Supplementary Material

AlphaGrad: Normalized Gradient Descent for Adaptive Multi-loss Functions in EEG-based Motor Imagery Classification

Rattanaphon Chaisaen[1], Phairot Autthasan[1], Apiwat Ditthapron[2], and Theerawit Wilaiprasitporn[1] *Senior Member, IEEE*

[1]Bio-inspired Robotics and Neural Engineering (BRAIN) Lab,
School of Information Science and Technology (IST),
Vidyasirimedhi Institute of Science and Technology (VISTEC),
Rayong, Thailand

[2]Department of Innovation Information Systems,
Faculty of Business Administration,
Rajamangala University of Technology Krungthep,
Bangkok, Thailand

Corresponding authors
Apiwat Ditthapron (e-mail: apiwat.d@mail.rmutk.ac.th), and
Theerawit Wilaiprasitporn (e-mail: theerawit.w@vistec.ac.th)

S1 Multi-Task Learning-Based MI-EEG Classification Architecture

We implemented MIN2Net [1] and FBMSNet [2] based on the descriptions provided in their respective publications. One notable modification made to MIN2Net (our previous work) is the migration from the TensorFlow framework to PyTorch. In this transition, we introduced minor changes to the decoder layers to align with PyTorch's implementation standards. The design philosophies of both models are discussed in Section II (Related Work) of the main manuscript. Here, we provide implementation-level details of MIN2Net in Table S1 and FBMSNet in Table S2. Additionally, the model complexity, including comparisons with single-task state-of-the-art baseline models, is summarized in Table S3. The complete source code and implementation used in our experiments are available at: `https://AlphaGrad.github.io`.

Table S1: MIN2Net architecture.

Block	**Layer**	**Filter**	**Size**	**Stride**	**Activation**	**Options**	**Output**
Encoder	Input		$(C,T,1)$				$(C,T,1)$
	Conv2D	C	$(64,1)$	1		padding=same	$(C,T,1)$
	Activation				ELU		
	BatchNorm2D						$(C,T,1)$
	AvgPool2D		$(T//FS,1)$				$(C,FS,1)$
	Conv2D	$C//2$	$(32,1)$	1		padding=same	$(C//2,FS,1)$
	Activation				ELU		
	BatchNorm2D						$(C//2,FS,1)$
	AvgPool2D		$(4,1)$				$(C//2,FS//4,1)$
	Flatten						$(C//2)*(T//(T//FS))//4$
Latent	FC		(z)				(z)
Decoder	FC		(z)				$(C//2)*(FS//4)$
	Unflatten		$(C//2,FS//4,1)$			dim=1	$(C//2,FS//4,1)$
	Conv2D	$C//2$	$(64,1)$	4		padding=same	$(C//2,FS//4,1)$
	Upsampling2D		$(FS,1)$		ELU		$(C//2,FS,1)$
	Activation						
	Conv2D		$(32,1)$			padding=same	$(C,FS,1)$
	Upsampling2D		$(T,1)$		ELU		$(C,T,1)$
	Activation						
Supervised Learning	Latent						(z)
	FC	N_c			Softmax		(N_c)

C is number of EEG channels, T is number of time points, FS is EEG sampling rate, z is latent size, and Nc is number of output classes.

Table S2: FBMSNet architecture.

Block	**Layer**	**Filter**	**Size**	**Stride**	**Activation**	**Options**	**Output**
Filtering	Input		(N_b,C,T)				(N_b,C,T)
	Conv2D	9	$(1,15)$	1	Linear	padding=same	$(9,C,T)$
	Conv2D	9	$(1,31)$	1	Linear	padding=same	$(9,C,T)$
	Conv2D	9	$(1,63)$	1	Linear	padding=same	$(9,C,T)$
	Conv2D	9	$(1,125)$	1	Linear	padding=same	$(9,C,T)$
Feature Extraction	Concatenate					dim=1	$(36,C,T)$
	BatchNorm						$(36,C,T)$
	DepthwiseConv2D	288	$(C,1)$	1	Linear	group=36	$(228,1,T)$
	BatchNorm						$(228,1,T)$
	Activation				swish		$(228,1,T)$
	Reshape						$(228,T//w,w)$
	Variance		(w)				$(228,T//w,1)$
Latent	Flatten						$(228*(T//w))$
	FC		$(228*(T//w))$				$(228*(T//w))$
Supervised Learning	FC	N_c			LogSoftmax		(N_c)

C is number of EEG channels, T is number of time points, Nb is the number of frequency bands, w is the window length of variance layer, and Nc is number of output classes.

S2 Model Complexity

To evaluate the real-time feasibility of the resulting MTL networks (MIN2Net and FBMSNet) in comparison with single-task deep learning models commonly used for EEG-based BCIs, we assessed both architectural complexity (in terms of total trainable parameters) and inference time, including the preprocessing step of band-pass filtering. Table S3 summarizes the number of filtering bands, total trainable parameters, and the average inference time per sample for each model. Importantly, the reported inference time includes the band-pass filtering step, thereby providing a realistic estimate of the end-to-end latency in practical BCI systems. The low latency observed in both MTL models indicates their viability for real-time applications, where fast and accurate decoding is critical for user feedback.

Table S3: Model Complexity.

Model	**#Filtring Bands**	**Total Parameters**	**Inference time (ms/sample)**
DeepConvNet	1	153,427	3.8925 ± 0.5336
EEGNet	1	4,402	2.9769 ± 0.5393
FBCNet	9	8,930	13.3224 ± 3.0026
MIN2Net	1	55,232	3.8065 ± 0.5934
FBMSNet	9	13,349	10.5056 ± 1.5961

References

[1] P. Autthasan, R. Chaisaen, T. Sudhawiyangkul, P. Rangpong, S. Kiatthaveephong, N. Dilokthanakul, G. Bhakdisongkhram, H. Phan, C. Guan, and T. Wilaiprasitporn, "Min2net: End-to-end multi-task learning for subject-independent motor imagery eeg classification," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 6, pp. 2105–2118, 2022.

[2] K. Liu, M. Yang, Z. Yu, G. Wang, and W. Wu, "Fbmsnet: A filter-bank multi-scale convolutional neural network for eeg-based motor imagery decoding," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 2, pp. 436–445, 2023.