Integrating Transient and Long-term Physical States for Depression Intelligent Diagnosis

Ke Wu, Han Jiang, Li Kuang, Yixuan Wang, Huaiqian Ye, Yuanbo He

KeyWords: EEG, Tongue, MultiModal, Traditional Chinese Medicine

Motivations:
- 1. Traditional Depression Diagnosis: Subjective and Scale-Based Approaches
- 2. Unimodal EEG Diagnosis: Lacks Long-term Analysis
- 3. TCM Tongue Observation: Awaiting Intelligent Diagnostic Methods

Contributions
- 1. Comprehensive Multimodal Diagnosis Using EEG and Tongue Images
- 2. New Architecture for EEG Spatio-temporal feature extraction
- 3. New Methods for TCM Intelligent Diagnosis Based on Tongue Images

Results
EEG: SOTA on MODMA
- Tongue: First Depression-related Work
Fusion: Correlation Validation and Acc Improvement

Table 4: Comparison on MODMA

<table>
<thead>
<tr>
<th>Model</th>
<th>MODMA Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trans. EEGNet</td>
<td>90.4% ± 2.6%</td>
</tr>
<tr>
<td>MR4S</td>
<td>92.3±0.008</td>
</tr>
<tr>
<td>GE (SSE)</td>
<td>98.9±0.000</td>
</tr>
<tr>
<td>CNN (GER-4TNES4I)</td>
<td>99.3±0.004</td>
</tr>
<tr>
<td>FTJPS (SPE)</td>
<td>99.4±0.001</td>
</tr>
<tr>
<td>SPE (SPE)</td>
<td>99.5±0.002</td>
</tr>
</tbody>
</table>

An overview of MMTV which consists of three modules:
1) EEG module with Trans_EEGNet featuring the dual-stream input and self-attention mechanism.
2) Tongue module with multi-step pre-training method (loading pretrained weight on ImageNet, training Siamese net as meta-learning, training with unpaired tongue images)
3) Fusion module with best result selected from results of methods such as concat, DCCA et al.