, 2023

Outline

- What is P300? & Challenge in detecting P300
- Cutting-edge CNNs for P300 detection
- BCNN architecture & outstanding feature
- Experimental results
- Insights from Explainable AI

P300 ERP



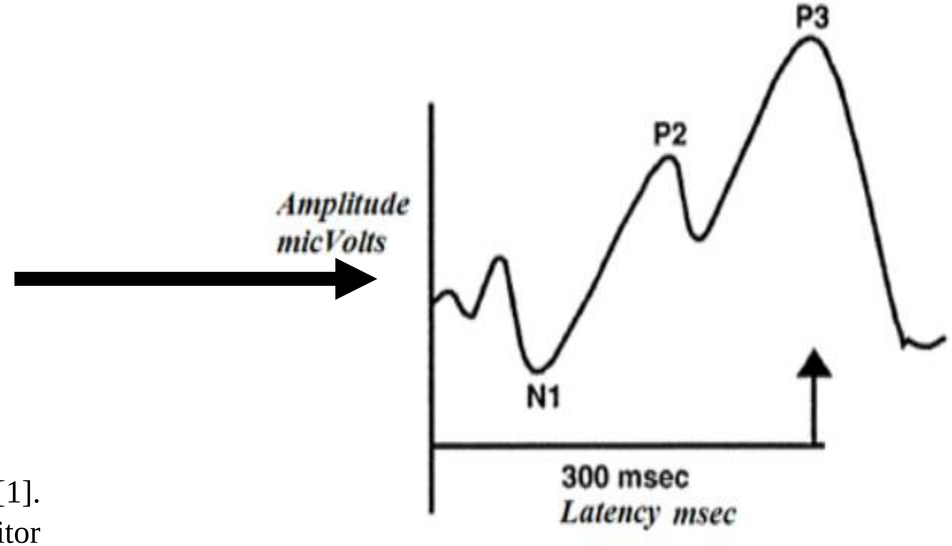
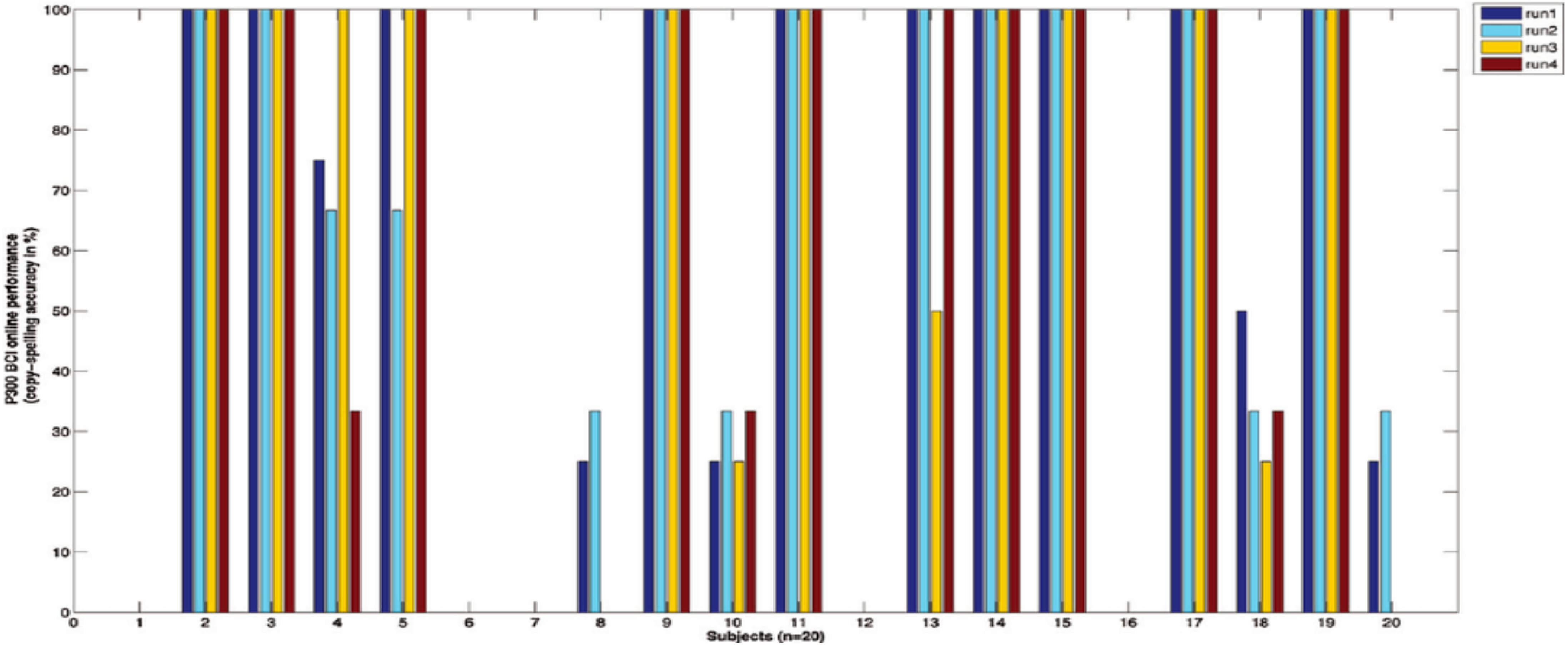


Figure 1. Example of ERP experiment using the oddball paradigm [1]. The subject viewed frequent Xs and infrequent Os on a computer monitor while the EEG was recorded from the active electrode Pz. Separate averages were computed for the X and O epochs. The amplitude of the bottom wave elicited by the uncommon stimuli Os, is obviously greater than the amplitude of the top wave elicited by the common stimuli Xs.

Experiment: The Consciousness Detector - EEG, Oddball Task, and P300 (backyardbrains.com)
A Study of Late Auditory Event Related Potentials Related to P300 Responses in Individuals with Learning Disabilities - (saera.eu)

Challenge in P300 detection (cross-subject detection)



P300 BCI online performance of 20 ALS subjects (four testing runs per subject: runs 1–4). 48.75% of runs achieved 100% online accuracy; 30% of runs had online accuracy of 0%.

P300 BCI online performance of 20 ALS subjects (four testing runs per... | Download Scientific Diagram (researchgate.net)

Cutting-edge CNNs for P300 detection

Table 1. Cutting-edge CNNs for P300 detection (S: spatial, T: temporal, Se: separable, D: depth-wise, F: filters, L: layers) [4].

Architecture	No. Conv filters	No. Conv layers	No. Dense filters & layers	No. Batch layers	No. Params	AUC value Dataset1, Dataset2	Training epochs Dataset1, Dataset2	Time
CNN1	10 S, 50 T	1 S, 1 T	102 F, 2 L	0	1,036,922	0.82±0.05, 0.78±0.04	97±33, 71±14	2010
UCNN1	10 S, 50 T	1 S, 1 T	102 F, 2 L	0	1,036,922	0.84±0.06, 0.78±0.05	88±27, 76±24	2010
CNN3	1 S, 50 T	1 S, 1 T	102 F, 2 L	0	1,031,009	0.78±0.11,0.73±0.08	111±37, 93±31	2010
UCNN3	1 S, 50 T	1 S, 1 T	102 F, 2 L	0	1,031,009	0.83±0.06, 0.76±0.07	114±42, 87±30	2010
CNN-R	96 S, 256 T	1 S, 2 T	6146 F, 3 L	0	19,848,098	0.83±0.06, 0.79±0.04	61±2, 64±2	2015
DeepConvNet	25 S, 375 T	1 S, 4 T	2 F, 1 L	4	139,877	0.84±0.06, 0.79±0.04	122±40, 106±24	2017
ShallowConvNet	40 S, 40 T	1 S, 1 T	2 F, 1 L	1	12,082	0.82±0.07, 0.79±0.03	177±29, 157±33	2017
BN^3	16 S, 16 T	1 S, 1 T	1 F, 3 L	2	44,589	0.83±0.06, 0.78±0.04	113±21, 95±9	2018
EEGNet	8 T, 16 Se	1 T, 1 D, 1 Se	2 F, 1 L	3	1,394	0.84±0.06, 0.80±0.03	200±3, 198±7	2018
OCLNN	16 T	1 T	2 F, 1 L	0	1,842	0.83±0.06, 0.79±0.04	199±5, 161±26	2018
FCNN	None	None	3 F, 2 L	0	2,477	0.83±0.06, 0.75±0.04	197±7, 132±12	2021
SepConv1D	4 Se	1 Se	1 F, 1 L	0	225	0.84±0.06, 0.78±0.04	199±5, 183±24	2021

BCNN architecture

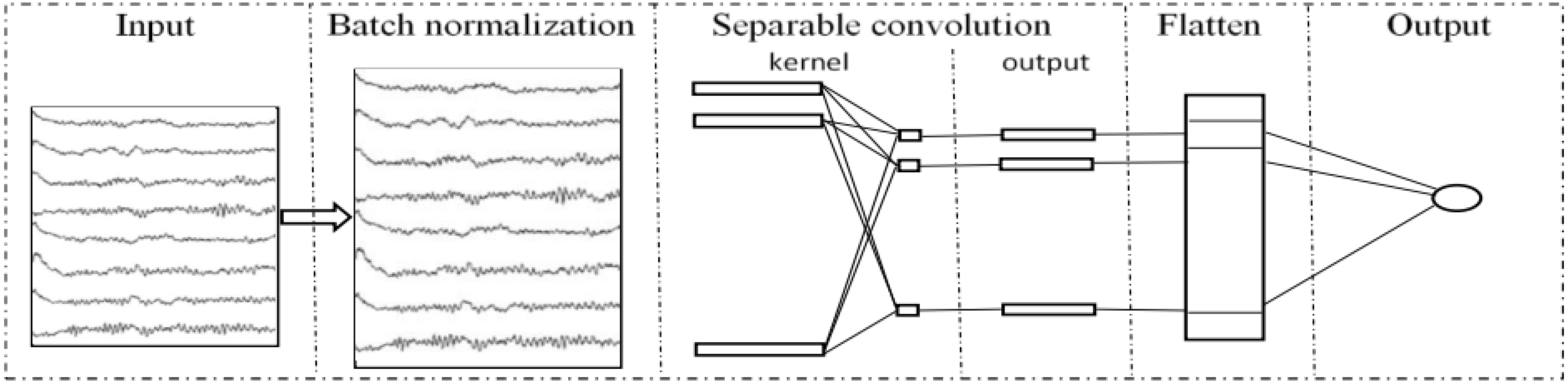


Figure 3. Visualization of BCNN architecture.

Table 2. BCNN architecture.

Block	Layer	# Filters	Size	# Parameters	Output	Activation	Options
1	Input		$T \times C$				
	BatchNorm		$T \times C$	$2 \times C$	(T, C)		
	ZeroPadding				$(T+2\times p, C)$		Padding = p
	Separable Convolution	F	Kernel = k Stride = s	$k\times C+F\times C+F$	$(1+(T+2\times p-k)/s, F)$		
	Activation				$(1+(T+2\times p-k)/s, F)$	*Tanh*	
	Flatten				$(1+(T+2\times p-k)/s)$		
Classifier	Dense	1		$1+F\times(1+(T+2\times p-k)/s)$	(1)	*sigmoid*	

Outstanding feature

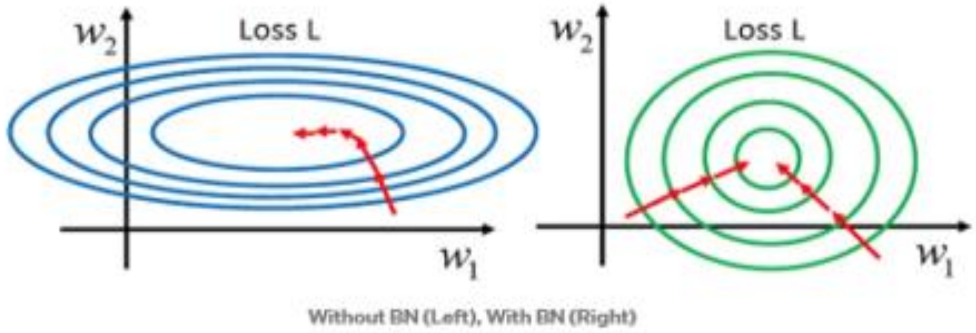


Figure 2. Batch normalization can smooth the loss landscape, thereby enabling large learning rate [6].

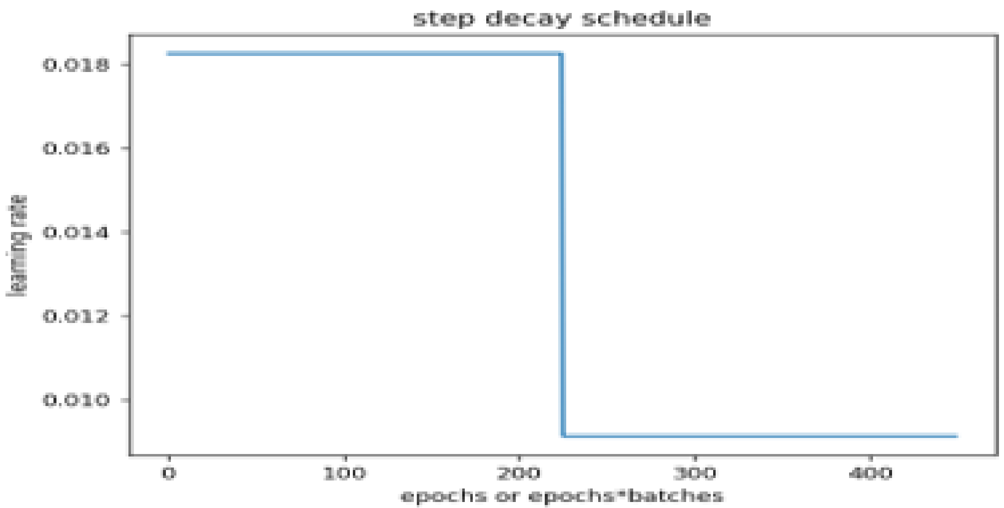


Figure 4. The step learning rate schedule.

Table 4. Number of parameters on Dataset1.

Parameters	SepConv1D (4 filters)	SepConv1D (1 filter)	BCNN (1 filter)
Trainable	225	141	153
Non-trainable	0	0	12
Total	225	141	165

Table 6. Number of parameters on Dataset2.

Parameters	SepConv1D (4 filters)	SepConv1D (1 filter)	BCNN (1 filter)
Trainable	265	163	179
Non-trainable	0	0	16
Total	265	163	195

Experimental results

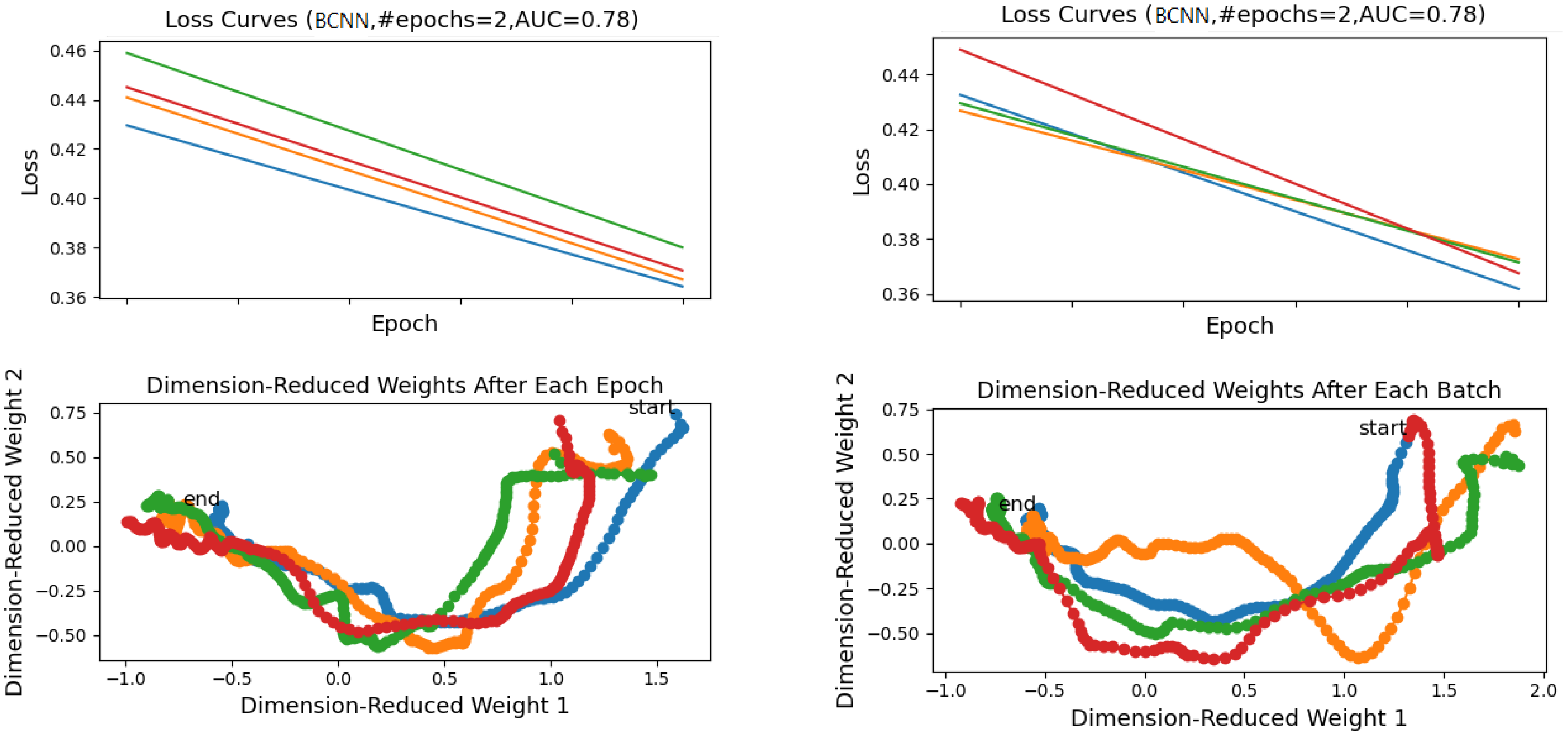
Table 5. AUC values on Dataset2.

Test subject	SepConv1D 4 filters 2 epochs	SepConv1D 1 filter 2 epochs	BCNN 1 filter 2 epochs
0	0.65	0.50	0.77
1	0.68	0.64	0.74
2	0.63	0.68	0.82
3	0.57	0.62	0.70
4	0.49	0.51	0.80
5	0.54	0.53	0.82
6	0.75	0.55	0.78
7	0.71	0.55	0.81
Mean	**0.63±0.09**	**0.57±0.07**	**0.78±0.04**

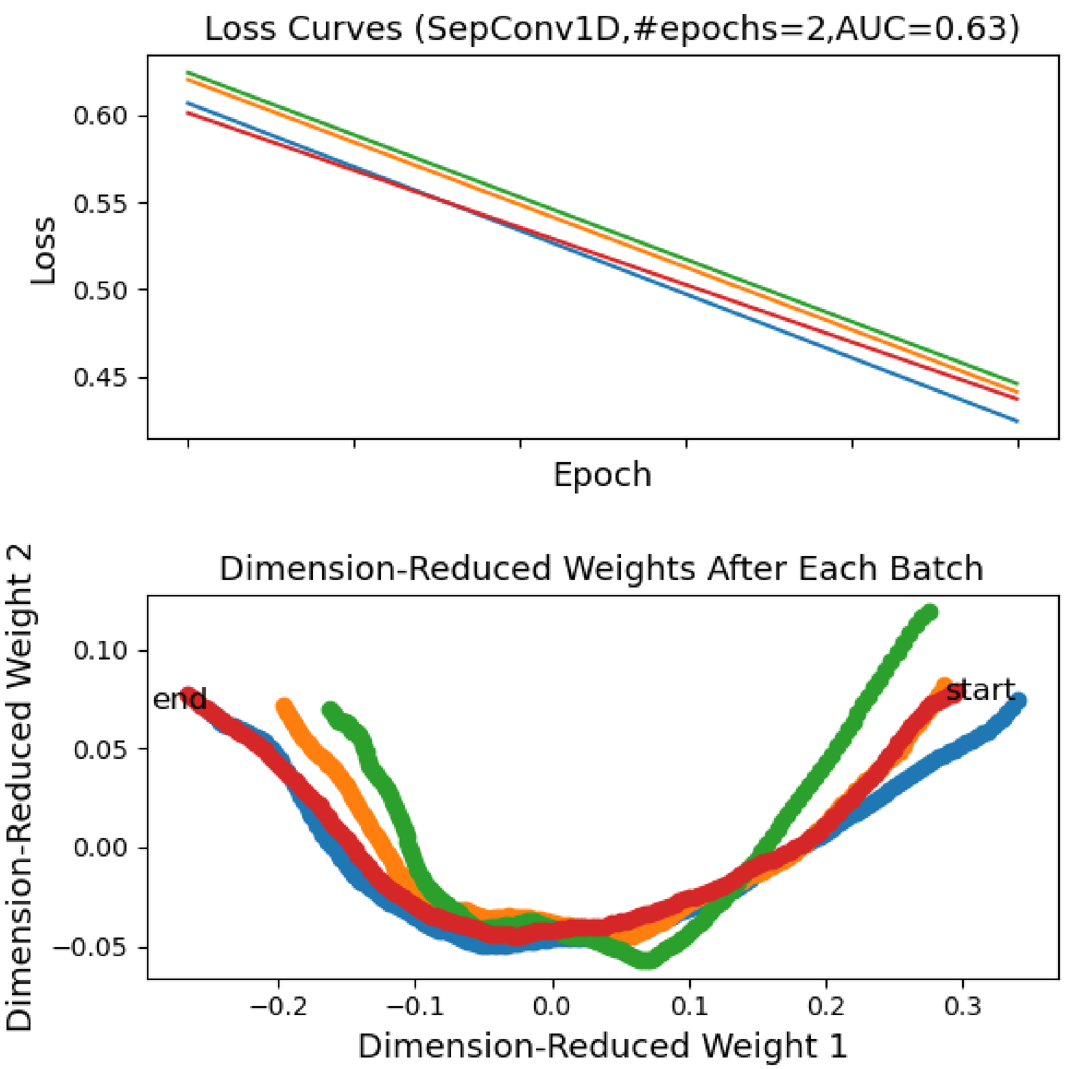
Table 3. AUC values on Dataset1.

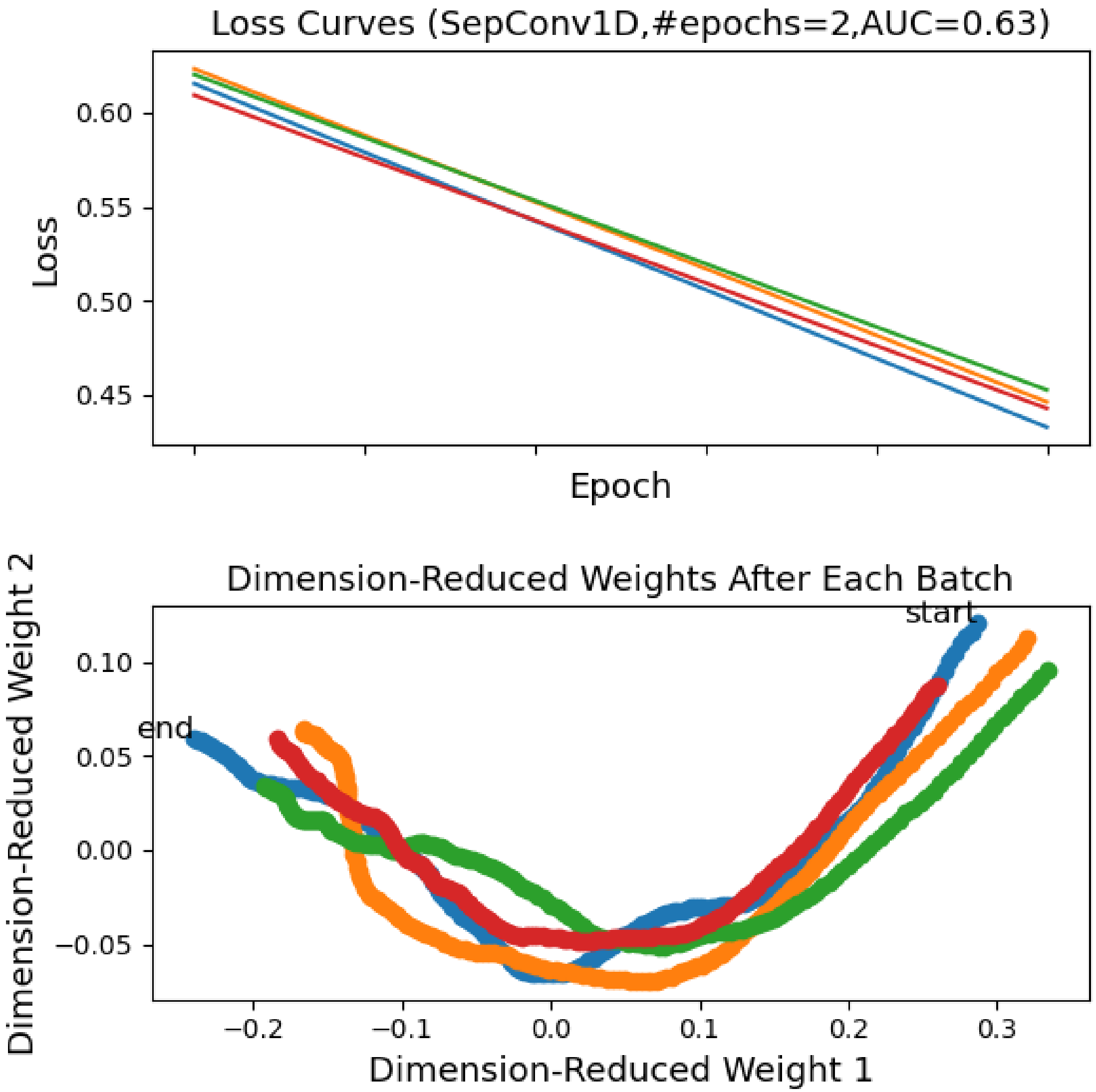
Test subject	SepConv1D 4 filters 2 epochs	SepConv1D 1 filter 2 epochs	BCNN 1 filter 2 epochs	SepConv1D 4 filters 33 epochs	BCNN 3 filters 33 epochs
0	0.80	0.77	0.85	0.86	0.87
1	0.86	0.65	0.86	0.87	0.86
2	0.87	0.83	0.87	0.87	0.88
3	0.69	0.71	0.73	0.73	0.73
4	0.63	0.69	0.71	0.7	0.73
5	0.84	0.84	0.89	0.89	0.89
6	0.72	0.64	0.78	0.78	0.79
7	0.74	0.70	0.78	0.77	0.80
8	0.85	0.80	0.86	0.86	0.88
9	0.83	0.80	0.85	0.83	0.85
10	0.81	0.77	0.77	0.77	0.82
11	0.78	0.67	0.77	0.78	0.79
12	0.80	0.77	0.84	0.83	0.83
13	0.84	0.65	0.85	0.84	0.87
14	0.80	0.74	0.82	0.79	0.83
15	0.86	0.79	0.83	0.84	0.84
16	0.67	0.64	0.73	0.71	0.74
17	0.88	0.81	0.90	0.9	0.91
18	0.91	0.88	0.89	0.9	0.91
19	0.85	0.87	0.88	0.87	0.88
20	0.87	0.77	0.90	0.89	0.90
21	0.78	0.71	0.82	0.82	0.83
Mean	**0.80±0.07**	**0.75±0.07**	**0.83±0.06**	**0.82±0.06**	**0.84±0.06**

Insights from Explainable AI: weight visualization

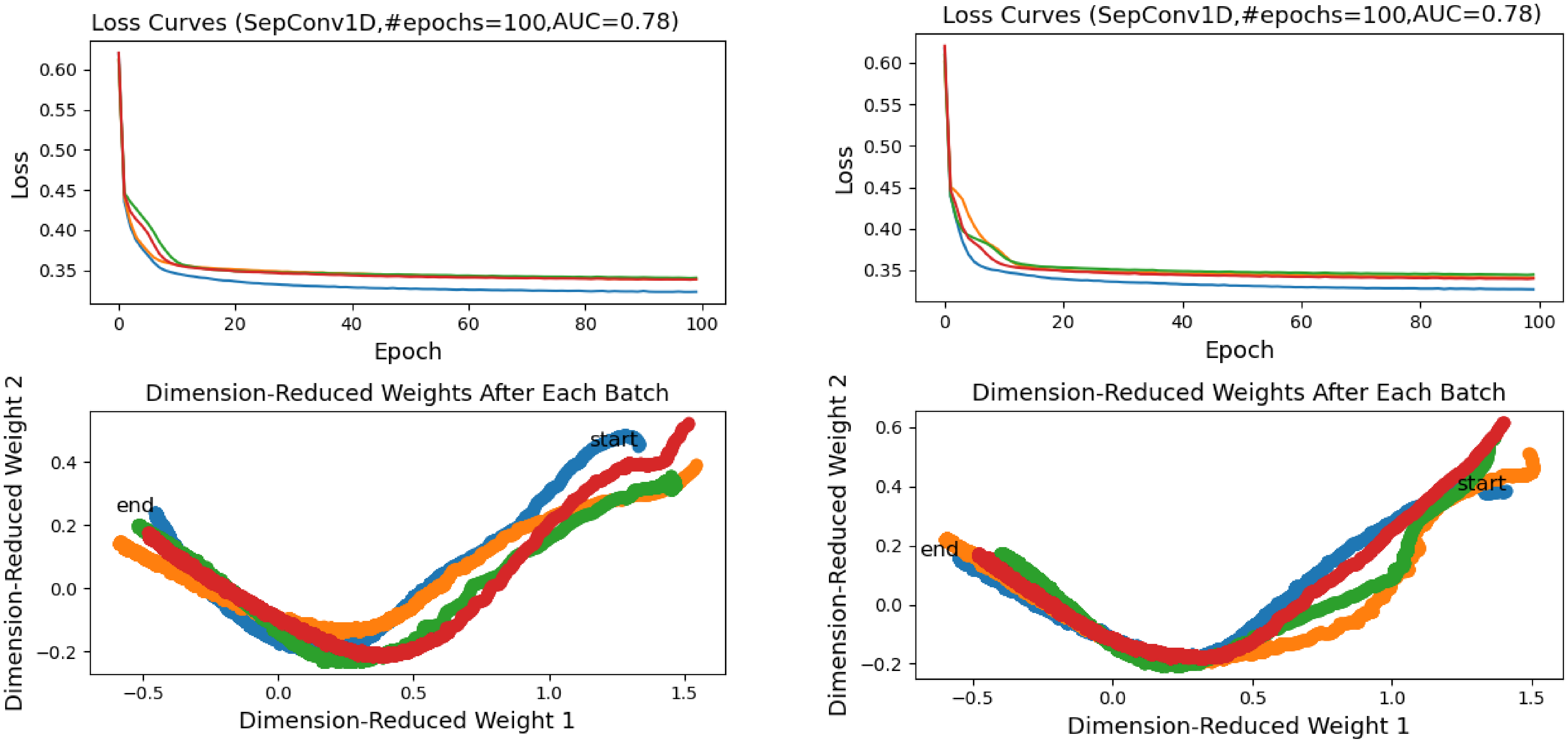


Visualization of the weight-updating process





Visualization of the weight-updating process



Insights from Explainable AI: saliency maps

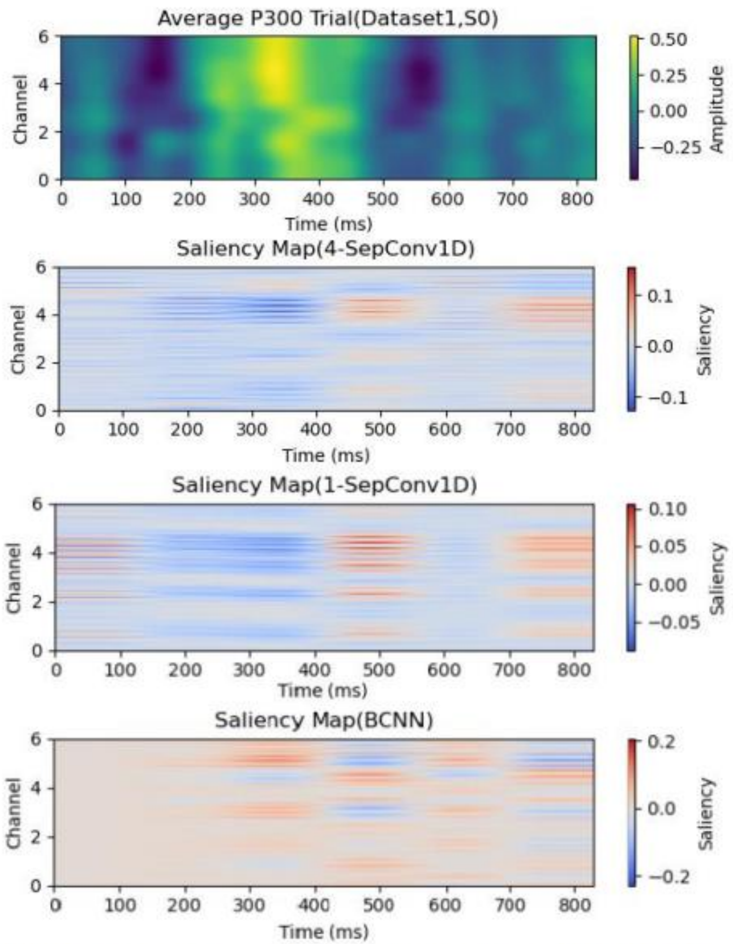


Figure 5. The average saliency maps of three CNNs (4-filter SepConv1D, 1-filter SepConv1D and BCNN) on the average P300 trial of subject 0 in Dataset1.

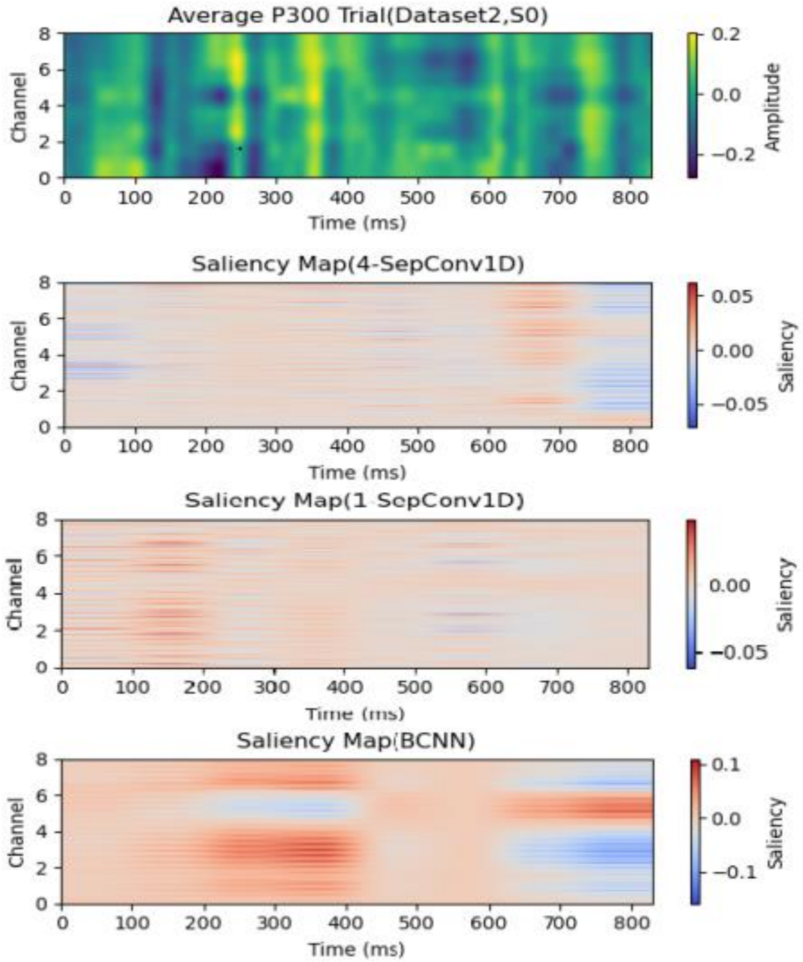


Figure 6. The average saliency maps of three CNNs (4-filter SepConv1D, 1-filter SepConv1D and BCNN) on the average P300 trial of subject 0 in Dataset2.

Summary

- **Proposes Bantamweight CNN (BCNN) for P300 Detection**
- **Separable Conv & Large learning rate**
- **1 filter & 165 parameters & 2 epochs**
- **State-of-the-art performance**
- **Insights from Explainable AI**

Thanks !
Any question ?