Undergraduate Group

Final report –Experimental results of sEMG

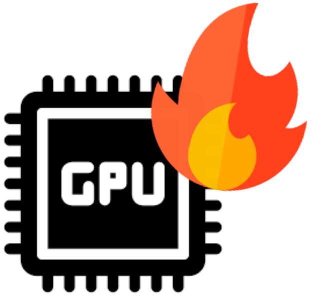
Presenter: Howard

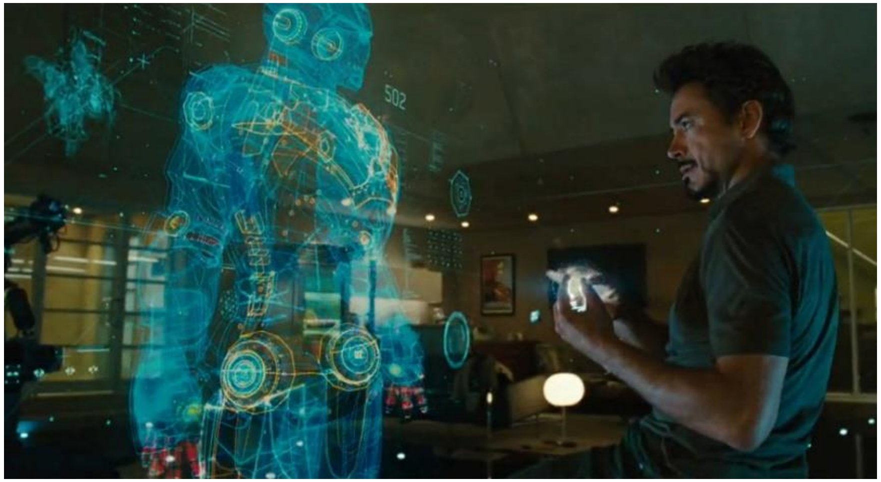
Advisor: Prof. An-Yeu (Andy) Wu

Date: 2024/1/2

Outline

- DNN experiment
  - Basic concept
  - Test on different normalization and filter
  - Test on dropout rate
- Transformer – Tra
  - Basic concept
  - Test on different learning rate
  - Test scheduler
  - Test on different p value
- Conclusion
- Future Work

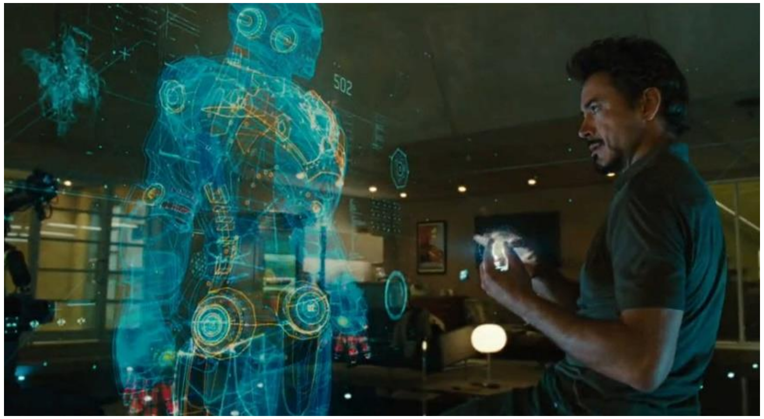




I. Introduction to EMG-based HGR

HCI – Human Computer Interaction

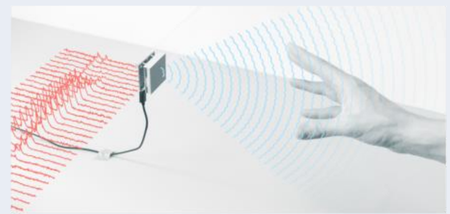
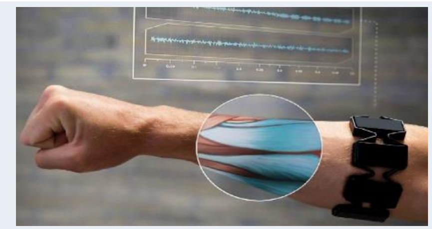
- Why Spatial Computing (VR/AR)
  - Next-gen computer
  - Immersive and Intuitive
- How to interact with our devices
  - Hand gestures !



Apple Vision Pro

Meta Quest 3

Solution to Hand Gesture Differentiation

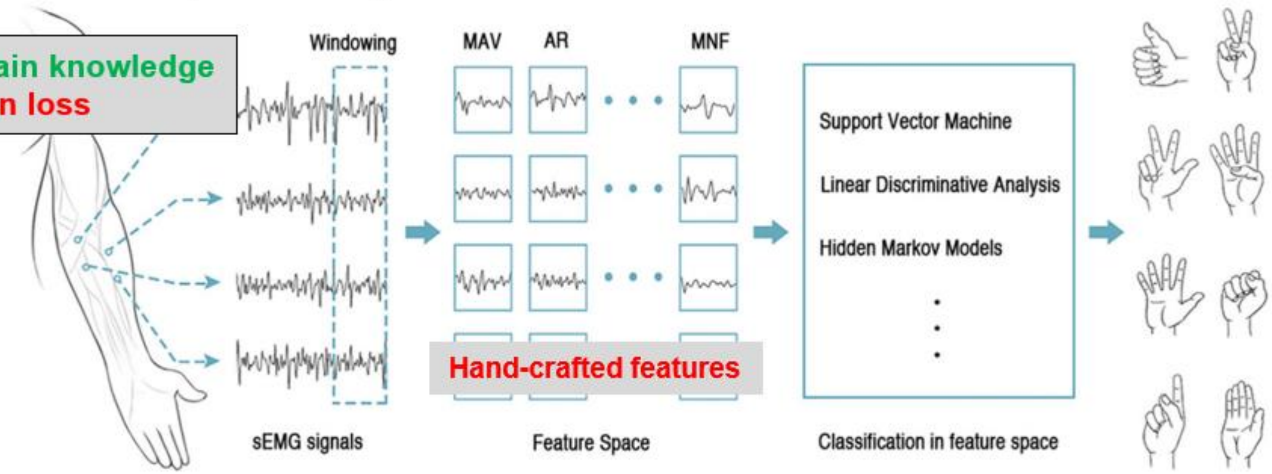
	Vision	Radar	EMG
Definition	Cameras to capture hand movements	Radio waves to detect hand movements	Muscles' electrics during hand movements
Mechanism			
Accuracy	High	Moderate	Moderate
Computation	High	Moderate	Low
Lighting	Highly affected	Not affected	Not affected
Occlusion	Highly affected	Less affected	Not affected
Privacy	No	Yes	Yes
Calibration	Initial	Initial	Frequently
Product	**Apple** Vision Pro **Meta** Quest 3	**Google** Soli	**Meta** Wristband

EMG: Electromyography (肌電圖)

Hand Gesture Recogntion

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

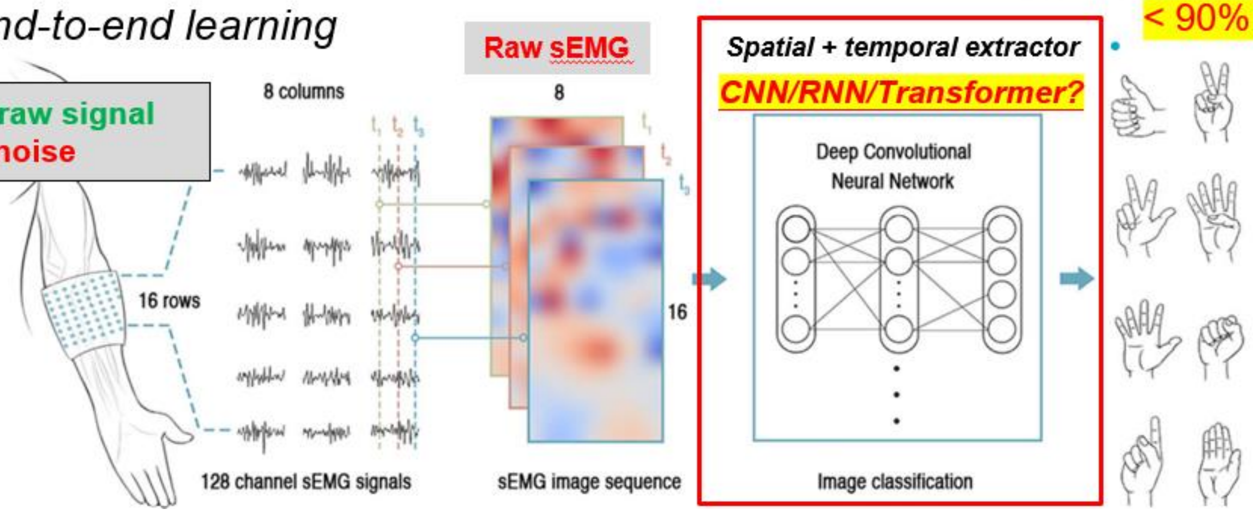
- Prior domain knowledge
- Information loss

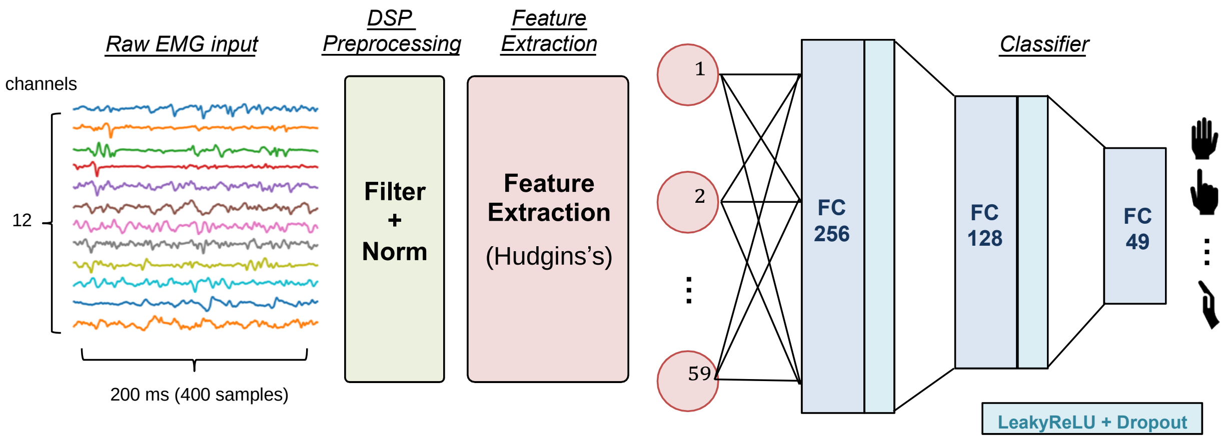


DL-based: *End-to-end learning*

< 90% over 17 gestures

- Informative raw signal
- Redundant noise

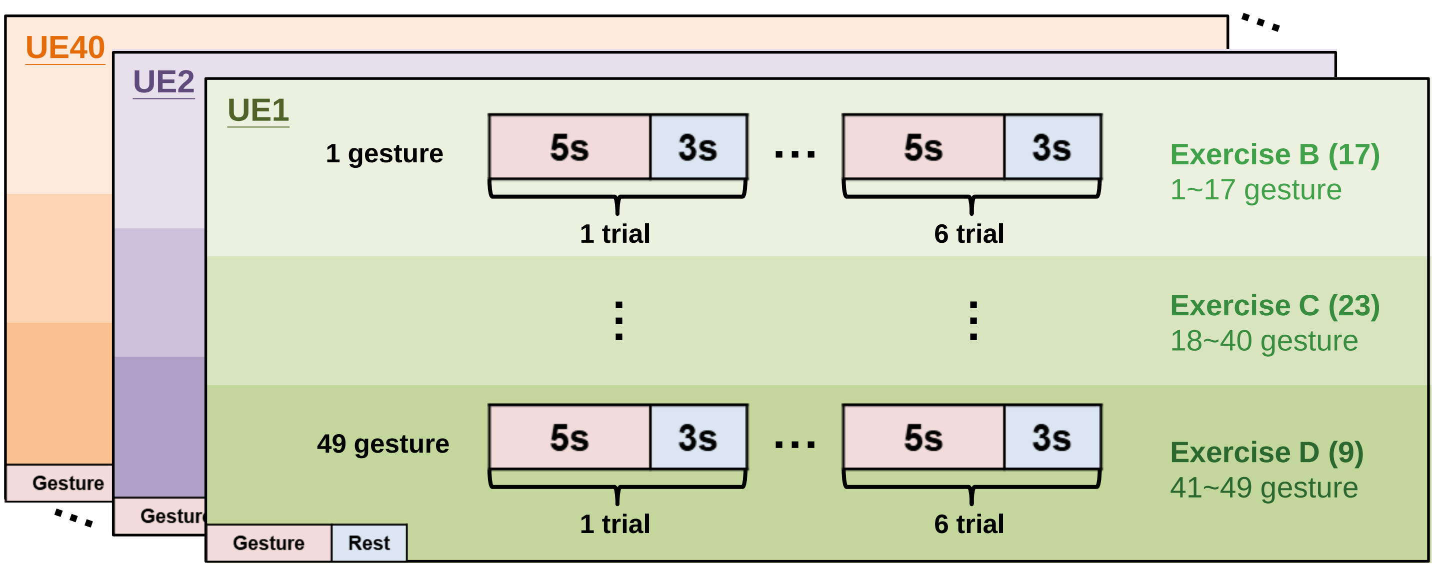




II. Feature-based DNN

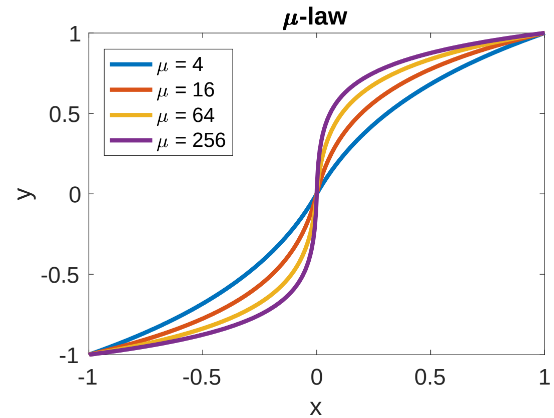
Intra Subject Test

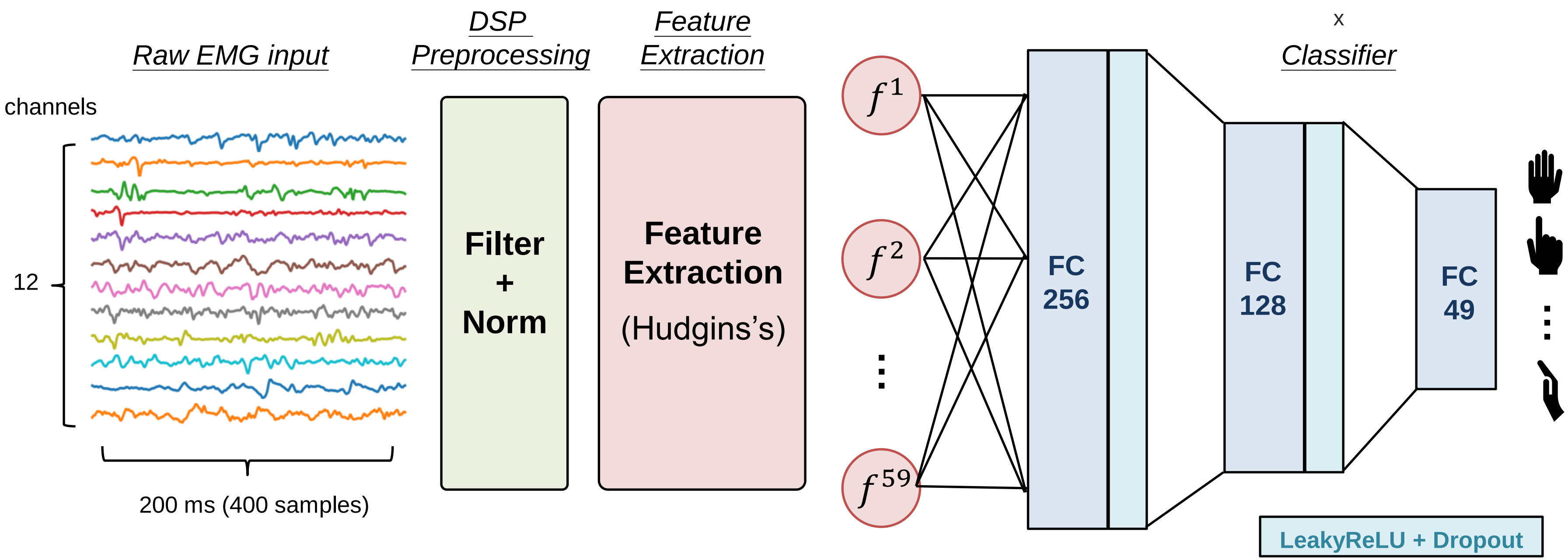
- **Device**: Delsys, 12 channels, 2048 sps
- **Setup**: 40 users / 49 gestures (B,C,D) / 6 trials
  - Train: each subject's trial 1,3,4,6
  - Valid: each subject's 1/2/3/4-th quarter of trial 1/3/4/6
  - Test : each subject's trial 2,5
- Acc = average over each subject's accuracy



Feature-based: DNN

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





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

Simulation Results – Norm & Filter

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

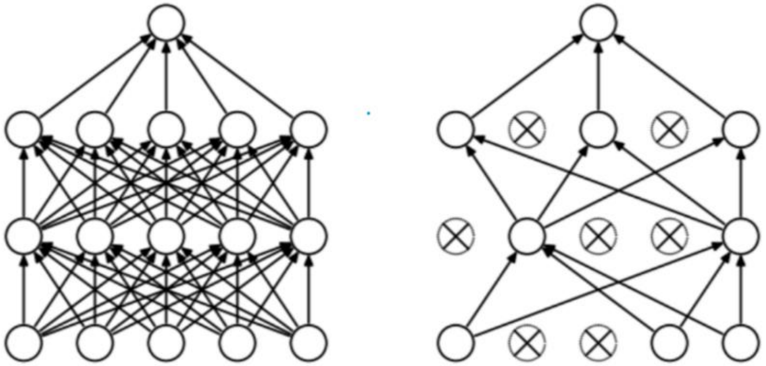
Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
window step	100 ms
Dropout	0.4

type_norm / type_filter	none	min_max	Z-score	μ-law(16)	μ-law(64)	μ-law(256)	μ-law(2048)	AVG
none	77.66	34.67	85.29	77.44	77.89	78.58	78.35	72.84
BPF(10,200)	74.56	24.28	77.86	77.86	75.57	76.22	77.99	69.19
BPF(10,500)	73.45	28.8	81.93	77.01	74.43	81.93	78.46	70.86
BPF(10,700)	76.92	33.28	84.21	77.17	77.16	77.64	78.52	72.13
LPF(1)	34.73	4.08	10.32	41.67	43.11	45.41	47.91	32.46
LPF(10)	69.47	9.22	70.75	70.45	71.23	70.05	69.7	61.55
LPF(20)	71.46	17.43	62.74	72.72	72.85	72.79	70.15	62.88
AVG	68.32	21.68	67.59	70.62	70.32	71.80	71.58	-

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

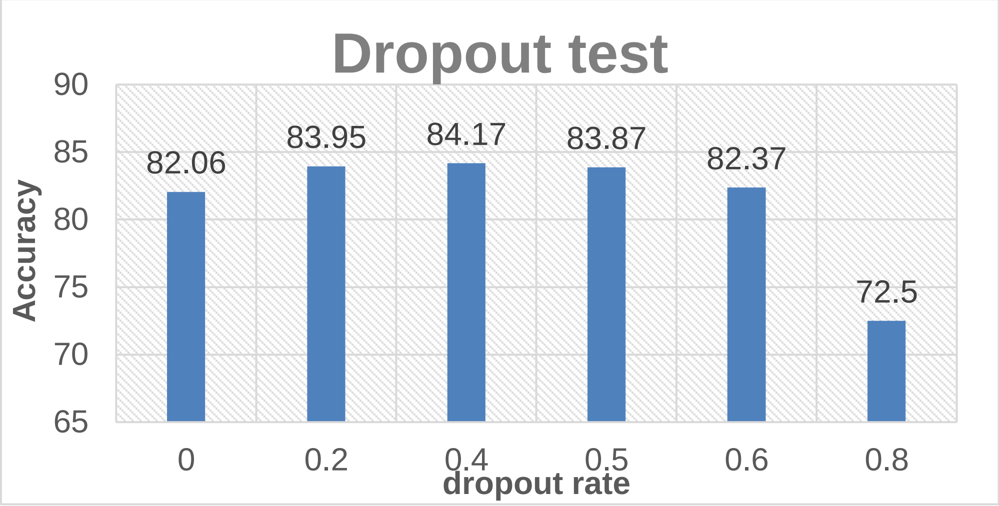
Simulation Results – Dropout

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

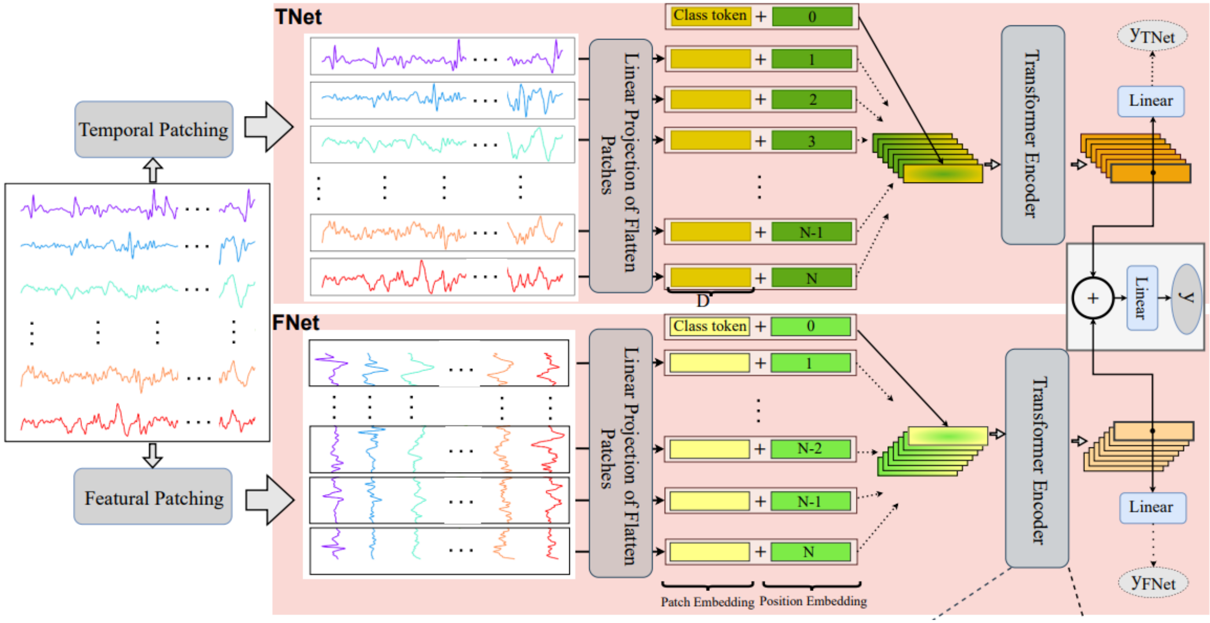


(a) Standard Neural Net (b) After applying dropout.

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



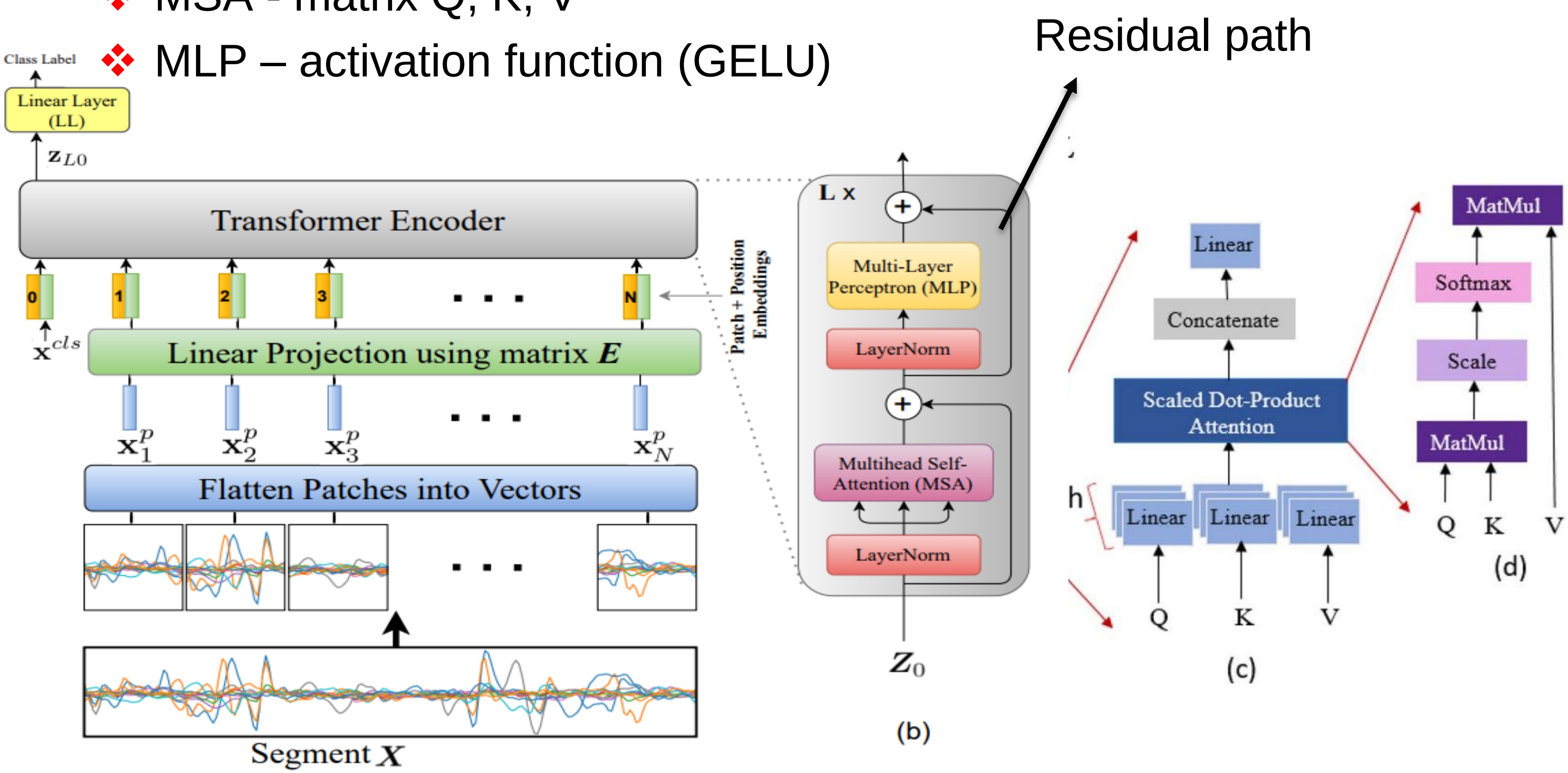
Dropout rate = 0.4 → Best performance



III. DL-based ViT

TEMG : Basic Transformer Concept[2][3]

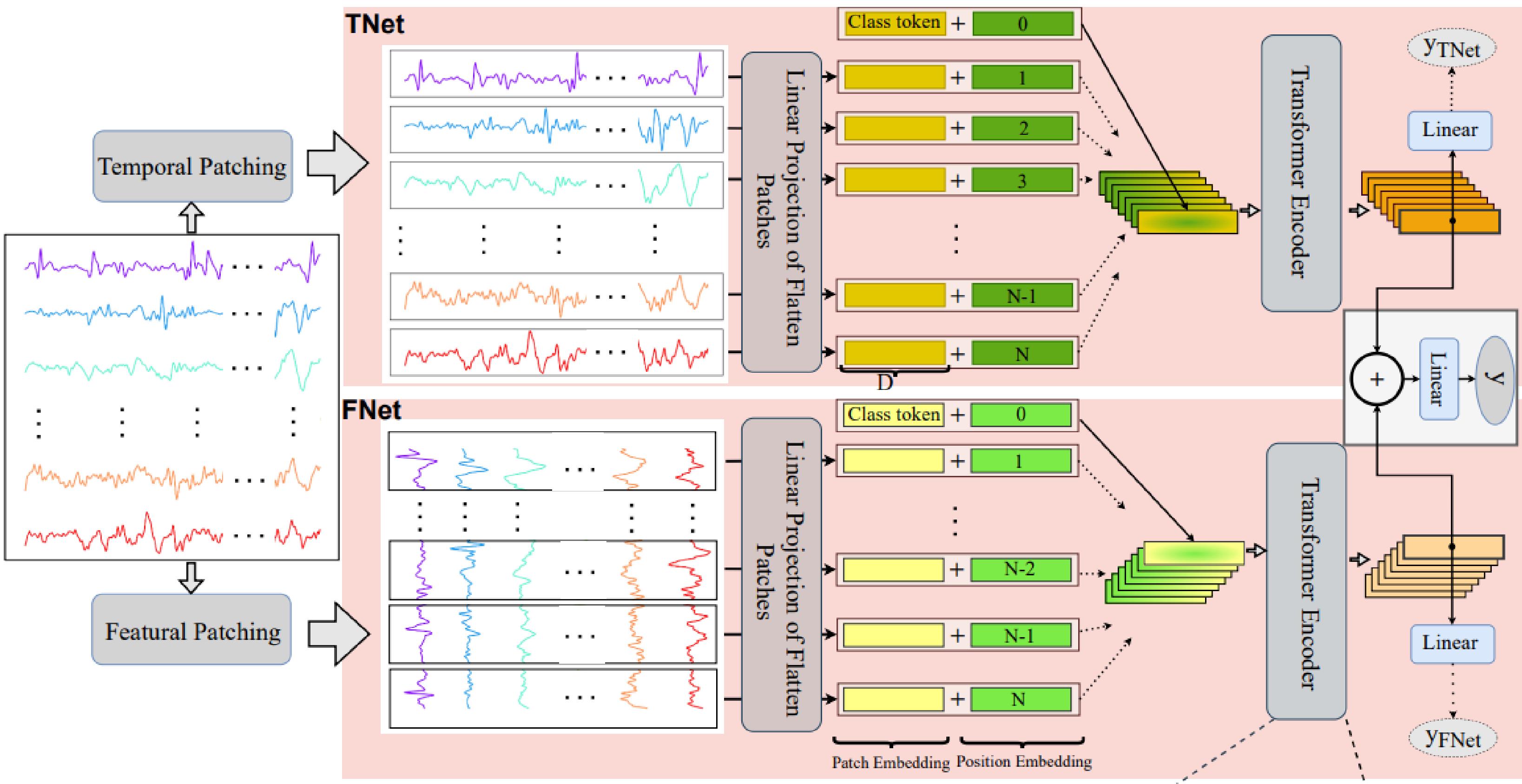
- x^{cls} - a trainable token
- Position embedding → encode the order of the input sequence
- Transformer encoder
  - MSA - matrix Q, K, V
  - MLP – activation function (GELU)



DL-based Transformer - Tra

- **TNet** extracts temporal features
- **Fnet** extracts spatial-temporal features
- **Tra** consists of two path (multi-view)

Model	200ms
TraHGR-Base	78.60 ± 6.03
TraHGR-large	83.58 ± 5.48
TraHGR-Huge	**86.18 ± 4.99**
TNet	83.39 ± 5.44
FNet	80.72 ± 5.82



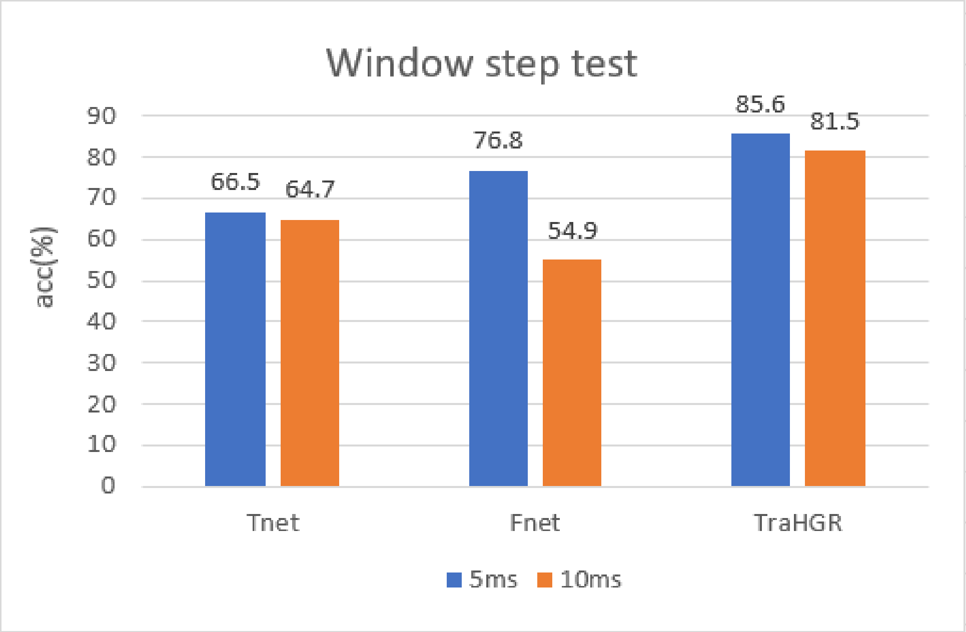
ViT 細節

- 可拿其中報告的內容來介紹

Simulation Results – window step

- Experiment method
  - Control variable : epoch, dropout, Filter/Norm
  - Experimental variable : window step
  - An alternative of data augmentation

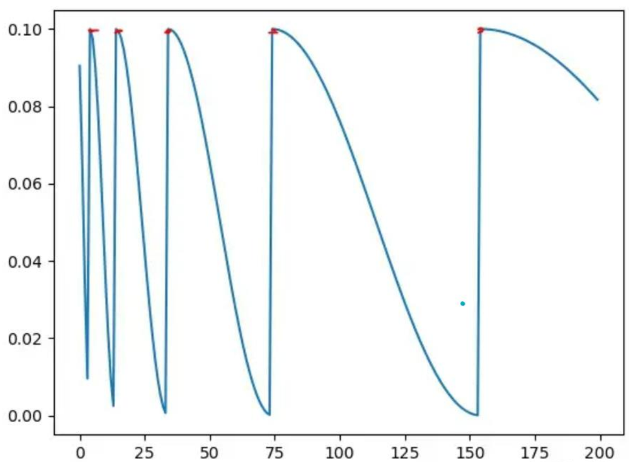
Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	intra-subject
window size	200 ms
Dropout	0.4
Filter	none
Norm	Z-score

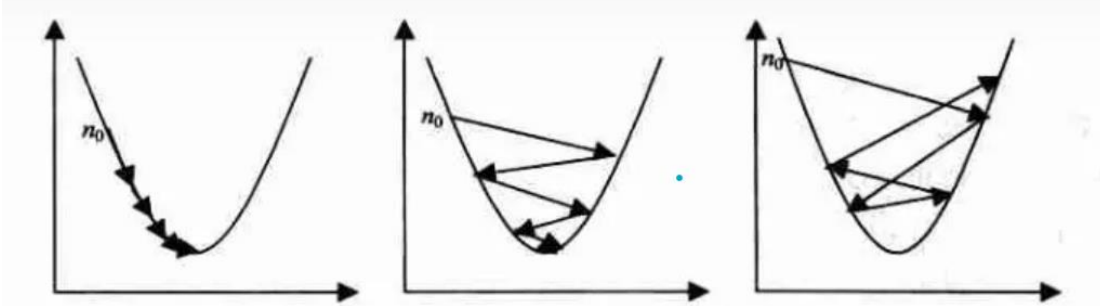


5ms→ Better performance but time consuming

Scheduler - CosineAnnealingRestarts

- T_0 : the first time learning rate back to initial
- T_mult : control the speed of lr back to initial
- Schedule the decay of learning rate



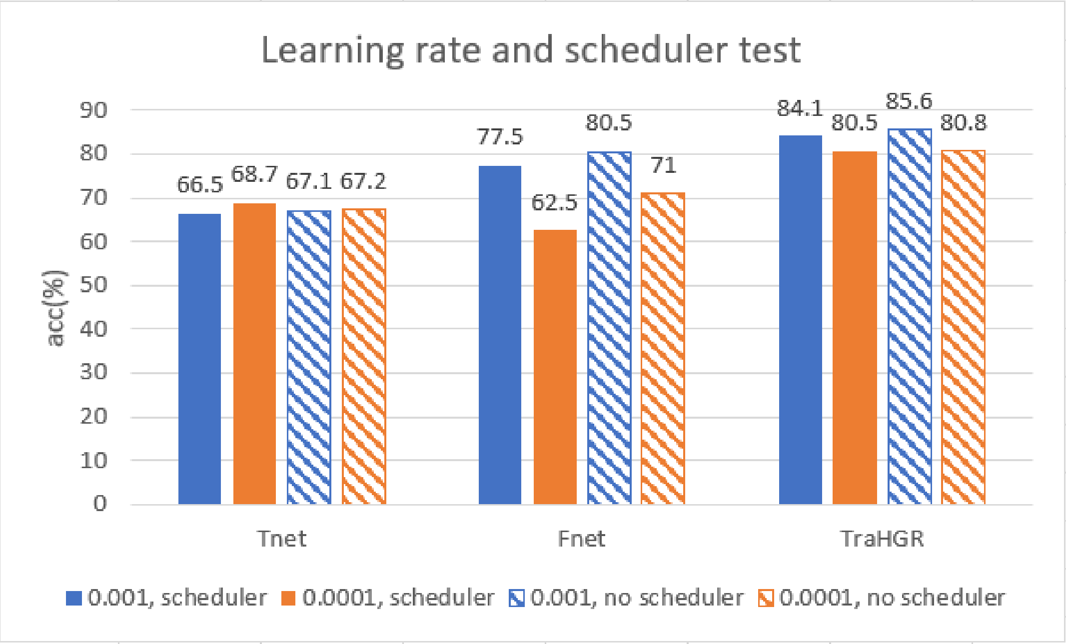


(a) η很小时，可以保证收敛 (b) 较大的η，训练将振荡收敛 (c) η过大，系统无法收敛

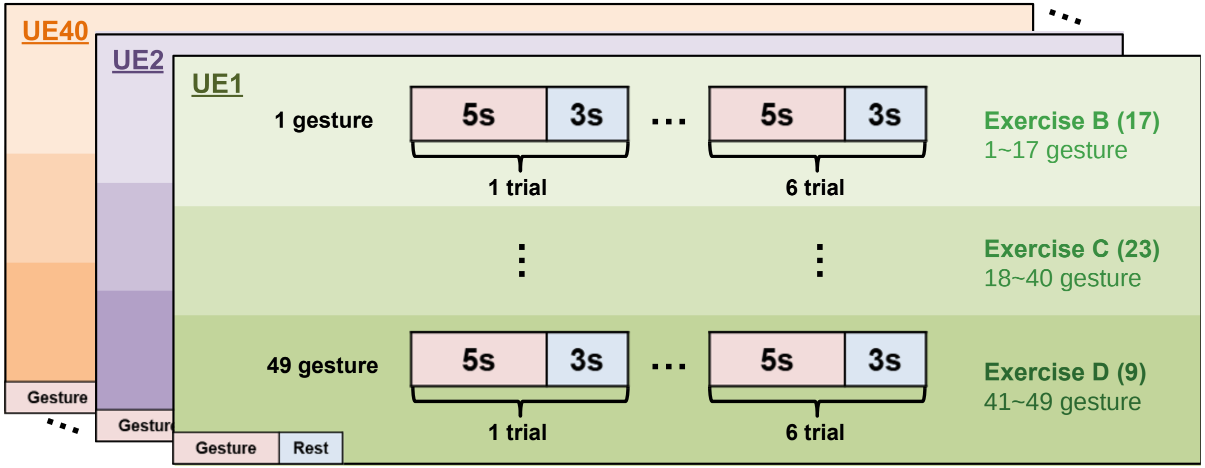
Simulation Results – learning rate

- 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
Filter	none
Norm	Z-score



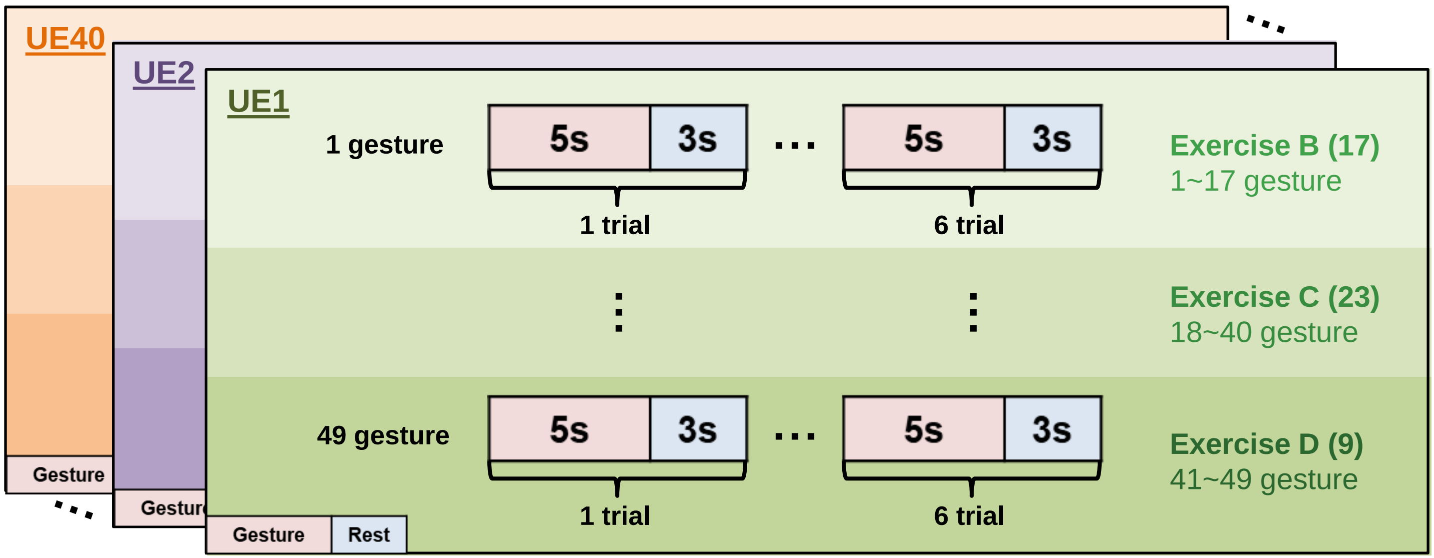
Lr = 0.001 / no scheduler→ Better performance



IV. Inter-subject Test

Inter Subject Test

- **Device**: Delsys, 12 channels, 2048 sps
- **Setup**: 40 users / 49 gestures (B,C,D) / 6 trials / choosing subject 1-5
  - 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
- Acc = one model trained on all subjects, and test on all subjects



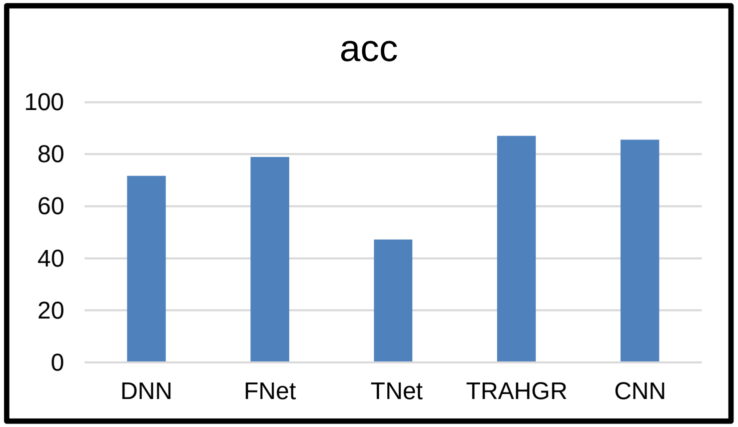
Simulation Results – Inter Subject

- Experiment method
  - Control variable : Dropout , Filter / Norm
  - Experimental variable : Transformer / DNN
- Transformer has better result
  - Transformer beats NN model as the data is adequate

Parameter	Setups
# gesture	49 (B,C,D)
# subject	1~5
scenario	Inter subject
window size	200 ms
Step size	10 ms
Dropout	0.4

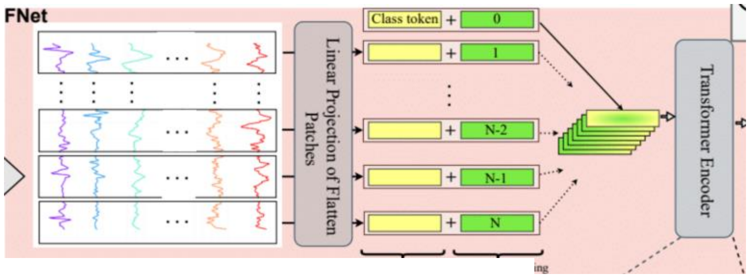
	DNN	FNet	TNet	TraHGR	CNN
Elapsed time (s)	986	41256	32328	43044	19515
Acc	71.7	79.0	47.2	87.0	85.7

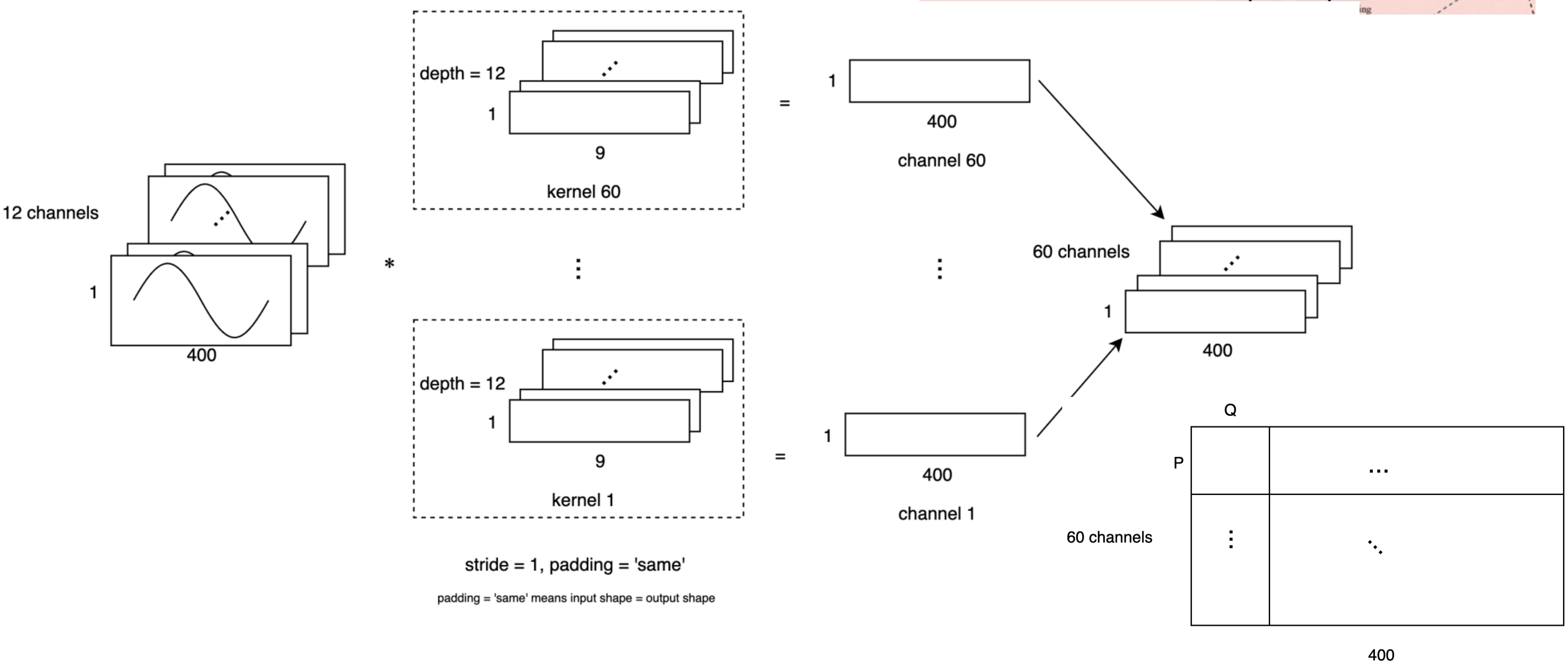
TraHGR performs the best



FNet Parameters

- P, Q : splitting parameters
  - Different patch size and # patches
- Test different P (common factor of 60)

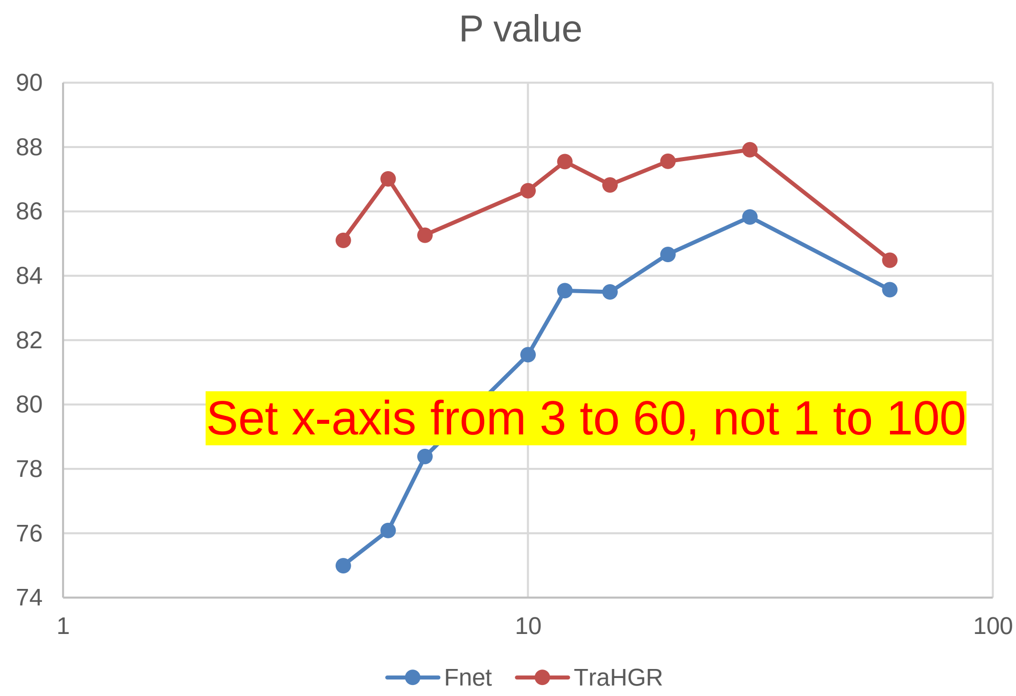




Simulation Results – Inter subject FNet

- Experiment method
  - Control variable : step size = 0.01, scheduler on, subject : 1-5, lr = 0.001
  - Experimental variable : P

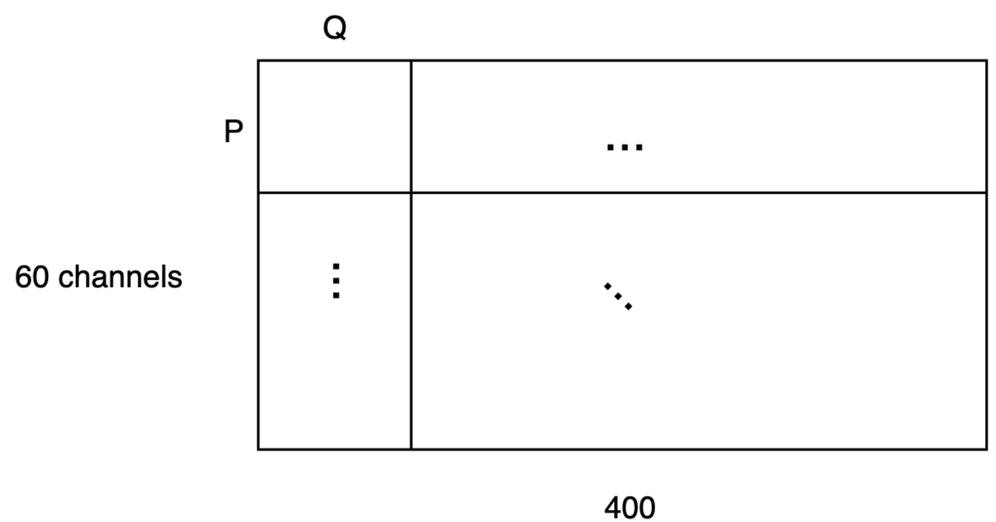
P	FNet	TraHGR
3	2.5	86.29
4	75.00	85.11
5	76.09	87.02
6	78.39	85.27
10	81.55	86.65
12	83.54	87.55
15	83.50	86.83
20	84.67	87.56
30	85.83得	87.92得
60	83.57	84.49



Only the Acc of FNet tends to increase when increasing P
Both of them perform the best when P=30

Discussion of Testing different P on FNet

- Based on common sense, ViT needs more input patch for attention
- But FNet with P=30, which provides less inputs than those FNets with lower P, performs the best among all possible cases.
- One explanation is that when choosing small P
  - patches are too small, and # patch are too much
  - “self-attention” does not perform well



Discussion of Testing different P on TraHGR

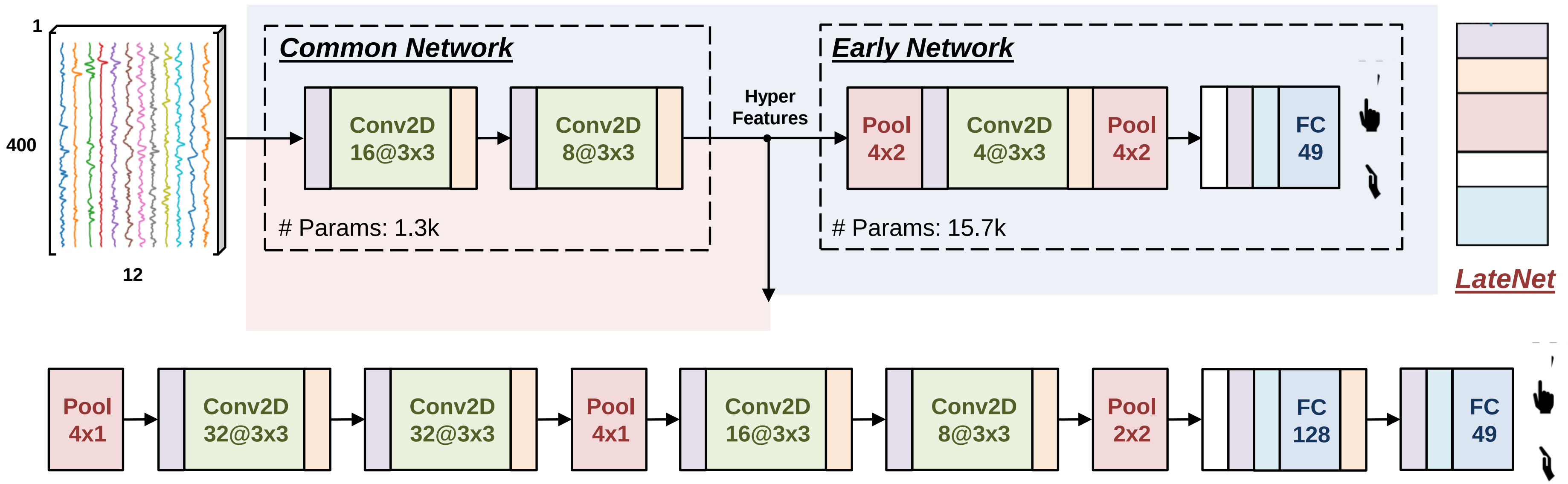
- TBD

Conclusion

- DNN
  - Dropout = 0.4
  - Filter = none, Normalization = z-score
  - Best performance = 84.19 for intra test
- Transformer beats DNN in inter subject test (87% vs 71%)
  - Window step = 10 ms
  - Learning rate = 0.001
  - No need for scheduler
- TraHGR is the combination of TNet and FNet
  - TNet does not perform well
  - FNet reaches its highest acc when P = 30
  - TraHGR can combine the advantages of both nets and achieve higher acc

Future Work

- Further research on other transformer models (CViT, LST-ViT, …)
- Use the EMG device to conduct experiment on human being
- Reduce the complexity by early exit



Reference – Overview

[1] S. Shen, X. Wang, F. Mao, L. Sun and M. Gu, "Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning," in *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13318-13325, 1 July1, 2022, doi: 10.1109/JSEN.2022.3179535.

[2] Rahimian, E.et al.Temgnet: Deep transformer-based decoding of upperlimb semg for hand gestures recognition.arXiv pre-printarXiv:2109.12379(2021).16.Toledo-Peral, C. L.et al.semg

[3] Montazerin, M., Rahimian, E., Naderkhani, F. et al. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals. Sci Rep 13, 11000 (2023). https://doi.org/10.1038/s41598-023-36490-w

[4] PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation

Qitao Zhao, Ce Zheng, Mengyuan Liu, Pichao Wang, Chen Chen.In CVPR 2023

[5] Dario Farina, Senior Member IEEE, and Alesˇ Holobar, Member IEEE,” Characterization of Human Motor Units From Surface EMG Decomposition”,

[6] Mattia Orlandi∗ , Marcello Zanghieri∗ , Victor Javier Kartsch Morinigo∗ , Francesco Conti∗ , Davide Schiavone† , Luca Benini∗‡, Simone Benatti§ , “sEMG Neural Spikes Reconstruction for Gesture Recognition on a Low-Power Multicore Processor”, 2022 IEEE Biomedical Circuits and Systems Conference (BioCAS)

[7]Xiangjun Zhu and Yingchun Zhang, “High-Density Surface EMG Decomposition based on a Convolutive Blind Source Separation Approach *”, 34th Annual International Conference of the IEEE EMBS San Diego, California USA, 28 August - 1 September, 2012

[8] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli and W. Geng, "Surface-Electromyography-Based Gesture Recognition by Multi-View Deep Learning," in IEEE Transactions on Biomedical Engineering, vol. 66, no. 10, pp. 2964-2973, Oct. 2019, doi: 10.1109/TBME.2019.2899222.