

Improving surgical motor skill assessment and acquisition via neuromodulation, neuroimaging, and machine learning

Yuanyuan Gao

Committee members:

Dr. Suvranu De, Chair

Dr. Xavier Intes, Co-Chair

Dr. Pingkun Yan

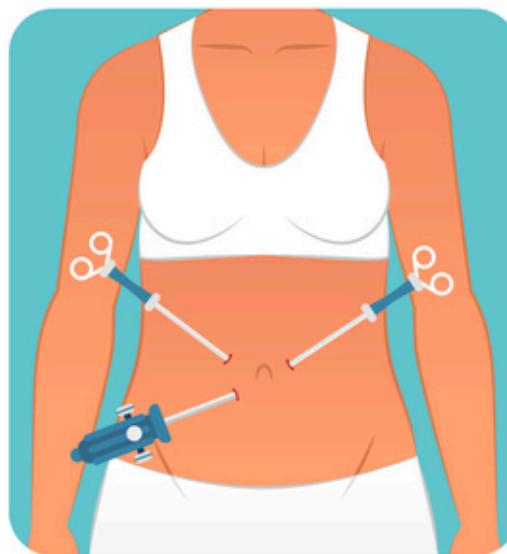
Dr. Lucy T. Zhang

Dr. Emily Liu

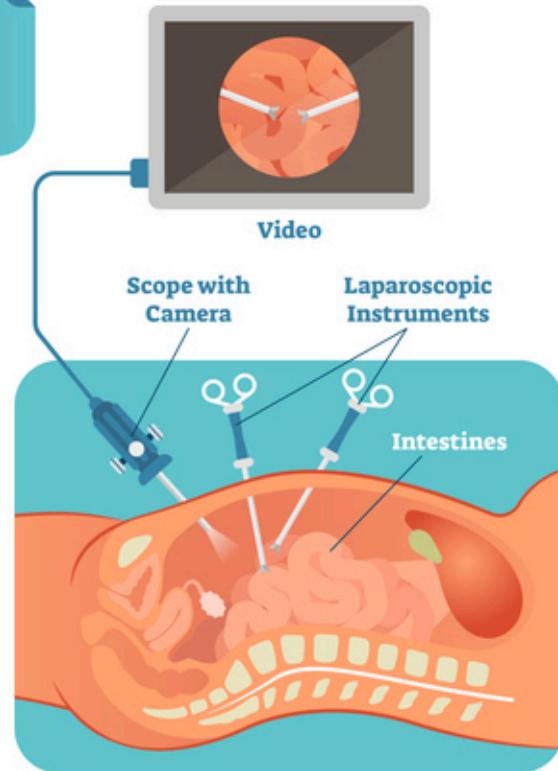


CeMSIM
THE CENTER FOR MODELING, SIMULATION
& IMAGING IN MEDICINE

LAPAROSCOPIC SURGERY



Patient Front



Patient Side View

craigrancho.com

- + Minimal incision
- + Reduced pain
- + Shorter recovery time.
- Complex bimanual motor skills and hand-eye coordination
- Variable adverse event rate¹
- Most errors during the learning phase

1. Flum DR, Koepsell T, Heagerty P, Sinanan M, Dellinger EP. Common Bile Duct Injury During Laparoscopic Cholecystectomy and the Use of Intraoperative Cholangiography: Adverse Outcome or Preventable Error? *Arch Surg.* 2001;136(11):1287–1292.

International accredited program

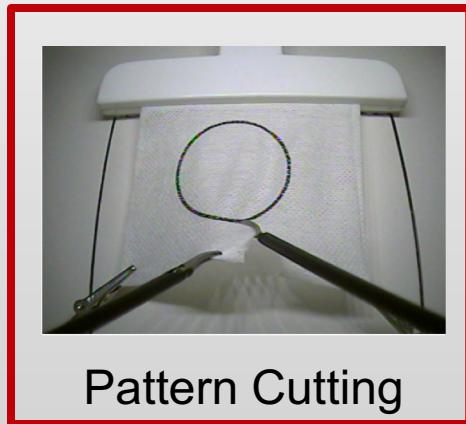


Society of American Gastrointestinal and Endoscopic Surgeons (SAGES)

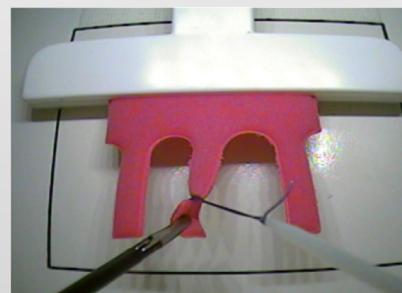
- **Fundamentals of Laparoscopic Surgery (FLS)** is a pre-requisite for Board certification to every general and Ob/Gyn surgeon².
- Two components – cognitive (high stakes exam) + psychomotor (trainer box)
- FLS trainer box is effective in teaching technical motor skills².
- Performance is assessed by FLS score, which is formulated by completion time and performance error.



Peg Transfer



Pattern Cutting



Ligating Loop



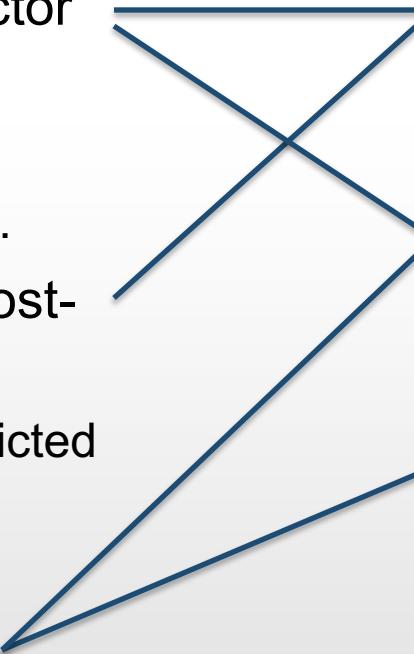
Intracorporeal & Extracorporeal Suture

(www.flspogram.org)

2. H. Peters, G. M. Fried, L. L. Swanstrom, N. J. Soper, L. F. Sillin, B. Schirmer, and K. Hoffman, "Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery," *Surgery*, vol. 135, no. 1, pp. 21–27, Jan. 2004.

Challenges:

- FLS score is manually calculated by proctor (www.flaprogram.org).
 - Time consuming- two to three weeks;
 - Labor intensive – trained proctor needed.
- Skill acquisition procedure is analyzed post-hoc.
 - Learning curve factors could not be predicted before the completion of the training.
 - Impede training protocol customization.
- The training protocol relies on repetition.
 - Time consuming.



Tools:

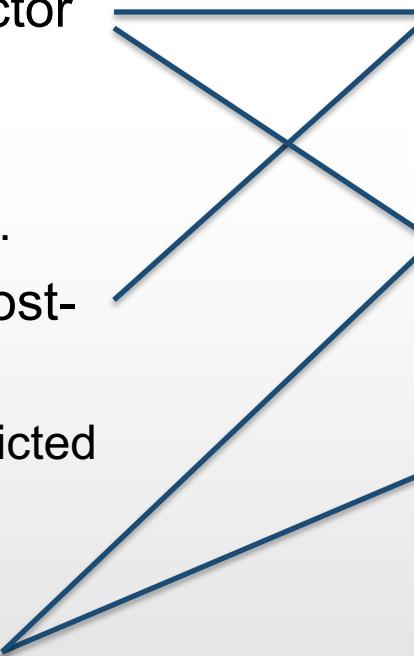
- Machine learning
 - High power in pattern recognition;
 - Automation in feature extraction.
- Neuroimaging
 - Instant measurement;
 - Insight in brain activities.
- Neuromodulation
 - Facilitated motor skill learning process.

Specific aim 1

“.. is to predict learning curve factors in the early stage of training”

Challenges:

- FLS score is manually calculated by proctor (www.flaprogram.org).
 - Time consuming- two to three weeks;
 - Labor intensive – trained proctor needed.
- Skill acquisition procedure is analyzed post-hoc.
 - Learning curve factors could not be predicted before the completion of the training.
 - Impede training protocol customization.
- The training protocol relies on repetition.
 - Time consuming.



Tools:

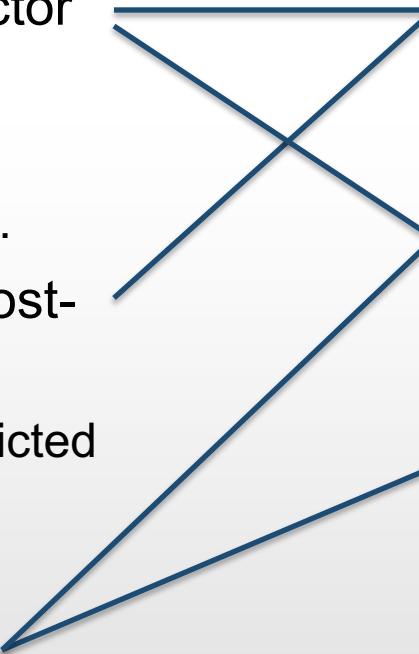
- Machine learning
 - High power in pattern recognition;
 - Automation in feature extraction.
- Neuroimaging
 - Instant measurement;
 - Insight in brain activities.
- Neuromodulation
 - Facilitated motor skill learning process.

Specific aim 2

“.. is to predict FLS scores via neuroimaging”

Challenges:

- FLS score is manually calculated by proctor (www.flaprogram.org).
 - Time consuming- two to three weeks;
 - Labor intensive – trained proctor needed.
- Skill acquisition procedure is analyzed post-hoc.
 - Learning curve factors could not be predicted before the completion of the training.
 - Impede training protocol customization.
- The training protocol relies on repetition.
 - Time consuming.



Tools:

- Machine learning
 - High power in pattern recognition;
 - Automation in feature extraction.
- Neuroimaging
 - Instant measurement;
 - Insight in brain activities.
- Neuromodulation
 - Facilitated motor skill learning process.

Specific aim 3

“.. is to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.”

Specific aim 1

“.. is to predict learning curve factors in the early stage of training”

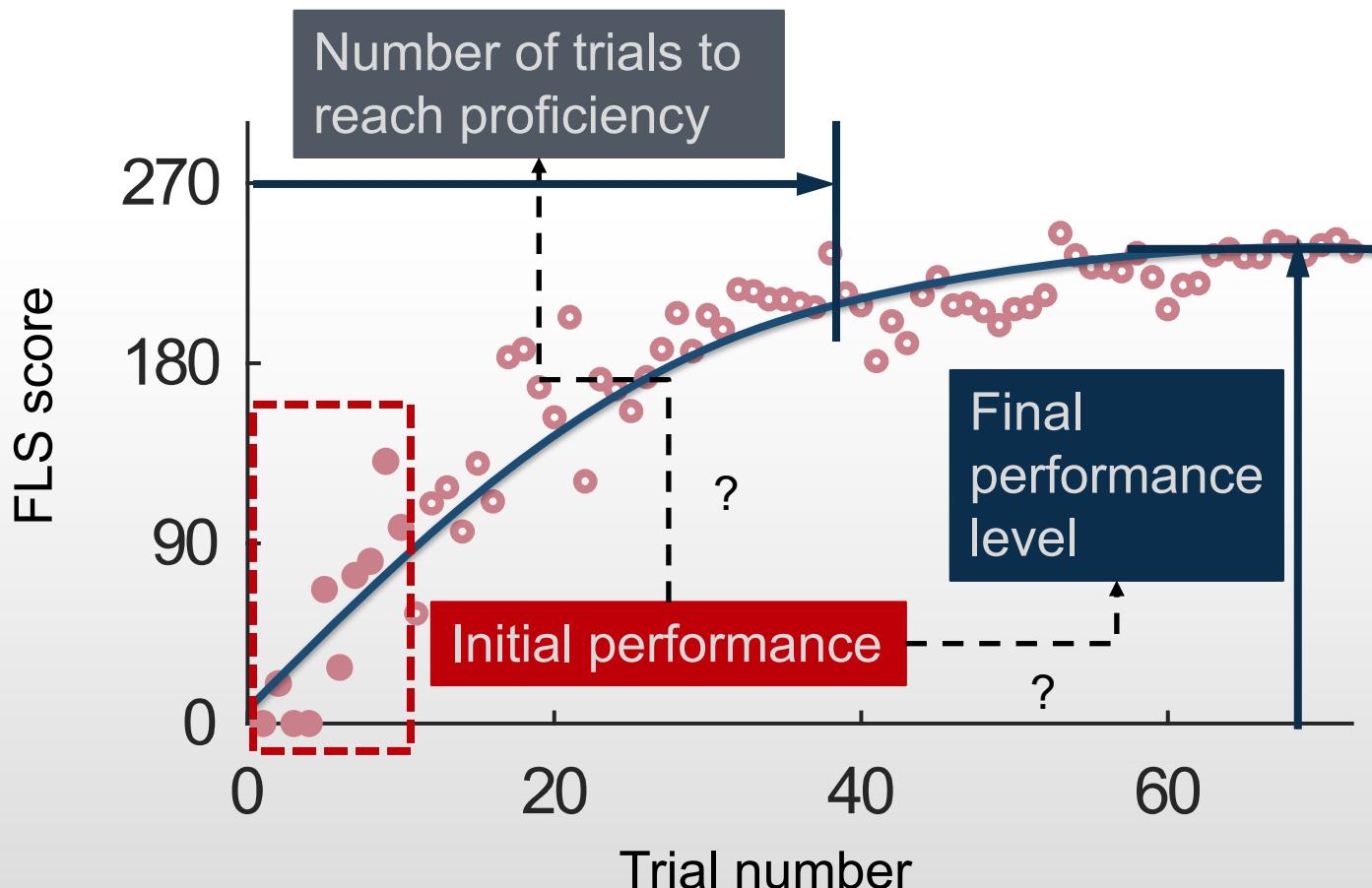
Specific aim 2

“.. is to predict FLS scores via neuroimaging”

Specific aim 3

“.. is to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.”

Background and motivation



- Predicting learning curve features ahead allows training protocol customization;
- The existing learning curve modeling methods
 - Log-linear
- $$Y = Y_0 N^\theta$$
- No approach is to predict the learning curve characteristics from the initial performance.

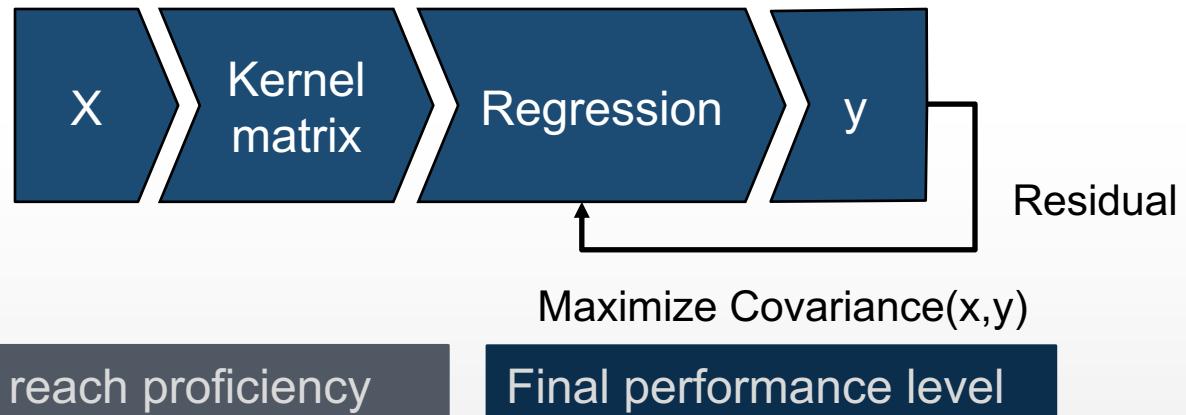
Question: Can the number of trials to reach proficiency and final performance level be predicted from the initial performance information?

N is the number of trials, Y is the performance, θ is the learning rate, and Y_0 is the initial performance level.

Learning curve analysis

Hypothesis #1: The initial performance of a trainee can predict the number of trials required to achieve proficiency and the final proficiency level.

- Kernel Partial Least Squares (KPLS)
 - Small sample size
 - High dimensional variables



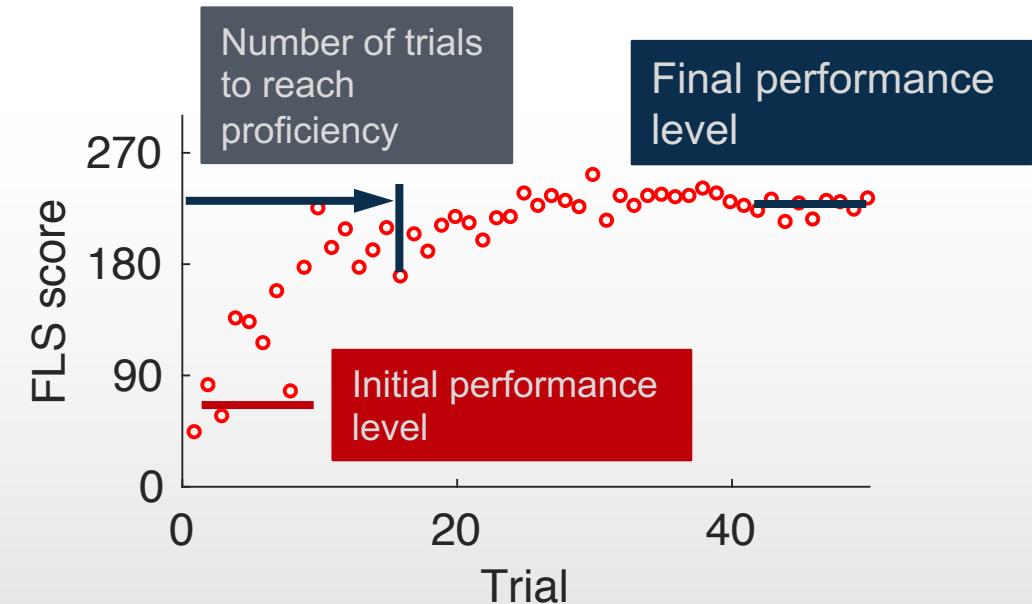
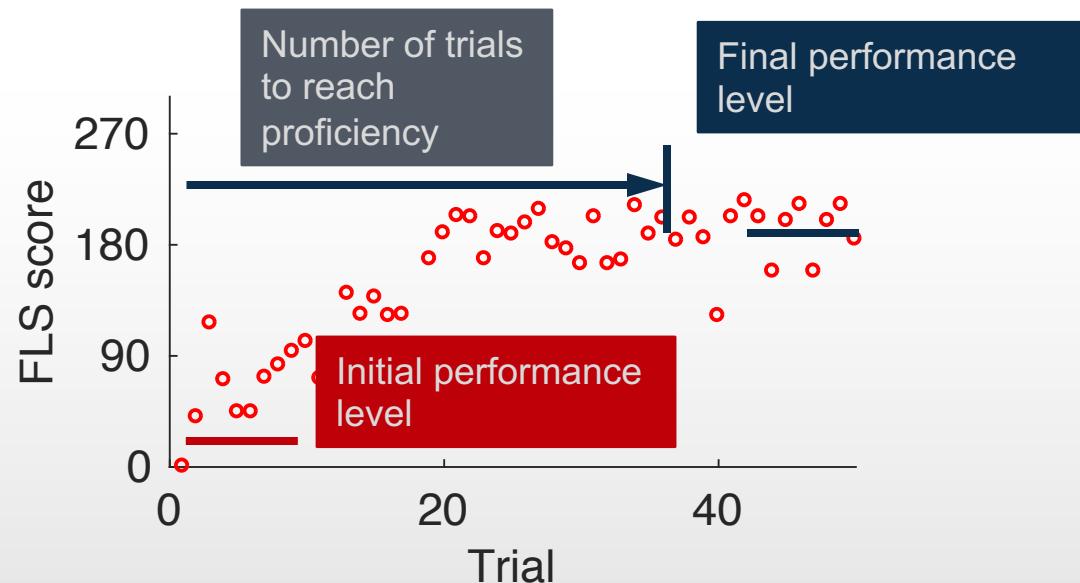
- The coefficient of determination:
$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$
- Cross-validation: Leave-one-out

Prediction	KPLS	Log-linear model
Initial performance → Number of trials to reach proficiency	$R^2 = 0.72$	$R^2 = -109.55$
Initial performance → Final performance level	$R^2 = 0.89$	$R^2 = -3.36$

Significance: enables the customization of the training protocol.

Background and motivation

- The learning procedure is different between trainees

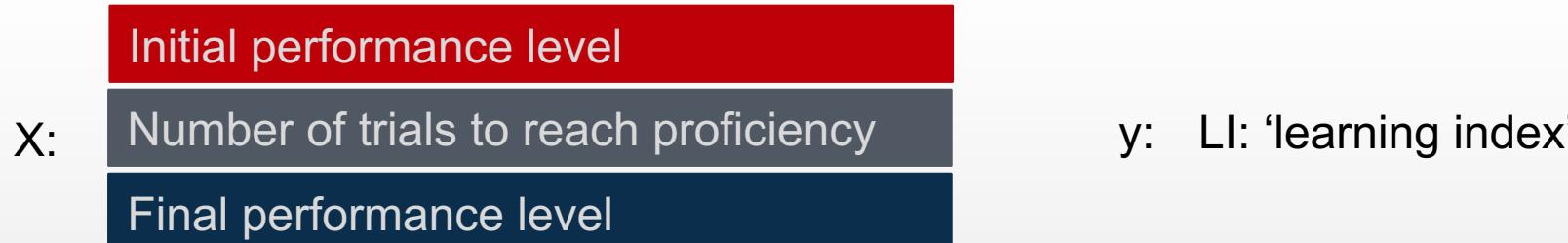


Question: Can we represent the difference by one factor?

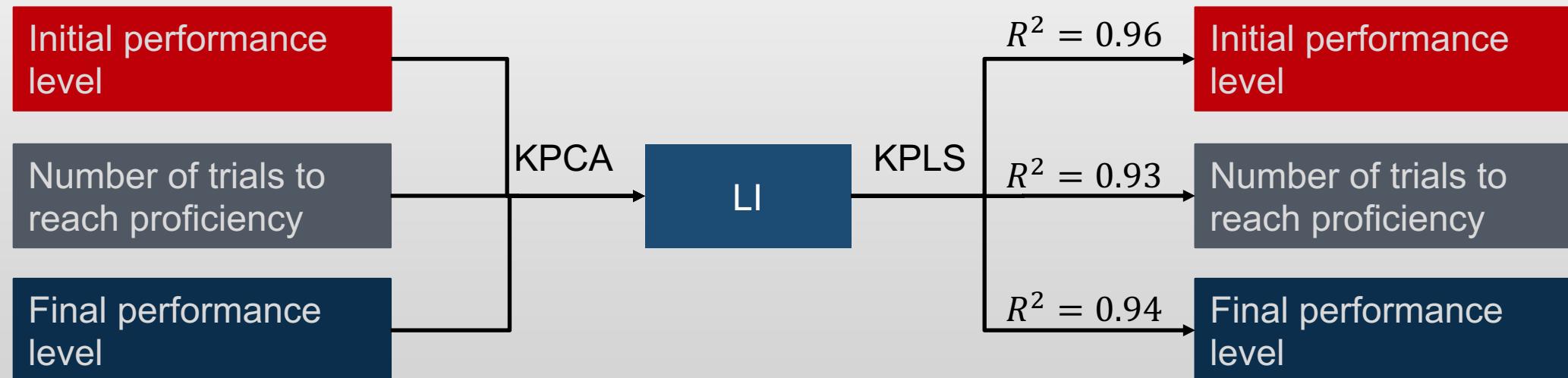
Learning curve analysis

Hypothesis #2: A single factor can describe the learning curve factors

- Kernel Principle Component Analysis (KPCA)
 - Small sample size
 - High dimensional variables

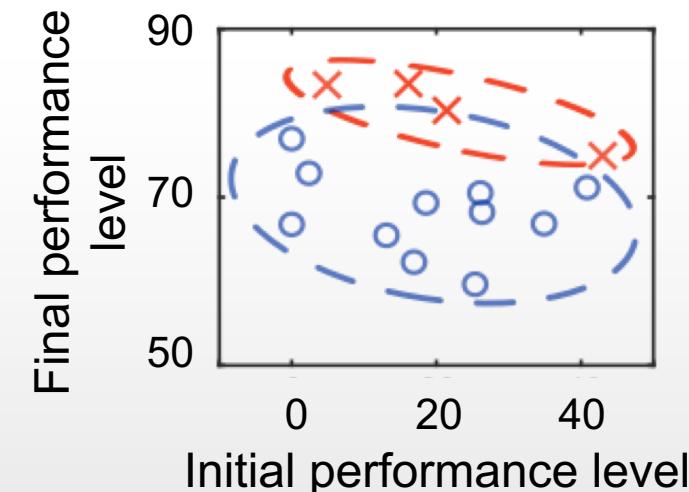
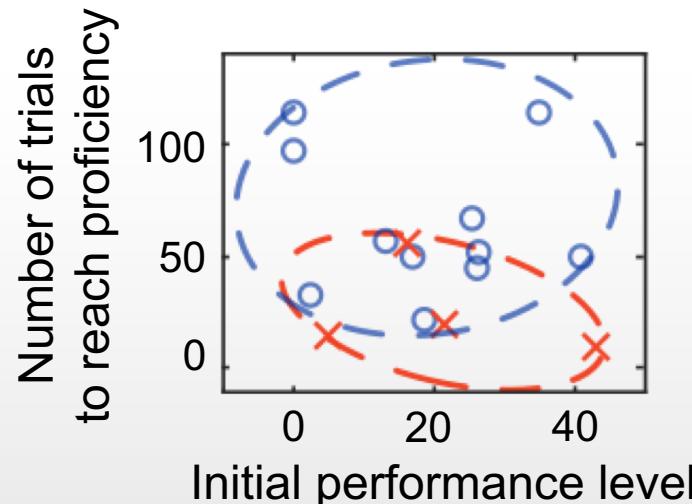


- Could LI reflect the learning curve factors?

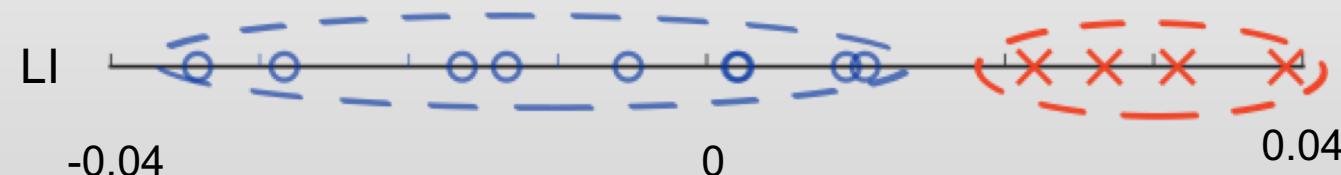


Learning curve analysis

- K-means grouping results
 - Using the learning curve factors



- Using the extracted LI value



Significance: provide a quantitative understanding in different learning abilities.

Specific aim 1

“.. is to predict learning curve features in the early stage of training” **(Completed)**

We established that:

- Learning curve factors can be predicted from the initial performance;
- Single factor can represent the learning curve factors.

Impact:

- Enabled the surgical training customization;
- Understanding of different learning abilities.

Specific aim 1

“.. is to predict learning curve features in the early stage of training”

Specific aim 2

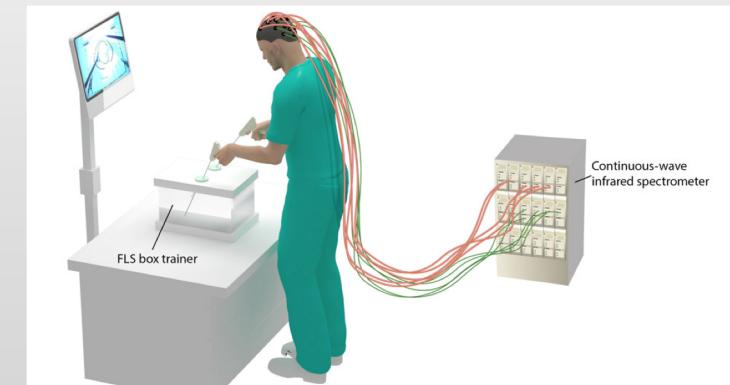
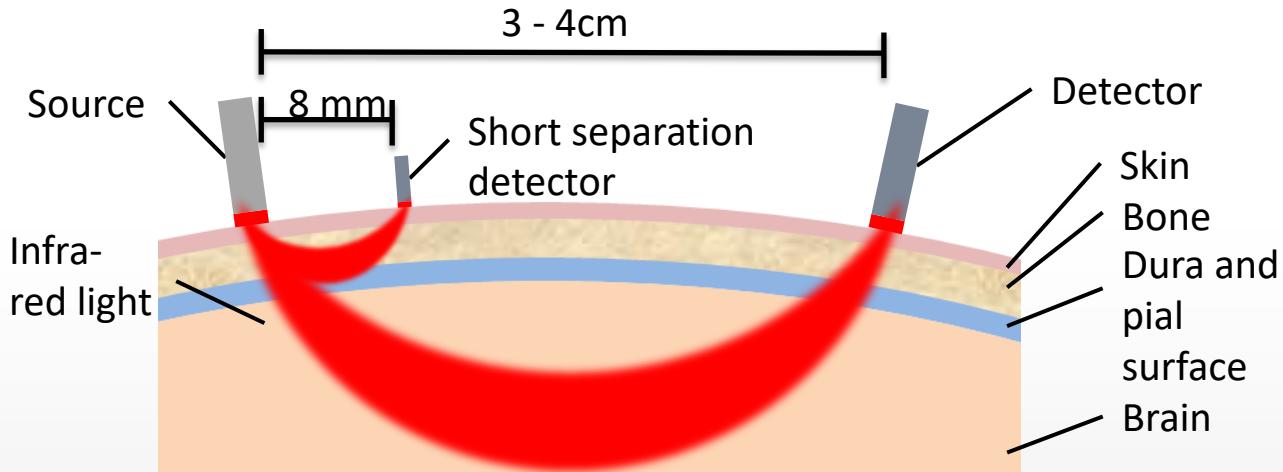
“.. is to predict FLS scores via neuroimaging”

Specific aim 3

“.. is to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.”

Background

- Functional Near infrared spectroscopy (fNIRS)
 - Non-invasive imaging technique
 - Delivers infrared light on the surface of the scalp via source probes.
 - Infrared light scatters through turbid tissue and the backscattered light is detected
- Since attenuated light is related to functional chromophores (such as oxy-HbO₂ and deoxy-HbO₂), the relative concentration of these chromophores can be determined and finally be correlated with brain activity.
- Why should we use NIRS to measure brain activity?
 - High temporal resolution (~ 100Hz)
 - High depth penetration (~1.5 cm)
 - Non-invasive and allows for complex tasks to be performed

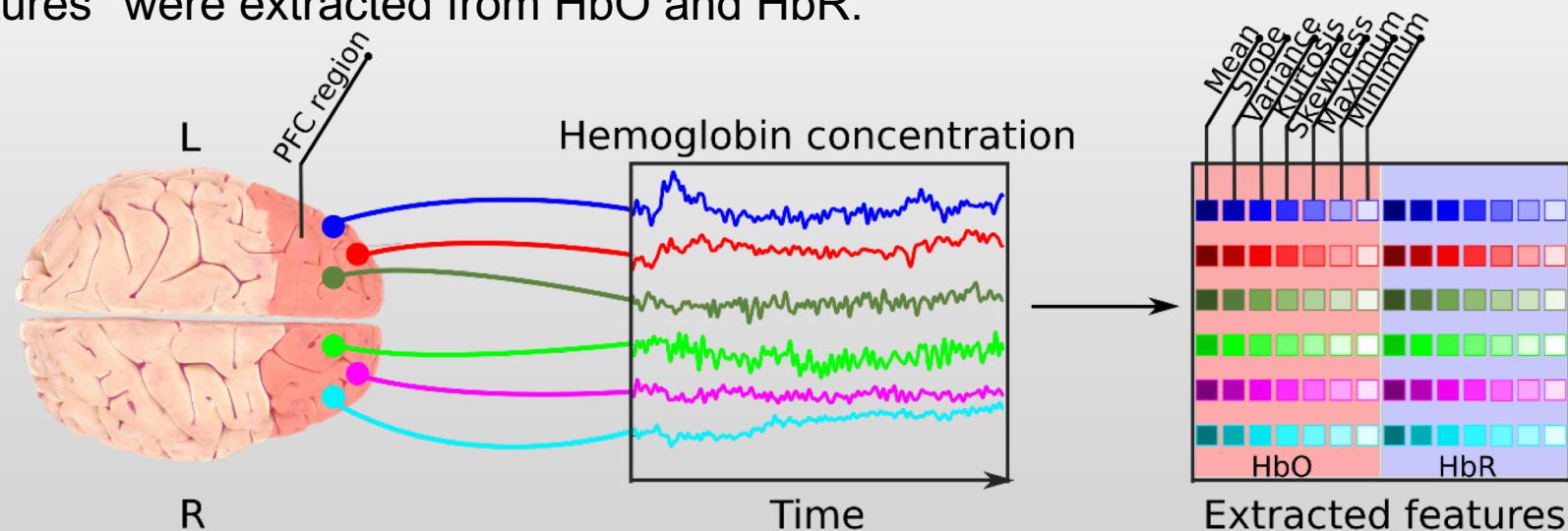


- fNIRS data could classify motor tasks.
 - Simple vs. complex motor tasks (Holper and Wolf 2011).
 - Left vs. right hand motion (Fazli et al. 2012, Naseer and Hong 2013).
 - Arm lifting vs. knee extension (Shin and Jeong 2014).
- fNIRS data could classify surgical levels.
 - Surgical skill levels could be classified by fNIRS (Nemani et al. 2018).
 - Transfer skill levels could be classified by fNIRS (Nemani et al. 2019).

Question: Can fNIRS data predict FLS score?

Method

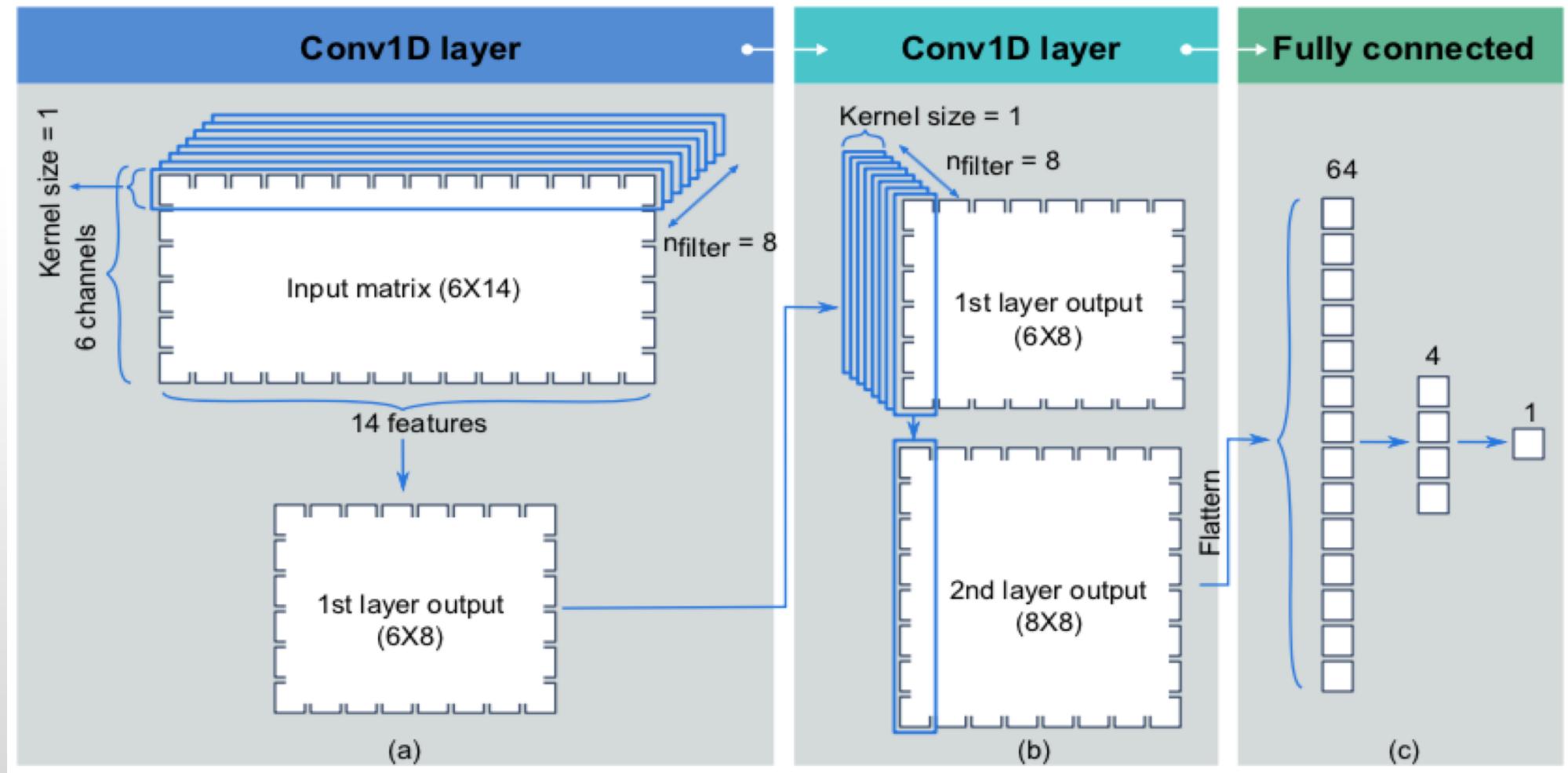
- Deep learning approaches
 - Seizer detection from EEG data via Convolutonal neural network (CNN) (Yuan et al. 2019).
 - Cardiac disease detection from ECG data via CNN (Fan et al. 2018).
 - Human activity classification from kinematic data via CNN (Yao et al. 2017).
- Aim: to predict FLS score from fNIRS data via deep learning approaches.
- Data acquisition:
 - fNIRS data was acquired from 13 medical students during the execution of pattern cutting task;
- Feature extraction:
 - Oxy- (HbO) and deoxy-hemoglobin (HbR) concentration from PFC region;
 - Seven features* were extracted from HbO and HbR.



*Note: Mean, variance, slope, skewness, kurtosis, minimum and maximum.

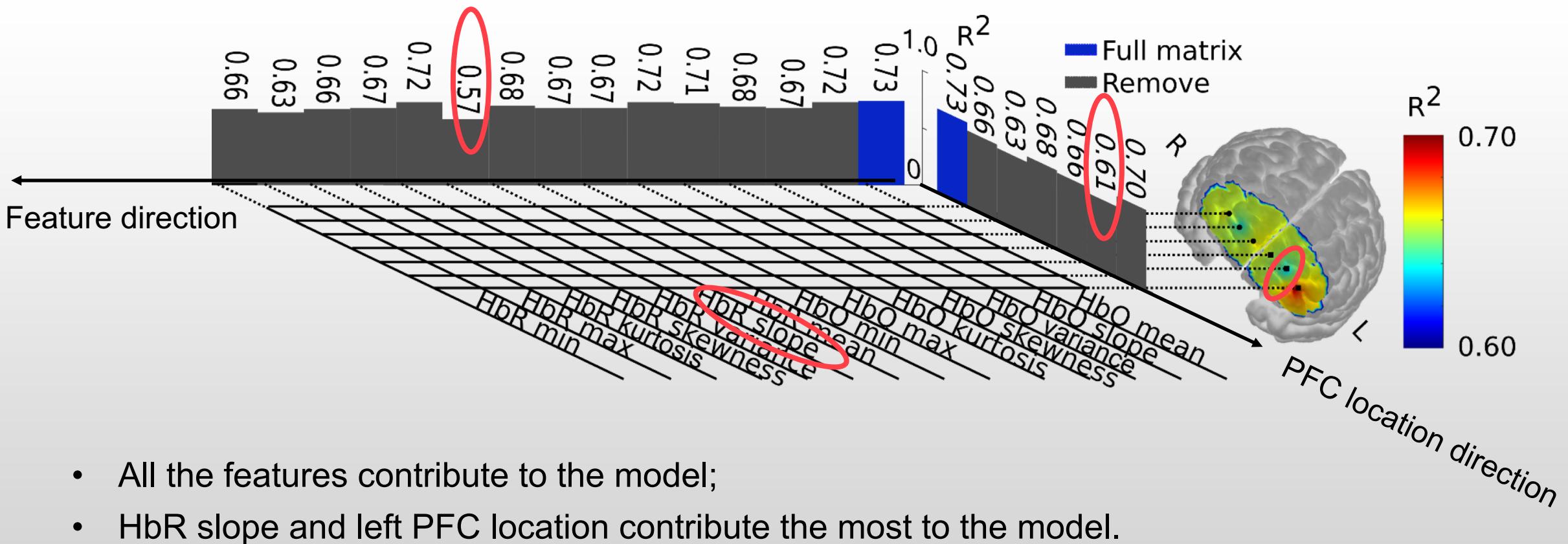
Method

- Brain-NET:



Results

- Question: Do all the features contribute to the model? Which features contribute the most?
 - Backward feature selection



Results

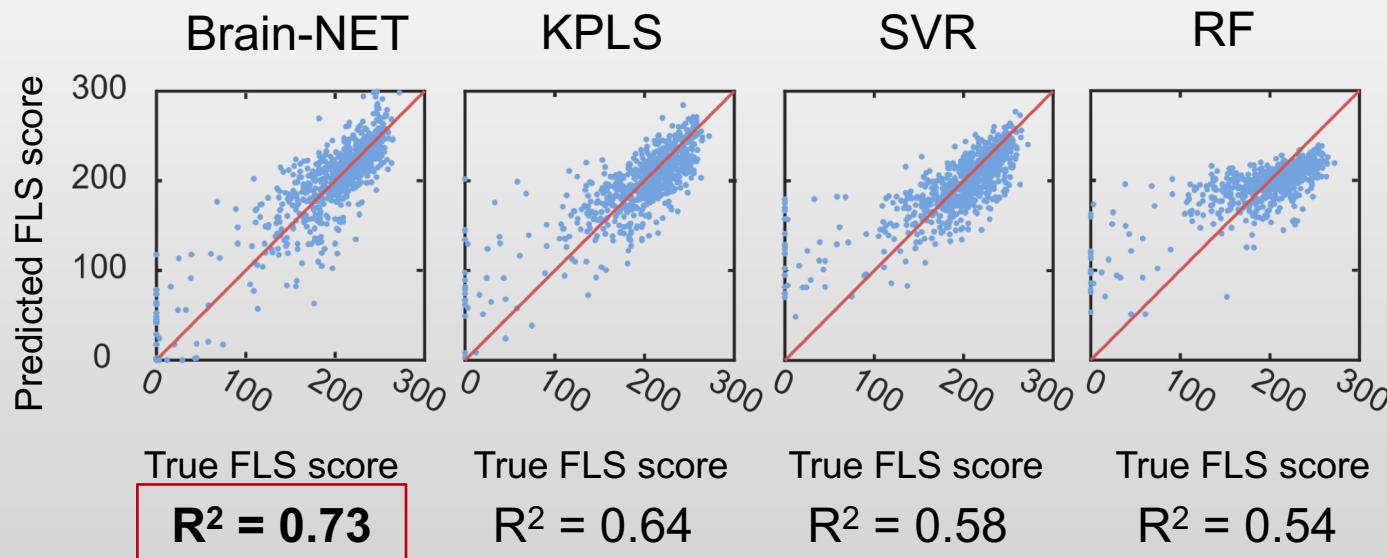
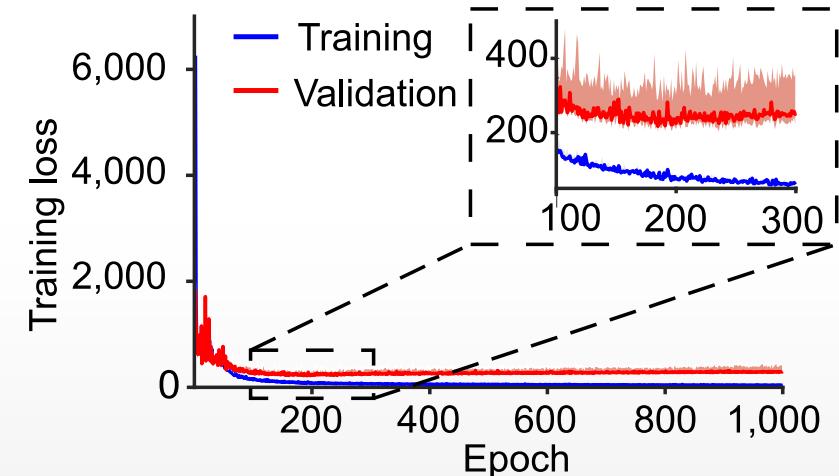
- **Training**

- The model is evaluated with 30 rounds of ten-fold cross validation using randomly shuffled samples for each round.

- **Results**

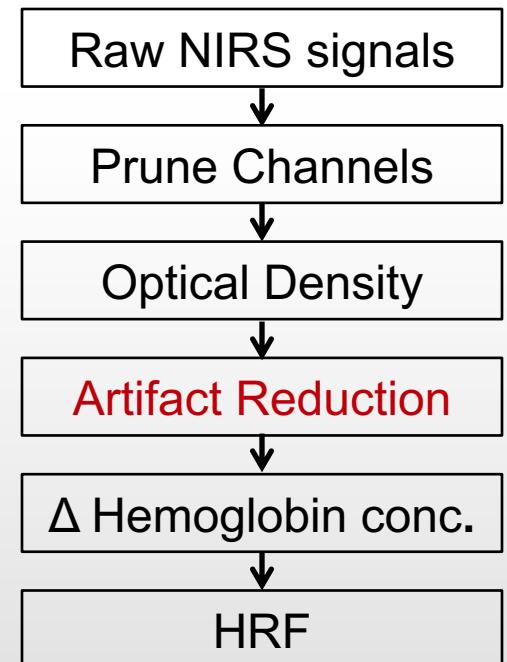
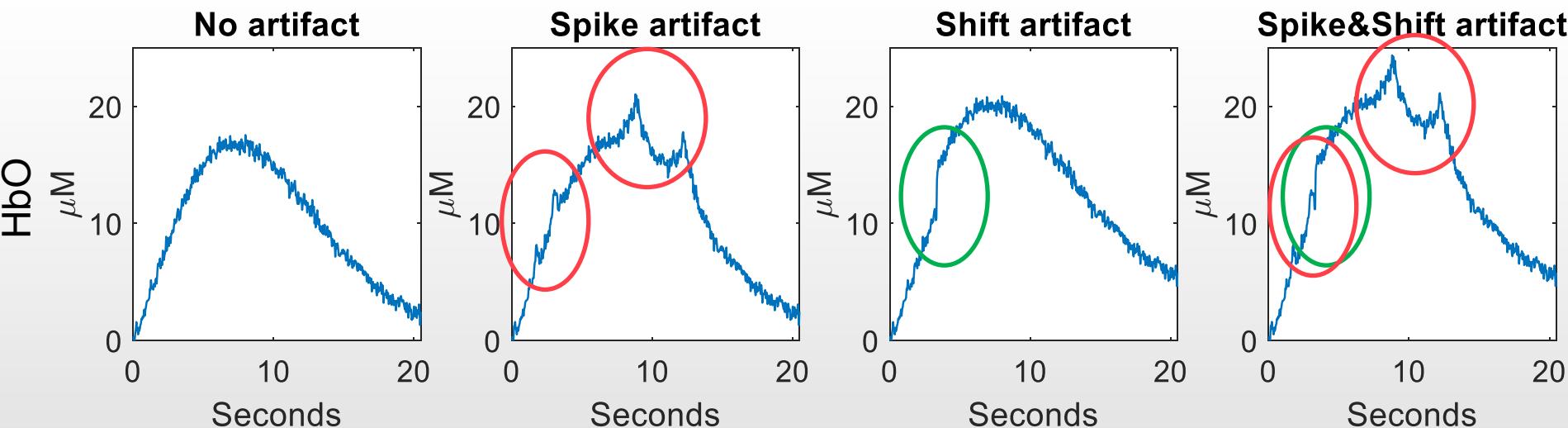
- Since the relationship between fNIRS data and FLS score is unknown, curve fitting regression methods such as linear regression or polynomial regression are not suitable.
- KPLS, SVR and RF are curve fitting free methods:
 - KPLS - Kernel partial least squares
 - SVR – Support vector regression
 - RF – Random forest
- We compared Brain-NET with KPLS, SVR, RF:

Significance: establish fNIRS as a fast, cost-effective way to assess FLS score.



Limitation

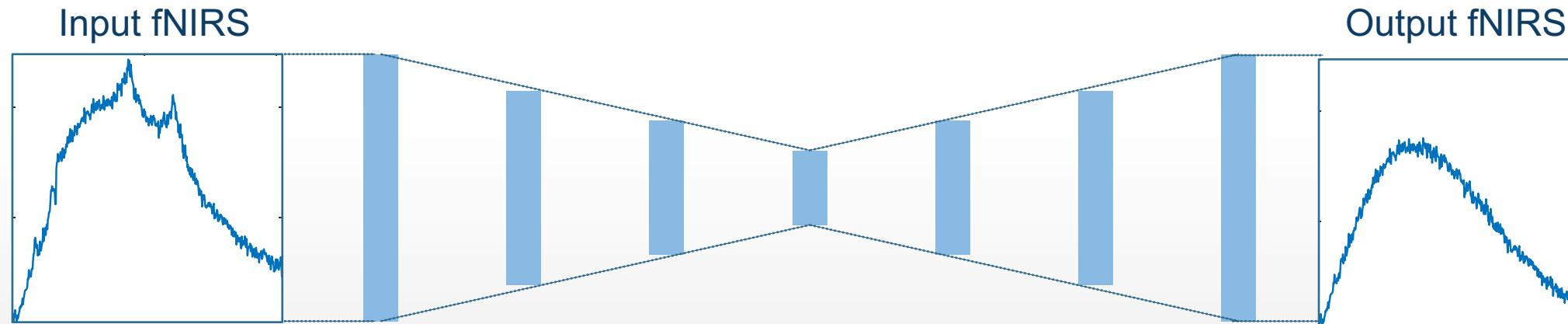
- Small sample size
 - Experimental limitation.
 - Trial discard due to **motion artifacts**.
- fNIRS data processing flow



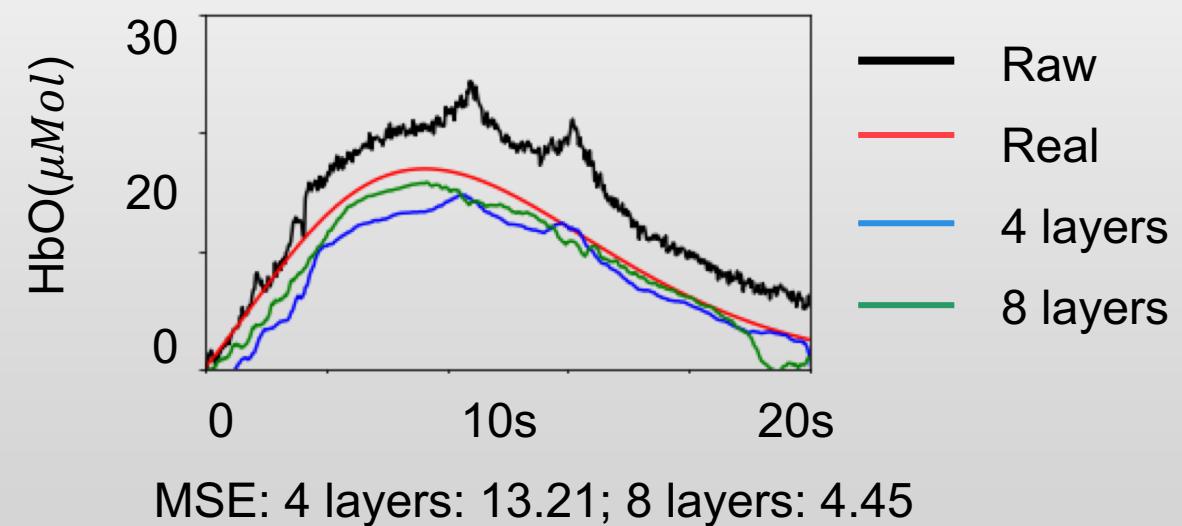
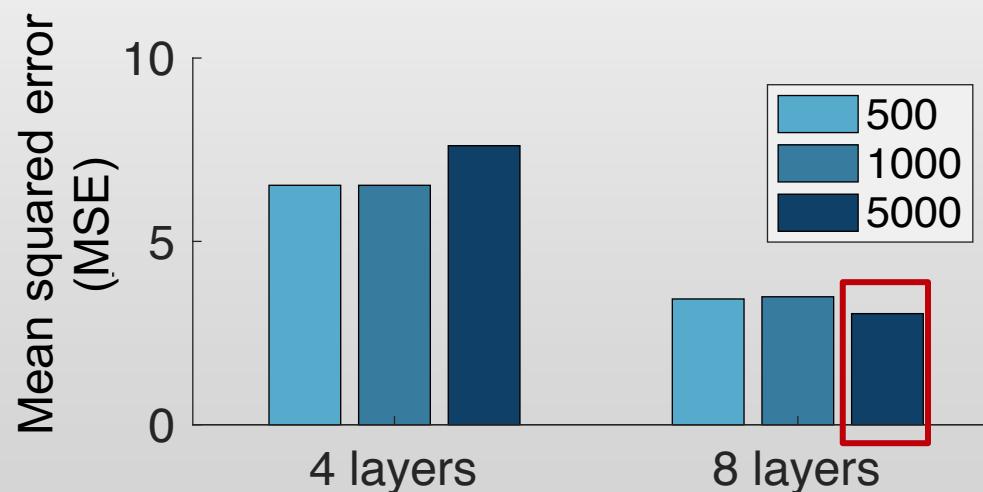
- Motion artifact removal methods in fNIRS
 - Discard
 - Spline (Scholkmann et al, 2010)
 - Kalman (Izzetoglu et al. 2010)
 - Wavelet (Molavi and Dumont 2012)
 - PCA (Zhang et al. 2005)

Proposal

- ‘Denoising autoencoder’ (DAE)



- Preliminary results



Specific aim 2

“.. is to predict FLS scores via neuroimaging”

We established that

- FLS score can be predicted by neuroimaging via deep learning.

Significance:

- Provide a fast, cost-effective method to assess FLS score.

Future work:

- fNIRS motion artifact removal by DAE
 - Train with larger data sample size;
 - Test on experimental data;
 - Comparison with other denoising models;

Expected outcome:

- Provide a model enabling removing motion artifact in fNIRS with high accuracy.

Specific aim 1

“.. is to predict learning curve features in the early stage of training”

Specific aim 2

“.. is to predict FLS scores via neuroimaging”

Specific aim 3

“.. is to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.”

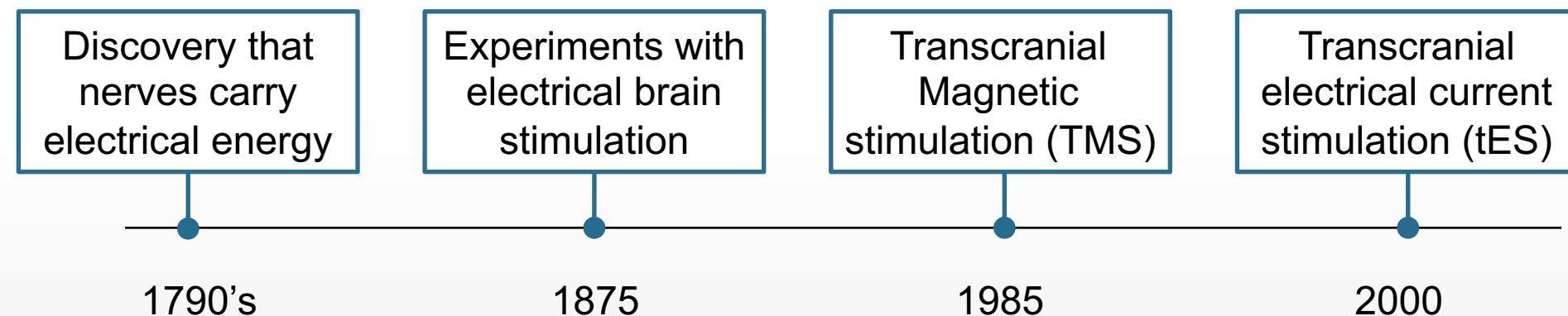
- Current training protocol relies on repetition of tasks.
 - Proficiency-based training protocol requires repetitions until the proficiency is detected (Ritter et al. 2007).
- There is no approaches to enhance the surgical skill training protocol.
- Studies on enhancing surgical learning through actuating neurophysiology are limited.
 - Neuron in cortex and brain stem will descend through motor pathways to control muscles (Bear et al. 2016).
 - Specific brain regions function for motor performance (Hikosaka et al. 2002).
- Neuromodulation could enhance motor skill learning procedure.
 - Basic motor skill learning process was enhanced by neuromodulation (Paulus et al. 2011).

Specific aim 3

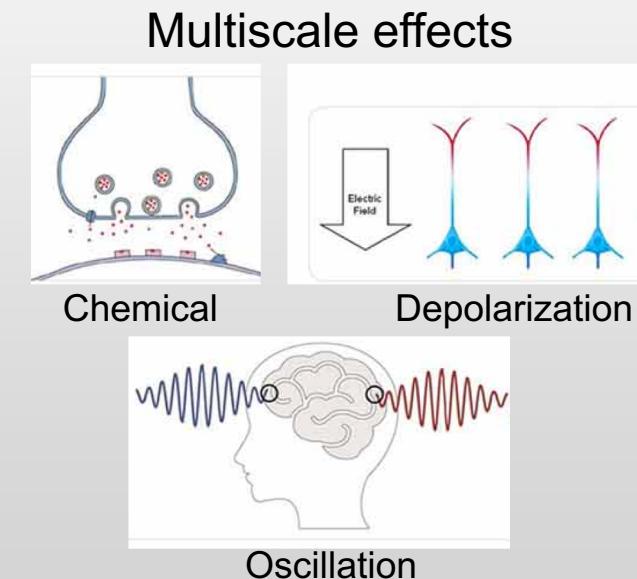
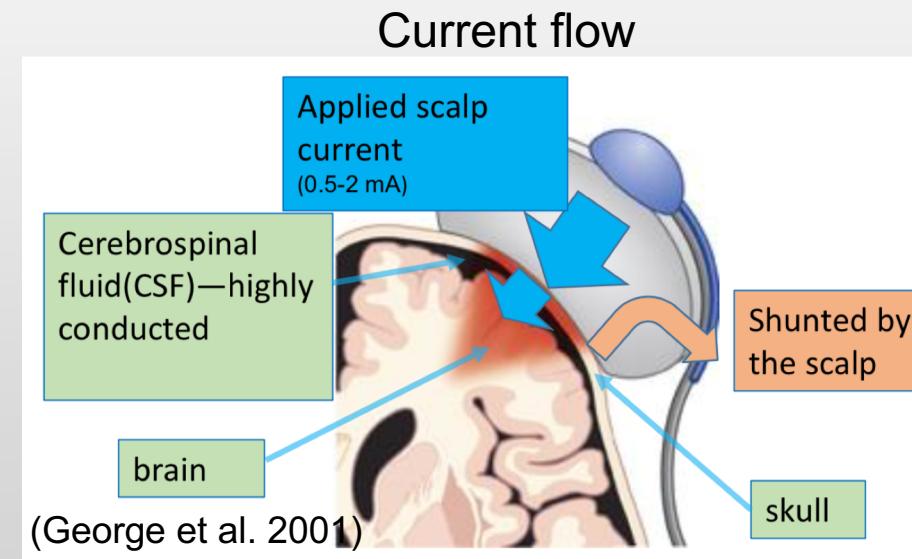
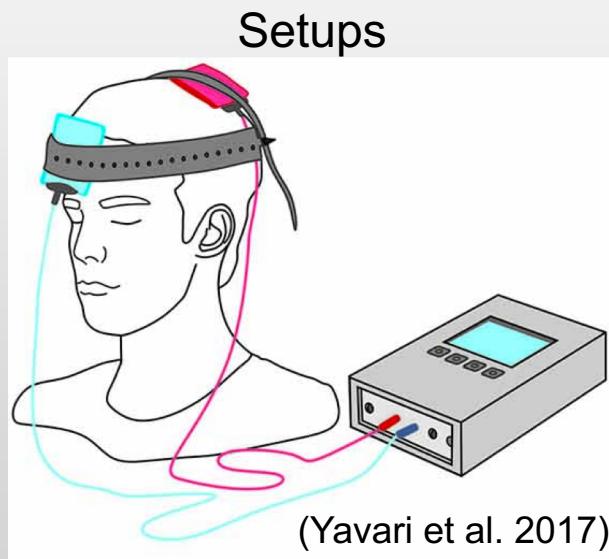
“.. is to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.”

Background

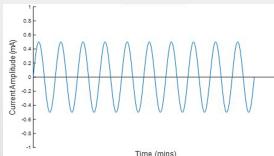
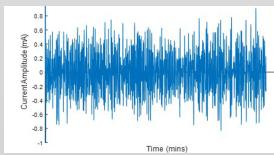
- Noninvasive brain stimulation



- Neuromodulation

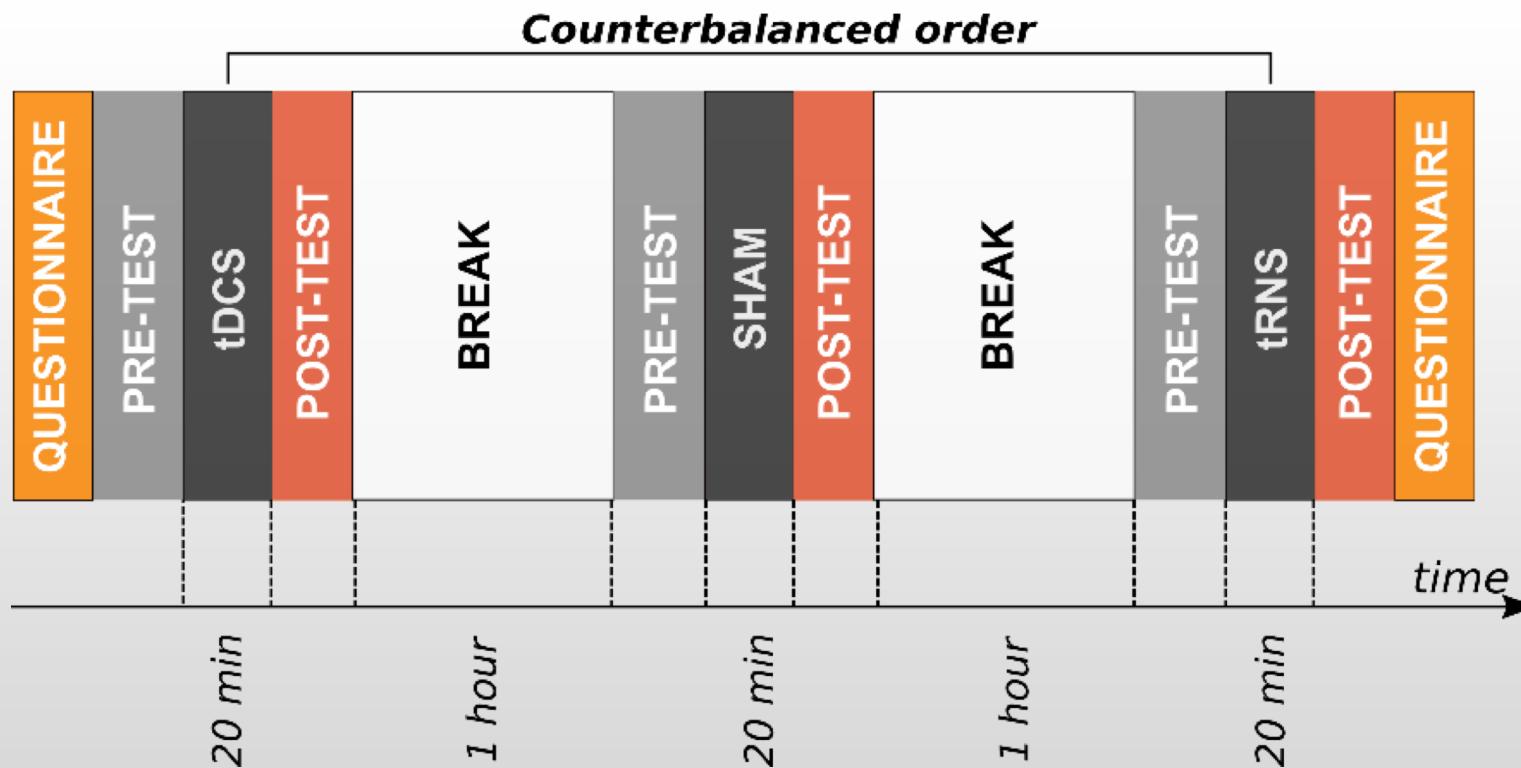


Gap

tES type	Basic motor skill learning	Fine motor skill learning
Transcranial direct current stimulation (tDCS) 	<ul style="list-style-type: none"> Muscle strength Visual-motor Implicit learning Explicit learning <ul style="list-style-type: none"> Adaption Lower limb Rehabilitation 	<ul style="list-style-type: none"> Dexterity Bimanual skill Professional performance
Transcranial alternating current stimulation (tACS) 	<ul style="list-style-type: none"> Muscle strength Visual-motor coordination Implicit learning Rehabilitation 	<ul style="list-style-type: none"> Dexterity Bimanual skill
Transcranial random noise stimulation (tRNS) 	<ul style="list-style-type: none"> Visual-motor coordination Implicit learning Rehabilitation 	(To date, no evidence)

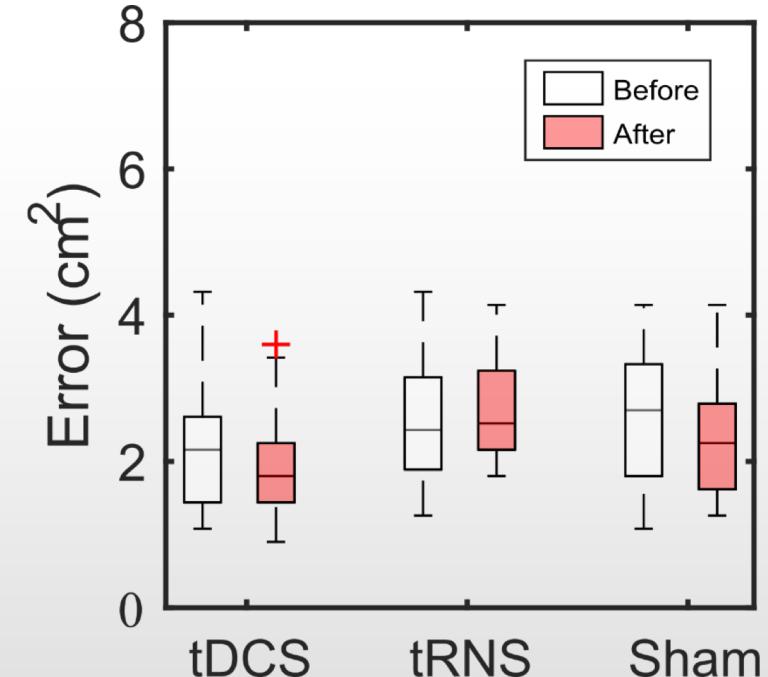
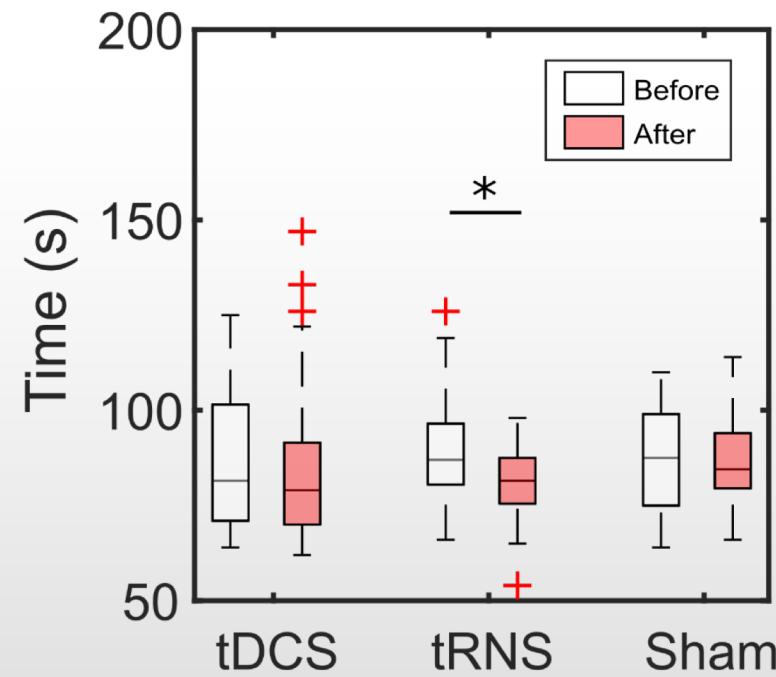
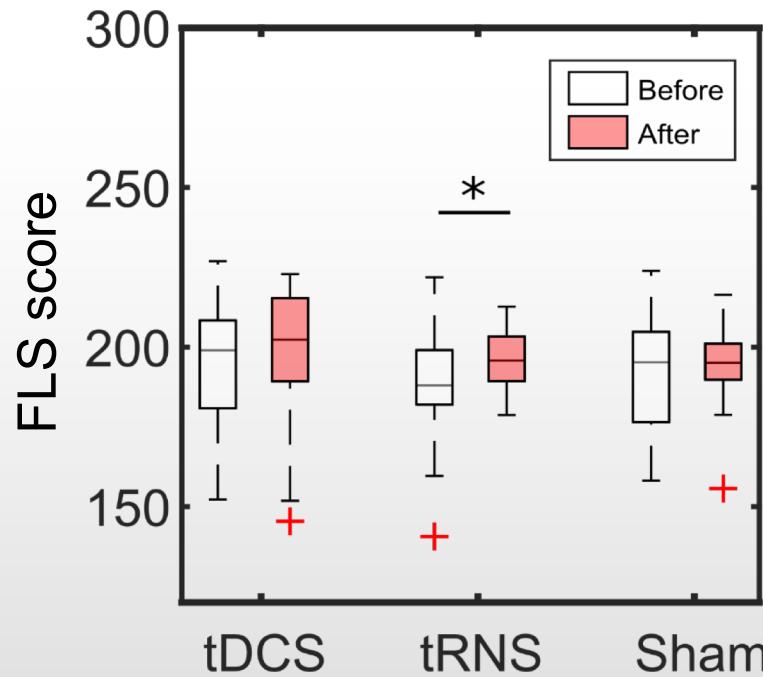
Pilot study

- **Hypothesis:** tES, including tDCS and tRNS will enhance the surgical bimanual task performance.
- **Experimental design:**



Motor task	Pattern cutting
Task repetition	4
N of trainees	6

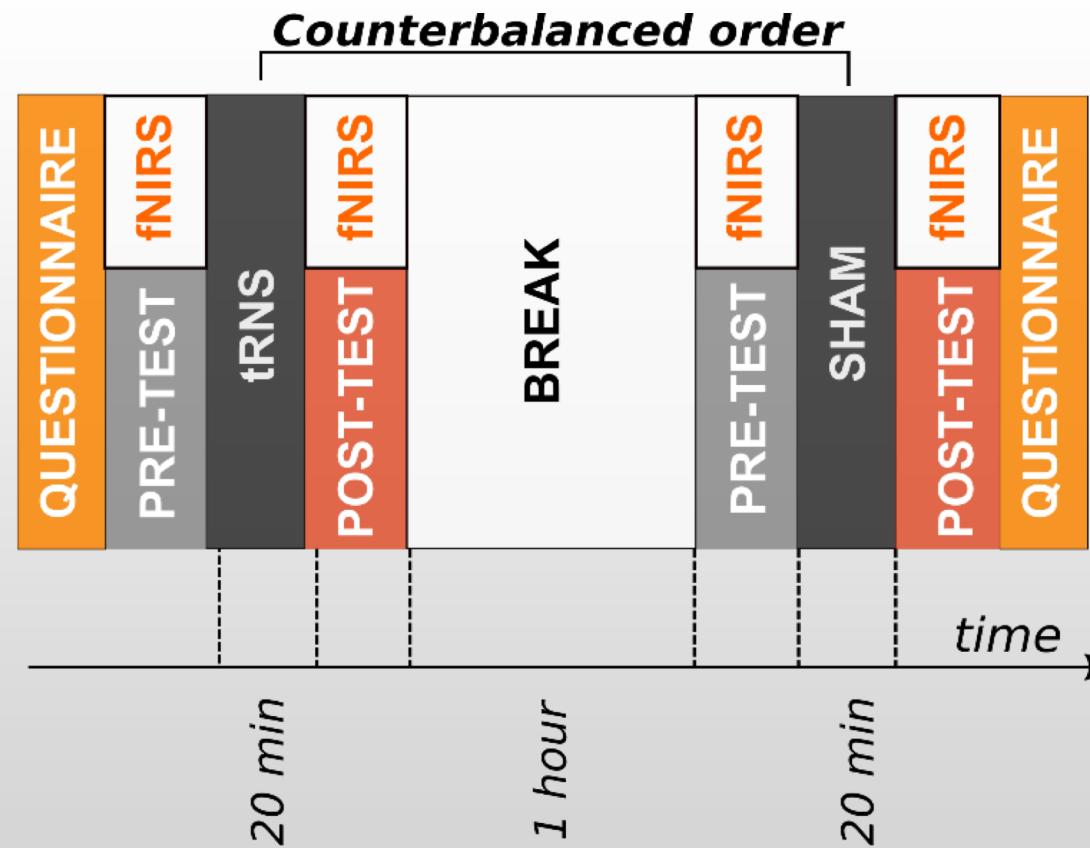
Pilot study



- The FLS score increased significantly under tRNS condition.

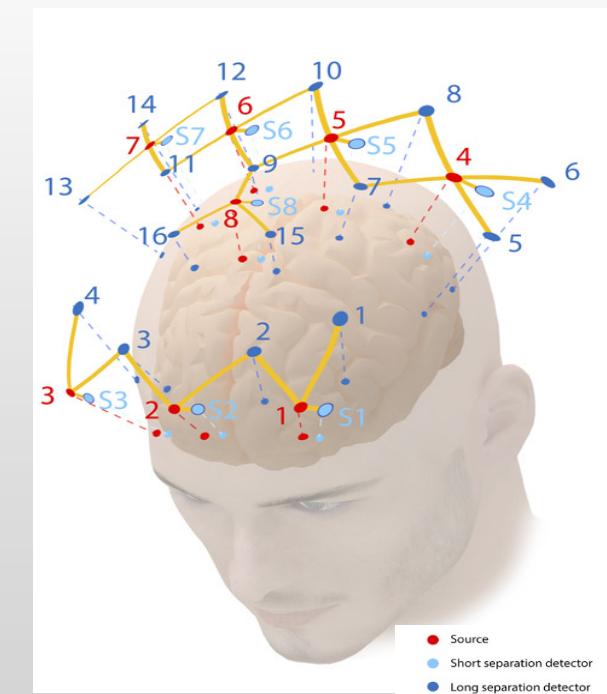
Experiment #1

- **Hypothesis:** Medical students that are subjected to tRNS perform better than those subjected to sham stimulation.
- **Experimental design:**

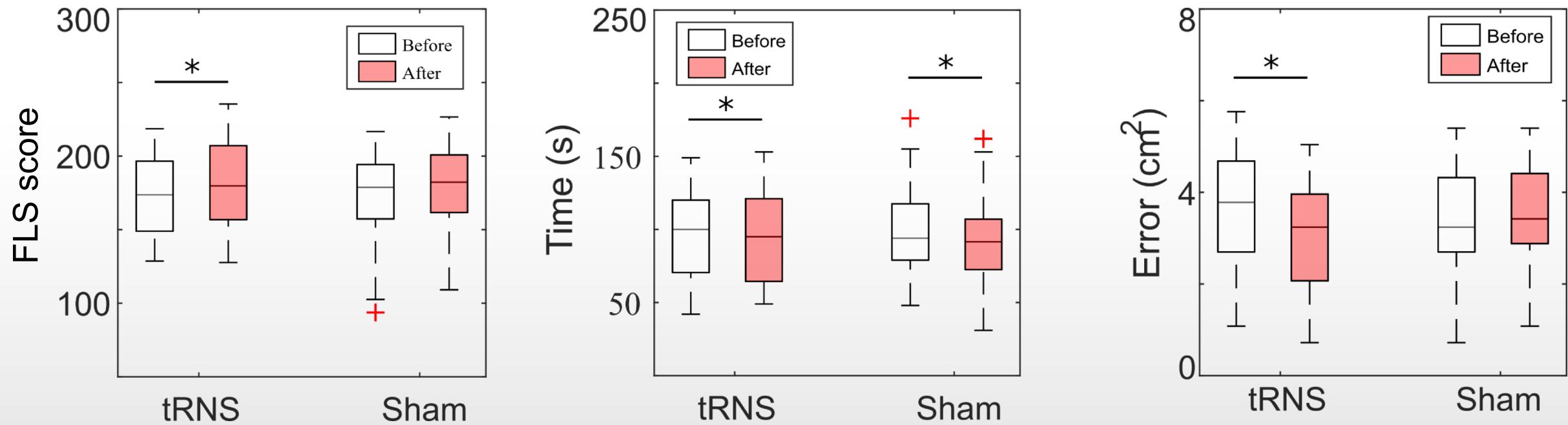


Motor task	Pattern cutting
Task repetition	4
N of trainees	12

Probe geometry:

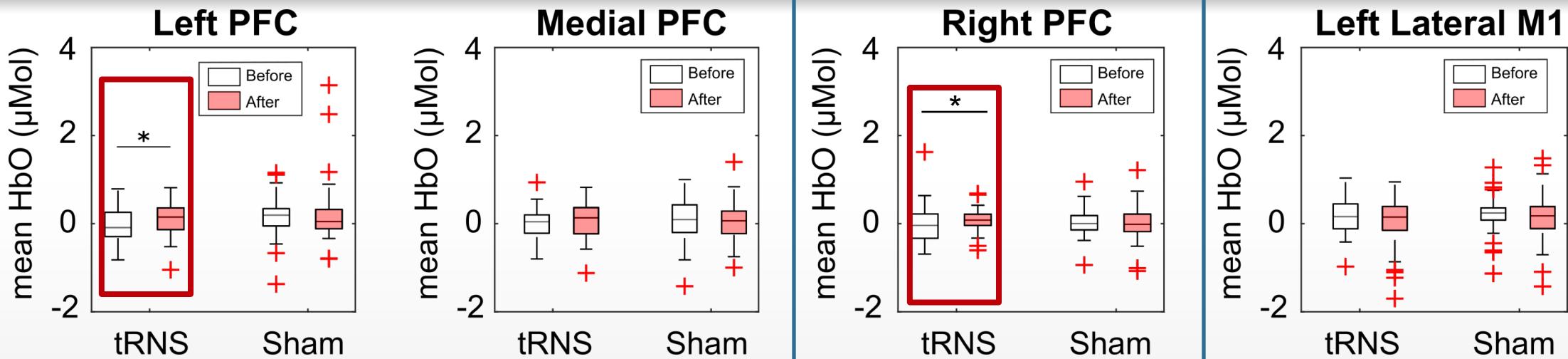


Experiment #1

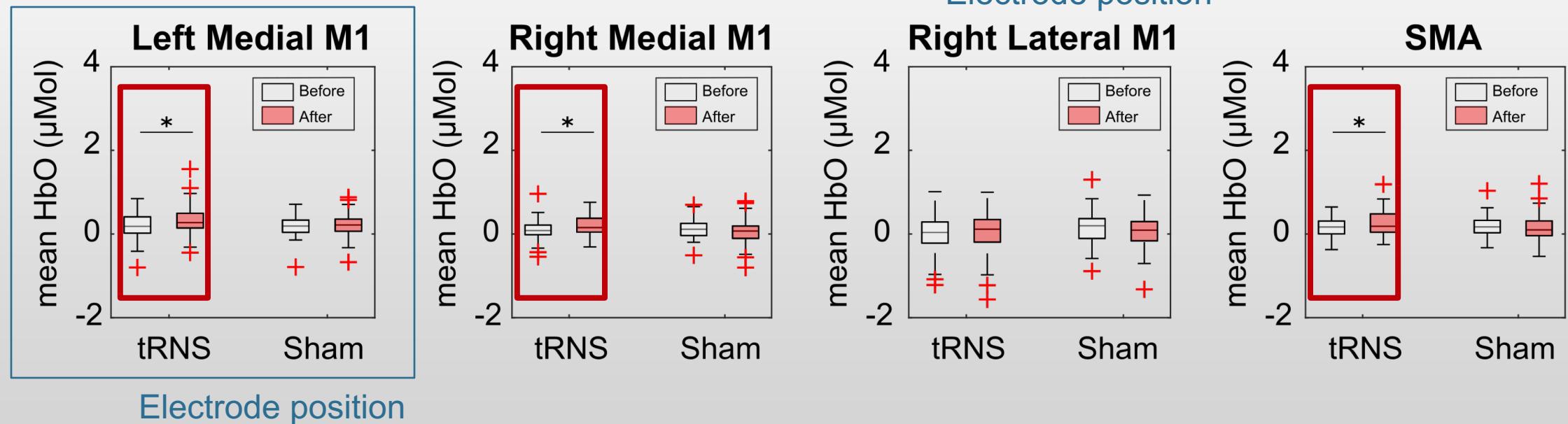


- The FLS score increased under tRNS condition;
- Time decreased for both conditions;
- Error decreased under tRNS condition.

Experiment #1



Electrode position



Electrode position

Specific aim 3 : to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.

Conclusion

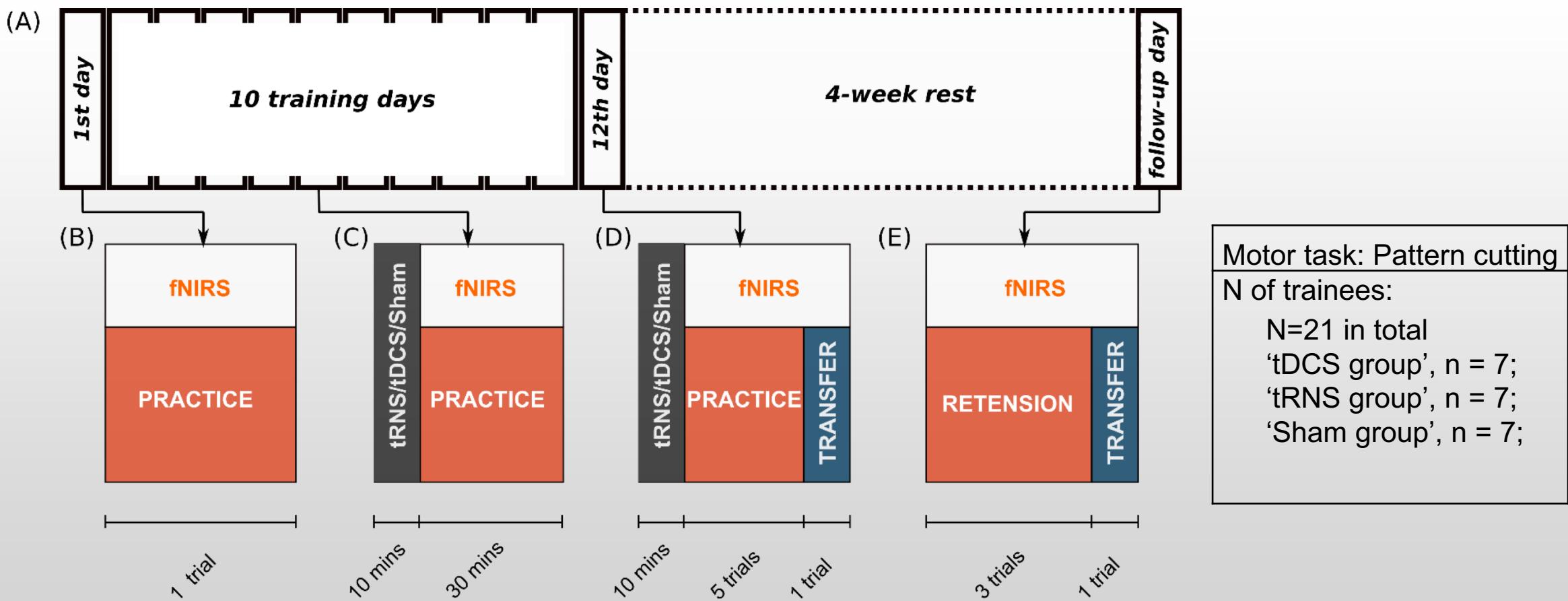
- tRNS enhanced surgical task performance level;
- Correlated HbO level increased in left and right PFC, medial M1 and SMA.

Significance

- tRNS could potentially decrease surgical error, to enhance patient safety;

Experiment #2

- **Hypothesis:** Novice participants subjected to tRNS learn at a faster rate than those subjected to sham stimulation.
- **Experimental design:**



Specific aim 3 : to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.

Future work

- The data analysis of Experiment #2
 - Learning curve data;
 - Learning curve features;
 - fNIRS data;
 - ANOVA test.

Expected outcome

- tDCS and tRNS could accelerate the learning acquisition, retention, and transfer.

Specific aim 1 : to predict learning curve features in the early stage of training. (Completed)

- Learning curve features can be predicted from the initial performance;
- Single factor can represent the learning curve features;

Specific aim 2 : to predict FLS scores via neuroimaging.

- Neuroimaging can predict FLS score; (Completed)
- fNIRS denoise model (To be done)

Specific aim 3 : to investigate whether surgical skill performance, acquisition, retention, and transfer can be enhanced via neuromodulation.

Pilot Study & Experiment #1 (Completed)

- tRNS enhanced surgical skill performance level; (Completed)
- HbO level increased in left and right PFC, medial M1 and SMA. (Completed)

Experiment #2 (Completed)

- FLS score and fNIRS data analysis (To be done)

Timeline

	December				January				February				March				
S.A. #2																	
DAE model																	
Paper manuscript																	
S.A. #3																	
Data processing																	
Paper manuscript																	
Dissertation + Defense																	

Journal papers

Yuanyuan Gao, Uwe Kruger, Steve Schwartzberg, Xavier Intes and Suvranu De, “A Machine Learning approach to predict surgical learning curves”, *Surgery*, accepted in October 2019. (S.A. #1)

Yuanyuan Gao, Pingkun Yan, Uwe Kruger, Suvranu De and Xavier Intes, “Predicting surgical skill levels via functional brain imaging”, *IEEE-TBME, under review* (S.A. #2)

Yuanyuan Gao, Lora Cavuoto, Suvranu De and Xavier Intes, “A comprehensive review of experimental neuroimaging studies of the effect of transcranial electrical stimulation on human motor skills”, *Science Translational Medicine, in preparation.* (review for S.A. #3)

To be done:

1. fNIRS denoise work (S.A. #2)
2. Experiment #1 & #2 (S.A. #3))

Conference papers

Yuanyuan Gao, Pingkun Yan, Uwe Kruger, Suvranu De and Xavier Intes, “Neuroimaging biomarkers for surgical skill level prediction”, *SPIE.bios, San Francisco, CA, February 2019.*

Yuanyuan Gao, Pingkun Yan, Uwe Kruger, Suvranu De and Xavier Intes, “fNIRS as a quantitative tool to asses and predict surgical skills”, *OSA Biophotonics Congress: Optics in the Life Sciences, Tucson, AZ, April 2019.*

Conference papers

Yuanyuan Gao, Lora Cavuoto, L., Pingkun Yan, Uwe Kruger, Steven Schwartzberg, Suvranu De, and Xavier Intes, “A deep learning approach to remove motion artifacts in fNIRS data analysis”. *In Optics and the Brain, Optical Society of America, 2020, submitted*

Yuanyuan Gao, Lora Cavuoto, L., Pingkun Yan, Uwe Kruger, Jessica Silvestri, Steven Schwartzberg, Xavier Intes and Suvranu De, “Monitoring the effect of transcranial Electric current Stimulation (tES) during a bimanual motor task via functional Near-InfraRed Spectroscopy (fNIRS)”. *In Optics and the Brain, Optical Society of America, 2020, submitted*

Yuanyuan Gao, Pingkun Yan, Uwe Kruger, Lora Cavuoto, Steven Schwartzberg, Suvranu De and Xavier Intes, “Deep neural network evaluation of surgical skills”, *American Surgical Association Annual 14th meeting (2020), submitted.*

Thank you!

In this study, we adopted a kernel PLS algorithm to model the non-linear relationship between predictor X ($n \times N$) and y ($n \times 1$). First, the standard PLS algorithm steps in this study are listed from step (1) to (8), as shown below:

- (1) $w = X^T y;$
- (2) $t = Xw;$
- (3) $q = \frac{t^T y}{t^T t};$
- (4) $p = \frac{x^T t}{t^T t};$
- (5) $X = X - tp^T;$
- (6) $y = y - tq^T;$
- (7) *repeat step (1) to (6) until converge*
- (8) $B = W(P^T W)^{-1} Q^T;$

KPLS algorithm is a non-linear counterpart of standard PLS algorithm. KPLS is able to model the non-linear relationship by mapping the predictor into a high-dimensional non-linear space, $x_i \rightarrow \Phi(x_i)$. We adopted radial basis kernel $k(x, y) = \exp(-\frac{\|x-y\|^2}{c})$ as the kernel Gram matrix $K = \Phi\Phi^T$ in KPLS algorithm.

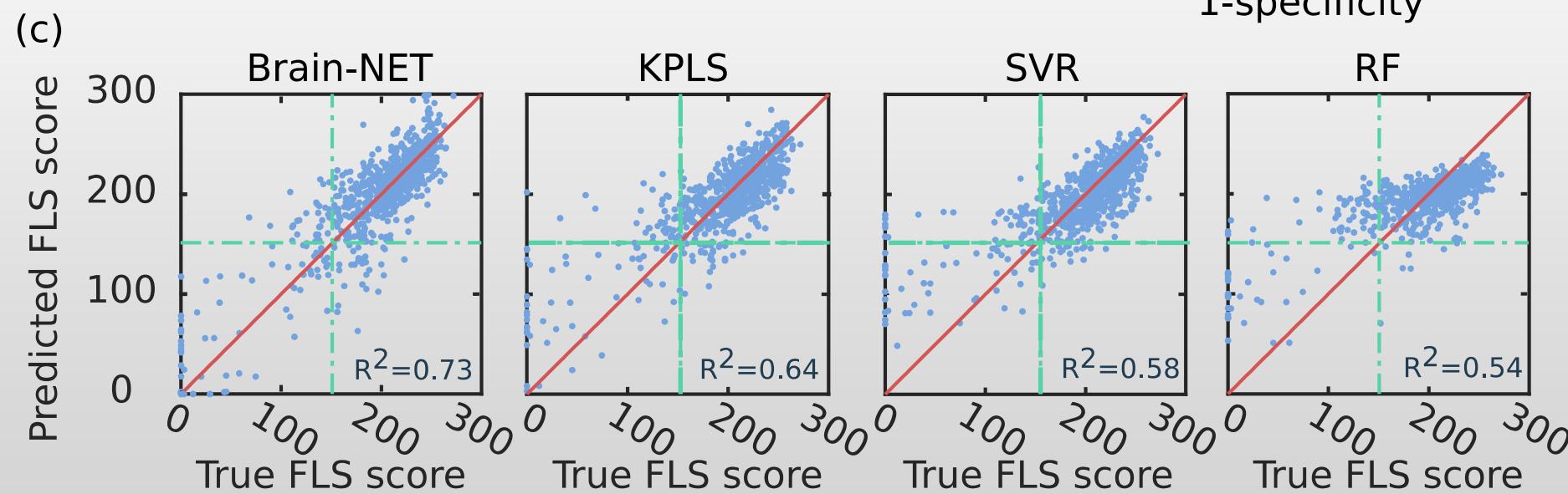
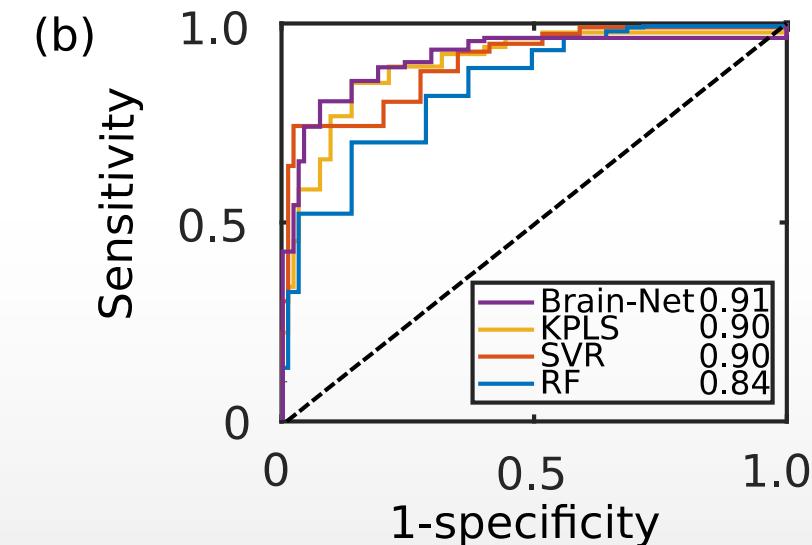
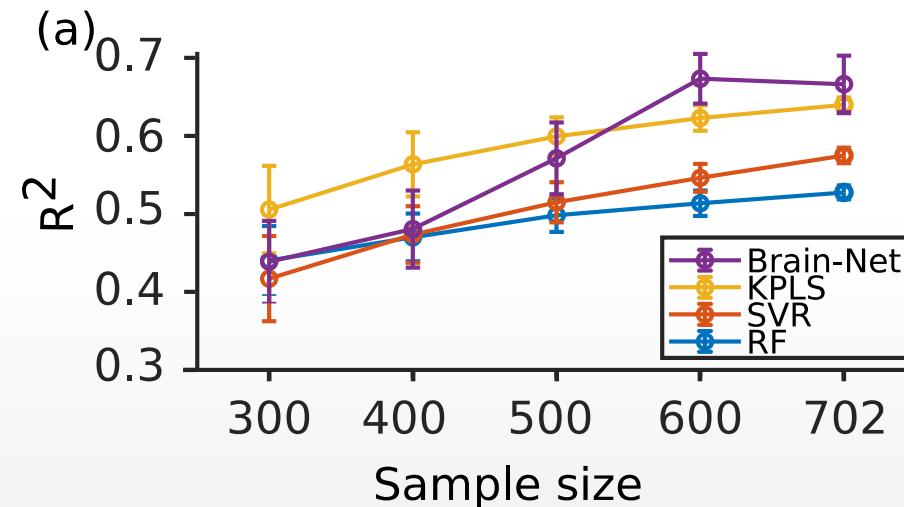
As a non-linear counterpart of PCA, the Kernel principle component analysis (KPCA) extracts primary component in a non-linear sense. We adopted singular value decomposition method to process PCA in this paper. The steps of KPCA are shown below:

1. As we did in KPLS, projected $x_i \rightarrow \varphi(x_i)$
2. Define a kernel matrix $K = \varphi\varphi^T$ (Here we use $K(x, y) = \exp(-\frac{\|x-y\|^2}{c})$)
3. Centered Gram matrix $G = K - \frac{1}{N}(KE_N)\mathbf{1}_N^T - \frac{1}{N}E_N(KE_N)^T + \frac{1}{N^2}\mathbf{1}_N(\mathbf{1}_N^T KE_N)E_N^T$.
4. eigendecomposition of G, the KPCA score vector t is calculated as

$$t = U^T \bar{\varphi} = A^T \varphi^T \left(\varphi - \frac{1}{N} \varphi E_N \right) = A^T (K(x, x') - \frac{1}{N} KE_N)$$

Where $\bar{\varphi}$ is centered φ , $A = \left[I - \frac{1}{N} E_N \right] V$, U and V represent the n largest eigenvalues and corresponding eigenvectors.

1. Choose the number of clusters(k) and obtain the data points
2. Place the centroids c_1, c_2, \dots, c_k randomly
3. Repeat steps 4 and 5 until convergence or until the end of a fixed number of iterations
4. For each data point x_i :
 - Find the nearest centroid (c_1, c_2, \dots, c_k) by Euclidean distance
 - Assign the point to that cluster
5. For each cluster $j = 1..k$
 - New centroid = mean of all points assigned to that cluster
6. End



Model: 4layers

Layer	Output Shape
Input	(512,1)
Conv1D	(512,32)
MaxPooling1D	(256,32)
Conv1D	(256,32)
MaxPooling1D	(128,32)
Conv1D	(128,32)
UpSampling1D	(256,32)
Conv1D	(256,32)
UpSampling1D	(512,32)
Conv1D	(512,1)

Total params: 9,793

Learning rate

0.00001, decay 0.5 per 100 epochs

Epochs

1000 (save the best)

Batch size

32

Loss

Mean squared error

Optimizer

Adam

Model: 4layers+dropout

Layer	Output Shape
Input	(512,1)
Conv1D	(512,32)
MaxPooling1D +dropout(0.1)	(256,32)
Conv1D	(256,32)
MaxPooling1D+ dropout(0.1)	(128,32)
Conv1D	(128,32)
UpSampling1D+	
Conv1D	(128,32)
UpSampling1D+dropout(0.1)	(256,32)
Conv1D	(256,32)
UpSampling1D+dropout(0.1)	(512,32)
Conv1D	(512,1)

Total params: 9,793

Model: 8layers

Layer	Output Shape
Input	(512,1)
Conv1D	(512,32)
MaxPooling1D +dropout(0.1)	(256,32)
Conv1D	(256,32)
MaxPooling1D	(128,32)
Conv1D	(128,32)
MaxPooling1D	(64,32)
Conv1D	(64,32)
MaxPooling1D	(32,32)
Conv1D	(32,32)
UpSampling1D	(64,32)
Conv1D	(64,32)
UpSampling1D	(128,32)
Conv1D	(128,32)
UpSampling1D	(256,32)
Conv1D	(256,32)
UpSampling1D	(512,32)
Conv1D	(512,1)

Total params: 30,401

Model: 8layers+dropout

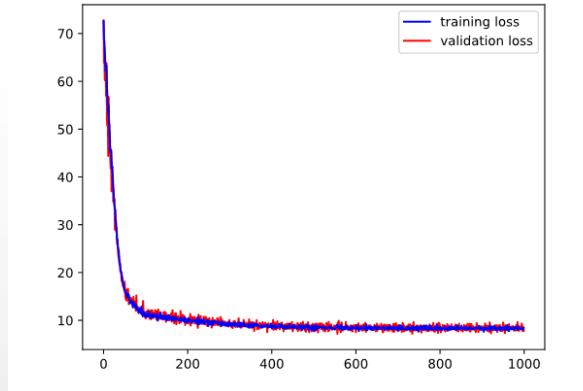
Layer	Output Shape
Input	(512,1)
Conv1D	(512,32)
MaxPooling1D +dropout(0.1)	(256,32)
Conv1D	(256,32)
MaxPooling1D +dropout(0.1)	(128,32)
Conv1D	(128,32)
MaxPooling1D +dropout(0.1)	(64,32)
Conv1D	(64,32)
MaxPooling1D +dropout(0.1)	(32,32)
Conv1D	(32,32)
UpSampling1D +dropout(0.1)	(64,32)
Conv1D	(64,32)
UpSampling1D	(128,32)
Conv1D	(128,32)
UpSampling1D +dropout(0.1)	(256,32)
Conv1D	(256,32)
UpSampling1D +dropout(0.1)	(512,32)
Conv1D	(512,1)

Total params: 30,401

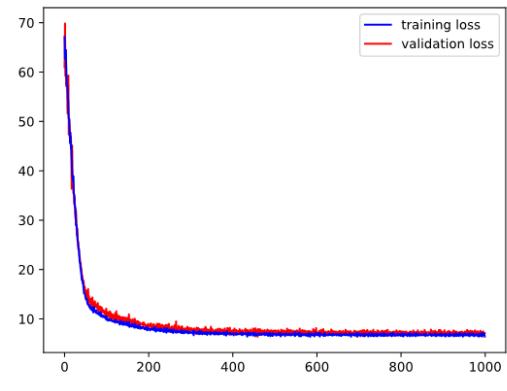
MSE loss

4 layers

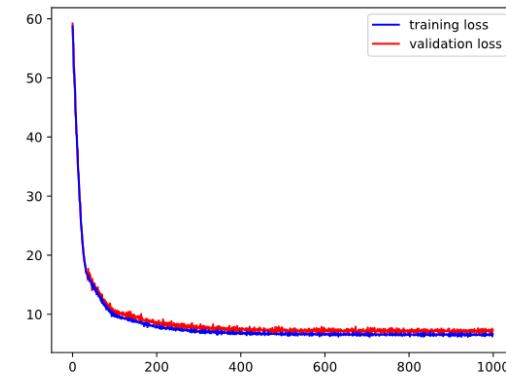
5000



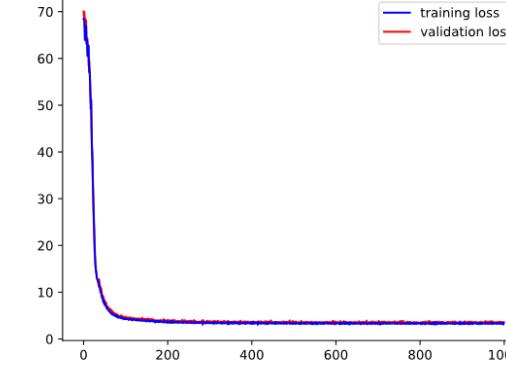
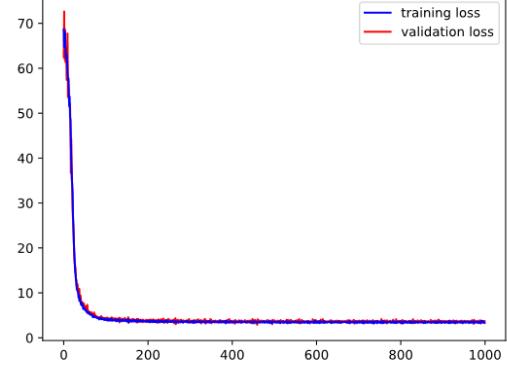
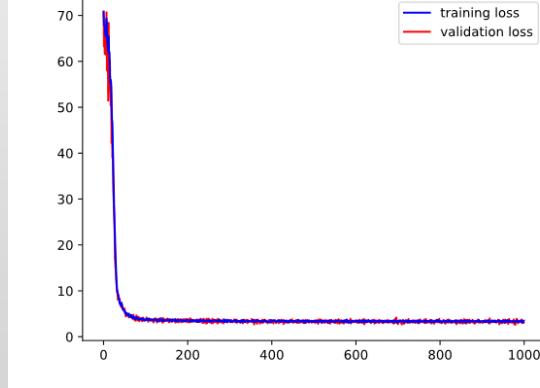
1000



500



8 layers



Shapiro-Wilk normality test results

Pilot Study

p value	tDCS	tRNS	Sham
Score	0.900	0.015	0.180
Time	0.847	0.023	0.187
Error	0.551	0.241	0.216

Study #1

p value	tRNS	Sham
Score	0.005	0.134
Time	0.816	0.065
Error	0.018	0.495

Shapiro-Wilk normality test results

Study #1

p value	tRNS	Sham
LPFC	0.000	0.272
MPFC	0.021	0.795
RPFC	0.000	0.850
LLM1	0.000	0.055
LMM1	0.000	0.018
RMM1	0.052	0.002
RLM1	0.009	0.751
SMA	0.127	0.002