Multimodel Emotion Recognition Using Deep Learning

Bao-Liang Lu
Center for Brain-Like Computing and Machine Intelligence
Department of Computer Science & Engineering
Key Lab. of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering
Brain Science and Technology Research Center
Shanghai Jiao Tong University, China
Outline

- Background and Current Research Situations
- Multimodal Emotion Recognition
  - Emotion Recognition Using EEG and Eye Movement Data
  - Personalizing Affective Models with Transfer Learning
  - EEG Data Augmentation Using GAN
  - Other Work in my lab and SEED Data Sets
- Toward Emotional Intelligence: Four Phases
- Summary
In this mind-expanding book, scientific pioneer Marvin Minsky continues his groundbreaking research, offering a fascinating new model for how our minds work. He argues persuasively that emotions, intuitions, and feelings are not distinct things, but different ways of thinking.
20 Big Questions about the Future of Humanity

1. Does humanity have a future beyond Earth?
2. When and where do you think we will find extraterrestrial life?
3. Will we ever understand the nature of consciousness?
4. Will robots replace humans in the workforce?
5. Will we find evidence of intelligent life in our solar system?
6. Will we ever understand the rules of the universe?
7. Will we be able to reverse aging and extend human lifespan?
8. Will we find a cure for cancer and other diseases?
9. Will we ever stop wars and conflict?
10. Will we be able to change the course of human history?
11. Will we be able to create a sustainable future for all?
12. Will we be able to solve the problem of climate change?
13. Will we ever be able to control the weather?
14. Will we be able to prevent nuclear war?
15. Will we use wearable technologies to detect our emotions?
16. Will we be able to control the effects of gravity in space?
17. Will we ever find a way to cure Alzheimer’s disease?
18. Will we ever be able to cure all diseases?
19. Will gender equality be achieved in the sciences?
20. Do you think we will one day be able to predict natural disasters such as earthquakes with warning times of days or hours?
“Emotions involve biochemical and electrical signals that reach every organ in our bodies—allowing, for example, stress to impact our physical and mental health.

Wearable technologies let us quantify the patterns in these signals over long periods of time.

In the coming decade wearables will enable the equivalent of personalized weather forecasts for our health.

Over the next 20 years, wearables, and analytics derived from them, can dramatically reduce psychiatric and neurological disease.”
“By 2022, your personal device will know more about your emotional state than your own family.”

Annette Zimmermann, research vice president at Gartner
Artificial Emotional Intelligence or Emotion AI

- Artificial emotional intelligence, which is also known as emotion AI or affective computing, is already being used to develop systems and products that can recognize, interpret, process, and simulate human affects.
- In psychology, an “affect” is a term used to describe the experience of feeling or emotion.
Affective Brain-Computer Interaction (aBCI)

- A brain-computer interface (BCI) is a direct communication pathway between a brain and an external device.
- A brain-computer interaction does not only rely on direct measurement of brain activity, but also includes signals from other physiological activity such as EOG, EMG, or ECG.
- **Affective Brain-Computer Interaction (aBCI)** aims to enhance brain-computer interaction systems with the ability to detect, process, and respond to users affect, emotion or mood.
BNCI
Horizon 2020
Roadmap of BCI
(http://bnci-horizon-2020.eu/)
(15 Jan. 2015)
BNCI Horizon 2020 Roadmap
Germanwings Flight 9525: Emotion AI is demanded!

- On 24 March 2015, the aircraft, an Airbus A320-211, crashed 100 kilometres northwest of Nice in the French Alps. All 144 passengers and six crew members were killed!

- The crash was deliberately caused by the copilot, Andreas Lubitz, who had previously been treated for suicidal tendencies and been declared "unfit to work" by a doctor.

- The European Federation of Psychologists' Associations (EFPA) issued a statement supporting psychological testing in the selection of pilots.
What is an Emotion?

- **An informal definition:**
  Emotion is a set of complex interactions between subjects and objects, which is regulated by nervous system and hormone. These interactions can trigger some emotional feelings or produce some cognition.

- **Classic model:**
  The classic model divides emotions into a group of different and discrete emotion state sets, such as anger, fear, sadness, happiness, surprise, disgust etc.

- **Dimension Model**
  Arousal refers to the quantitative activation level ranging from calm to excited; Valence represents the quality of an emotion ranging from unpleasant to pleasant;
Spatiotemporal Extent of an Emotion State

Emotion Can be Inferred from Various Data

Emotions are Functional States

Limitations of Face, Voice and Gesture Measures

- Sometimes the emotional states remain internal and cannot be detected by external expression.
- Considering extreme cases where people do not say anything but actually they are angry, even smile during negative emotional states due to the social masking.
- In these cases, the external emotional expression can be controlled subjectively and these external cues for emotion recognition may be in inadequate.
Outline

- Background and Current Research Situations
- Multimodal Emotion Recognition
  - Emotion Recognition Using EEG and Eye Movement Data
  - Personalizing Affective Models with Transfer Learning
  - EEG Data Augmentation Using GAN
  - Other Work in my lab and SEED Data Sets
- Toward Emotional Intelligence: Four Phases
- Summary
Emotion Recognition Using Physiological Signals

- Human brains are difficult to measure “in real life”. Behavior is often used as a proxy for physiological processes.
- Physiological signals can be employed as **Ground Truth Data** to validate behavioral algorithms.

D. Ravi, D. et al, 2017
Multimodal Emotion Recognition

- EEG
- Physiological Signals: EOG, SC, etc.
- Visual Signals: Facial Expression
Framework of Multimodal Emotion Recognition

Signal Acquisition → Artifact Removal → Temporal Filtering → Spatial Filtering

Feature Extraction → Dimension reduction → Classification or clustering → Cross Validation
Emotion Recognition Using EEG and Eye Movement Data

Why We Choose EEG and Eye Movement Signals?

- Single modal may not work well.
- The signals from different modalities have complementary information of emotions.
- Combining EEG and eye movements respectively from internal cognitive states and external subconscious behaviors can probably enhance the performance of emotion recognition.
- The more information, the better.
We extract 33 eye movement features in total.

<table>
<thead>
<tr>
<th>Eye movement parameters</th>
<th>Extracted features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupil diameter (X and Y)</td>
<td>Mean, standard deviation and PSD (or DE) in four bands: 0-0.2 Hz, 0.2-0.4 Hz, 0.4-0.6 Hz, 0.6-1 Hz</td>
</tr>
<tr>
<td>Dispersion (X and Y)</td>
<td>Mean, standard deviation</td>
</tr>
<tr>
<td>Fixation duration (ms)</td>
<td>Mean, standard deviation</td>
</tr>
<tr>
<td>Blink duration (ms)</td>
<td>Mean, standard deviation</td>
</tr>
<tr>
<td>Saccade</td>
<td>Mean, standard deviation of saccade duration (ms) and saccade amplitude (°)</td>
</tr>
<tr>
<td>Event statistics</td>
<td>Blink frequency, fixation frequency, fixation duration maximum, fixation dispersion total, fixation dispersion maximum, saccade frequency, saccade duration average, saccade amplitude average, saccade latency average(^1).</td>
</tr>
</tbody>
</table>
Experiment Setup for Eliciting Three Emotions

Positive | Neutral | Negative

- **Hint of Start**
  - 5 sec

- **Movie clip**
  - 4 min

- **Self-assessment**
  - 45 sec

- **Rest**
  - 15 sec

- **Session 1**
- **Session 2**
- **Session 3**
- **...**
- **Session N**
Data Fusion Strategies and Recognition Accuracy

- **Modality Fusion Strategies**
  - Feature level fusion (FLF): concatenating two feature vectors
  - Decision level fusion (DLF): maximal rule, sum rule, fuzzy integral

- **SVMs with linear kernel are used as classifiers**

Lu, Zheng, Li, and Lu, *IJCAI'15*, 2015
Complementary Characteristics of EEG and EM (1)

- Eye movement modality is not so good at classifying happy state as EEG modality.
- Neutral and sad states can be recognized with higher accuracy in eye movement modality than in EEG modality.
- Fuzzy integral is much better at classifying neutral and sad states.

Lu, Zheng, Li, and Lu, *IJCAI'15*
Complementary Characteristics of EEG and EM (2)


Various findings of activated patterns associated with different emotions have been reported.

The stability of neural patterns over time has not been fully investigated yet.

Identifying neural patterns that are both common across subjects and stable across sessions can provide valuable information for emotion recognition.

Are there any stable EEG patterns or brain regions for representing emotions?
The lateral temporal areas activate more for positive emotion than negative one in beta and gamma bands.

Neural patterns of neutral emotion are similar to that of negative emotion.

Demo of Stable Neural Patterns over Time

Eye Tracking

Positive

Negative
Stability of Emotion Recognition over Time

- A comparative mean classification accuracy of **79.28%** is achieved with training and test datasets from different sessions.
- The performance of the emotion recognition model is better with training data and test data obtained from sessions performed in nearer time.
- The relation between the variation of emotional states and the EEG signals is stable for one person in a period of time.

<table>
<thead>
<tr>
<th>Stats.</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First</td>
<td>Second</td>
</tr>
<tr>
<td>Mean</td>
<td>90.83</td>
<td>72.55</td>
</tr>
<tr>
<td></td>
<td>75.86</td>
<td>88.22</td>
</tr>
<tr>
<td></td>
<td>76.28</td>
<td>78.17</td>
</tr>
<tr>
<td>Std.</td>
<td>8.64</td>
<td>10.29</td>
</tr>
<tr>
<td></td>
<td>7.71</td>
<td>8.59</td>
</tr>
<tr>
<td></td>
<td>11.47</td>
<td>13.41</td>
</tr>
</tbody>
</table>
How One Intelligent Machine Learned to Recognize Human Emotions

Nobody knew how to identify people’s emotional states by looking at their brain waves. Then a machine learning algorithm stepped in.

January 23, 2016
# Investigating Critical Channels

---

<table>
<thead>
<tr>
<th>#</th>
<th>Feat.</th>
<th>Stat.</th>
<th>$\delta$</th>
<th>$\theta$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>PSD</td>
<td>Mean</td>
<td>39.23</td>
<td>39.19</td>
<td>38.77</td>
<td>39.90</td>
<td>41.22</td>
<td><strong>60.22</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>8.36</td>
<td>8.00</td>
<td>7.35</td>
<td>9.15</td>
<td>10.77</td>
<td>16.18</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Mean</td>
<td>36.84</td>
<td>37.18</td>
<td>39.25</td>
<td>39.23</td>
<td>41.21</td>
<td><strong>64.24</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>6.76</td>
<td>7.01</td>
<td>9.63</td>
<td>9.61</td>
<td>9.59</td>
<td>15.39</td>
</tr>
<tr>
<td>4</td>
<td>PSD</td>
<td>Mean</td>
<td>47.30</td>
<td>47.57</td>
<td>50.73</td>
<td>52.83</td>
<td>51.26</td>
<td><strong>58.98</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>12.29</td>
<td>12.22</td>
<td>15.65</td>
<td>13.26</td>
<td>15.04</td>
<td>17.85</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Mean</td>
<td>45.21</td>
<td>47.99</td>
<td>50.65</td>
<td>54.41</td>
<td>57.19</td>
<td><strong>67.02</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>14.64</td>
<td>12.69</td>
<td>13.68</td>
<td>14.44</td>
<td>17.59</td>
<td>15.87</td>
</tr>
<tr>
<td>6</td>
<td>PSD</td>
<td>Mean</td>
<td>49.74</td>
<td>49.52</td>
<td>48.95</td>
<td>57.58</td>
<td>54.84</td>
<td><strong>57.70</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>17.48</td>
<td>13.47</td>
<td>15.66</td>
<td>20.92</td>
<td>18.46</td>
<td>16.94</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Mean</td>
<td>50.47</td>
<td>48.88</td>
<td>51.99</td>
<td>60.70</td>
<td>60.57</td>
<td><strong>70.33</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>17.62</td>
<td>14.61</td>
<td>15.10</td>
<td>19.88</td>
<td>16.89</td>
<td>14.45</td>
</tr>
<tr>
<td>62</td>
<td>PSD</td>
<td>Mean</td>
<td>53.68</td>
<td>57.13</td>
<td>60.57</td>
<td><strong>63.60</strong></td>
<td>58.47</td>
<td>56.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>13.40</td>
<td>14.23</td>
<td>15.54</td>
<td>18.78</td>
<td>18.60</td>
<td>14.54</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>Mean</td>
<td>57.58</td>
<td>57.98</td>
<td>61.22</td>
<td>66.66</td>
<td>66.34</td>
<td><strong>70.58</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Std.</td>
<td>12.64</td>
<td>12.30</td>
<td>16.46</td>
<td>18.80</td>
<td>17.49</td>
<td>17.01</td>
</tr>
</tbody>
</table>

Data Shortage in EEG-based Tasks

- Comparing with images and speech signals, it is hard to collect large-scale EEG data.
  - The prices of EEG acquisition devices for research are quite high.
  - These experiments can not last for a long time because the subjects may feel uncomfortable with wearing EEG acquisition device.
  - The raw EEG data is usually mixed with noise and various artifacts, and researchers have to discard some bad channels and data.
  - It is difficult to collect precise labeled data since the subjects may not evoke emotion well in emotion recognition experiments.
A Framework for EEG Data Augmentation

Positive

Neutral

Negative

Raw data of SEED (DEAP)

Feature extraction

DE (PSD)

Original training data set

Feature extraction

DE (PSD)

Test data set

Feature extraction

Generated data

Augmented training data set

Our models

cWGAN, sWGAN

Training

Predicting

Emotion

SVM

DNN

Classifiers

Support vectors

Marginal hyperplane

Feature 1

Feature 2
Two Data Augmentation Strategies

(a) All of the generated data are used to augment the training data set.

(b) Only data with high quality are used to augment the training data set.
Three Indicators for Assessing the Quality of Data

- In computer vision, the qualities of generated images can be assessed directly by users. However, the qualities of the generated high-dimensional EEG data are impossible to be visually evaluated.

- We use three indicators to evaluate the qualities of generated EEG data:
  - Discriminator loss: The loss of discriminator represents EMD between $X_r$ and $X_g$ when the network converges.
  - Maximum Mean Discrepancy (MMD): MMD is frequently used as a measurement of the distance between two distributions.
  - Two-dimensional mapping: The high-dimensional $X_g$ are mapped into two-dimensions by t-SNE.
Two-dimensional Mapping of Generated Data

- Two-dimensional visualizations of the real and generated DE data of one subject in SEED dataset.
- Data points with red, green and blue colors represent three emotions of negative, neutral and positive, respectively. The lines represent the real data and the thin points represent the generated data.
The experimental results demonstrate that using the EEG data generated by CWGAN significantly improves the accuracies of emotion recognition models.

<table>
<thead>
<tr>
<th>Data Appended</th>
<th>SEED Mean</th>
<th>SEED Std.</th>
<th>DEAP-Arousal Mean</th>
<th>DEAP-Arousal Std.</th>
<th>DEAP-Valence Mean</th>
<th>DEAP-Valence Std.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0×Dataset</td>
<td>0.8399</td>
<td>0.0972</td>
<td>0.6902</td>
<td>0.1361</td>
<td>0.5376</td>
<td>0.1308</td>
</tr>
<tr>
<td>1×Dataset</td>
<td>0.8696</td>
<td>0.1272</td>
<td>0.7817</td>
<td>0.0958</td>
<td>0.7389</td>
<td>0.1082</td>
</tr>
<tr>
<td>2×Dataset</td>
<td>0.8659</td>
<td>0.1283</td>
<td>0.7437</td>
<td>0.1227</td>
<td>0.6543</td>
<td>0.1506</td>
</tr>
<tr>
<td>3×Dataset</td>
<td>0.8638</td>
<td>0.1276</td>
<td>0.7479</td>
<td>0.1133</td>
<td>0.6625</td>
<td>0.1477</td>
</tr>
<tr>
<td>4×Dataset</td>
<td>0.8623</td>
<td>0.1275</td>
<td>0.7450</td>
<td>0.1013</td>
<td>0.6496</td>
<td>0.1504</td>
</tr>
<tr>
<td>5×Dataset</td>
<td>0.8600</td>
<td>0.1294</td>
<td>0.7434</td>
<td>0.1151</td>
<td>0.6467</td>
<td>0.1372</td>
</tr>
</tbody>
</table>
The confusion matrixes (SEED dataset) trained by (a) original training set and (b) appended training set (1 time).

<table>
<thead>
<tr>
<th></th>
<th>positive</th>
<th>neutral</th>
<th>negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive</td>
<td>0.99</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>neutral</td>
<td>0.08</td>
<td>0.89</td>
<td>0.03</td>
</tr>
<tr>
<td>negative</td>
<td>0.17</td>
<td>0.21</td>
<td>0.62</td>
</tr>
</tbody>
</table>

(a) Original training set

<table>
<thead>
<tr>
<th></th>
<th>positive</th>
<th>neutral</th>
<th>negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>neutral</td>
<td>0.06</td>
<td>0.93</td>
<td>0.01</td>
</tr>
<tr>
<td>negative</td>
<td>0.13</td>
<td>0.20</td>
<td>0.67</td>
</tr>
</tbody>
</table>

(b) Appended training set (1 time)
Visualization of Generated Data for Positive Emotion

(a) Real Data
(b) Generated Data
Visualisation of Generated Data for Negative Emotion

(a) Real Data

(b) Generated Data
Personalizing Affective Models with Transfer Learning

The feature distributions of EEG data of source subjects and target subjects are not independently and identically distributed (i.i.d.) due to:

- The structural and functional variability between subjects
- The non-stationary nature of EEG signals
- The inherent changes of environmental variables
Transfer Learning Framework

Labeled Training Data

Source Domain Data

Transfer Learning Algorithms

Predictive Models

Output

Target Domain Data

Unlabeled Data/a few labeled data for adaptation

Testing

Target Domain Data

New Subject
Domain Shift and Domain Adaptation

- Training data are drawn from source domain \( \{X_S, \phi_S\} \), and test data are drawn from target domain \( \{X_T, \phi_T\} \).
- Both domains have the same set of features \( (X_S, X_T \in \mathbb{R}^m) \).
- Domain shift \( (P(X_S) \neq P(X_T)) \) makes ordinary learning methods degenerate.
- We can use domain adaptation methods to eliminate or reduce the domain shift.

\[
P(\phi(X_S)) \approx P(\phi(X_T)) \\
P(Y_S|\phi(X_S)) \approx P(Y_T|\phi(X_T))
\]
Transductive Parameter Transfer

Labeled Source Data

Training Phase

Feature Distributions

Parameter Regression Model

Individual Classifier

Parameter Transfer $\pi$

$\theta_i$

$\theta_2$

$\theta_N$

$D_i$

$D_2$

$D_N$

$X_i$

Personalized Affective Model

Unlabeled Target Data

Testing Phase

# Experimental Results on Three Emotions

Zheng and Lu, *IJCAI 2016*

---

### Table of Results

<table>
<thead>
<tr>
<th>Stats.</th>
<th>Generic</th>
<th>KPCA</th>
<th>TCA</th>
<th>T-SVM</th>
<th>TPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.5673</td>
<td>0.6128</td>
<td>0.6364</td>
<td>0.7253</td>
<td><strong>0.7631</strong></td>
</tr>
<tr>
<td>Std.</td>
<td>0.1629</td>
<td>0.1462</td>
<td>0.1488</td>
<td><strong>0.1400</strong></td>
<td>0.1589</td>
</tr>
</tbody>
</table>
# New Domain Generalization Methods

## Session 1: 11.00-13.00 Friday 13 December

<table>
<thead>
<tr>
<th>1.3 Cutler Room</th>
<th>Adversarial Networks and Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session Chair: Ting Fang</td>
<td></td>
</tr>
<tr>
<td>• B-DCGAN: Evaluation of Binarized DCGAN for FPGA</td>
<td></td>
</tr>
<tr>
<td><strong>Hideo Terada</strong> and <strong>Hayaru Shouno</strong></td>
<td></td>
</tr>
<tr>
<td>• Reducing the Subject Variability of EEG Signals with Adversarial Domain Generalization</td>
<td></td>
</tr>
<tr>
<td><strong>Bo-Qun Ma</strong>, <strong>He Li</strong>, <strong>Wei-Long Zheng</strong> and <strong>Bao-Liang Lu</strong></td>
<td></td>
</tr>
</tbody>
</table>
Other Research Work Performed in my Lab

- Gender Differences in Emotion Recognition
- Ethnicity Differences in Emotion Recognition (Chinese; German; Frenchman)
- Emotion Recognition under Sleep Deprivation
- Object Assessment of Sleep Quality Using EEG
- Multimodal Fatigue Driving Detection Using EEG and EOG
SJTU Emotion EEG Dataset (SEED)

- Publicly available emotion dataset: EEG and eye tracking data
- No. experiments: 45, No. Channels: 62
- Three emotional states (happy, sad, and neutral)
- Total number of applications: 558 (Oct. 2019)

http://bcmi.sjtu.edu.cn/~seed/index.html

SJTU Emotion EEG Dataset (SEED-V)

- Publicly available emotion dataset: EEG and eye tracking data
- No. experiments: 45, No. Channels: 62
- Four emotional states (happy, sad, fear, disgust, and neutral)
- Total number of applications: 110 (Oct. 2019)

Vigilance Estimation Dataset (SEED-VIG)

- Publicly available multimodal dataset: EEG and forehead EOG
- No. experiments: 23 (11 males, 12 females)
- Released: April 2017
- Total number of applications: 50 (Oct. 2019)

Statistics of Citations of SEED Data Sets

Year | Citations
--- | ---
2015 | 2
2016 | 21
2017 | 50
2018 | 98
2019 | 140
Outline

- Background and Current Research Situations
- Multimodal Emotion Recognition
  - Emotion Recognition Using EEG and Eye Movement Data
  - Personalizing Affective Models with Transfer Learning
  - EEG Data Augmentation Using GAN
  - Other Work in my lab and SEED Data Sets
- Toward Emotional Intelligence: Four Phases
- Summary
Brain–machine interfaces (BMIs) create closed-loop control systems that interact with the brain by recording and modulating neural activity and aim to restore lost function, most commonly motor function in paralyzed patients. Moreover, by precisely manipulating the elements within the control loop, motor BMIs have emerged as new scientific tools for investigating the neural mechanisms underlying control and learning. Beyond motor BMIs, recent work highlights the opportunity to develop closed-loop mood BMIs for restoring lost emotional function in neuropsychiatric disorders and for probing the neural mechanisms of emotion regulation. Here we review significant advances toward functional restoration and scientific discovery in motor BMIs that have been guided by a closed-loop control view. By focusing on this unifying view of BMIs and reviewing recent work, we then provide a perspective on how BMIs could extend to the neuropsychiatric domain.
Brain-Machine Interface from Motor to Mood
## Comparison of Challenges in Motor and Mood BMIs

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Motor BMI</th>
<th>Mood BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural measurements</td>
<td>Motor cortical networks (including premotor, primary motor, and posterior parietal cortex)</td>
<td>Distributed multisite corticolimbic networks, whose functional organization is not as well characterized</td>
</tr>
<tr>
<td>Behavioral measurements</td>
<td>Continuous in time (movements)</td>
<td>Infrequent and discrete in time (for example, self-reports)</td>
</tr>
<tr>
<td>Time-scale of behavioral dynamics</td>
<td>Milliseconds (movement dynamics)</td>
<td>Minutes to days and longer (mood dynamics)</td>
</tr>
<tr>
<td>Behavioral assessment</td>
<td>Relatively easy and accurate</td>
<td>Difficult and less accurate with self-reports being common measurement instruments</td>
</tr>
<tr>
<td>Modeling the effect of direct brain stimulation</td>
<td>In general not needed, unless artificial sensory feedback is provided in bidirectional BMIs</td>
<td>Needed, and should be modeled across distributed multisite corticolimbic networks</td>
</tr>
</tbody>
</table>
“This book tries to explain how minds work. How can intelligence emerge from non-intelligence? To answer that, we’ll show that you can build a mind from many little parts, each mindless by itself”
Toward Emotion AI: Four Phases

1. Logical intelligence
   - No any emotion AI

2. Emotion quantization and modeling human emotion
   - Emotion recognition and emotion regulation

3. Self learning and emergence mechanism
   - Comprehensive perception of the objective world

4. Formation of selfhood and values
   - Consciousness and creativity
Artificial General Intelligence (AGI)

- Alpha-GO has powerful logic intelligence, but it doesn’t have any Emotion AI!

- Unlike industrial robots, service robots must have ability to interact with humans and have emotion AI

- AGI should combine logic Intelligence and emotion AI
Summary

☐ Multimodal approaches are effectiveness for emotion recognition
☐ Identifying stable neural patterns over time for positive, neural, and negative emotions.
☐ Tackling individual differences with transfer learning.
☐ GAN-based data augmentation method can improve the accuracy of emotion recognition models.
☐ Challenge: from emotion recognition to emotion regulation
☐ Toward Emotion AI: Four Phases
Will Robots Ever Have Emotions? YES!
Acknowledgements

- National Natural Science Foundation of China
- National Key Research and Development Program of China (NSFC)
- The Science and Technology Commission of Shanghai Municipality
- My students: Wei-Long Zheng, Wei Liu, Yun Luo, Li-Ming Zhao, Xue Yan, Si-Yuan Wu, Jia-Jun Tong, He Li, Xun Wu, Bo-Qun Ma, Hao Tang, Tian-Hao Li, Lu Gan, Ji-Lin Qiu, Lan-Qing Bao
Thank you for your attention!