DMSACNN: Deep Multiscale Attentional Convolutional Neural Network for EEG-Based Motor Decoding

Ke Liu, Xin Xing, Tao Yang, Zhuliang Yu, *Member, IEEE*, Bin Xiao, Guoyin Wang*, Wei Wu*, *Senior Member, IEEE*

I. Experiments

A two-stage training strategy was adopted to prevent overfitting and reduce required epochs. The training procedure is detailed in Algorithm 1, while the trainable parameters for all deep learning algorithms are shown in Table I. Specifically, we implemented Deep ConvNet, EEGNet, and FBCNet using code from `https://github.com/ravikiran-mane/FBCNet`, EEG-Inception from `https://github.com/esantamariavazquez/EEG-Inception`, and EEG Conformer from `https://github.com/eeyhsong/EEG-Conformer`.

Algorithm 1: Two-stage Training Strategy

```
Input: $N$ trials of training data $X_{\text{train}} \in \mathbb{R}^{N \times C \times T}$, true labels $Y_{\text{train}} \in \mathbb{R}^{N \times 1}$, initialized parameters of DMSACNN $\Theta$ and the maximum number of training epochs.
Output: The parameters of DMSACNN $\Theta$.
Stage 1:
$X_{sub}, X_{val}$ = func_split($X_{train}$);
$Y_{sub}, Y_{val}$ = func_split($Y_{train}$);
while $epoch < max\_epochs\_1st$ do
    Train Network $(X_{sub}, Y_{sub}, \Theta)$;
    $pred_{sub}, loss_{sub}$ = func_predict($X_{sub}, Y_{sub}, \Theta$);
    $pred_{val}, loss_{val}$ = func_predict($X_{sub}, Y_{sub}, \Theta$);
    if $loss_{val} < min\_loss$ then
        $min\_loss = loss_{val}$;
        best_model = $\Theta$;
    end
    $doStop$ = func_stopCheck($loss_{val}, min\_loss$);
    if $doStop$ then
        $\epsilon = loss_{sub}$;
        break;
    end
    $epoch = epoch + 1$;
end
Stage 2:
while $epoch < max\_epochs\_2st$ do
    Train Network $(X_{train}, Y_{train}, \Theta)$;
    $pred_{val}, loss_{val}$ = func_predict($X_{val}, Y_{val}, \Theta$);
    if $loss_{val} < min\_loss$ then
        $min\_loss = loss_{val}$;
        best_model = $\Theta$;
    end
    $doStop$ = func_stopCheck($loss_{val}, \epsilon$);
    if $doStop$ then
        $\Theta_{best}$ = best_model;
        break;
    end
    $epoch = epoch + 1$;
end
```

TABLE I
Number of learnable parameters for different algorithms across three datasets.

Dataset	Deep ConvNet	EEGNet	EEG-Inception	FBCNet	EEG Conformer	DMSACNN
BCI-IV-2a	282879	4028	29816	11812	789572	35884
HGD	296629	4380	30872	18148	824772	38524
OpenBMI	278827	3002	29706	8930	786306	20762