華中科技大學
HUAZHONG UNIVERCITY OF SCIENCE AND TECHNOLOGY

脑机接口与机器学习实验室
BRAIN-COMPUTER INTERFACE AND MACHINE LEARNING LABORATORY

Affective Brain-Computer Interface (aBCI): A Tutorial

Dongrui Wu

School of Artificial Intelligence and Automation

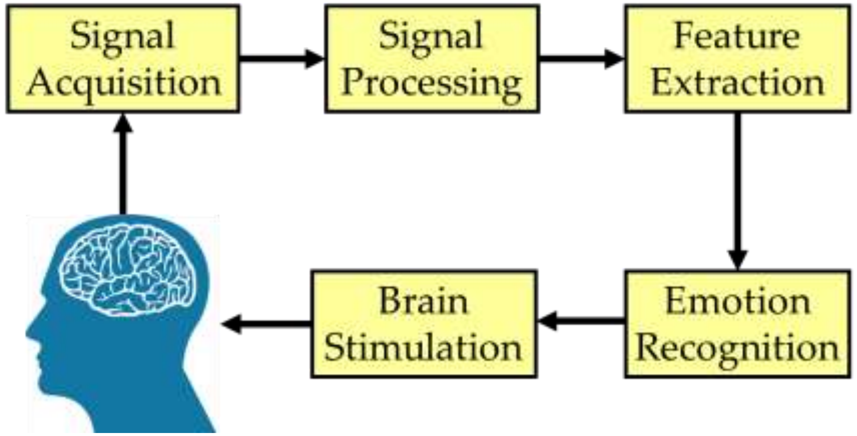
Huazhong University of Science and Technology

drwu@hust.edu.cn

http://lab.bciml.cn/

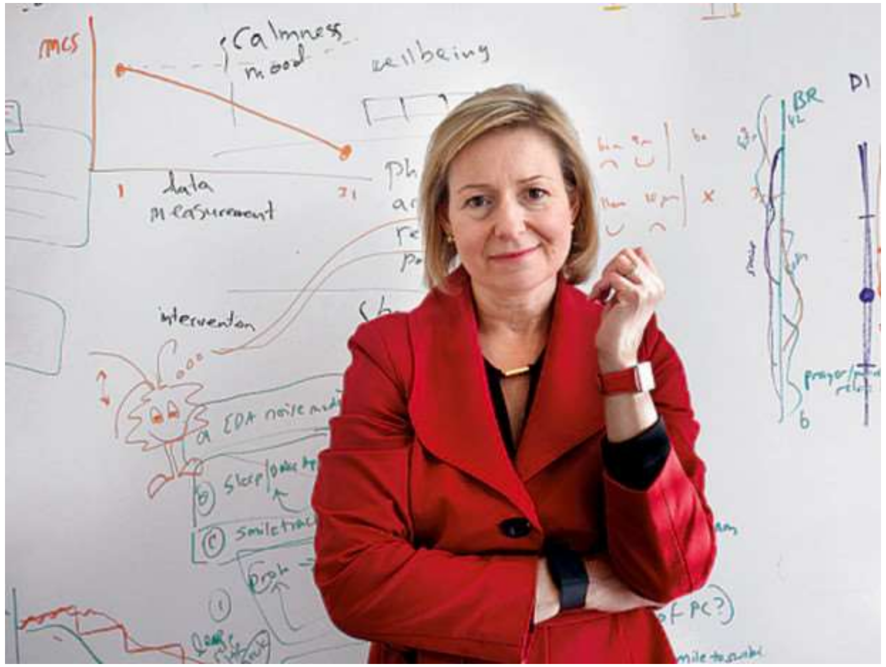
Outline

- **Affective Brain-Computer Interface (aBCI)**
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- Feature Extraction in aBCI
- Emotion Recognition in aBCI
- Brain Stimulation
- aBCI Applications
- Challenges and Opportunities



Affective Computing

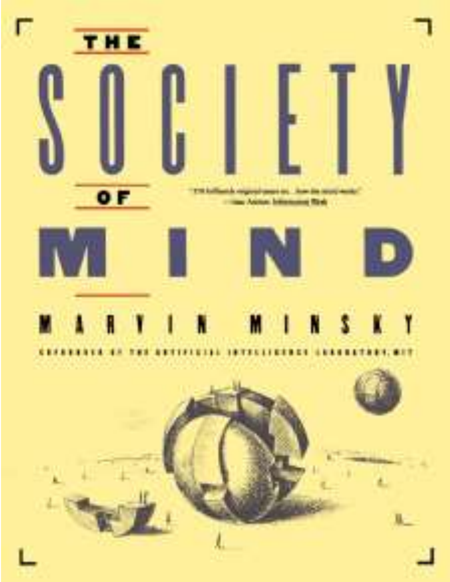
- Prof. **Rosalind Picard**, MIT Media Lab, 1995
- Computing that relates to, arises from, or influences **emotions**



Affect & Artificial Intelligence (AI)

"The question is not whether intelligent machines can have any **emotions**, but whether machines can be intelligent without **emotions**."

- Minsky, The Society of Mind, 1986



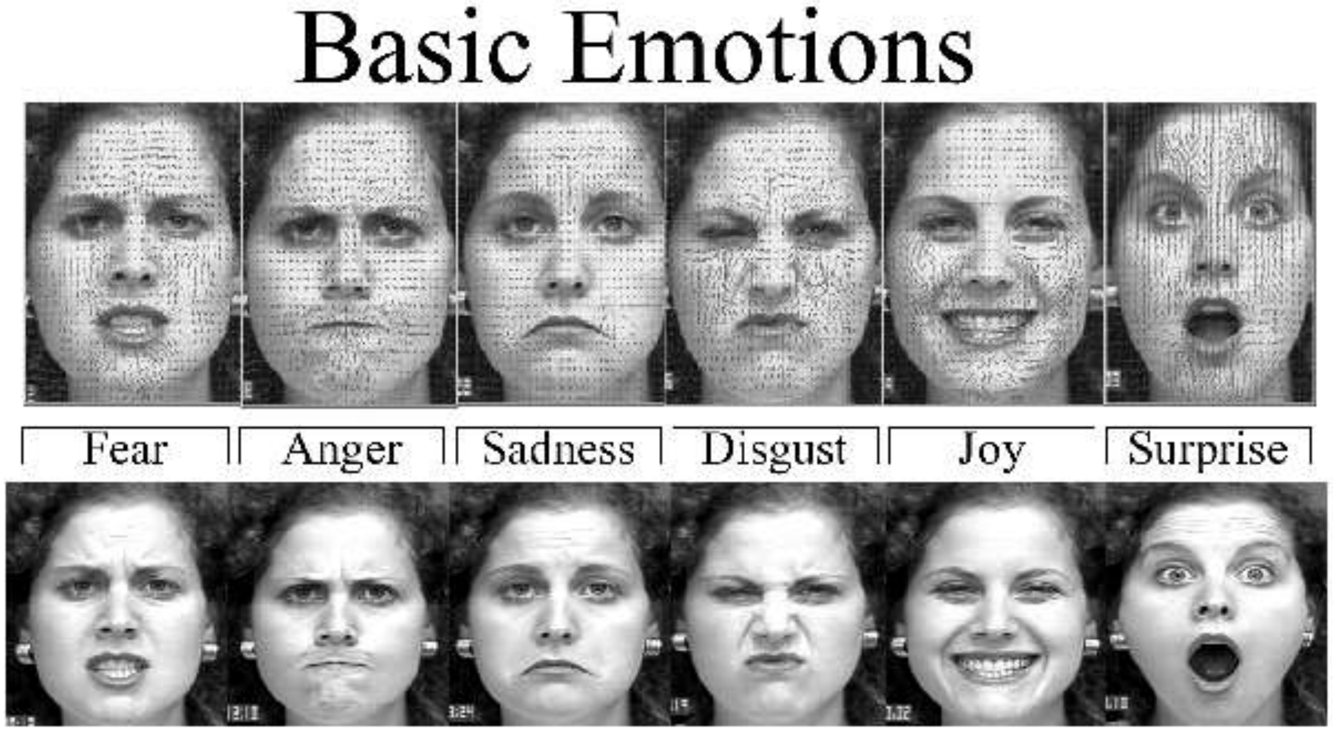
Marvin Minsky

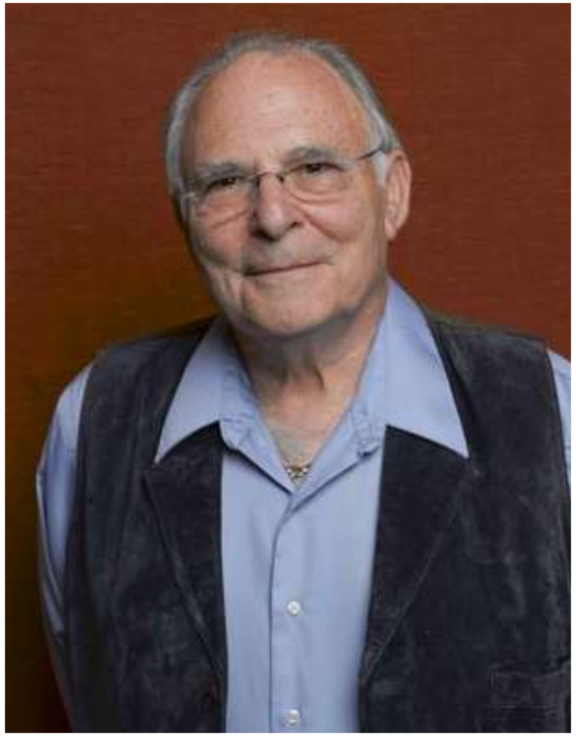
1927-2016

MIT Professor

Turing Award

Six Basic Emotions

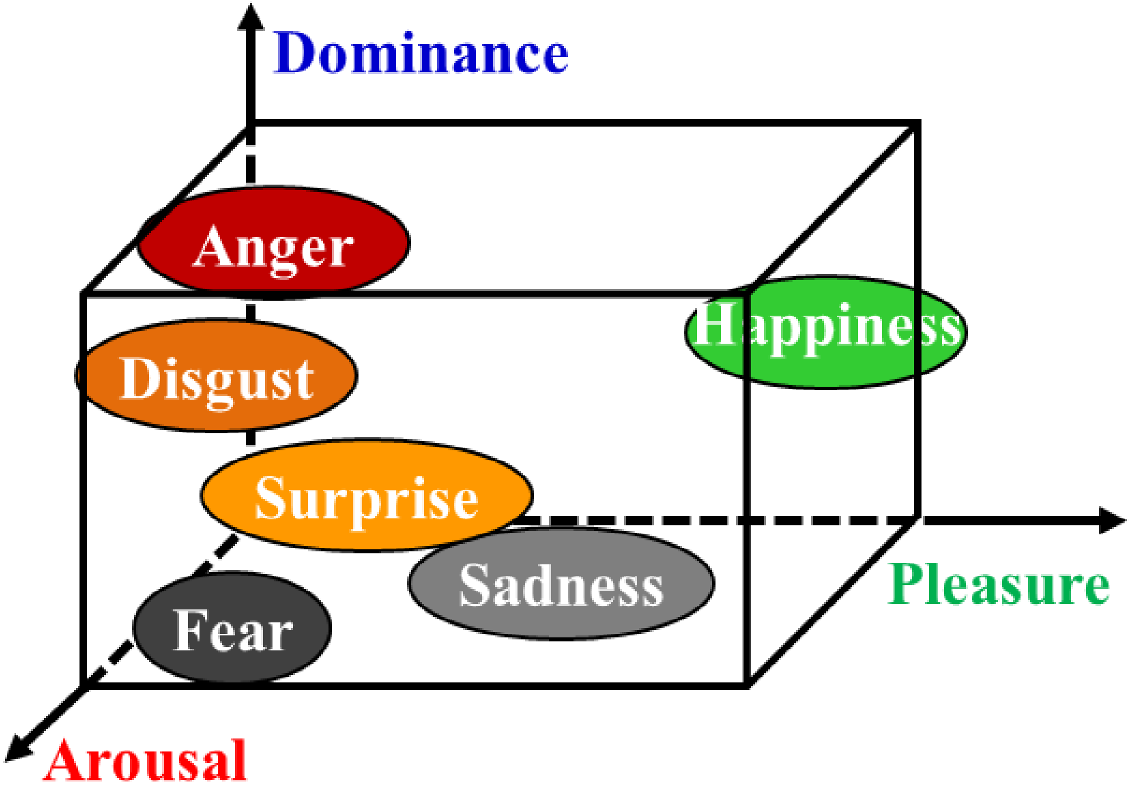


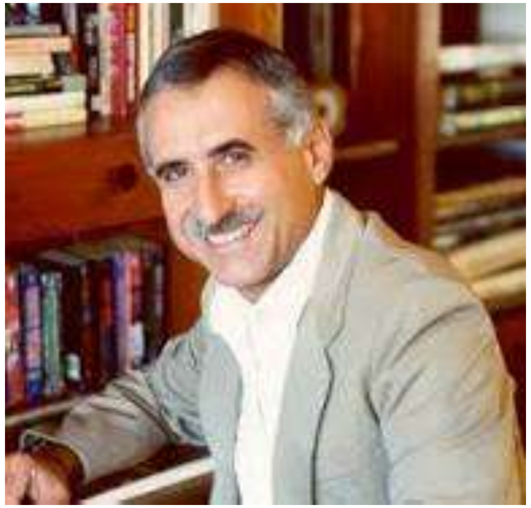


Paul Ekman

UCSF, 1971

3D Representation of Emotions



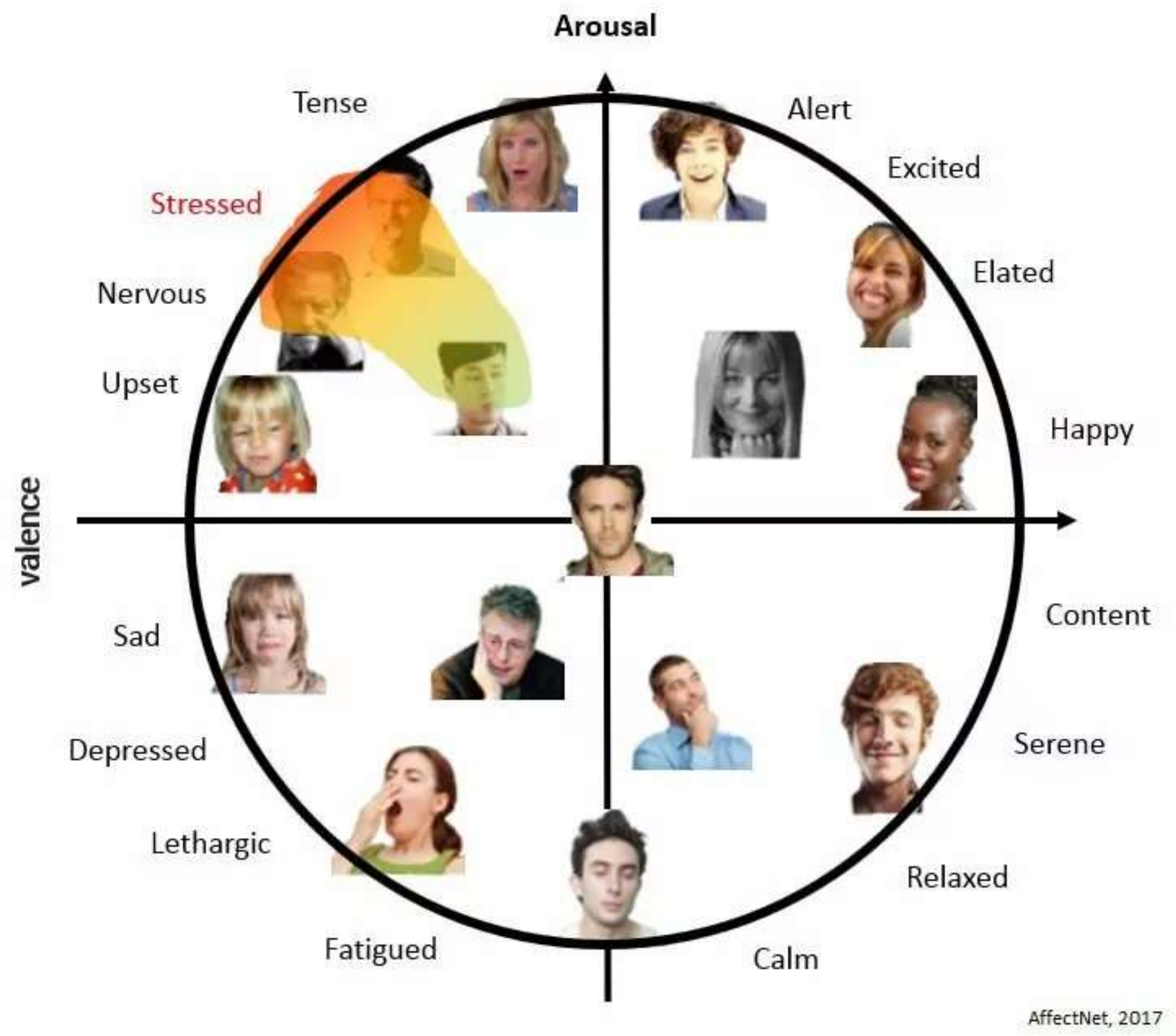


Albert Mehrabian
UCLA, 1974

James Russell
UCLA, 1974

A. Mehrabian and J. A. Russell, **An Approach to Environmental Psychology**, MIT Press: Cambridge, MA, 1974

2D Representation of Emotions



James Russell

UBC, 1980

James Russell, A circumplex model of affect, Journal of Personality and Social Psychology, 39 (6): 1161–1178, 1980.

Brief History of Affective Computing

1971, Ekman, UCSF,
Six Basic Emotions

1986, Minsky, MIT,
The Society of Mind

humaine

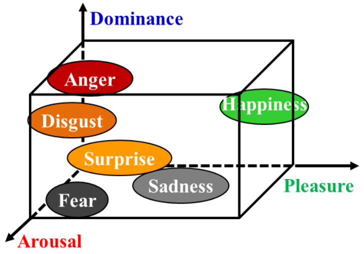
2004-2007, **Human-Machine Interaction Network on Emotion**, EU Project, 27 Universities

2010, **IEEE Trans. on Affective Computing**

Microsoft

2014, **Microsoft Xiaoice**, V1

1974, Mehrabian & Russell, UCLA:
3D Representation of Emotion



1997, Picard, MIT:
Affective Computing

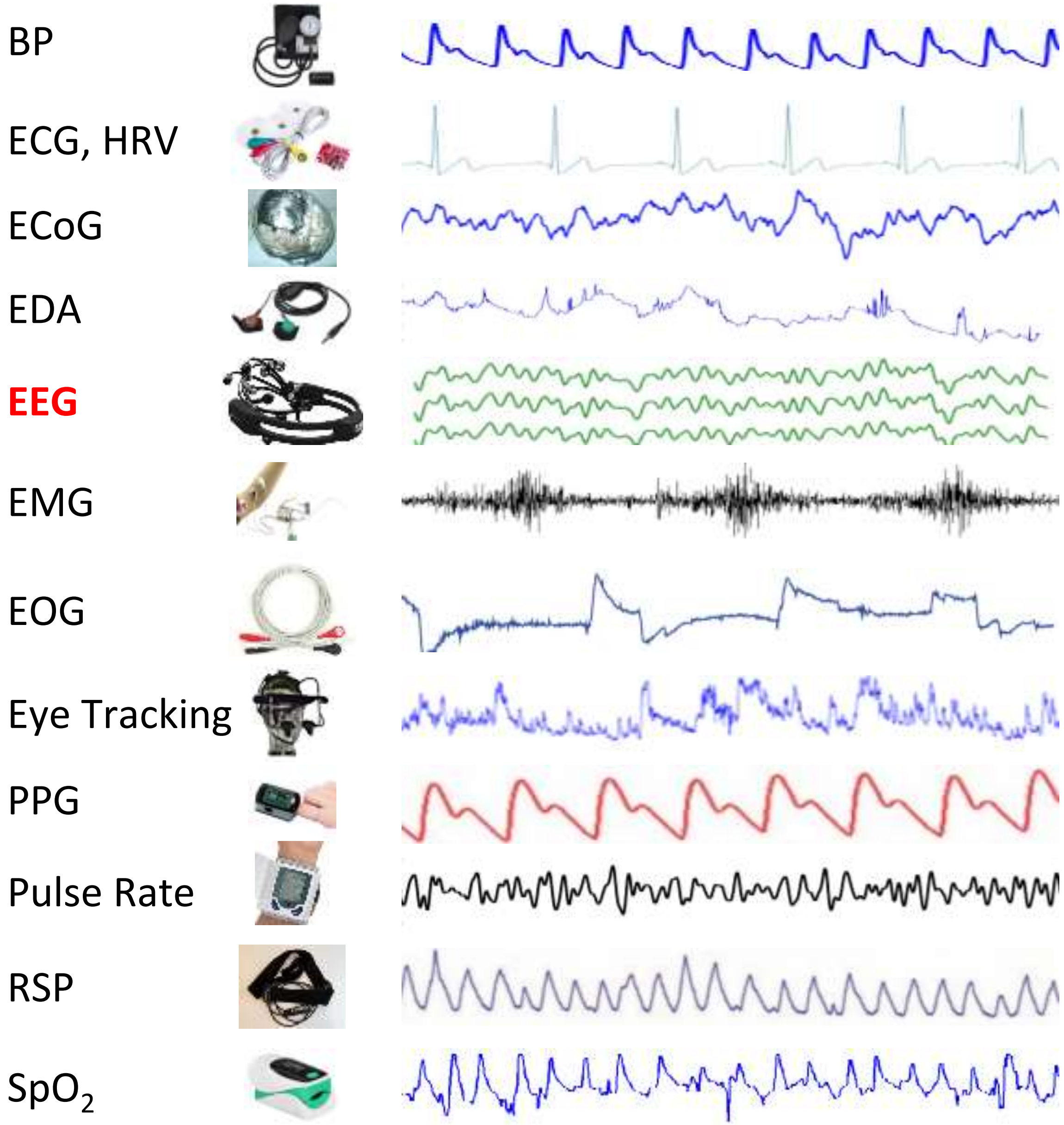
Int'l. Conf. on Affective Computing and Intelligent Interaction (ACII): 2005, Beijing; 2007, Lisbon; 2009, Amsterdam; 2011, Memphis; 2013, Geneva; 2015, Xi'An; 2017, San Antonio; 2019, Cambridge; 2021, Virtual

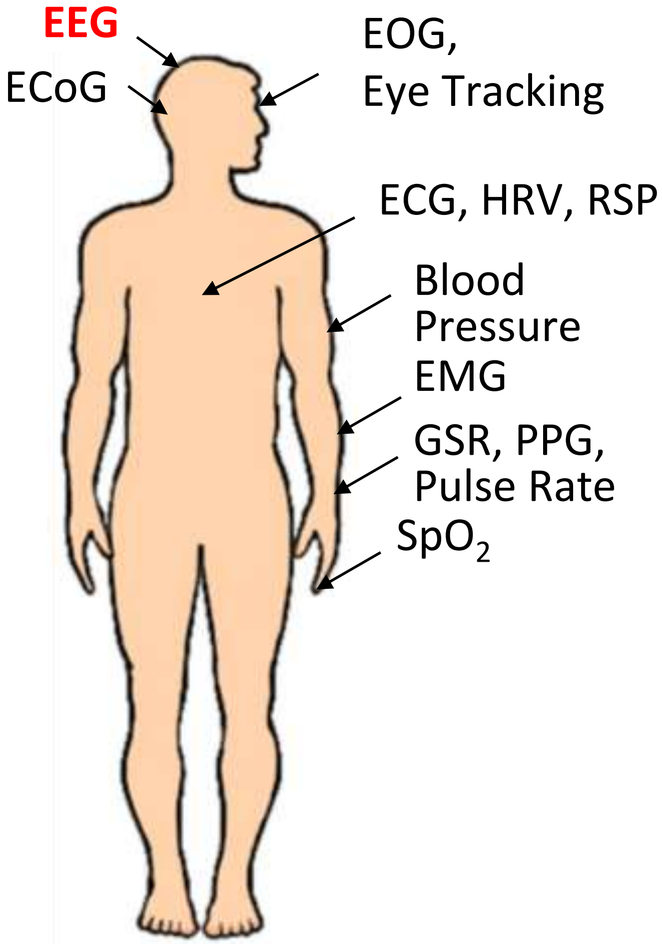
AAAC

2011, **Association for the Advancement of Affective Computing (AAAC)**

2020, **Microsoft Xiaoice**, V8

Physiological Signals in Affective Computing

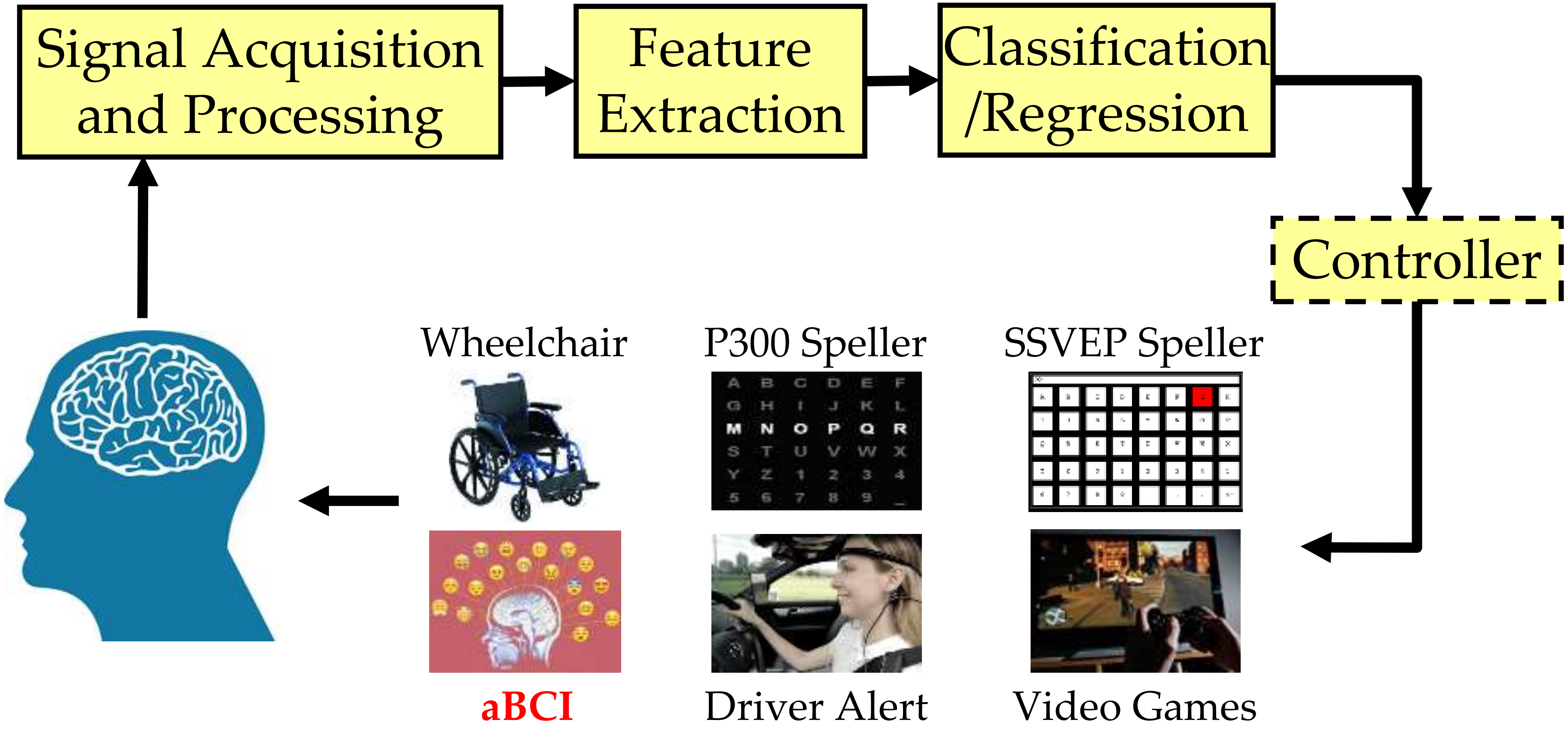




D. Wu, J. Xu, W. Fang, Y. Zhang, L. Yang, X. Xu*, H. Luo* and X. Yu*, Adversarial Attacks and Defenses in Physiological Computing: A Systematic Review, National Science Open, 2023.

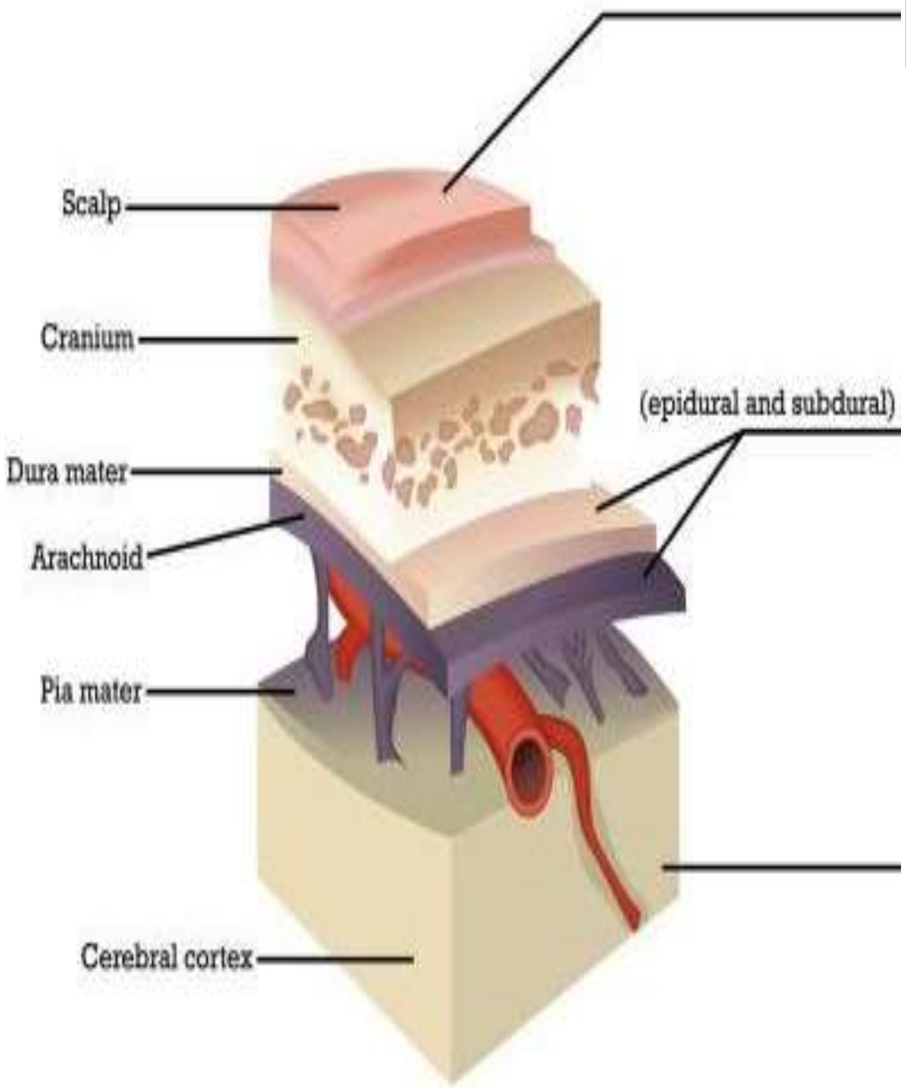
Brain-Computer Interface (BCI)

- A direct communication pathway between the brain and an external device
- Research, assist, augment, or repair human cognitive or sensory-motor functions



Krucoff, M. O., Rahimpour, S., Slutzky, M. W., Edgerton, V. R., & Turner, D. A. (2016). Enhancing nervous system recovery through neurobiologics, neural interface training, and neurorehabilitation. Frontiers in Neuroscience, 10, 584.

Three Types of BCIs



- **Non-invasive BCIs (EEG)**:
  - ✓ Easy to wear; No surgery.
  - ✓ Relatively poor spatial resolution.
  - ✓ Cannot effectively use high-frequency signals.
  - ✓ Require calibration prior to usage.
- **Partially invasive BCIs (ECoG)**:
  - ✓ Implanted inside the skull but rest outside the brain.
  - ✓ Better signal resolution than non-invasive BCIs
  - ✓ Lower risk of forming scar-tissue in the brain than fully invasive BCIs.
- **Invasive BCIs (LFP & Spikes)**:
  - ✓ Repair damaged sight; provide new functionality for paralyzed people.
  - ✓ Implanted directly into the gray matter of the brain.

1. Rao, R. P. (2013). *Brain-Computer Interfacing: An Introduction*. New York, NY: Cambridge University Press.
2. Martini, M. L., Oermann, E. K., Opie, N. L., Panov, F., Oxley, T., &Yaeger,K. (2020). Sensormodalities for brain-computer interface technology: a comprehensive literature review. Neurosurgery, 86, E108–E117.

Affective Brain-Computer Interface (aBCI)

aBCI monitors and/or regulates the emotional state of the brain

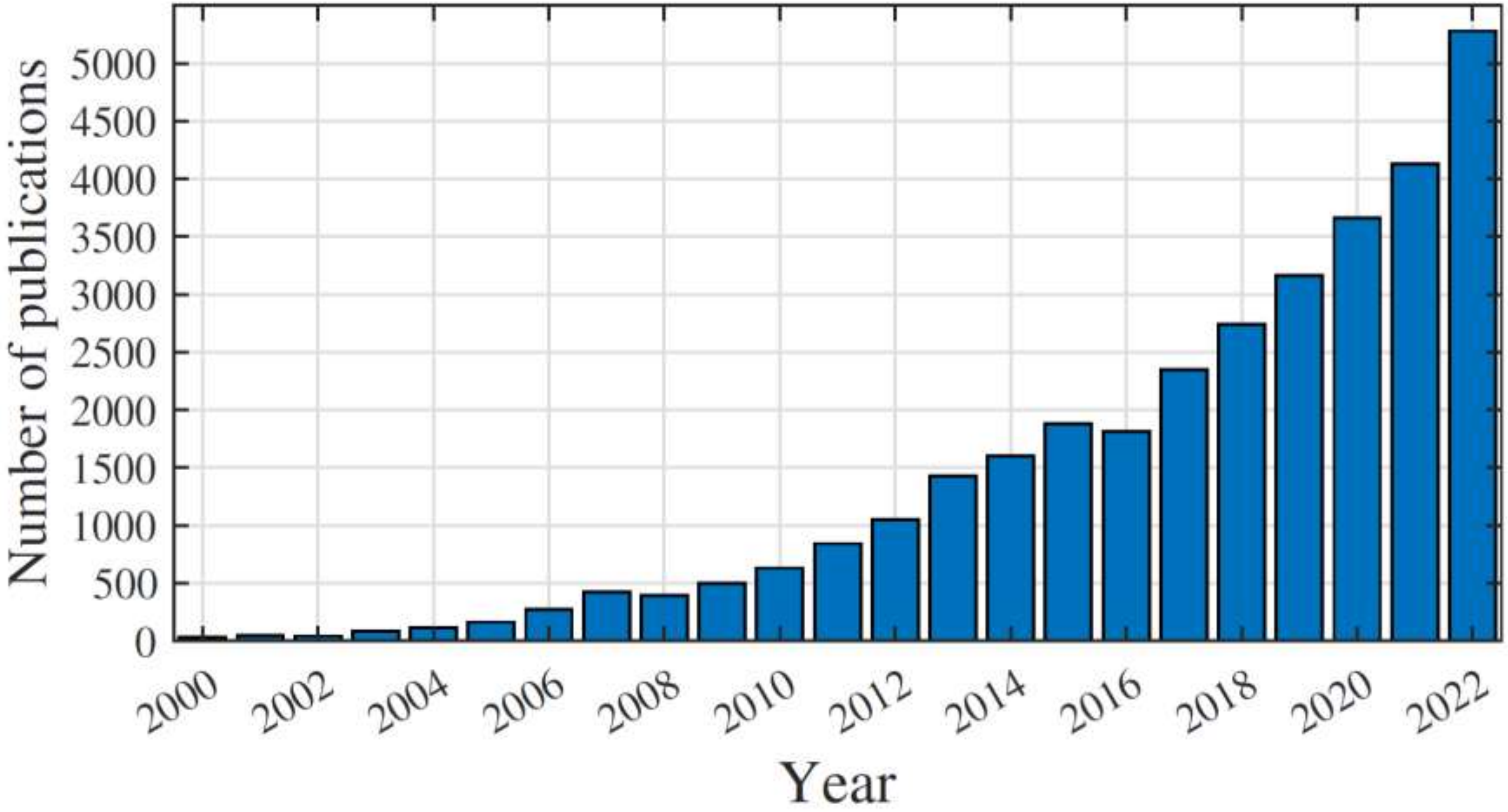
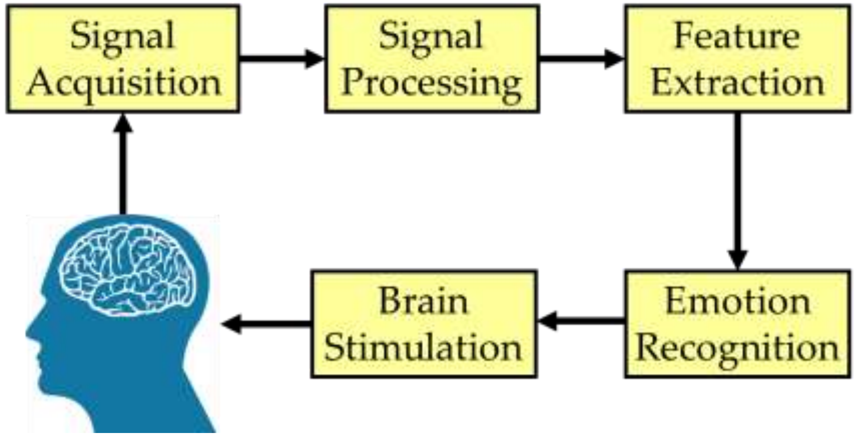


Fig. 2. Number of publications per year, returned by query "*Emotion OR Affect 'brain computer interface'*" on Google Scholar on April 15, 2023.

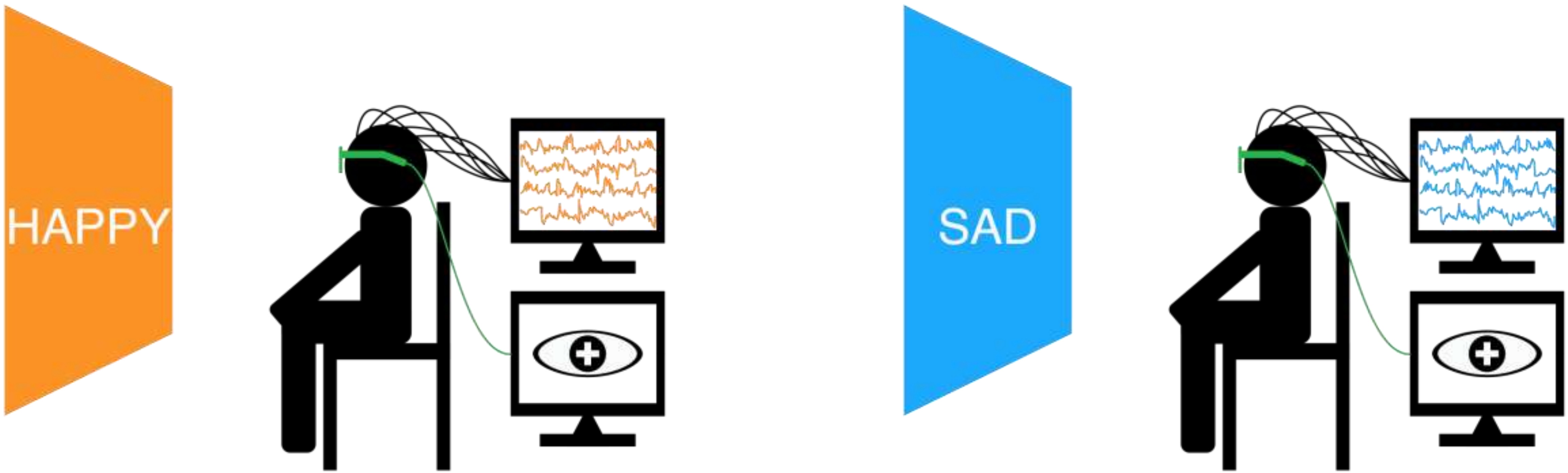
Outline

- Affective Brain-Computer Interface (aBCI)
- **Signal Acquisition in aBCI**
- Signal Processing in aBCI
- Feature Extraction in aBCI
- Emotion Recognition in aBCI
- Brain Stimulation
- aBCI Applications
- Challenges and Opportunities



Emotion Elicitation

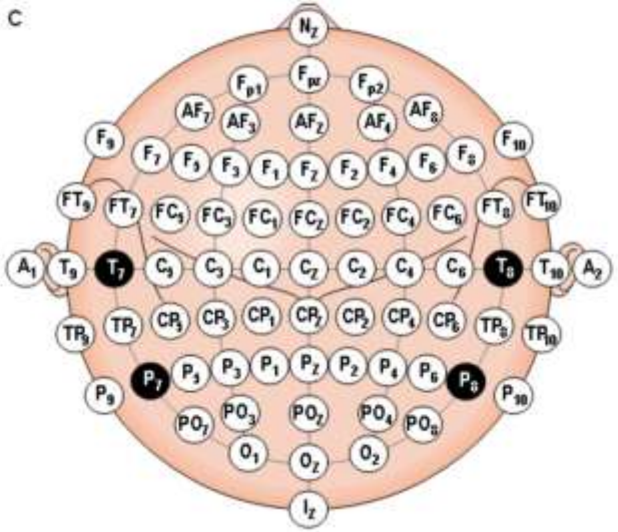
- Most performed in controlled laboratory environments, using deliberately designed settings to elicit specific emotions
- **Assumption**: A `happy' movie clip rated by multiple evaluators elicits a happy emotion for the subject



EEG Acquisition

- Typical **amplitude**: 5-300 uV
- Useful **frequency**: <100 Hz

c

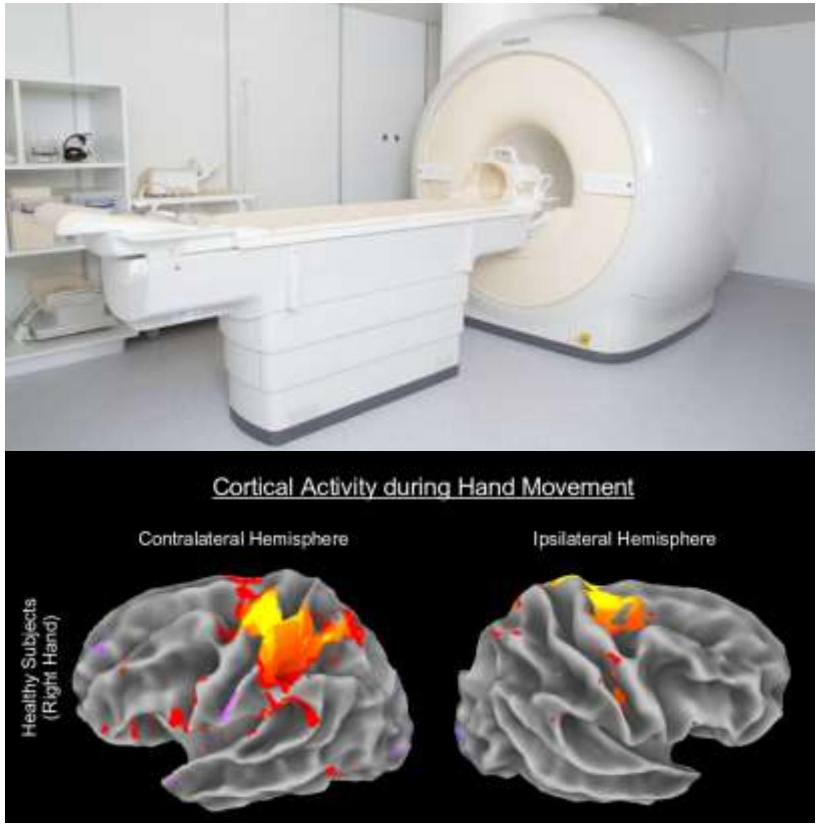


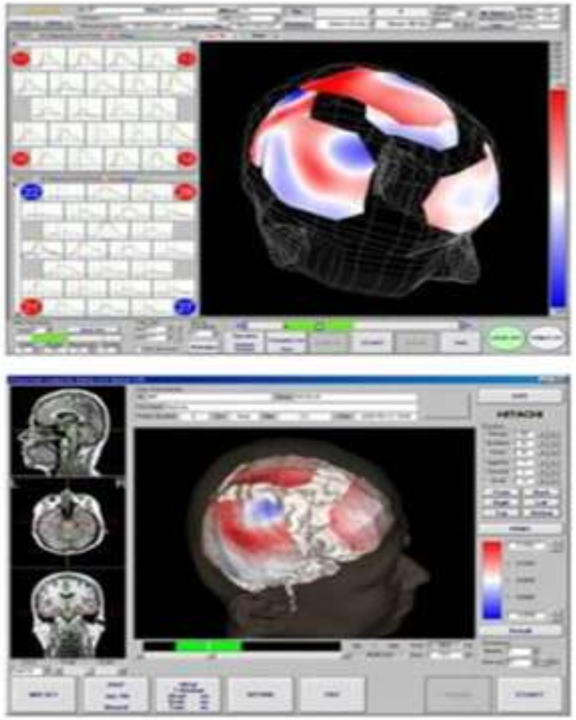
EEG Frequency Bands

Band	Typical Lobes	Frequency (Hz)	Emotions/States
Delta	Frontal (adults) or posterior (children)	0.5-4	Deep sleep
Theta	Lobes unrelated to task at hand	4-8	Drowsiness, idling
Alpha	Posterior	8-13	Relaxed, reflecting, meditation, elation
Beta	Frontal	13-30	Active thinking, focus, high alert, anxious
Gamma	Somatosensory	30-42	Cognitive processes, e.g., attention and memory
Mu	Somatosensory cortex	8-12	Sadness, sympathy

fMRI and fNIRS

- **fMRI:** Measure the neural activities by detecting changes associated with the brain blood flow
- **fNIRS:** Use near-infrared light to estimate cortical hemodynamic activities associated with the neural activities





Public aBCI Datasets

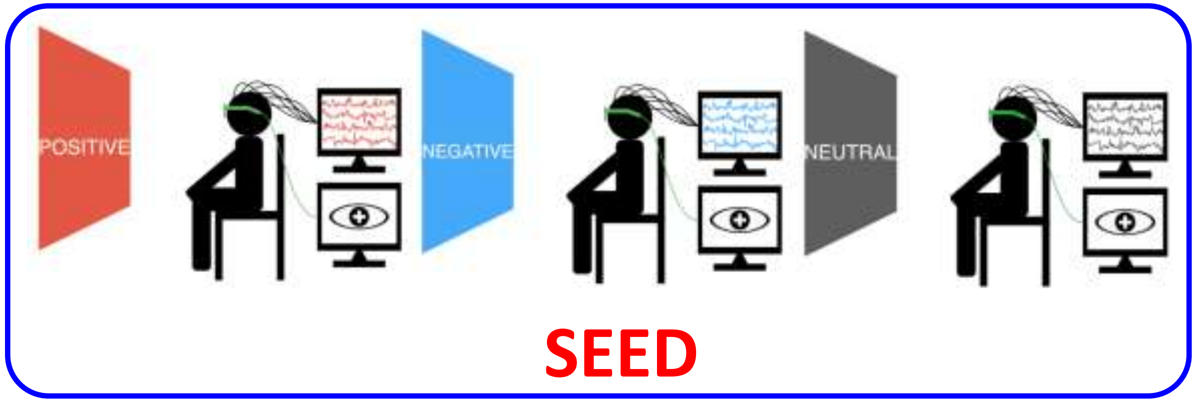
Dataset	Year	Number of Subjects	Ages	Stimuli	Modalities	Dimensional States	Discrete States	Number of Citations
MAHNOB-HCI [43]	2012	30	19-40 (mean 26.06)	20 video clips (34.9-117 seconds)	EEG, ECG, RSP, ST, gaze, video, audio	Valence (1-9), arousal (1-9), dominance (1-9), predictability (1-9)	Anger, anxiety, fear, sadness, disgust, neutrality, surprise, amusement, joy	1,304
DEAP [33]	2012	32	19-37 (mean 26.9)	40 music video clips (1 minute)	EEG, EOG, EMG, GSR, RSP, BP, ST	Valence (1-9), arousal (1-9), dominance (1-9), liking (1-9), familiarity (1-9)	None	3,430
SEED [44]	2015	15	Not reported (mean 23.27)	15 film clips (4 minutes)	EEG, EM	None	Positive, negative, neutral	1,163
ASCERTAIN [45]	2016	58	Not reported (mean 30)	36 movie clips (51-127 seconds)	EEG, ECG, GSR, facial activity	Valence (-3-3), arousal (0-6), engagement, liking, familiarity	None	359
DEAMER [46]	2017	23	22-33 (mean 28.3)	18 film clips (65-393 seconds)	EEG, ECG	Valence (1-5), arousal (1-5), dominance (1-5)	Anger, fear, sadness, disgust, calmness, surprise, amusement, happiness, excitement	513
AMIGOS [47]	2018	40	21-40 (mean 28.3)	16 video clips (51-150 seconds)	EEG, ECG, GSR, full body video	Valence (1-9), arousal (1-9), dominance (1-9), familiarity (1-9), liking (1-9)	Anger, fear, sadness, disgust, neutrality, surprise, happiness	363
RCLS [48]	2019	14	21-26 (mean 23.3)	15 video clips	EEG	None	Happy, sad, neutral	64

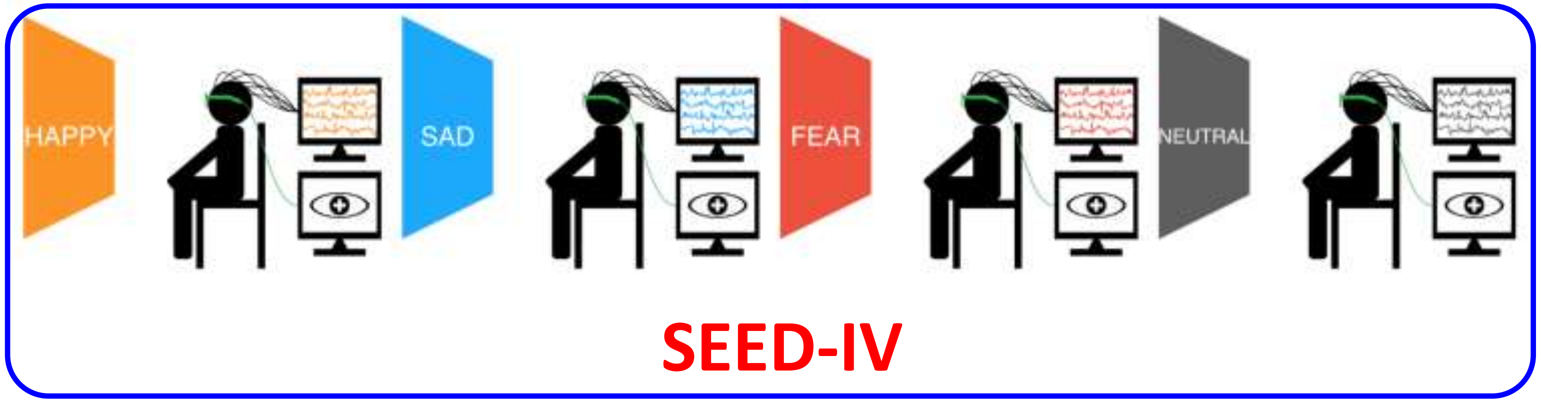
Public aBCI Datasets

MPED [49]	2019	23	18-24 (mean 21.46)	28 video clips (2.5-5 minutes)	EEG, ECG, RSP, GSR	Valence, arousal, differential emotion scale	Joy, funny, disgust, anger, fear, sadness, neutrality	148
SEED-IV [50]	2019	15	20-24 (Not reported)	72 film clips (2 minutes)	EEG, EM	Valence, arousal	Happy, sad, fear, neutral	435
HR-EEG4EMO [51]	2020	27	Not reported (mean 35.0)	13 Video clips (40-360 seconds)	EEG, GSR, ECG, RSP, SpO2, pulse rate	Valence	Positive, negative	80
SEED-V [52]	2022	16	19-28 (mean 23.27)	15 film clips (2-4 minutes)	EEG, EM	None	Happy, sad, disgust, fear, neutral	36
SEED-FRA [53]	2022	8	Not reported (mean 22.5)	21 film clips (1-4 minutes)	EEG, EM	None	Positive, negative, neutral	4
SEED-GER [53]	2022	8	Not reported (mean 22.25)	20 film clips (1-4 minutes)	EEG, EM	None	Positive, negative, neutral	4
OVPD [54]	2022	10	23-45 (Not reported)	32 film clips (40-140 seconds), odor	EEG	Valence, arousal	Positive, negative, neutral	3

ECG: Electrocardiogram; EMG: Electromyogram; EOG: electrooculogram; RSP: respiration; ST: skin temperature; GSR: galvanic skin response; EM: eye movement; BP: blood pressure; SpO2: peripheral oxygen saturation.

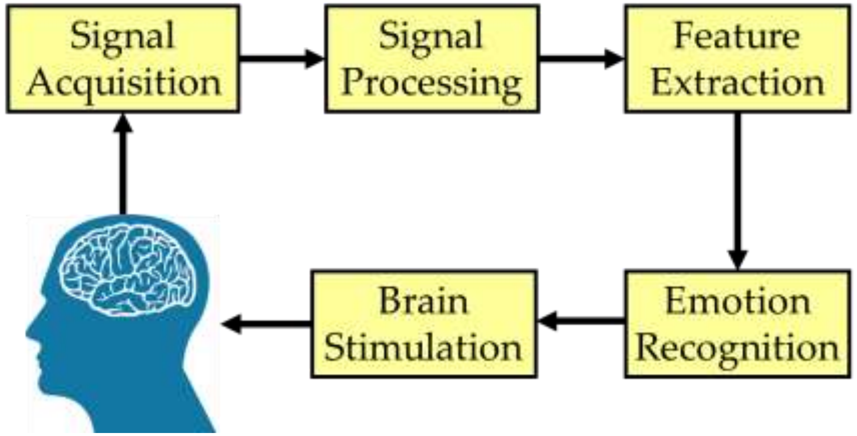
SJTU Emotion EEG Dataset (SEED)





Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- **Signal Processing in aBCI**
- Feature Extraction in aBCI
- Machine Learning in aBCI
- Brain Stimulation
- aBCI Applications
- Challenges and Opportunities

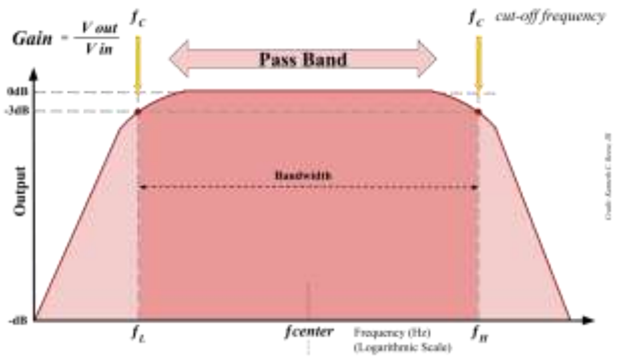


Signal Processing in aBCI

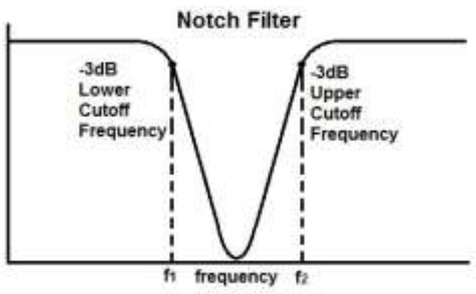
1. Temporal Filtering
2. Re-Referencing
3. Artifact Removal
4. Resampling
5. Epoching

Temporal Filtering

- Not all EEG frequencies are useful for emotion recognition: Very low frequencies may be DC drifts; Very high frequencies may be noise



- Usually EEG signals need to be band-pass filtered
- A commonly used passband is 4-45 Hz
- Delta band (0.5-4 Hz) power may be very useful in characterizing negative emotions, and high-pass filtering EEGs at 1 Hz was recommended for independent component analysis (ICA); so, a better pass-band choice may be 1-45 Hz

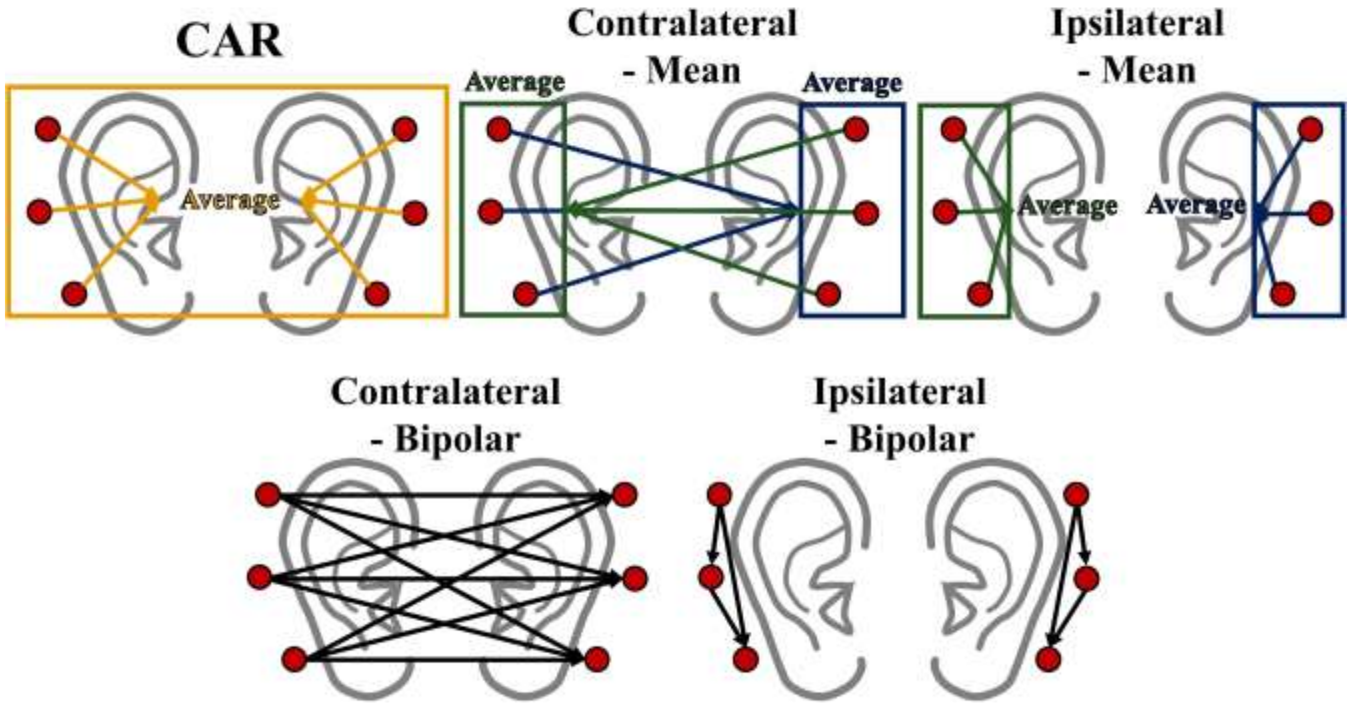


- Notch filtering to remove 50/60 Hz powerline interference

Re-Referencing

- EEG records the voltages with respect to a specific reference
- Typical reference electrodes include one mastoid (e.g., TP10), linked mastoids, the vertex electrode (Cz), single or linked earlobes, or the nose tip
- Headsets with active electrodes may record EEG reference-free

- **Re-referencing** is usually performed after data recoding and filtering to increase the signal-to-noise ratio
- **Re-referencing** post hoc can remove 40 dB unnecessary noise of active headsets.



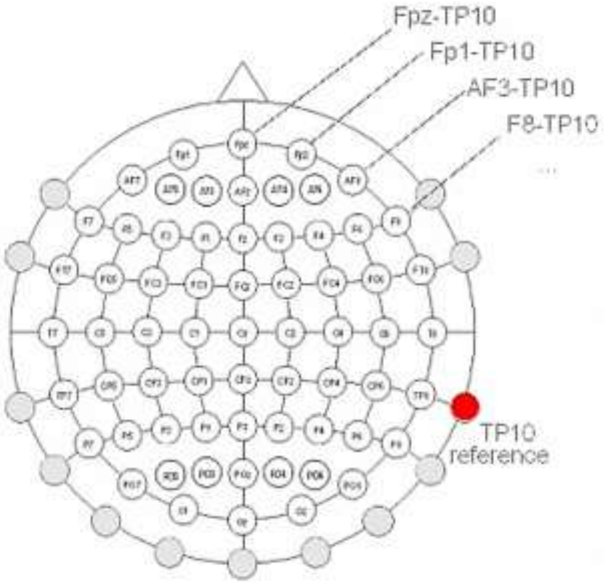
https://eeglab.org/tutorials/ConceptsGuide/rereferencing_background.html

Common Average Reference (CAR)

Most frequently used, which removes the mean of all channels from each individual channel:

- If the data were recorded with reference to nose tip or ear lobe, then these reference electrodes should be excluded from computing the average reference

- If N-channel EEG data were recorded with reference to a particular electrode, e.g., TP10, then the signal of TP10 can be recovered from the N-channel data first, i.e., TP10=(Sum of N electrode activities)/(N+1). Now there are N+1 electrodes, and their average can be computed and removed from each individual electrode



1. Average Reference assumption

   Fpz + Fp1 + AF3 + F8 + FT8 + ... + TP10 = 0

2. First recalculate the activity at reference TP10

   Sum of all electrode activity =

   Fpz + Fp1 + AF3 + F8 + ... - 64TP10
   minus Fpz + Fp1 + AF3 + F8 + ... + TP10 = 0

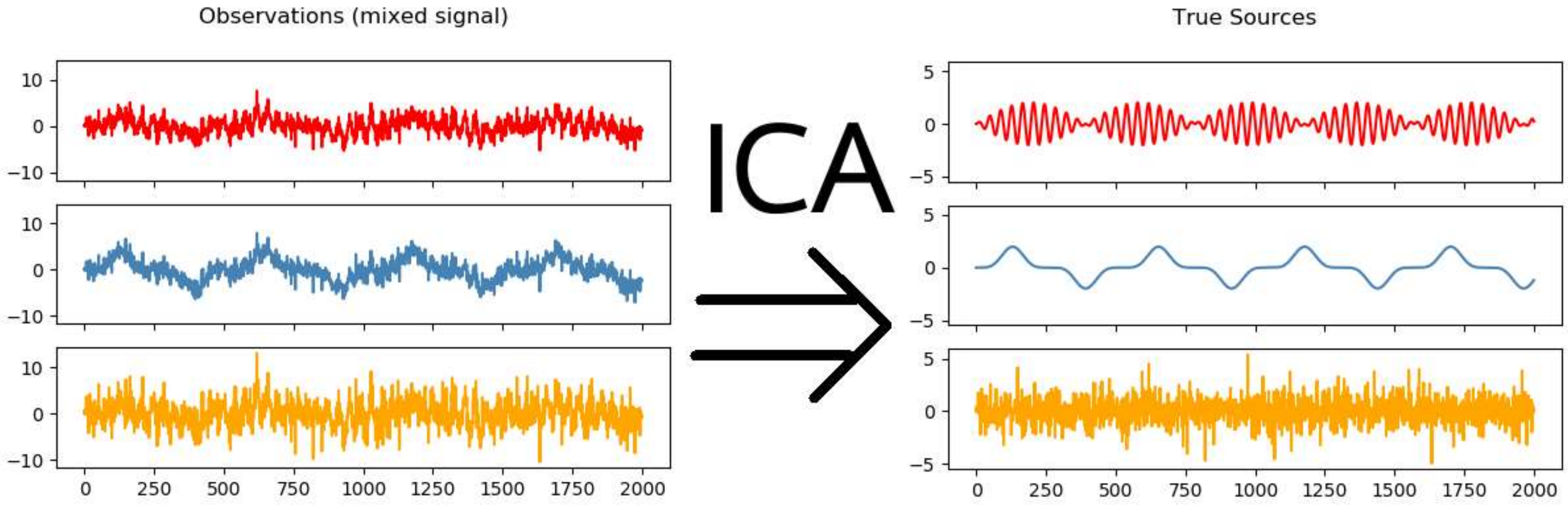
   TP10 = - (Sum of all electrode activity)/65

3. Add up the activity of TP10 to all channels

https://eeglab.org/tutorials/ConceptsGuide/rereferencing_background.html

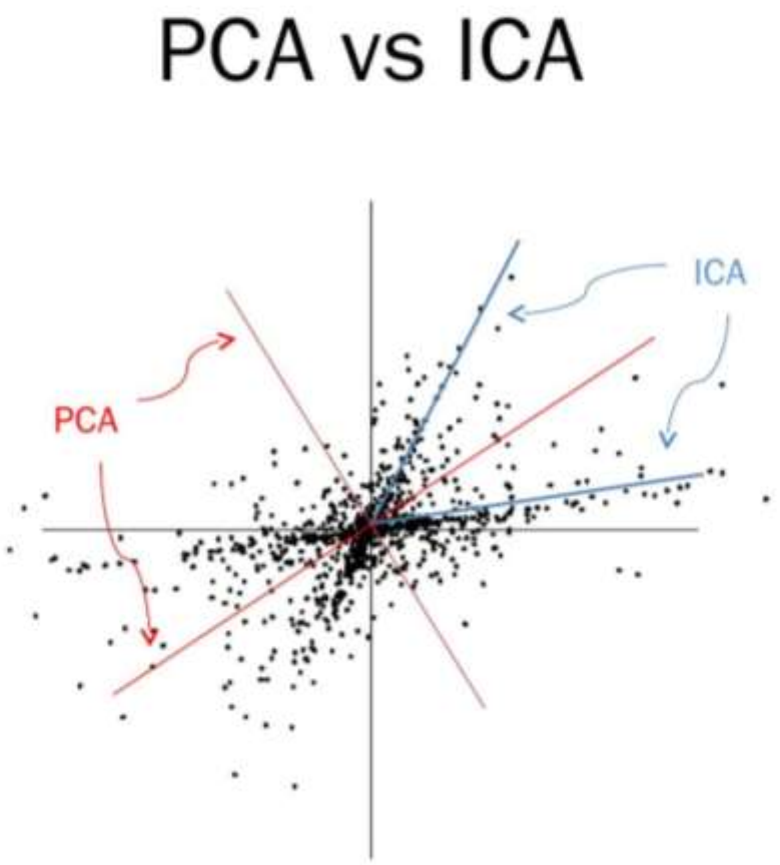
Artifact Removal

- Artifacts (e.g., eye blinks, muscle movements) may be removed manually
- Semi-automatic approaches, e.g., Independent Component Analysis (ICA), blind source separation, and principal component analysis (PCA), have also been used
- Find spatial filters to transform the original EEG channels into some “virtual channels”, some of which may be artifacts/noise and hence removed
- **ICA** identifies multiple independent component filters to produce maximally temporally independent signal sources available in the original data. A certain signal source may be eye blinks, and hence removed



ICA vs PCA

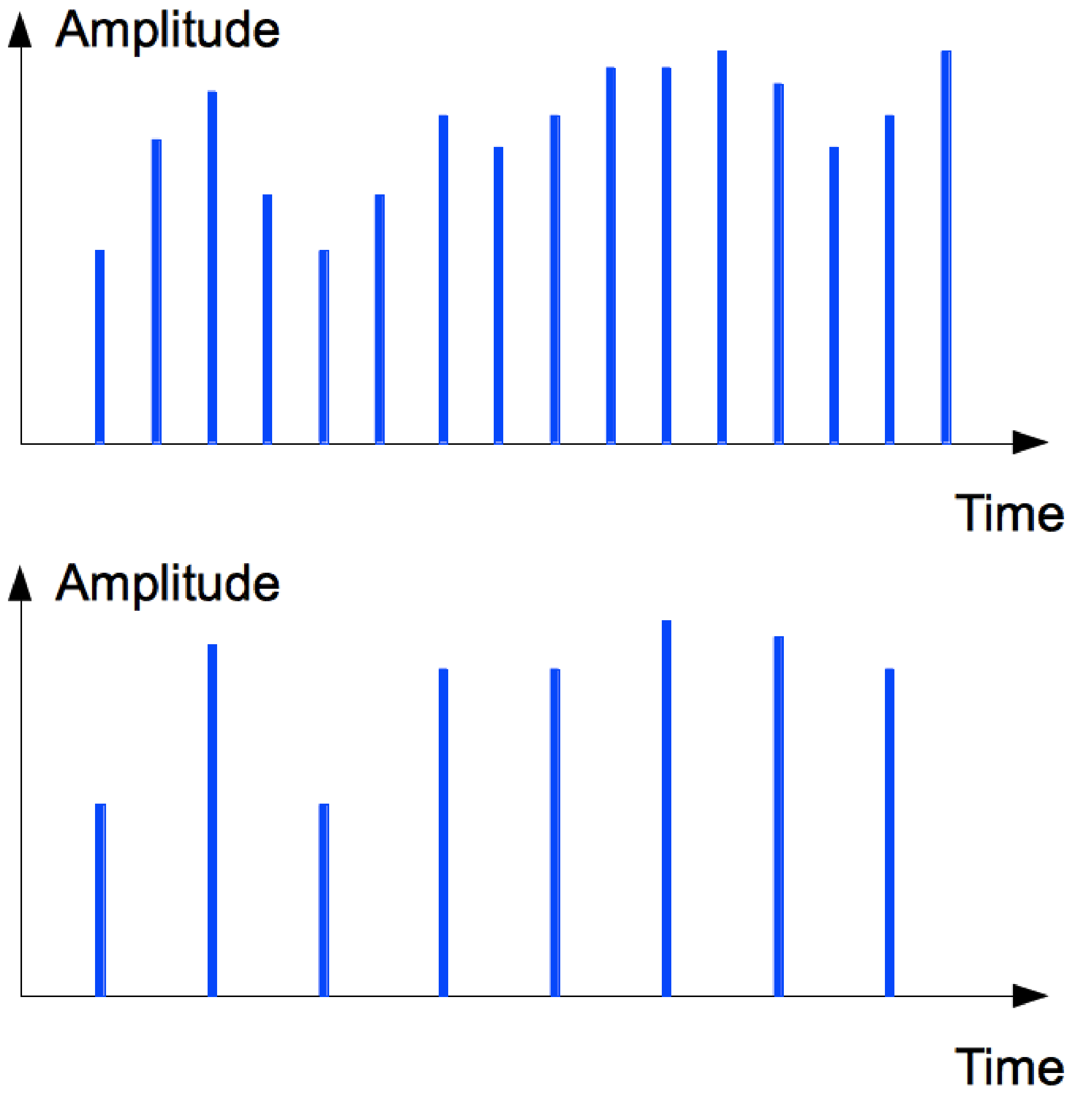
- Each successive PCA component accounts for as much as possible of the remaining activity not accounted for by previous PCA components, so different PCA components may have dramatically different contributions, with the first the maximum (could be more than 50%). ICA components have much more homogeneous contributions, ranging from roughly 5% down to ~0%, because ICA tries to identify maximally independent activity sources
- PCA components of EEG data are spatially or temporally orthogonal, depending on which dimension PCA is applied to. ICA components are maximally temporally independent, without spatial constraints



https://eeglab.org/tutorials/ConceptsGuide/ICA_background.html

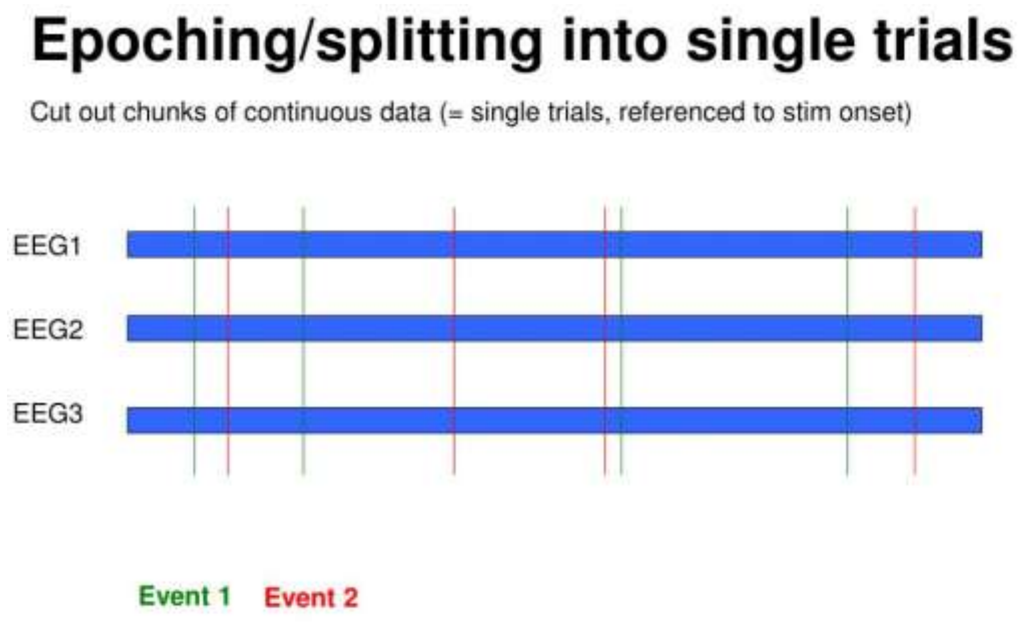
Resampling

- EEGs are typically recorded at very high sampling rate, e.g., 1024 Hz
- Down-sampling reduces the memory and computational cost
- Many studies down-sampled EEG data to 128 Hz



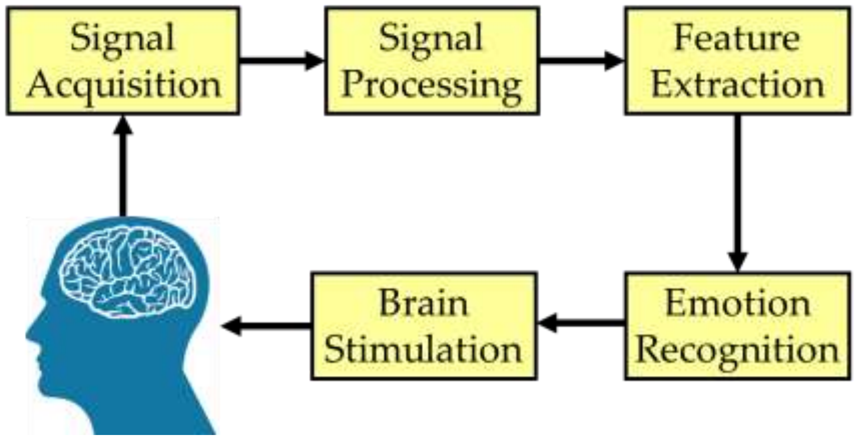
Epoching

- EEG signals are usually recorded continuously for each stimulus, which may last several minutes
- Each such piece of EEG data may be viewed as a block
- Usually each block is further partitioned into many overlapping/non-overlapping shorter (e.g., 10-second) epochs, to increase the number of trials in analysis



Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- **Feature Extraction in aBCI**
- Machine Learning in aBCI
- Brain Stimulation
- aBCI Applications
- Challenges and Opportunities



Feature Extraction in aBCI

- ✓ **Time domain features**
- ✓ **Frequency domain features**
- ✓ **Time-frequency domain features**
- ✓ **Brain connectivity features**
- ✓ **Feature combinations**

Feature Extraction in aBCI

Category	Features
Time domain	Mean, standard deviation, power, un-normalized and normalized 1st/2nd differences, Hjorth's activity/mobility/complexity, normalized non-stationarity index, fractal dimension, higher order crossings
Frequency domain	Band power derivatives (e.g., mean, minimum, maximum, variance, ratio of mean powers of different bands, differential entropy), differential asymmetry, rational asymmetry, DLAT, DCAU, bispectrum, bicoherence, higher order spectra (e.g., bispectra and bicoherence magnitudes)
Time-frequency domain	Short-time Fourier transform and its spectrograms, discrete wavelet transform, Cohen's class, Zhao-Atlas-Marks transform, empirical mode decomposition
Brain connectivity	Pearson correlation coefficient, phase locking value, phase lag index

Time Domain Features

Let $\boldsymbol{x}(t) \in \mathbb{R}^T$ be the time series of a single EEG channel, where T is the number of time samples.

- **Mean**: $\mu_{\boldsymbol{x}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}(t)$.
- **Standard deviation**: $\sigma_{\boldsymbol{x}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{x}(t) - \mu_{\boldsymbol{x}})^2}$.
- **Power**: $P_{\boldsymbol{x}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}^2(t)$.
- **1st difference** (or **normalized length density**): $\delta_{\boldsymbol{x}} = \frac{1}{T-1}|\boldsymbol{x}(t+1) - \boldsymbol{x}(t)|$.
- **Normalized 1st difference**: $\bar{\delta}_{\boldsymbol{x}} = \delta_{\boldsymbol{x}}/\sigma_{\boldsymbol{x}}$.
- **2nd difference**: $\gamma_{\boldsymbol{x}} = \frac{1}{T-2}|\boldsymbol{x}(t+2) - \boldsymbol{x}(t)|$.
- **Normalized 2nd difference**: $\bar{\gamma}_{\boldsymbol{x}} = \gamma_{\boldsymbol{x}}/\sigma_{\boldsymbol{x}}$.
- **Hjorth's activity**: $A_{\boldsymbol{x}} = \sigma_{\boldsymbol{x}}^2$.
- **Hjorth's mobility**: $M_{\boldsymbol{x}} = \sigma_{\dot{\boldsymbol{x}}}/\sigma_{\boldsymbol{x}}$, where $\dot{\boldsymbol{x}}$ is the first derivative of $\boldsymbol{x}$.
- **Hjorth's complexity**: $C_{\boldsymbol{x}} = M_{\dot{\boldsymbol{x}}}/M_{\boldsymbol{x}}$.

Time Domain Features

- **Normalized non-stationarity index** (NSI): $\boldsymbol{x}(t)$ is first normalized by the standard deviation, and then divided into multiple small segments. The mean of each segment is computed, and the NSI is the standard deviation of these means.
- **Fractal dimension**: First re-write $\boldsymbol{x}(t)$ as $\{\boldsymbol{x}(m), \boldsymbol{x}(m+k), \ldots, \boldsymbol{x}\left(m + k\lfloor\frac{T-m}{k}\rfloor\right)\}$, where k is the time interval, $m \in \{1, \ldots, k\}$ is the initial time, and $\lfloor\cdot\rfloor$ is the floor operation. Then, compute

$$L_m(k) = \frac{\sum_{k=1}^{\lfloor\frac{T-m}{k}\rfloor} |\boldsymbol{x}(m+ik) - \boldsymbol{x}(m+(i-1)k)|}{\frac{T-1}{k^2\lfloor\frac{T-m}{k}\rfloor}}.$$

The fractal dimension is finally computed as the negative slope of the log-log plot of $\frac{1}{k}\sum_{m=1}^{k} L_m(k)$ against k.

- **Higher order crossings** (HOC): Apply k different high-pass filters to a zero-mean time series to obtain k filtered time series, and extract the k HOC features as the number of zero-crossings of them.

Frequency Domain Features

- **Band power derivatives**, e.g.,
  - Mean
  - Minimum
  - Maximum
  - Variance
  - Ratio of mean powers of different bands
  - Differential entropy (DE)
- **Higher order spectra**, e.g.,
  - Bispectra
  - Bicoherence magnitudes

Frequency Domain Features - Differential Entropy

$$\mathrm{DE}_{\boldsymbol{x}} = -\int_{-\infty}^{+\infty} \frac{e^{-\frac{(x-\mu_{\boldsymbol{x}})^2}{2\sigma_{\boldsymbol{x}}^2}}}{\sqrt{2\pi\sigma_{\boldsymbol{x}}^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma_{\boldsymbol{x}}^2}} e^{-\frac{(x-\mu_{\boldsymbol{x}})^2}{2\sigma_{\boldsymbol{x}}^2}}\right) \mathrm{d}x$$

$$= \frac{1}{2}\log\left(2\pi e\sigma_{\boldsymbol{x}}^2\right)$$

For a fixed-length EEG sequence, DE is equivalent to the logarithmic energy spectrum in a certain frequency band

Frequency Domain Features – DASM & RASM

Differential asymmetry (DASM) and **rational asymmetry** (RASM), which are the difference and ratio between the DEs of a pair of hemispherically symmetric electrodes (e.g., O1 and O2, denoted as $\boldsymbol{x}_{left}$ and $\boldsymbol{x}_{right}$), respectively:

$$\mathrm{DASM} = \mathrm{DE}_{\boldsymbol{x}_{left}} - \mathrm{DE}_{\boldsymbol{x}_{right}}$$

$$\mathrm{RASM} = \frac{\mathrm{DE}_{\boldsymbol{x}_{left}}}{\mathrm{DE}_{\boldsymbol{x}_{right}}}$$

Frequency Domain Features – DLAT & DCAU

- **DLAT:** Differential spectral band powers (delta, theta, alpha, beta and gamma) for 12 left-right electrode pairs, e.g., Fp1-Fp2 and F7-F8

- **DCAU:** Differential spectral band powers for 12 fronto-posterior electrode pairs, e.g., Fp1-O1 and F7-P7

Frequency Domain Features – Bis & Bic

Bispectrum Bis is the Fourier transform of the 3rd order moment of $\boldsymbol{x}(t)$:

$$\mathrm{Bis}(f_1, f_2) = \mathrm{E}\left[FT(f_1) \cdot FT(f_2) \cdot FT^*(f_1 + f_2)\right],$$

where E is expectation, FT the Fourier transform of $\boldsymbol{x}(t)$, and * the complex conjugate.

Bicoherence Bic is the normalized Bis:

$$\mathrm{Bic}(f_1, f_2) = \frac{\mathrm{Bis}(f_1, f_2)}{\sqrt{P(f_1) \cdot P(f_2) \cdot P(f_1 + f_2)}},$$

where $P(f) = \mathrm{E}[FT(f) \cdot FT^*(f)]$ is the power spectrum.

Time-Frequency Domain Features

Usually 2D spectral representations of EEG signals in simultaneously time and frequency domains:

- Short-time Fourier transform (STFT)
- Spectrograms computed from STFT
- Discrete wavelet transform
- Cohen's class
- Zhao-Atlas-Marks transform
- Empirical mode decomposition (EMD; Hilbert-Huang spectrum)

Brain Connectivity Features

Connectivities between different brain regions (electrodes):

- ✓ Pearson correlation coefficient (PCC)
- ✓ Phase locking value (PLV)
- ✓ Phase lag index (PLI)

The **PCC** between two time series $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$ is:

$$\mathrm{PCC}_{\boldsymbol{x},\boldsymbol{y}} = \frac{\mathrm{cov}(\boldsymbol{x},\boldsymbol{y})}{\sigma_{\boldsymbol{x}}\sigma_{\boldsymbol{y}}},$$

where $\mathrm{cov}(\boldsymbol{x},\boldsymbol{y})$ is the covariance.

Brain Connectivity Features

Define $\boldsymbol{z}(t) = \boldsymbol{x}(t) + j\tilde{\boldsymbol{x}}(t)$, where $j = \sqrt{-1}$, and

$$\tilde{\boldsymbol{x}}(t) = \frac{PV}{\pi} \int_{-\infty}^{\infty} \frac{\boldsymbol{x}(\tau)}{t - \tau} d\tau$$

is the Hilbert transform of $\boldsymbol{x}(t)$, in which PV is the Cauchy principal value. The instantaneous phase $\phi(t)$ of $\boldsymbol{x}(t)$ is then

$$\phi(t) = \arctan \frac{\tilde{\boldsymbol{x}}(t)}{\boldsymbol{x}(t)}.$$

The **PLV** and **PLI** between Channels i and k are then computed as

$$\mathrm{PLV}(i,k) = \left| \frac{1}{T} \sum_{t=1}^{T} e^{j(\phi_i(t) - \phi_k(t))} \right|, \quad \mathrm{PLI}(i,k) \quad = \left| \frac{1}{T} \sum_{t=1}^{T} \mathrm{sign}\left(\phi_i(t) - \phi_k(t)\right) \right|.$$

Both PLV and PLI take values in $[0, 1]$. A larger value indicates better phase locking.

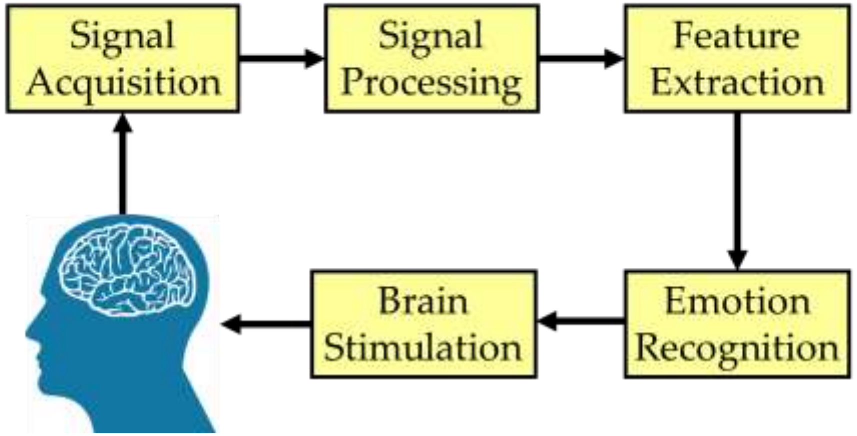
Feature Combinations

Combine different feature extraction approaches:

- **Extract more features from different frequency bands**, e.g., PSD, DE, DASM, RASM and so on in five different frequency bands
- **Extract more features (e.g., DE) from different intrinsic mode functions in EMD**

Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- Feature Extraction in aBCI
- **Emotion Recongition in aBCI**
- Brain Stimulation
- aBCI Applications
- Challenges and Opportunities



Machine Learning in aBCI

- Within-Subject Emotion Recognition
- Transfer Learning for Cross-Subject/ Cross-Session Emotion Recognition
- Deep Transfer Learning
- Multi-Modal Learning
- Cross-Modal Learning

Within-Subject Emotion Recognition

- Each subject is considered individually
- Training and test sets should come from different blocks, e.g., Blocks 1-12 for training and Blocks 13-15 for test
- If we mix all trials together and randomly select 80% of them for training and the remaining 20% for test, then there is **block-design pitfall**:

  The block design leads to classification of arbitrary brain states based on block-level temporal correlations that are known to exist in all EEG data, rather than stimulus-related activity. Because every trial in their test sets comes from the same block as many trials in the corresponding training sets, their block design thus leads to classifying arbitrary temporal artifacts of the data instead of stimulus-related activity

R. Li, J. S. Johansen, H. Ahmed, T. V. Ilyevsky, R. B. Wilbur, H. M. Bharadwaj, and J. M. Siskind, "The perils and pitfalls of block design for EEG classification experiments," IEEE Trans. on Pattern Analysis and Machine Intelligence, 43(1): 316–333, 2021.

Within-Subject Emotion Recognition

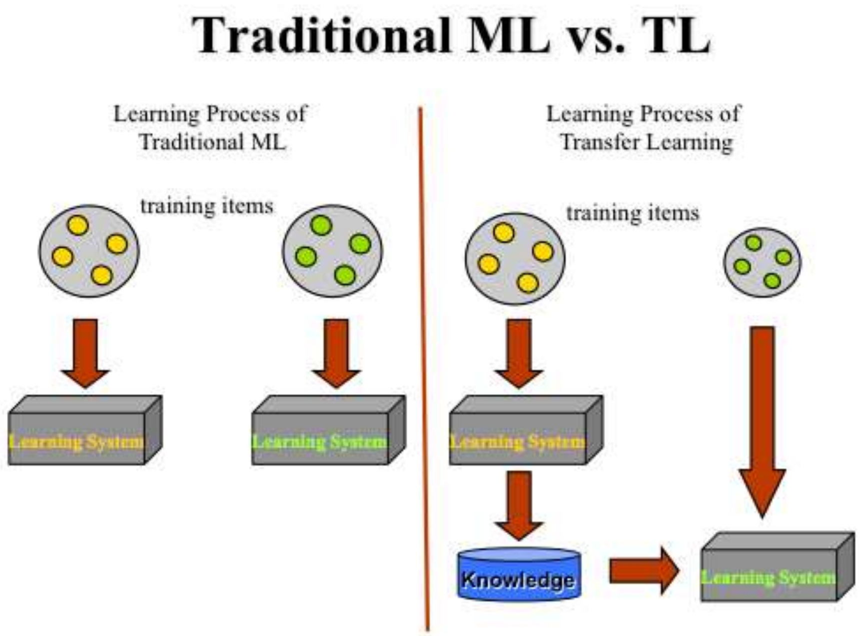
- The **block-design pitfall** becomes more significant if each EEG block is partitioned into multiple overlapping trials
- For example, if the [0,10] second trial in a certain block is used in training, and the subsequent [5,15] second trial from the same block is used in test, then there is **data leakage** since part of the test data have been seen in training, and hence the test results will be over-optimistic
- Very important to perform **cross-block** data partition and evaluations

Transfer Learning for Cross-Subject/Cross-Session Emotion Recognition

- A machine learning model may work well when:
  1) There are **adequate** training data
  2) Training and test data have the **same distribution**
- These two assumptions are not always satisfied in aBCI:
  1) Fast calibration requires collecting **very few calibration trials** from a new subject, i.e., inadequate subject-specific training data
  2) Due to large **individual differences**, generally it is not feasible to use data from existing subjects directly for new subject calibration

Transfer Learning for Cross-Subject/Cross-Session Emotion Recognition

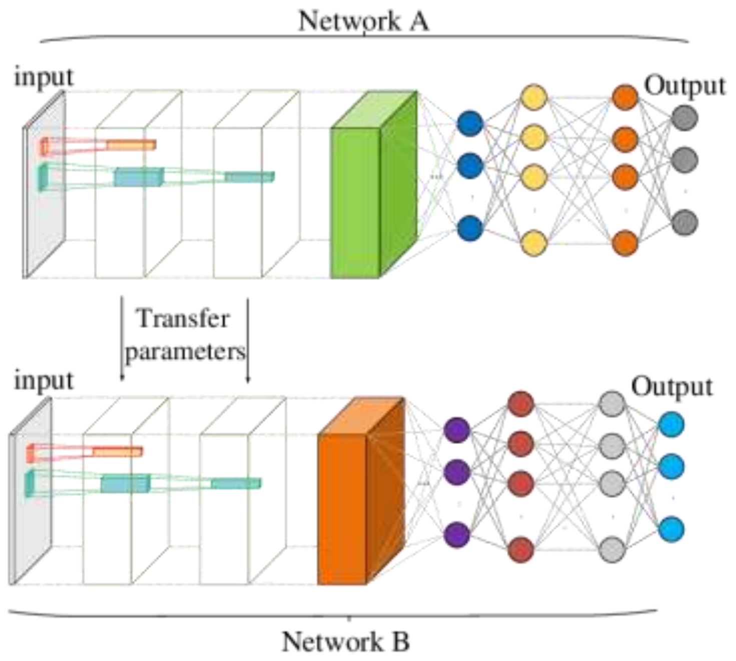
- Transfer learning uses data or knowledge from some source domains (existing subjects) to facilitate the model training in a target domain (new subject)
- Widely used in aBCI to reduce the subject-specific calibration effort:
  - **Instance Transfer:** Weight the source domain samples so that their distribution is more similar to that of the target domain
  - **Feature Transfer:** Perform feature transformations so that the feature distributions of the source and target domains are more similar in the new feature space
  - **Parameter Transfer:** Use the source models to regularize the target model



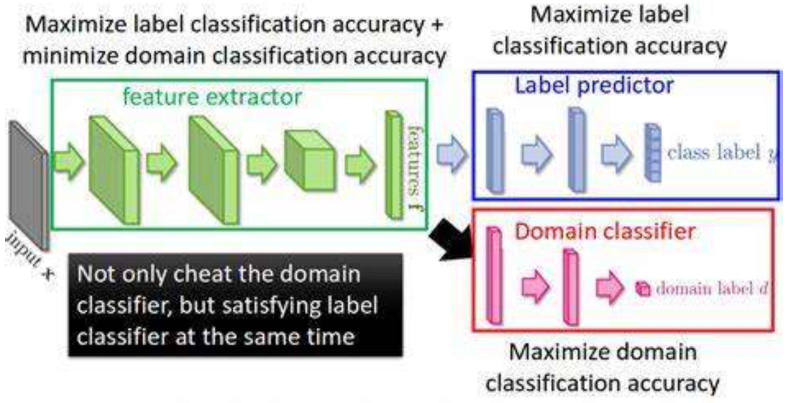
Deep Transfer Learning

- Feature extraction and classification/regression are integrated into a single neural network, and simultaneously optimized

- Deep transfer learning can be achieved through:
  - **Parameter Transfer:** Train a deep learning model using data from multiple auxiliary subjects, and then adapt it to the new subject by fixing the first few feature extraction layers and then using the subject-specific data to fine-tune the last few classification layers
  - **Adversarial Learning:** Bring the data distributions of the auxiliary and new subjects closer



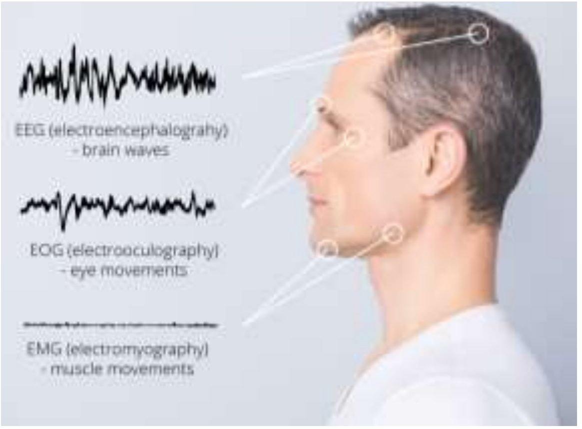
Domain-adversarial training

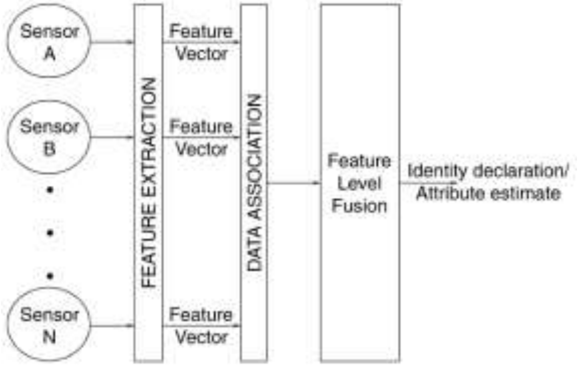


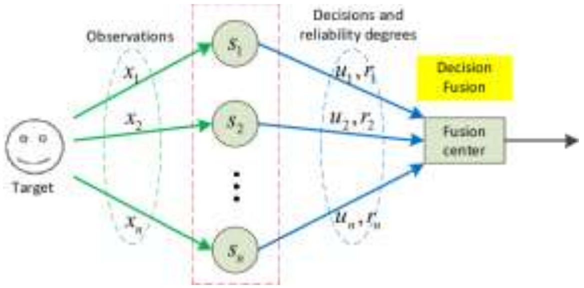
This is a big network, but different parts have different goals.

Multi-Modal Learning

- Multi-modal signals, e.g., EEG, ECG, eye movement and facial expressions, may be used together for more reliable emotion recognition
- Modality fusion approaches:
  - **Feature-level Fusion:** Features of each modality are extracted individually and independently, and then concatenated into a single larger feature vector for classification or regression
  - **Decision-level Fusion:** Builds a classifier for each modality and then aggregate their results







Cross-Modal Learning

- In real-world aBCIs, maybe not all input signals used in training are available in test, e.g., both EEG and eye movement signals are used in training, whereas only eye movement signals are available in test
- Using all available modalities in training may still be beneficial than using only one of them:
  - **Training Phase**: Train a multi-modal fusion network and an emotion classifier using both EEG and eye movement features, then a conditional GAN to learn the relationship between eye movements and the multi-modal features
  - **Test Phase**: Use eye movement features to regress multi-modal features for emotion classification

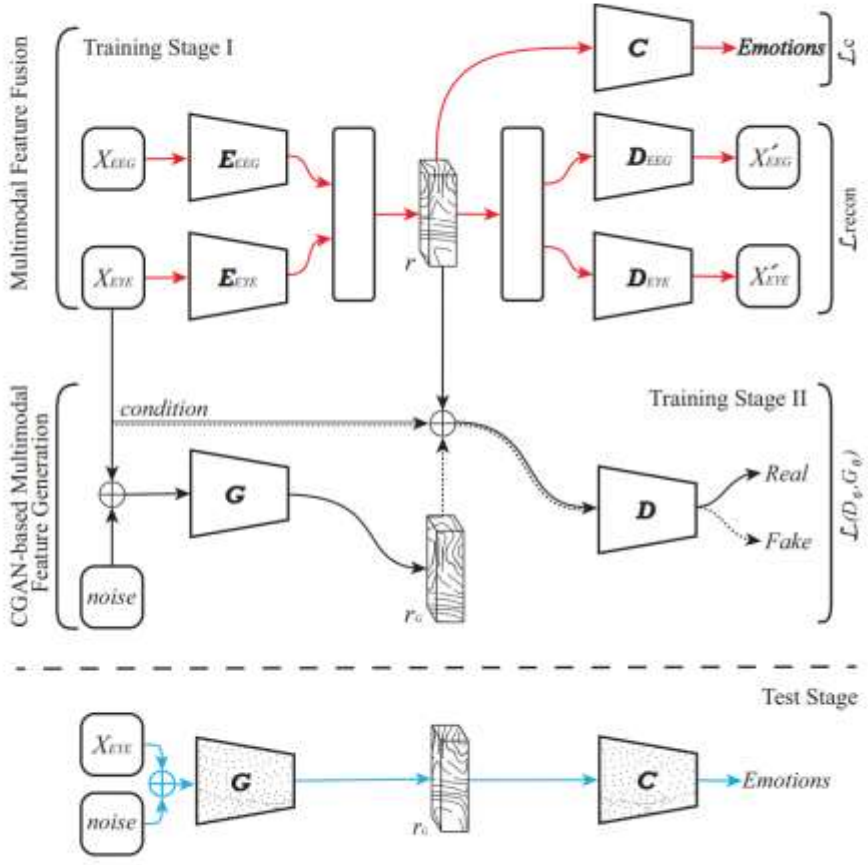
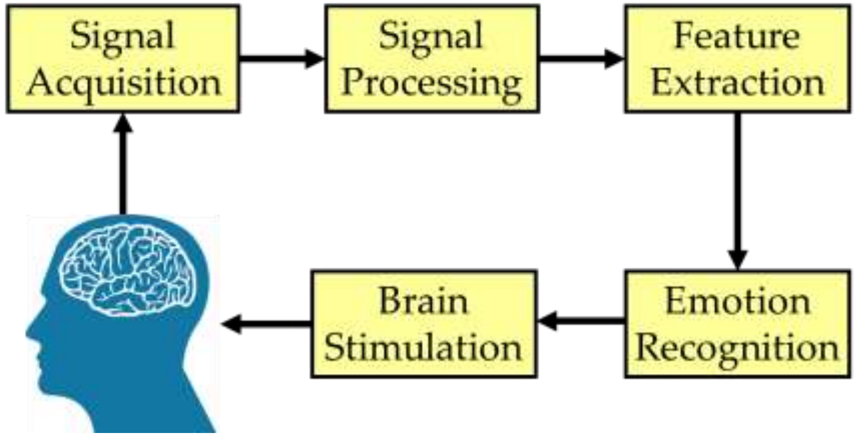


Figure 1: The framework of our system. The training stage can be divided into two parts as multimodal feature extraction and multimodal feature generation. In the test phase, only eye movement signals are needed. The shaded *G* and *C* indicate that they have been trained. Best viewed in color.

Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- Feature Extraction in aBCI
- Emotion Recongnition in aBCI
- **Brain Stimulation**
- aBCI Applications
- Challenges and Opportunities

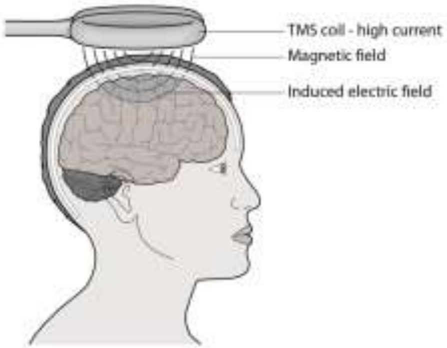


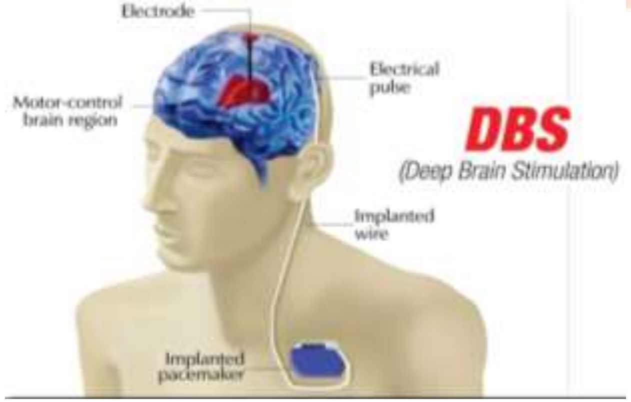
Brain Stimulation

- For emotion regulation, which refers broadly to implementation of a conscious or non-conscious goal to start, stop or otherwise modulate the trajectory of an emotion
- Multiple EEG-based emotion biomarkers have been proposed, e.g., increasing EEG Alpha activity reduced the anxiety, and vice versa
- fMRI and fNIRS based emotion biomarker: Regulate the activity of a certain region of interest, e.g., amygdala or dorsolateral prefrontal cortex

Brain Stimulation

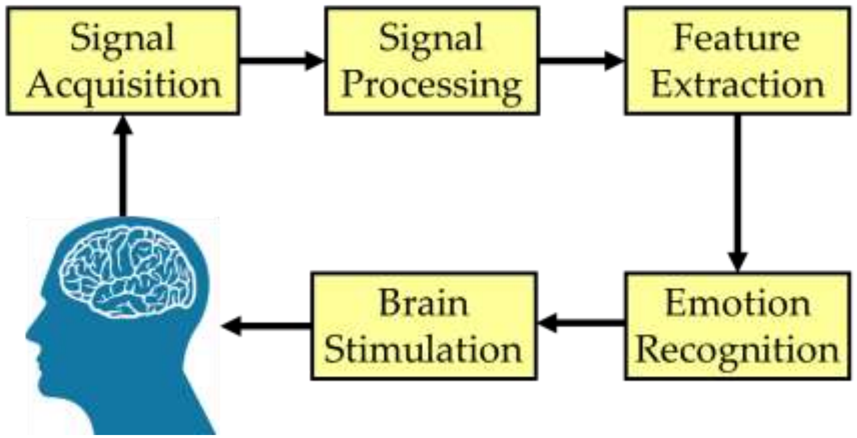
- Brain stimulation using electroconvulsive therapy (ECT), transcranial magnetic stimulation (TMS), vagal nerve stimulation (VNS) and deep brain stimulation (DBS) have also been used in treatment resistant depression (TRD) therapies
- ECT is widely available and its effects are relatively rapid in severe TRD, but its cognitive adverse effects may be cumbersome
- TMS is safe and well tolerated, and it has been approved by FDA for adults who have failed to respond to one antidepressant, but its use in TRD is still controversial as it is not supported by rigorous double-blind randomized clinical trials
- VNS and DBS require surgeries
- VNS has been FDA-approved for TRD, however it is not indicated for management of acute illness
- DBS for TRD is still an experimental area of investigation and double-blind clinical trials are underway





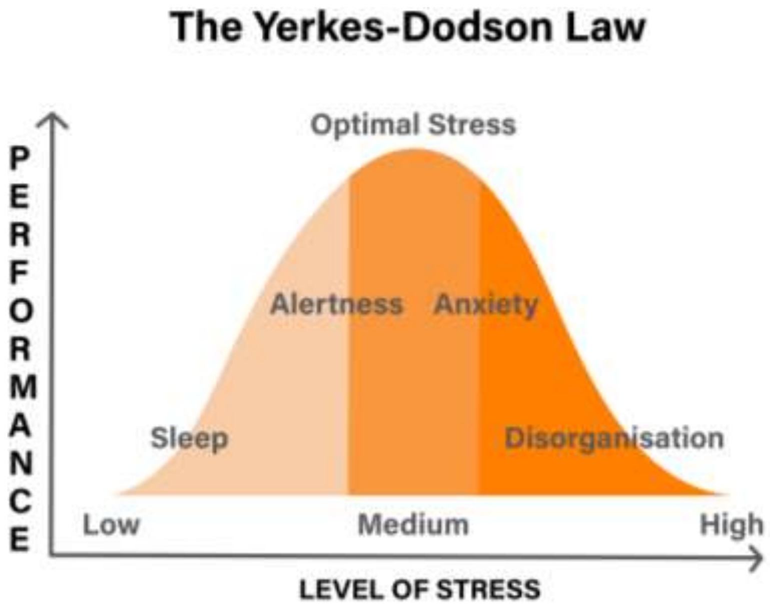
Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- Feature Extraction in aBCI
- Emotion Recognition in aBCI
- Brain Stimulation
- **aBCI Applications**
- Challenges and Opportunities



Cognitive Workload Recognition

- The relation between the function relating the mental resources demanded by a task and those resources available to be supplied by the human
- Maintaining an appropriate workload helps improve the operator's safety and efficiency
- There may be associations between workload and EEG frequency band power

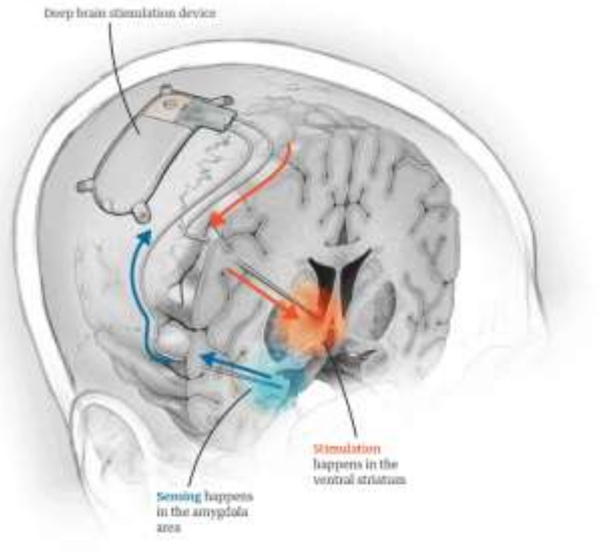


Fatigue Estimation

- Fatigue (drowsiness) is an important contributor to many accidents, particularly traffic accidents
- Different strategies, e.g., computer vision based and driving behavior based, have been used for driver fatigue estimation
- EEG-based driver fatigue (drowsiness) estimation, which is less dependent on overt behavior and less susceptible to deception, belongs to the latter
- Another advantage of EEG-based driver fatigue estimation is that fatigue may be detected from EEG signals earlier than from other modalities like facial expressions, as fatigue originates deep in the brain

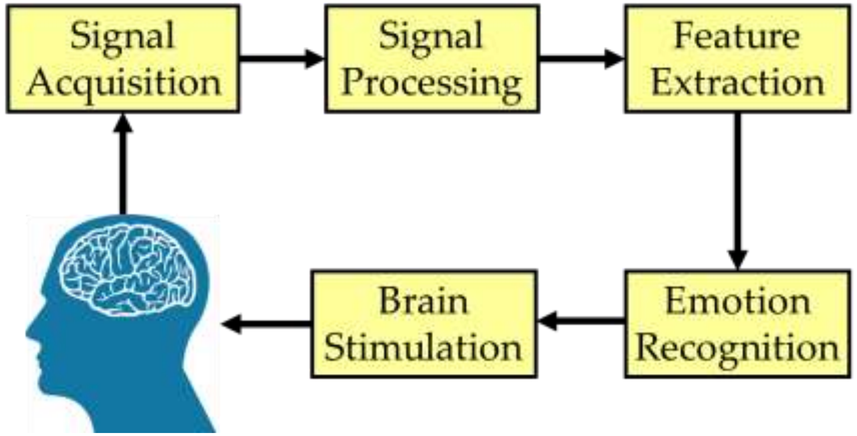
Depression Diagnosis and Treatment

- Depression is a common mental disorder
- Globally, it is estimated that 5% of adults suffer from depression
- Depression is a leading cause of disability worldwide and is a major contributor to the overall global burden of disease
- Multiple EEG-based biomarkers are proposed for depression diagnosis
- aBCIs have also been extensively used in treatment resistant depression therapies, e.g., DBS implants electrodes within certain areas of the brain to generate electrical impulses to regulate the emotion



Outline

- Affective Brain-Computer Interface (aBCI)
- Signal Acquisition in aBCI
- Signal Processing in aBCI
- Feature Extraction in aBCI
- Emotion Recognition in aBCI
- Brain Stimulation
- aBCI Applications
- **Challenges and Opportunities**



Brain Signal Acquisition

- Wet EEG electrodes use conductive gel to increase the conductivity:
  - **Pro**: Good EEG signal quality
  - **Con**: Gel injection is time-consuming and user-unfriendly, which may hinder the acceptance of wet EEG electrode based aBCIs to consumers

- Dry EEG electrodes do not need conductive gel:
  - **Pro**: Convenient to use
  - **Con**: Signal quality still needs improvements

- For broad real-world applications, important to develop cheap, convenient and high-fidelity brain signal acquisition devices

Emotion Labeling

- Video or music based emotion elicitation may have problems:
  1) Because of individual differences, the subject may not feel happy when watching a ‘happy’ movie
  2) Even though the subject may feel happy, the activation level may be too low to be reflected in his/her EEG signals
  3) The subject may feel happy for a short duration of the movie, but it is difficult to know which part it is, so it is usually assumed that the subject has a happy emotion during the entire duration of the movie
  4) People may exhibit multiple emotions simultaneously, e.g., a graduate may feel both happy and sad at the graduation ceremony, but most aBCI experiments assign only one emotion label or rating to each movie/music
- More accurate and realistic emotion labeling approaches are needed

Diversity and Size of aBCI Datasets

Problems with current public aBCI datasets:

- **Dataset size is very small**: The number of subjects is 8-58 (most <30), and the number of stimuli is 13-40. Larger aBCI datasets will also facilitate more rapid progresses
- **Demography of subjects is limited**: The mean age is 21-35, so children, teenagers and the elderly are not adequately represented. Dedicated datasets and models should be created for them

Negative Transfer in Emotion Recognition (1)

Possible reasons for **negative transfer**:

1) **Large domain divergence**, e.g., negative transfer in emotion recognition may occur when the source and target subjects have significantly different cultural backgrounds

2) **Poor source data quality**, e.g., the emotion recognition accuracy of the source subject is too low to be useful in helping the target subject

3) **Poor target data quality**, e.g., EEG data from the target subject contain too many artifacts/noise, or the EEG electrodes may be placed at scalp locations that are not responsible for emotions

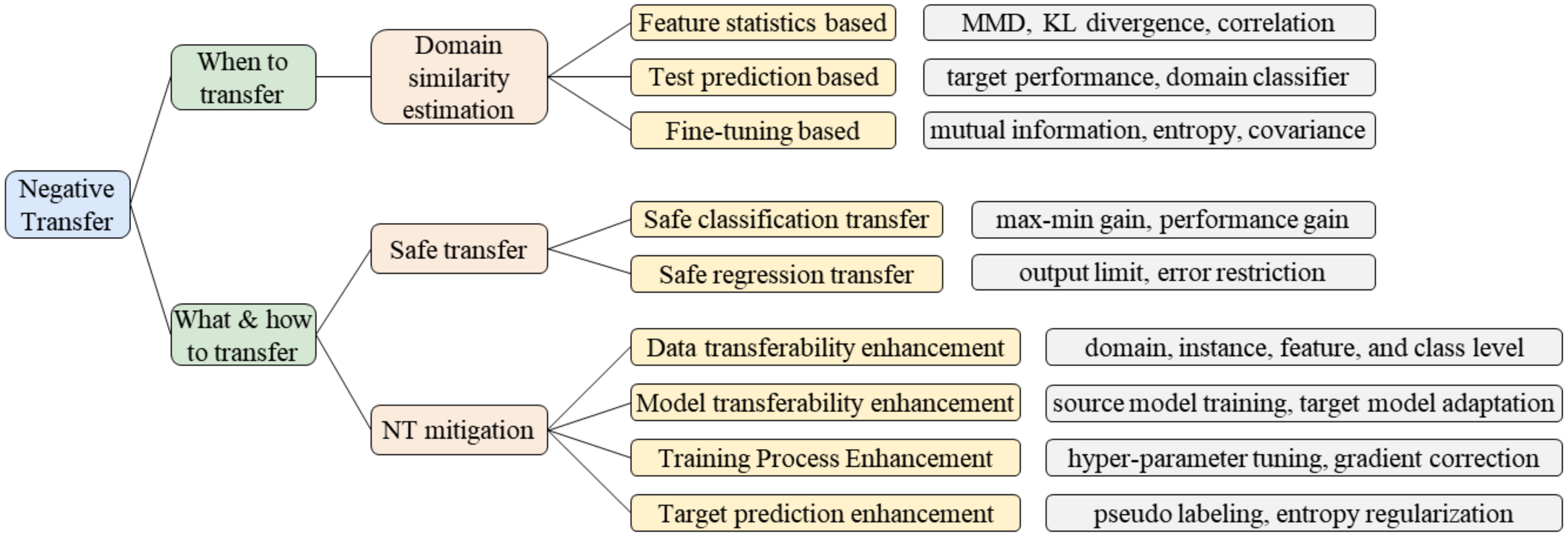
4) **Inappropriate transfer learning algorithm**. Each transfer learning algorithm has its assumptions and specific application scenarios

W. Zhang, L. Deng, L. Zhang and D. Wu*, A Survey on Negative Transfer, IEEE/CAA Journal of Automatica Sinica, 2023.

Negative Transfer in Emotion Recognition (2)

Mitigate **negative transfer** in aBCI:

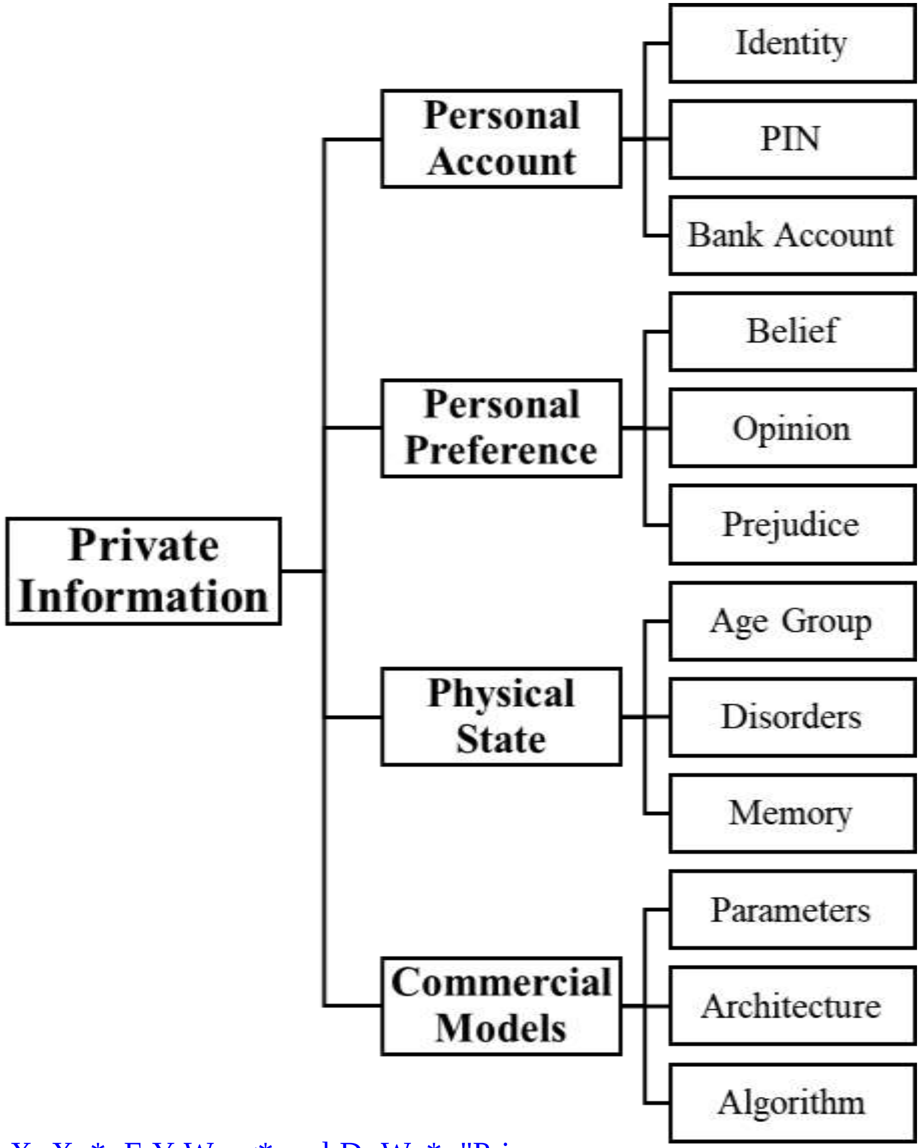
1) **Domain similarity estimation**, which is particularly useful in selecting the most similar source domains from multiple ones, i.e., to reduce the domain divergence
2) **Safe transfer**, which includes deliberately designed algorithms that can avoid negative transfer with theoretical guarantees, regardless of how the source and target subjects are different from each other
3) **Negative transfer mitigation**, which alleviates negative transfer using data/model transferability enhancements, training process enhancements, and/or target prediction enhancements



W. Zhang, L. Deng, L. Zhang and D. Wu*, A Survey on Negative Transfer, IEEE/CAA Journal of Automatica Sinica, 2023.

Privacy-Preserving aBCI (1)

- Many aBCIs use transfer learning to facilitate the calibration, making use of EEG data from source subjects
- In addition to emotions, EEG signals also contain other private information, e.g., user identify, health status, psychological state, etc., which may be easily revealed
- For example, Kong et al. performed EEG-based user identification on SEED, achieving over 99% accuracy using only seconds of EEG data
- User privacy protection in aBCIs is very important
- Recent laws and regulations, e.g., European GDPR (5/25/2018) and China Personal Information Protection Law (11/1/2021), enforce strict user privacy protection



K. Xia, W. Duch, Y. Sun, K. Xu, W. Fang, H. Luo, Y. Zhang, D. Sang, X. Xu*, F-Y Wang* and D. Wu*, "Privacy-Preserving Brain-Computer Interfaces: A Systematic Review," IEEE Trans. on Computational Social Systems, 2022

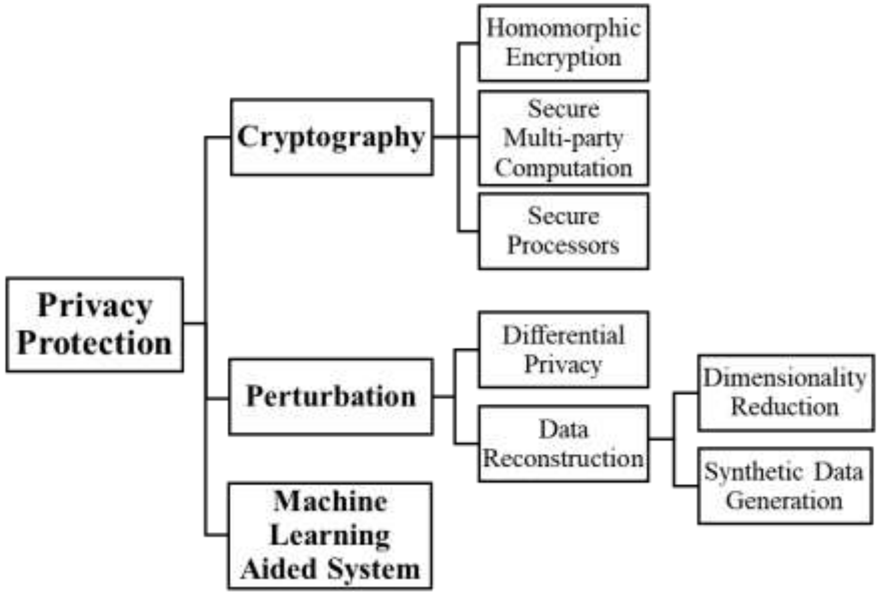
Privacy-Preserving aBCI (2)

Three different strategies to implement privacy-preserving BCIs:

1) **Cryptography**, which includes homomorphic encryption, secure multi-party computation, and secure processors

2) **Perturbation**, which transforms or adds noise to the original EEG data while maintaining their utility for emotion recognition

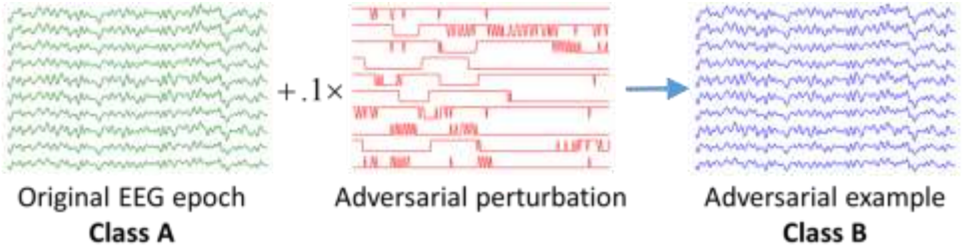
3) **Machine learning aided systems**, which may be used to help people better understand privacy policies, and inform them about the privacy risks when making privacy decisions



K. Xia, W. Duch, Y. Sun, K. Xu, W. Fang, H. Luo, Y. Zhang, D. Sang, X. Xu*, F-Y Wang* and D. Wu*, "Privacy-Preserving Brain-Computer Interfaces: A Systematic Review," IEEE Trans. on Computational Social Systems, 2022

Secure aBCI

- EEG-based BCIs are subject to adversarial attacks: Deliberately designed tiny perturbations can be used to fool a machine learning algorithm

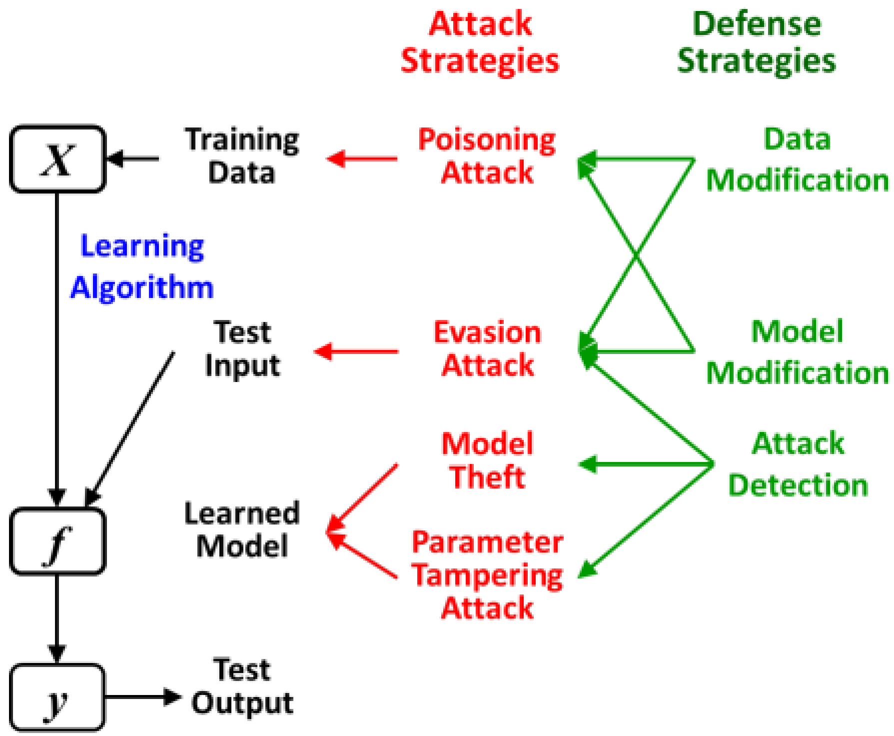


- Defense strategies:

  1) **Data Modification:** Modify the training data, e.g., through adversarial training, or the test data, e.g., through data compression or randomization

  2) **Model Modification:** Modify the target model directly, e.g., through regularization and defensive distillation

  3) **Auxiliary Tools:** Use auxiliary machine learning modules, e.g., adversarial detection



D. Wu, J. Xu, W. Fang, Y. Zhang, L. Yang, X. Xu*, H. Luo* and X. Yu*, Adversarial Attacks and Defenses in Physiological Computing: A Systematic Review, National Science Open, 2022

Thank you!