Exploring Spatial–Temporal Representations for fNIRS-based Intimacy Detection via Attention–enhanced Cascade Convolutional Recurrent Neural Network

Chao Li 1, Qian Zhang 1, Ziping Zhao 1, Li Gu 2,3, Björn Schuller 4,5

1 College of Computer and Information Engineering, Tianjin Normal University, China
2 School of Humanities and Management, Guangdong Medical University, China
3 Faculty of Psychology, Tianjin Normal University, China
4 ZD.B Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Germany
5 Group on Language, Audio & Music, Imperial College London, UK
Authors

- Chao Li
- Qian Zhang
- Ziping Zhao
- Li Gu
- Björn Schuller
Motivation

- Can we automatically detect the categories of intimacy by analysing the activity of certain brain regions?
- Considering the temporal dynamic characteristics of intimacy relationship on neural mechanism, how to model spatio-temporal dynamics for intimacy prediction effectively is still a challenge.
- Not all channel contributes to intimacy detection, how to find the most important channels to infer intimate relationship?

Given the advantages of time-frequency resolution in complex neuronal activities analysis, this paper utilizes functional near-infrared spectroscopy (fNIRS) to analyse and infer to intimate relationship.
Contribution

(1) A fNIRS-based database is collected to analyze human’s complex brain response pattern corresponding to intimacy. Forty-two-channel fNIRS signals are recorded from 44 subjects when they watching the pictures from his lover, friend and stranger;

(2) a fNIRS-based cascade deep learning architecture is utilized to detect three different intimacy classes, including lover, friend and stranger. Compared with hand-craft features, the proposed method can automatically extract spatial and temporal features from fNIRS signals by cascade convolutional recurrent neural network, which is capable of learning feature representations and modeling the spatial-temporal dependencies between their activation;

(3) we also investigate the usage of attention-based architectures to improve fNIRS-based intimacy prediction. The attention mechanism allows the network to focus on the salient parts of a sequence.
METHODOLOGY

(a) collects intimacy-induced fNIRS signals;
(b) removes fNIRS signal noise;
(c) captures high-level spatial representation by CNN model;
(d) captures high-level temporal representation by LSTM model;
(e) finds the salient parts of a sequence for intimacy by attention mechanism;
(f) learns the final representation for intimate relationship category
(g) predicts the final intimate relationship category by a softmax classifier
(h) outputs predicted results.

Fig. 1. The framework of attention-enhanced cascade convolutional recurrent neural network based on fNIRS signals for intimacy prediction.
EXPERIMENTS AND RESULTS

• Data acquisition

1) Participants: In order to effectively analyse and infer to intimacy, forty-four healthy subjects were recruited for the experiment, 25 males and 19 females with an average age of $22.12 \pm 2.51$ years old and $20.4 \pm 2.11$ years old. All of subjects are right-handed, with normal vision or corrected vision, no history of mental illness, and no major conflict with lovers during the week before the visit. Before the experiment, the principle of the instrument was introduced to ensure that it was harmless to the human body and does not involve any ethical issues. And the participants were asked to sign the experimental informed consent form.

2) Stimuli: We adopt the photos from subject’s lover, friends, and strangers to induce his/her brain imaging in different relationships. Each participant was asked to provide 20 photos (10 for lovers and 10 for friends). Thirty volunteers (15 male, 15 female, unrelated to the experiment) provided 60 photos (2 per person) as the induction of stranger photos. During the experiment, each subject views the photos from his/her lover, friends and strangers by random selection.
EXPERIMENTS AND RESULTS

- Data acquisition

3) Instrumentation & 4) Experimental protocol:

Fig. 2. Experimental paradigm for inducing subject’s fNIRS response with different intimate relationships
EXPERIMENTS AND RESULTS

- Network structure of our proposed method

<table>
<thead>
<tr>
<th>Layer (type)</th>
<th>Output Shape</th>
<th>Param #</th>
</tr>
</thead>
<tbody>
<tr>
<td>conv2d_1 (Conv2D)</td>
<td>(None, 110, 42, 32)</td>
<td>320</td>
</tr>
<tr>
<td>conv2d_2 (Conv2D)</td>
<td>(None, 110, 42, 16)</td>
<td>4624</td>
</tr>
<tr>
<td>max_pooling2d_1 (MaxPooling2)</td>
<td>(None, 55, 42, 16)</td>
<td>0</td>
</tr>
<tr>
<td>reshape_1 (Reshape)</td>
<td>(None, 55, 672)</td>
<td>0</td>
</tr>
<tr>
<td>lstm_1 (LSTM)</td>
<td>(None, 55, 128)</td>
<td>410112</td>
</tr>
<tr>
<td>dropout_1 (Dropout)</td>
<td>(None, 55, 128)</td>
<td>0</td>
</tr>
<tr>
<td>attention_1 (Attention)</td>
<td>(None, 128)</td>
<td>183</td>
</tr>
<tr>
<td>dense_1 (Dense)</td>
<td>(None, 32)</td>
<td>4128</td>
</tr>
<tr>
<td>dropout_2 (Dropout)</td>
<td>(None, 32)</td>
<td>0</td>
</tr>
<tr>
<td>dense_2 (Dense)</td>
<td>(None, 3)</td>
<td>99</td>
</tr>
</tbody>
</table>

Fig. 4 The network structure and hyper-parameters of our proposed method
EXPERIMENTS AND RESULTS

- Results and Discussion:
  - Performance on different methods for fnirs-based intimacy prediction with different instance length

<table>
<thead>
<tr>
<th>Classifier</th>
<th>1 second</th>
<th></th>
<th></th>
<th>5 second</th>
<th></th>
<th></th>
<th>10 second</th>
<th></th>
<th>30 second</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC(%)</td>
<td>Loss</td>
<td>ACC(%)</td>
<td>Loss</td>
<td>ACC(%)</td>
<td>Loss</td>
<td>ACC(%)</td>
<td>Loss</td>
<td>ACC(%)</td>
<td>Loss</td>
</tr>
<tr>
<td>SVM</td>
<td>69.2</td>
<td>-</td>
<td>60.5</td>
<td>-</td>
<td>45.7</td>
<td>-</td>
<td>34.6</td>
<td>-</td>
<td>33.7</td>
<td>-</td>
</tr>
<tr>
<td>RF</td>
<td>41.5</td>
<td>-</td>
<td>40.8</td>
<td>-</td>
<td>38.6</td>
<td>-</td>
<td>37.1</td>
<td>-</td>
<td>34.4</td>
<td>-</td>
</tr>
<tr>
<td>LDA</td>
<td>37.8</td>
<td>-</td>
<td>37.1</td>
<td>-</td>
<td>35.0</td>
<td>-</td>
<td>34.3</td>
<td>-</td>
<td>33.6</td>
<td>-</td>
</tr>
<tr>
<td>1-layer CNN</td>
<td>73.7</td>
<td>0.656</td>
<td>53.8</td>
<td>0.974</td>
<td>38.0</td>
<td>1.135</td>
<td>35.6</td>
<td>1.572</td>
<td>33.1</td>
<td>1.971</td>
</tr>
<tr>
<td>2-layer CNN</td>
<td>86.3</td>
<td>0.401</td>
<td>83.0</td>
<td>0.496</td>
<td>54.2</td>
<td>1.837</td>
<td>40.1</td>
<td>2.854</td>
<td>30.7</td>
<td>3.248</td>
</tr>
<tr>
<td>3-layer CNN</td>
<td>67.4</td>
<td>0.821</td>
<td>60.9</td>
<td>0.866</td>
<td><strong>65.2</strong></td>
<td><strong>0.845</strong></td>
<td><strong>54.6</strong></td>
<td>2.312</td>
<td>31.8</td>
<td>7.572</td>
</tr>
<tr>
<td>1-layer LSTM</td>
<td>86.9</td>
<td>0.572</td>
<td>32.9</td>
<td>4.958</td>
<td>34.9</td>
<td>6.251</td>
<td>35.7</td>
<td>2.875</td>
<td>27.8</td>
<td>1.101</td>
</tr>
<tr>
<td>2-layer LSTM</td>
<td>50.1</td>
<td>0.877</td>
<td>33.8</td>
<td>9.465</td>
<td>34.7</td>
<td>7.263</td>
<td>32.4</td>
<td>3.821</td>
<td>28.4</td>
<td>0.991</td>
</tr>
<tr>
<td>CCRNN</td>
<td>95.7</td>
<td>0.178</td>
<td>80.1</td>
<td>0.539</td>
<td>47.1</td>
<td>1.013</td>
<td>38.6</td>
<td>4.567</td>
<td>32.9</td>
<td>2.802</td>
</tr>
<tr>
<td>Ours</td>
<td><strong>97.4</strong></td>
<td><strong>0.130</strong></td>
<td><strong>94.8</strong></td>
<td><strong>0.247</strong></td>
<td>38.6</td>
<td>3.903</td>
<td>36.2</td>
<td>1.126</td>
<td>27.9</td>
<td>1.167</td>
</tr>
</tbody>
</table>
Results and Discussion:

- Accuracy and loss curves during the testing for 200 epochs under different methods

(a) Accuracy curves

(b) Loss curves
Conclusions

The proposed cascade model in the paper is motivated by the existing progress on deep models, and takes advantage of CNN, LSTM, the attention mechanism achieves intimacy prediction. With the proposed model, we achieved a potent improvement in the current state-of-the-art for the task of intimacy prediction on the fNIRS-based dataset. The increase in performance in comparison to other existing models shows that attention mechanism can improve the performance of cascade convolutional recurrent neural network for intimacy prediction. An overall analysis of the performance of our proposed method is provided and compared to other techniques.

In the future, we will expend our dataset by increasing the number of subjects to make it publicly available to the research community, and multi-modal fusion method will also be investigated to further boost the performance of the intimacy prediction task.
References

References


References


Thank you!