Graduate Institute of Electronics Engineering, NTU

DSP Group

Human-Computer Interaction (HCI)

EMG-based Hand Gesture Recognition

Presenter: Howard

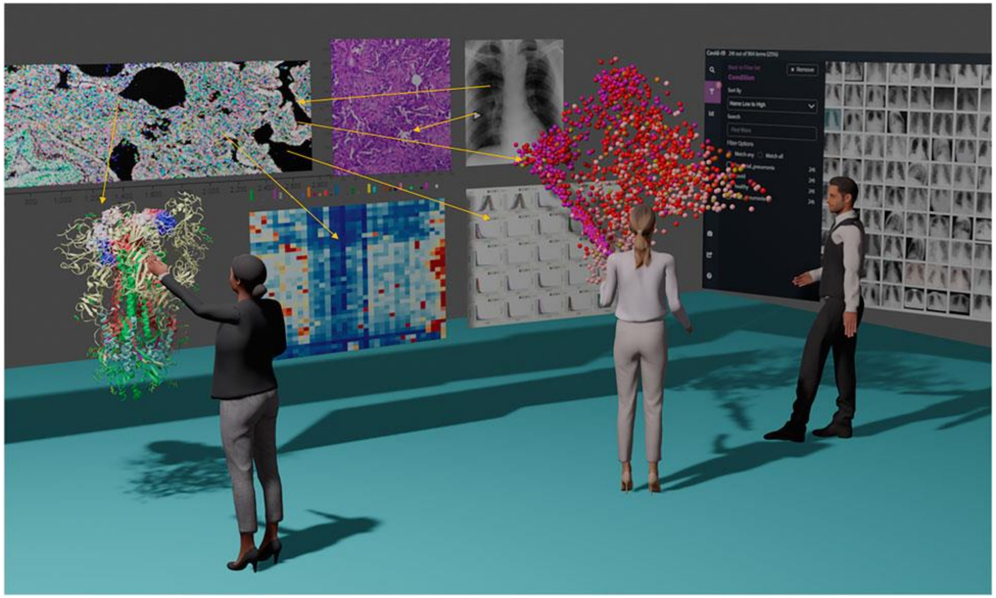
Teammates: Shawn,Miguel

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/12/4

ACCESS IC LAB

Human Computer Interaction (HCI)



***Meta** Quest Pro*

***Apple** Vision Pro*

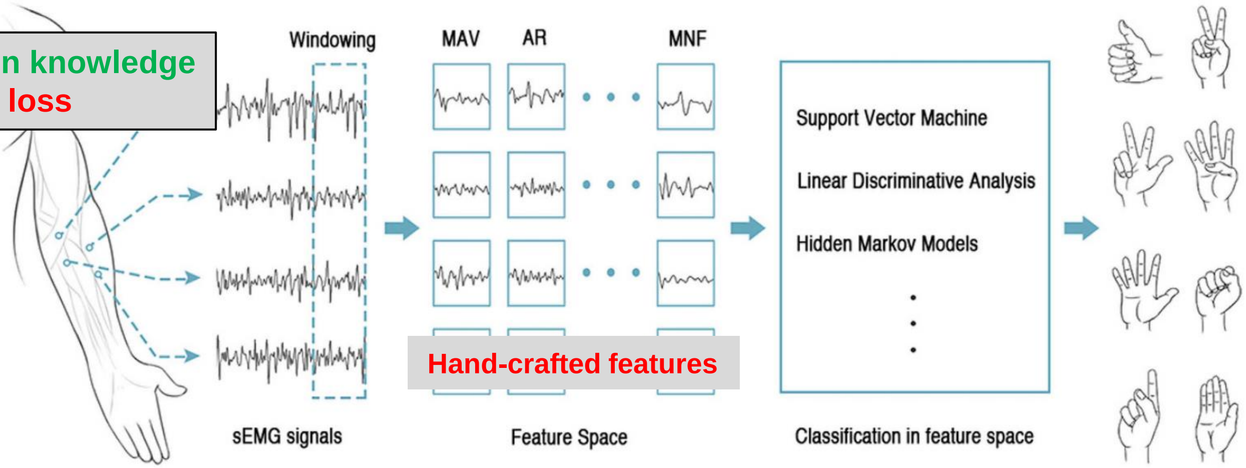
- Mega trend: **Metaverse (AR/VR)**
  - Immersive experience
  - Natural and intuitive interaction

Hand gesture is one of the most intuitive **interface**

EMG-based HGR Processing Flow

Feature-based: *Pre-processing + feature extractor + classifier*

- Prior domain knowledge
- Information loss



Hand-crafted features

DL-based: *End-to-end learning*

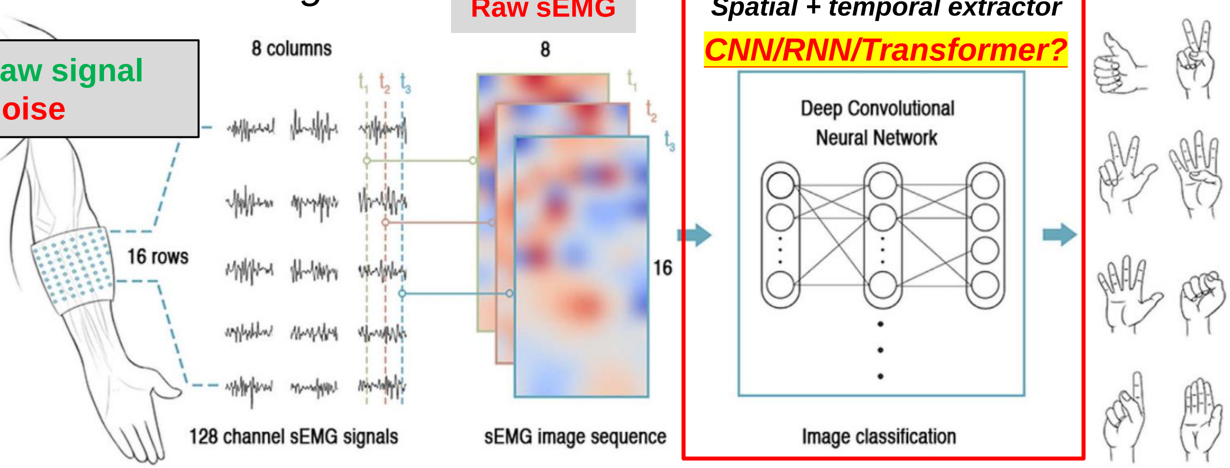
< 90% over 17 gestures

- Informative raw signal
- Redundant noise

Raw sEMG

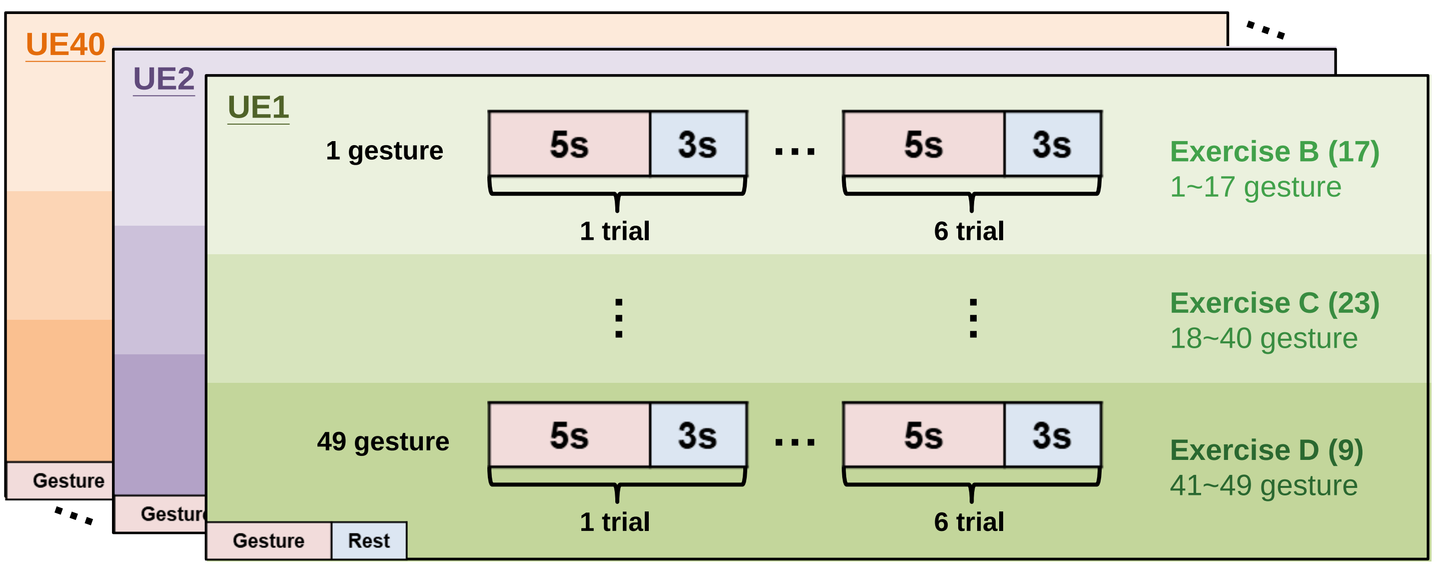
Spatial + temporal extractor

CNN/RNN/Transformer?



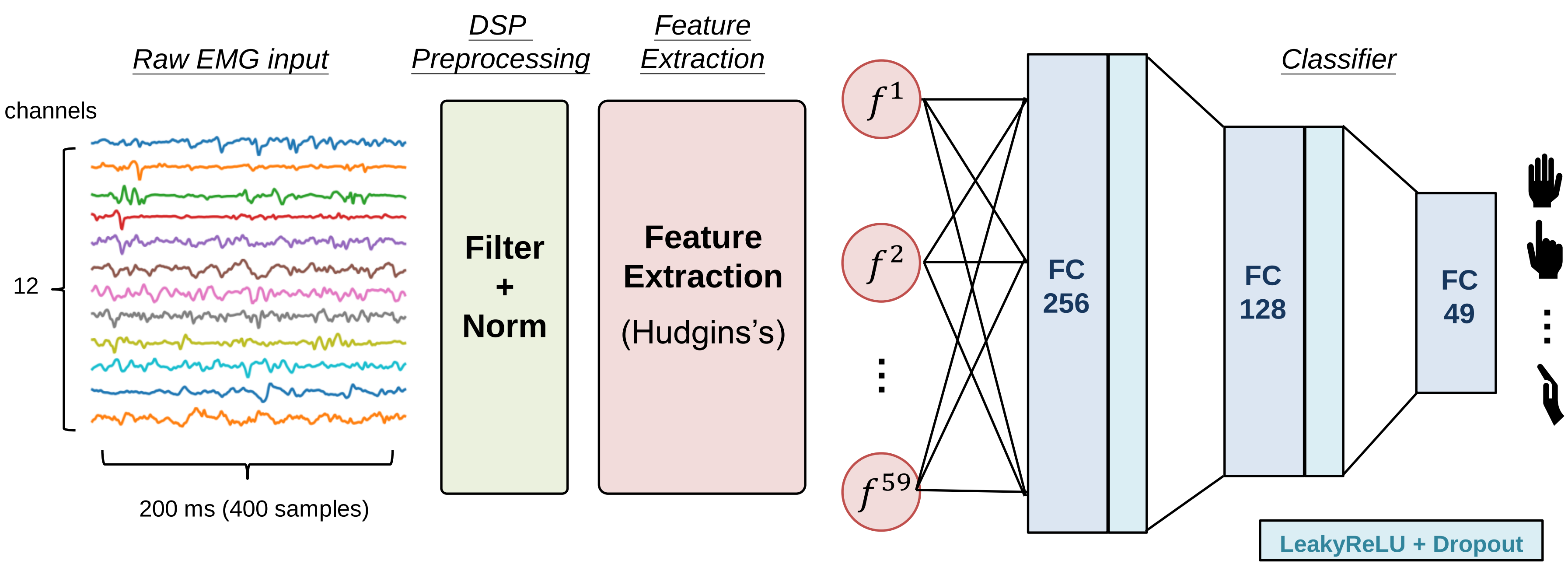
Open Dataset: NinaPro DB2

- **Device**: Delsys, 12 channels, 2048 sps
- **Setup**: 40 users / 49 gestures (B,C,D) / 6 trials
  - Train: trial 1,3,4,6
  - Valid: 1/2/3/4-th quarter of trial 1/3/4/6
  - Test : trial 2,5



DNN with Feature Extraction

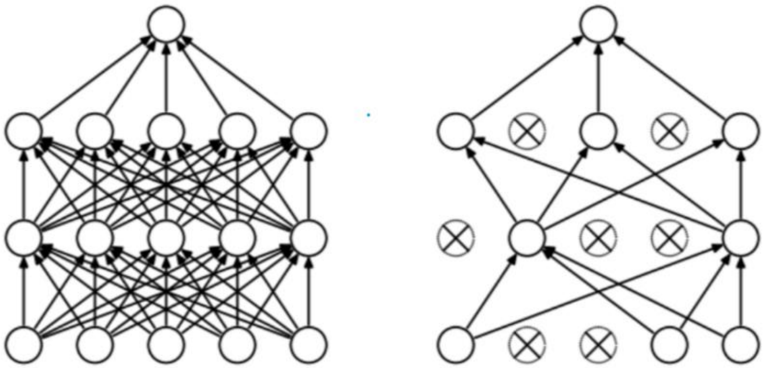
- **Filter**: 1-st order Butterworth filter LPF or BPF
- **Norm**: min-max, z-score, μ-law
- **Hudgins's** time domain feature set [ref]
  - MAV (12), MAVS (11), WL (12), SSC (12) , ZC (12) → 59 in totals



MAV: Mean absolute value, MAVS: MAV slope, WL: waveform length, SSC: slope sign change, ZC: zero crossing

Simulation Results – Dropout

- Complex Neural Networks + inadequate data → overfitting
- Dropout Layer !
- Experiment method
  - Control variable : Filter/Norm
  - experimental variable : dropout rate



(a) Standard Neural Net (b) After applying dropout.

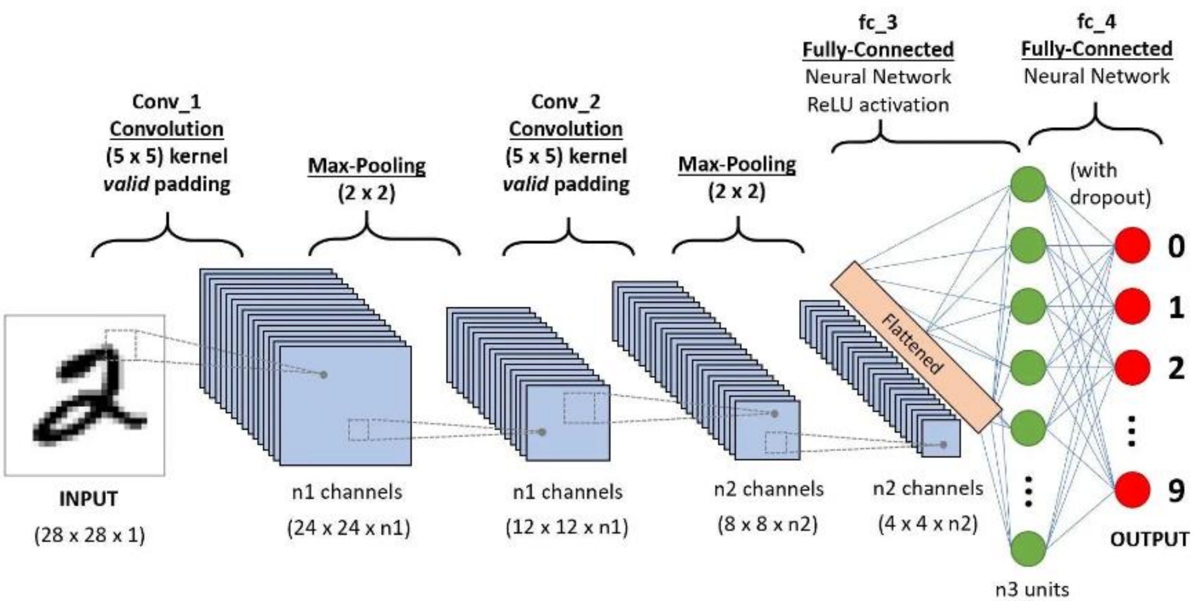
Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
window step	100 ms
Filter	none
Norm	Z-score

Dropout	AVG
0.0	82.1
0.2	84.0
0.4	84.2
0.5	83.9
0.6	82.4
0.8	72.5

Dropout rate = 0.4 → Best performance

CNN

- **CNN** is efficient to capture the temporal and spatial correlations
  - Convolution → Learned filters
  - Deeper CNN has higher cost (9431MiB CUDA out of memory ☹)




```python
class CNN(nn.Module):
    def __init__(self, number_gesture=49, class_rest=False, dropout=0.4):
        super(CNN, self).__init__()
        output_class = number_gesture + int(class_rest==True)
        self.layers = nn.Sequential(
            nn.BatchNorm2d(1),
            nn.Conv2d(1,64, kernel_size=(3,3), stride=(1,1), padding='same'),
            nn.ReLU(),
            nn.MaxPool2d((4,1)),
            # nn.BatchNorm2d(32),
            nn.Conv2d(64,64, kernel_size=(3,3), stride=(1,1), padding='same'),
            nn.ReLU(),
            nn.MaxPool2d((4,1)),
            # nn.BatchNorm2d(16),
            nn.Conv2d(64,8 , kernel_size=(3,3), stride=(1,1), padding='same'),
            nn.ReLU(),
            nn.MaxPool2d((2,2)),
            nn.Flatten(),
            nn.Dropout(dropout),
            nn.Linear(576,128),
            nn.ReLU(),
            nn.Dropout(dropout),
            nn.Linear(128, output_class)
            # nn.Linear(64, output_class)
        )
        self.conv = nn.Conv2d(1, 16, kernel_size=(3,3), stride=(1,1), padding='same')
        self.maxpool = nn.MaxPool2d((4,1))
```

Simulation Results – Norm

Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
window step	100 ms
Dropout	0.4

- Experiment method
  - Control variable : Dropout
  - Experimental variable : Filter / Norm
- Epoch : significant factor

Filter \ Norm	none	Min-max	z-score	μ-law (256)	μ -law (2048)	AVG
none	76.9	38.1	76.9	*69.8*	*70.3*	66.4
BPF [20,200]	*67.6*	27.1	74.8	68.3	75.9	66.1
none	-	-	80.9	76.5	77.5	77.8
none	-	-	83.2	77.1	77.6	78.7
LPF [1,]	-	-	10.7	41.0	45.8	33.9
LPF [10,]	-	-	*50.1*	60.3	70.1	62.7
LPF [20,]	-	-	60.9	72.1	72.3	69.2
AVG	72.3	32.6	63.6	68.6	71.0	-

Filter = none , Norm = Z-score → Best performance

Simulation Results – window step

- Experiment method
  - Control variable : epoch, dropout, Filter/Norm
  - Experimental variable : window step

Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
Dropout	0.4

Transform\window step	5ms	10ms
Tnet	Elapsed time = 9044.59 Acc = 64.7	Elapsed time = 18520.75 Acc = 66.5
Fnet	Elapsed time = Acc =	Elapsed time = Acc =

10ms→ Best performance

Optimizer

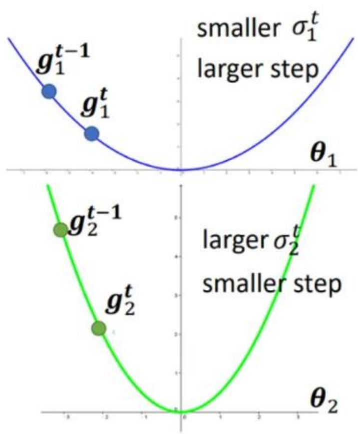
- Adagrad - Adaptive gradient algorithm
  - adjust the learning rate by past changes in each weight
  - more changes → learning rate smaller
- RMSprop - Adaptive gradient algorithm
  - adjust the learning rate by α (Gradient)
- Adam
  - RMSprop + Momentum

Root Mean Square

$$\boldsymbol{\theta}_i^{t+1} \leftarrow \boldsymbol{\theta}_i^t - \boxed{\frac{\eta}{\sigma_i^t}} \boldsymbol{g}_i^t$$

$$\sigma_i^t = \sqrt{\frac{1}{t+1}\sum_{i=0}^{t}\left(\boldsymbol{g}_i^t\right)^2}$$

Used in **Adagrad**



RMSProp

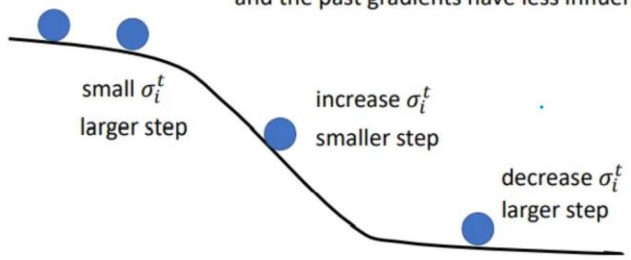
$$\boldsymbol{\theta}_i^{t+1} \leftarrow \boldsymbol{\theta}_i^t - \boxed{\frac{\eta}{\sigma_i^t}} \boldsymbol{g}_i^t$$

$\boldsymbol{g}_i^1 \; \boldsymbol{g}_i^2 \; \cdots\cdots \; \boldsymbol{g}_i^{t-1}$

$0 < \alpha < 1$

$$\sigma_i^t = \sqrt{\alpha\left(\sigma_i^{t-1}\right)^2 + (1-\alpha)\left(\boldsymbol{g}_i^t\right)^2}$$

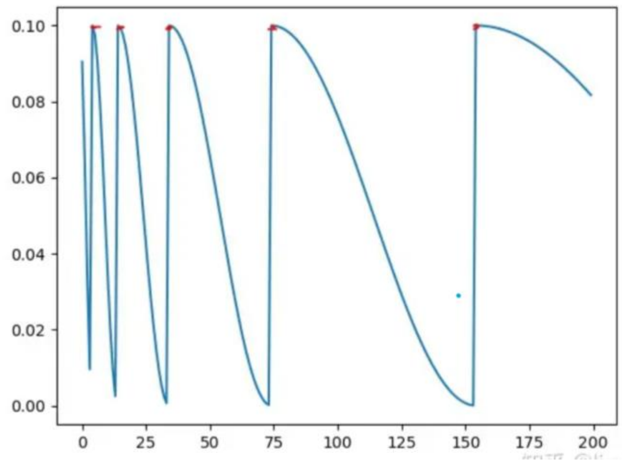
The recent gradient has larger influence, and the past gradients have less influence.

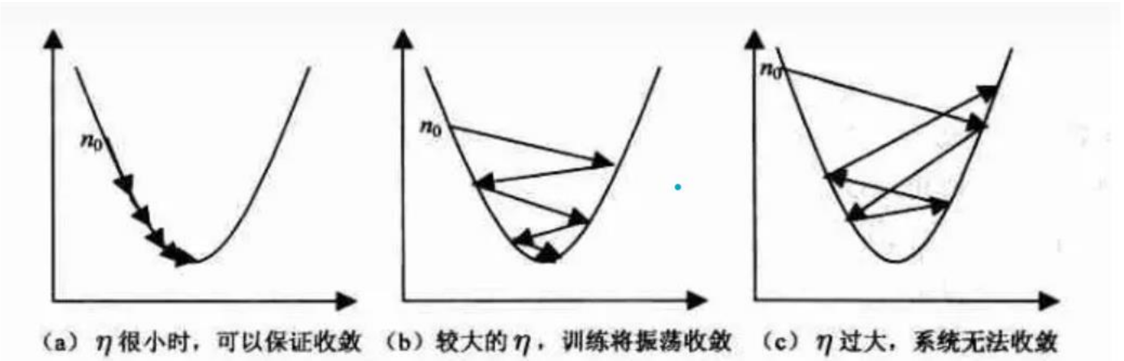


Scheduler - CosineAnnealingRestarts

- T_0 : the first time learning rate back to initial
- T_mult : control the speed of lr back to initial
- If T_mult=1, lr at T_0, 2*T_0, 3*T_0,...., i*T_0 back to initial
- If T_mult > 1, lr at T_0, (1+T_mult)*T_0, (1+T_mult+T_mult**2)T_0 ,.....,(1+T_mult+T_mult2+...+T_0i)*T0,

```
scheduler = torch.optim.lr_scheduler.CosineAnnealingWarmRestarts(
    optimizer, T_0=25, T_mult=2, eta_min=1e-5, verbose=False)
```





Simulation Results – learning rate

- Experiment method
  - Control variable : epoch, dropout, Filter/Norm
  - Experimental variable : Optimizer, lr

Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
window step	5 ms
Dropout	0.4

- Tnet

Scheduler \ Lr	0.001	0.0001
Y	Elapsed time = 23541.23 Acc = 66.5	Elapsed time = 59062.08 Acc = 68.7
N	Elapsed time = 29738.66 Acc = 67.1	Elapsed time = 61856.19 Acc = 67.2

- Fnet

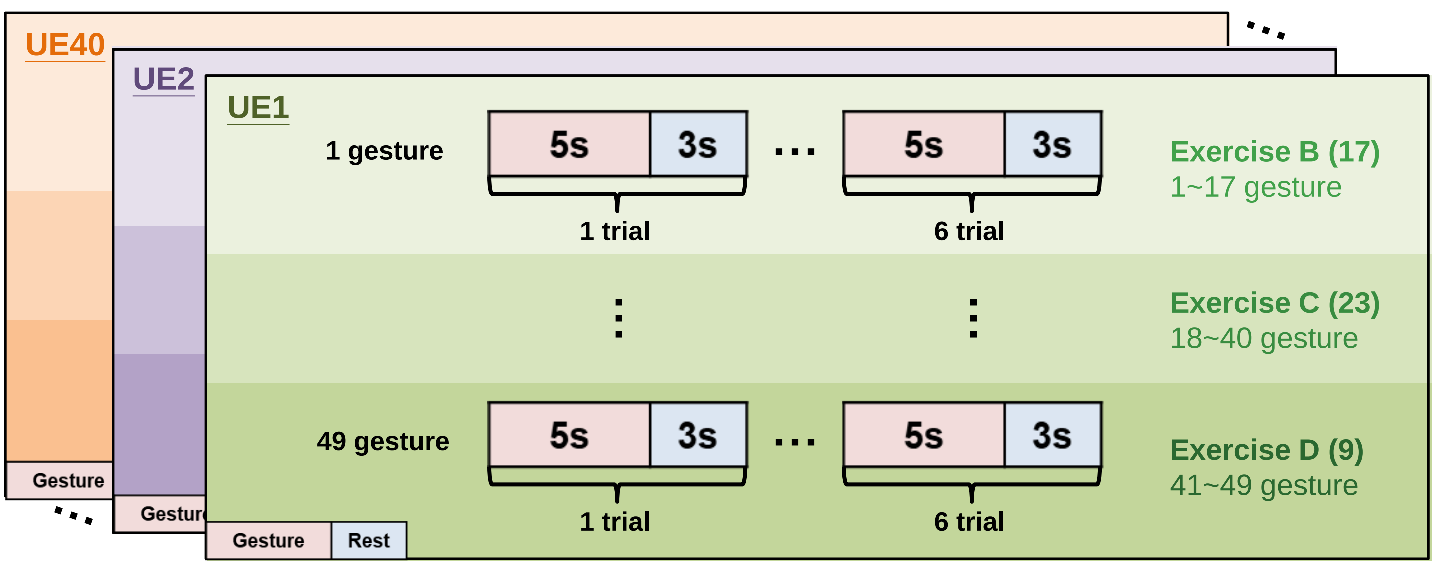
Scheduler \ Lr	0.001	0.0001
Y	Elapsed time = 71757.85 Acc = 77.5	Elapsed time = 103871.84 Acc = 62.5
N	Elapsed time = 59921.01 Acc = 80.5	Elapsed time = 100833.51 Acc = 71.0

- TraHGR

Scheduler \ Lr	0.001	0.0001
Y	Elapsed time = 112238.31 Acc = 84.1	Elapsed time = 129596.13 Acc = 80.5
N	Elapsed time = 104745.21 Acc = 85.6	Elapsed time = 135579.72 Acc = 80.8

Inter subject test

- **Device**: Delsys, 12 channels, 2048 sps
- **Setup**: 40 users / 49 gestures (B,C,D) / 6 trials
  - Train: 5 subjects trial 1,3,4,6
  - Valid: 5 subjects 1/2/3/4-th quarter of trial 1/3/4/6
  - Test : 5 subjects trial 2,5



Simulation Results – Inter subject

- Experiment method
  - Control variable : Dropout , Filter / Norm
  - Experimental variable : Transformer / DNN
- Epoch : significant factor

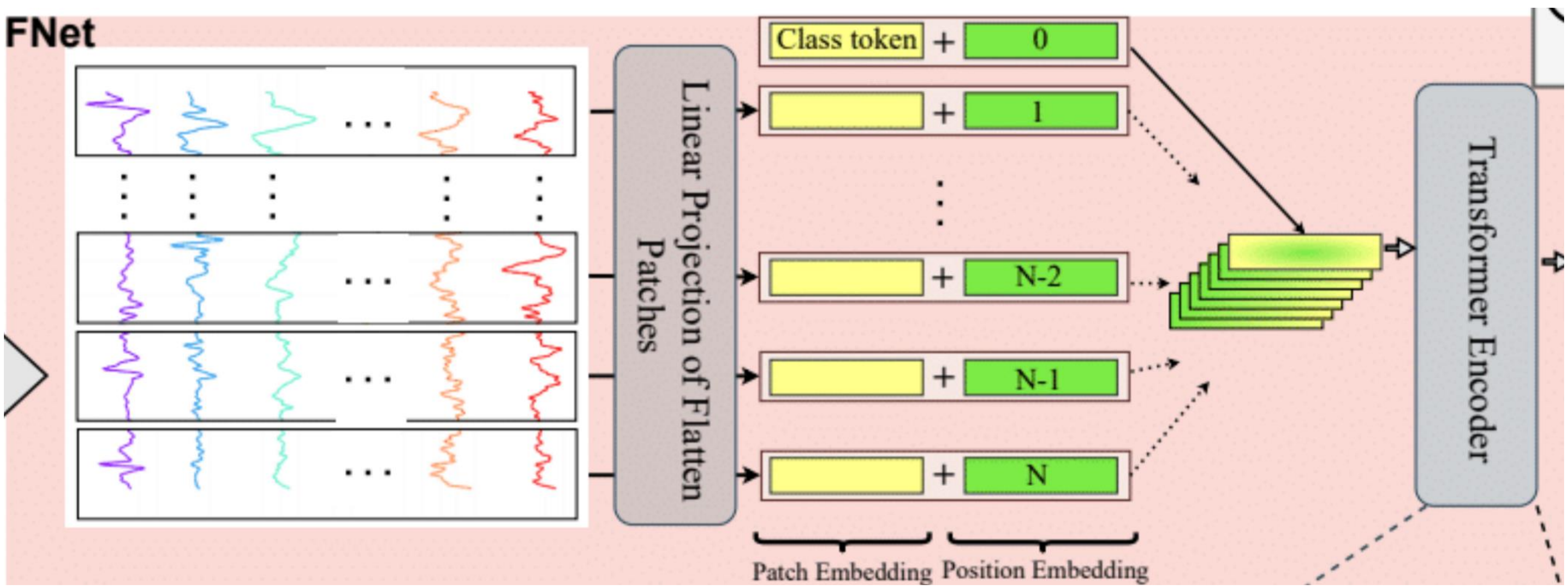
Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
window step	100 ms
Dropout	0.4

	DNN	FNet	TNet	TRAHGR	CNN
Elasped time	986.60	41256.29	32328.57	*43044.78*	*19515.69*
acc	*71.7*	79.0	47.2	87.02	85.65

TraHGR→ Best performance

Inter subject FNet

- Fnet model parameters
  - F, P, Q
  - Test different P values



Simulation Results – Inter subject FNet

- Step size : 0.01
- Epoch : 500
- Learning Rate : 0.0001 v.s. 0.001
- Sceduler : on

P value	Lr = 0.0001	Lr = 0.001
4	66.78	75.00 🀄
5	73.25	No time to train 🈵
6	71.97	78.39 🀄
10	79.52	81.55 🀄

Lower P => Better Performance, Try P = 15, 20 in the future

Reference – Overview

[1] T. Ganokratanaa and S. Pumrin, "The vision-based hand gesture recognition using blob analysis,“ *International Conference on Digital Arts, Media and Technology*, 2017, pp. 336-341.

[2] J. Lien et al., “Soli: Ubiquitous gesture sensing with millimeter wave radar,” *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–19, 2016.

[3] W. Jiang, Y. Ren, Y. Liu, Z. Wang and X. Wang, “Recognition of Dynamic Hand Gesture Based on Mm-Wave Fmcw Radar Micro-Doppler Signatures,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 4905-4909.

[4] Inside Facebook Reality Labs: Wrist-based interaction for the next computing platform - Tech at Meta

[5] ProComp Infiniti ™ Hardware Manual: Introduction (bio-medical.com)

[6] Côté-Allard, U., Gagnon-Turcotte, G., Laviolette, F., & Gosselin, B. “A low-cost, wireless, 3-D-printed custom armband for sEMG hand gesture recognition,” *Sensors*, 2019.

[7] Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., & Li, J. “Gesture recognition by instantaneous surface EMG images ,” *Scientific reports*, 2016.

[8] M. Atzori et al., “Electromyography data for non-invasive naturallycontrolled robotic hand prostheses,” *Sci. Data*, vol. 1, 2014, Art. no. 140053.

[9] S. Pizzolato et al., “Comparison of six electromyography acquisition setups on hand movement classification tasks,” *PLoS ONE*, vol. 12, no. 10, pp. 1–17, Oct. 2017.

[10] F. Palermo et al., “Repeatability of grasp recognition for robotic hand prosthesis control based on sEMG data,” *Proc. Int. Conf. Rehabil. Robot.*, 2017, pp. 1154–1159.

[11] A. Krasoulis et al., “Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements,” *J. Neuroeng. Rehabil.*, vol. 14, no. 1, pp. 71–84, 2017.