A Low-Cost, Wireless, 3-D-Printed Custom Armband for sEMG Hand Gesture Recognition

Presenter: Shawn

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/06

Myo armband vs 3DC armband

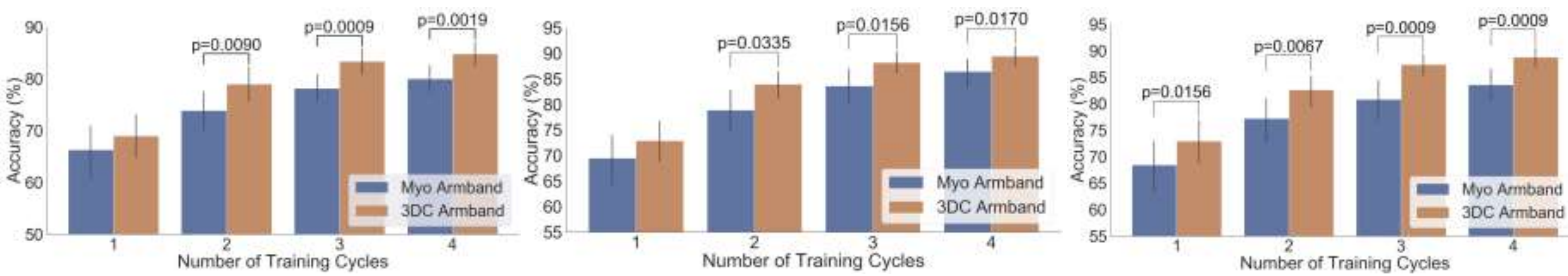
	Myo armband	3DC armband
sEMG channels	8	10
Sampling rate	200 sps	1000 sps
Contact Dimensions	100 mm^2	50 mm^2
Contact Material	Stainless steel silver coated	Electroless nickel immersion gold
Weight	93 g	62 g
Pricing	200 USD	Roughly 150 USD
Warm Up Period	Shorter	Longer

more affordable and widely accessible than clinical-grade systems currently available

Myo armband vs 3DC armband

Figure 7. The eleven hand/wrist gestures employed in the proposed dataset.

	LDA Classifier	RAW ConvNet	Spectrogram ConvNet
Myo armband	80.00%	86.41%	Roughly 84%
3DC armband	84.81%	89.47%	Roughly 89%



3DC armband was shown to significantly outperform Myo armband

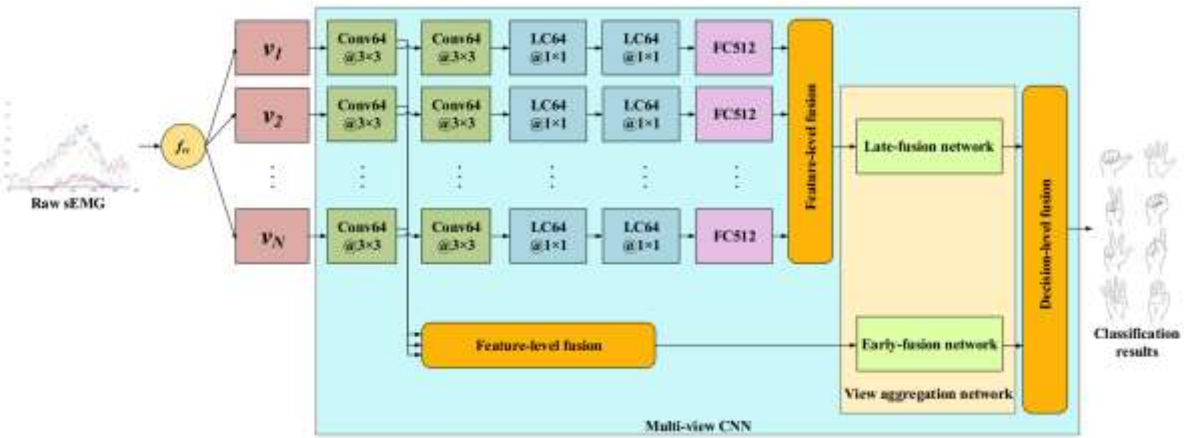
Surface-Electromyography-Based Gesture Recognition by Multi-View Deep Learning

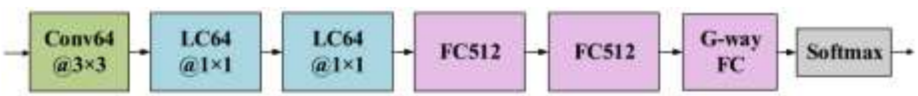
Presenter: Shawn

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/06

View Aggregation Network





a). Early-fusion network

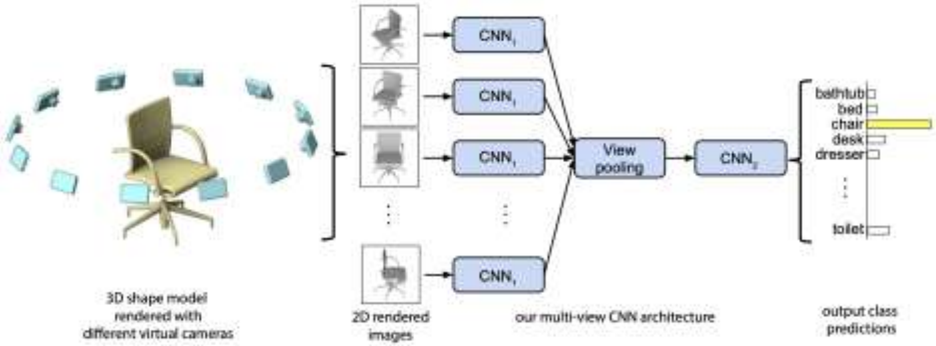
b). Late-fusion network

The view aggregation network achieved higher accuracy than did its two subnetworks

TABLE III
GESTURE RECOGNITION ACCURACIES ACHIEVED BY THE VIEW AGGREGATION NETWORK AND ITS TWO SUBNETWORKS ON NINAPRO. THE RESULTS IN BOLD TEXT INDICATE THE BEST PERFORMANCES

Method	Database	Accuracy
View aggregation network	NinaPro DB1	**88.2%**
Early-fusion network	NinaPro DB1	87.5%
Late-fusion network	NinaPro DB1	87.9%
View aggregation network	NinaPro DB2	**83.7%**
Early-fusion network	NinaPro DB2	83.3%
Late-fusion network	NinaPro DB2	82.5%
View aggregation network	NinaPro DB5	**90.0%**
Early-fusion network	NinaPro DB5	89.5%
Late-fusion network	NinaPro DB5	89.7%

View Aggregation Network



Score summation fusion: Elementwise summation of the softmax scores of all streams, i.e.,

$$y_{\texttt{final}} = \sum_{i=1,2,3} y_i \tag{7}$$

Score maximum fusion: Elementwise maximum of the softmax scores of all streams, i.e.,

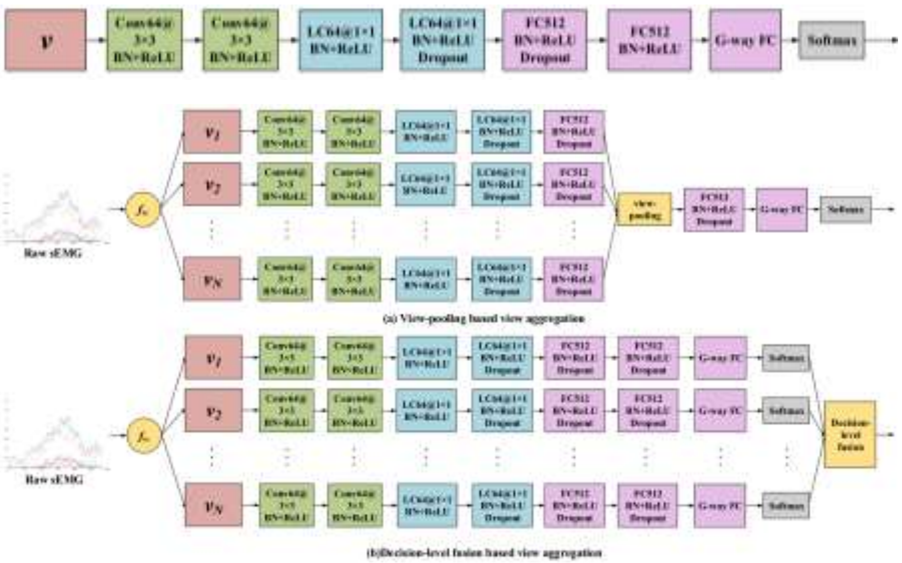
$$y_{\texttt{final}} = \max(y_i, i = 1, 2, 3) \tag{8}$$

On NinaPro DB1 and DB2:
View Aggregation Network achieves slightly higher
On NinaPro DB5:
Performances are quite similar

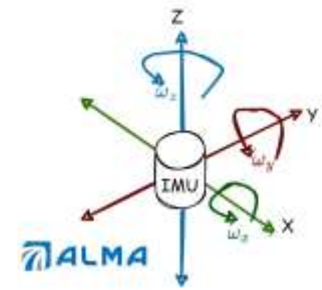
TABLE IV
PERFORMANCE COMPARISON OF THE VIEW AGGREGATION NETWORK, THE VIEW-POOLING BASED VIEW AGGREGATION APPROACHES, AND THE SCORE FUSION BASED VIEW AGGREGATION APPROACHES. THE RESULTS IN BOLD ENTRIES INDICATE THE BEST PERFORMANCES

Method	Database	Accuracy
View aggregation network	NinaPro DB1	**88.2%**
View-pooling at 1st Conv	NinaPro DB1	87.2%
View-pooling at 2nd Conv	NinaPro DB1	87.4%
View-pooling at 1st LC	NinaPro DB1	87.4%
View-pooling at 2nd LC	NinaPro DB1	87.6%
View-pooling at 1st FC	NinaPro DB1	87.5%
View-pooling at 2nd FC	NinaPro DB1	88.0%
Score summation fusion	NinaPro DB1	87.8%
Score maximum fusion	NinaPro DB1	87.7%
View aggregation network	NinaPro DB2	**81.4%**
View-pooling at 1st Conv	NinaPro DB2	80.7%
View-pooling at 2nd Conv	NinaPro DB2	80.1%
View-pooling at 1st LC	NinaPro DB2	80.7%
View-pooling at 2nd LC	NinaPro DB2	77.7%
View-pooling at 1st FC	NinaPro DB2	77.2%
View-pooling at 2nd FC	NinaPro DB2	76.3%
Score summation fusion	NinaPro DB2	77.0%
Score maximum fusion	NinaPro DB2	77.5%
View aggregation network	NinaPro DB5	90.0%
View-pooling at 1st Conv	NinaPro DB5	89.5%
View-pooling at 2nd Conv	NinaPro DB5	89.7%
View-pooling at 1st LC	NinaPro DB5	89.3%
View-pooling at 2nd LC	NinaPro DB5	89.7%
View-pooling at 1st FC	NinaPro DB5	89.7%
View-pooling at 2nd FC	NinaPro DB5	89.4%
Score summation fusion	NinaPro DB5	**90.1%**
Score maximum fusion	NinaPro DB5	90.0%

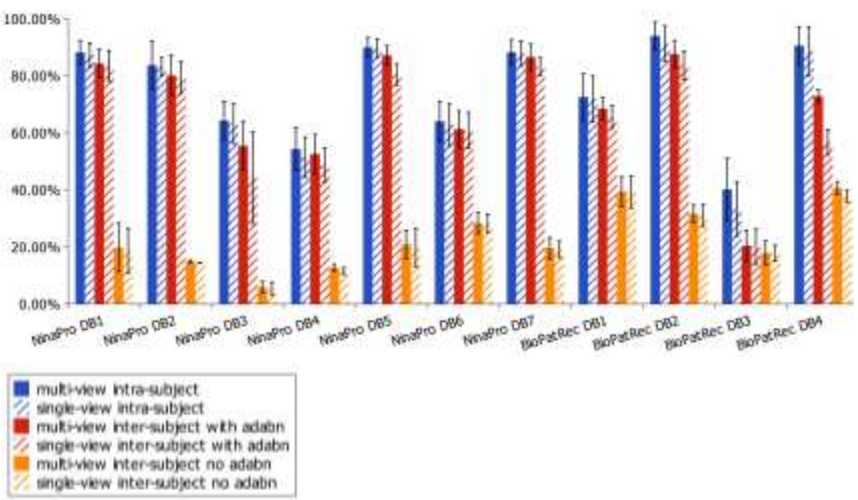
Multi-view vs Single-view



Classification results



Database	Method	Intra-subject gesture recognition accuracy	
		With IMU	Without IMU
Ninapro DB2	multi-view	**94.40%**	83.70%
	single-view	92.84%	83.30%
Ninapro DB3	multi-view	**87.06%**	64.30%
	single-view	84.97%	63.30%
Ninapro DB5	multi-view	**91.31%**	90.00%
	single-view	90.22%	89.60%
Ninapro DB6	multi-view	**77.10%**	64.10%
	single-view	73.99%	62.90%
Ninapro DB7	multi-view	**94.54%**	88.30%
	single-view	92.48%	87.80%



Intra-subject performs better than inter-

AdaBN is useful when optimizing inter-

MV-CNN performs slightly better than SV-

IMU data play significant role when training in both MV-CNN and SV-CNN cases

Undergraduate

Paper survey

DL model of EMG-based Hand Gesture Recognition

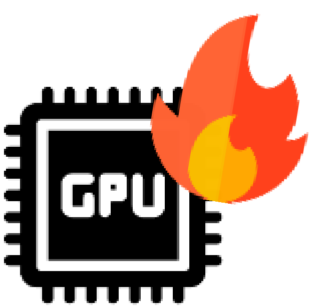
Presenter: Howard

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/06

Outline

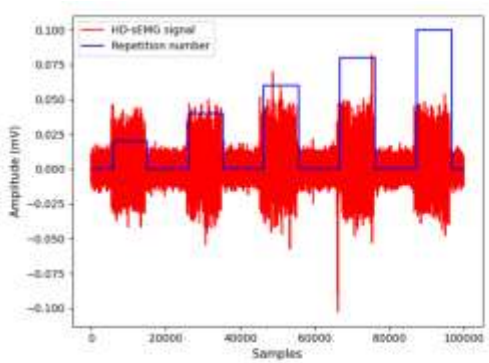
- Brief introduction of sEMG
- Review of three papers
  - Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning
  - TEMGNet: Deep Transformer-based Decoding of Upperlimb sEMG for Hand Gestures Recognition
  - Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals
- Discussion
- Summary



Surface Electromyography (sEMG)

- Without invading human body, get the information of the signal sent by upper limb muscle
- Armband with a 9-axis Inertial Measurement Unit and well-organized electrodes can get HD-sEMG datasets and further information and leverage
- Application
  - robotic arm control
  - medical rehabilitation
  - sign language recognition
  - virtual reality

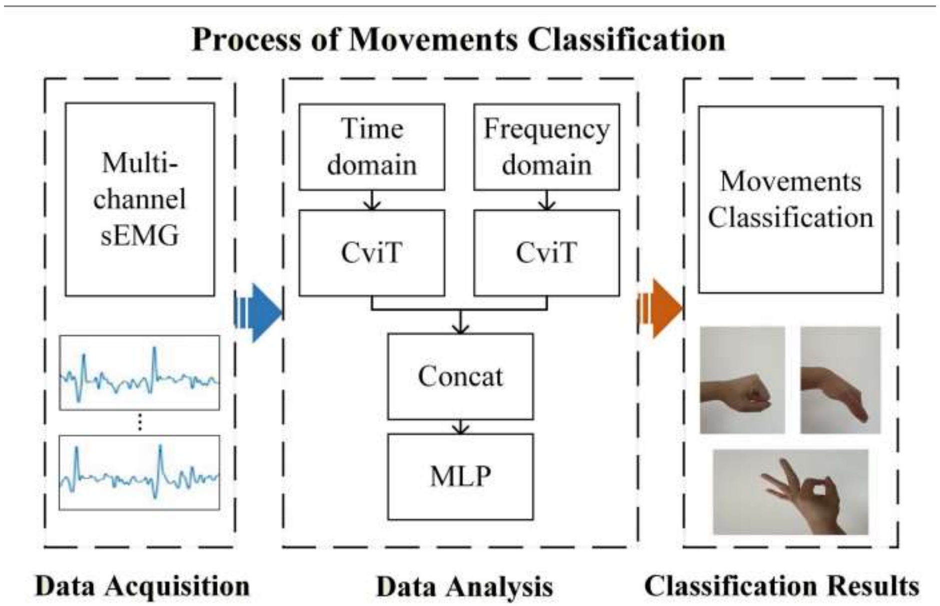
Power Grip

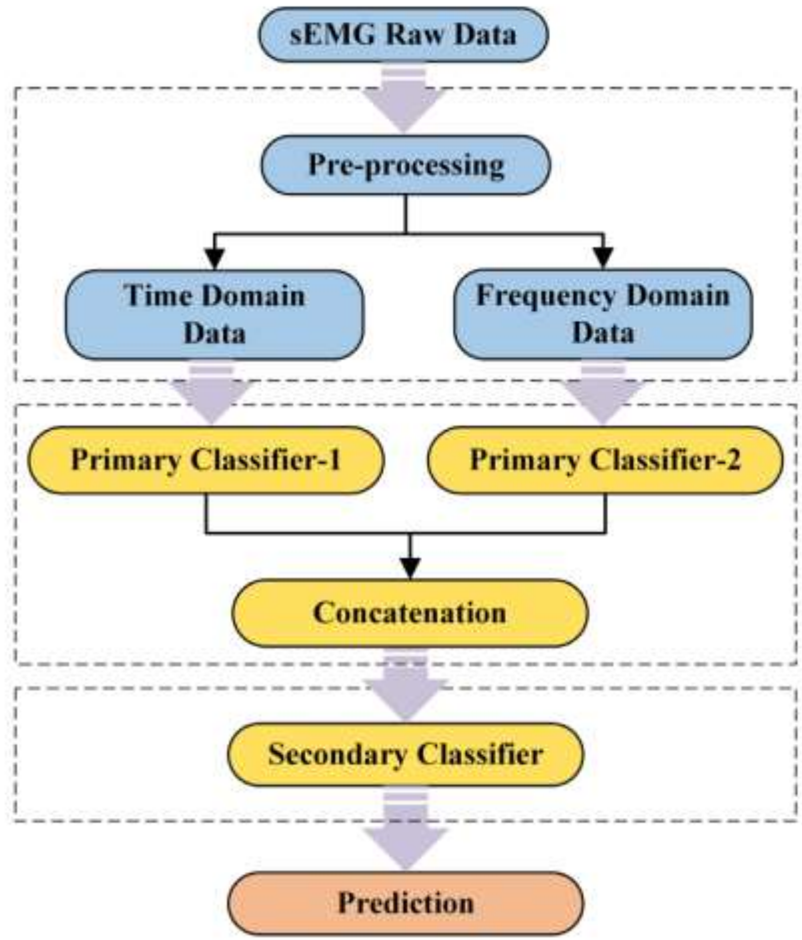


ML

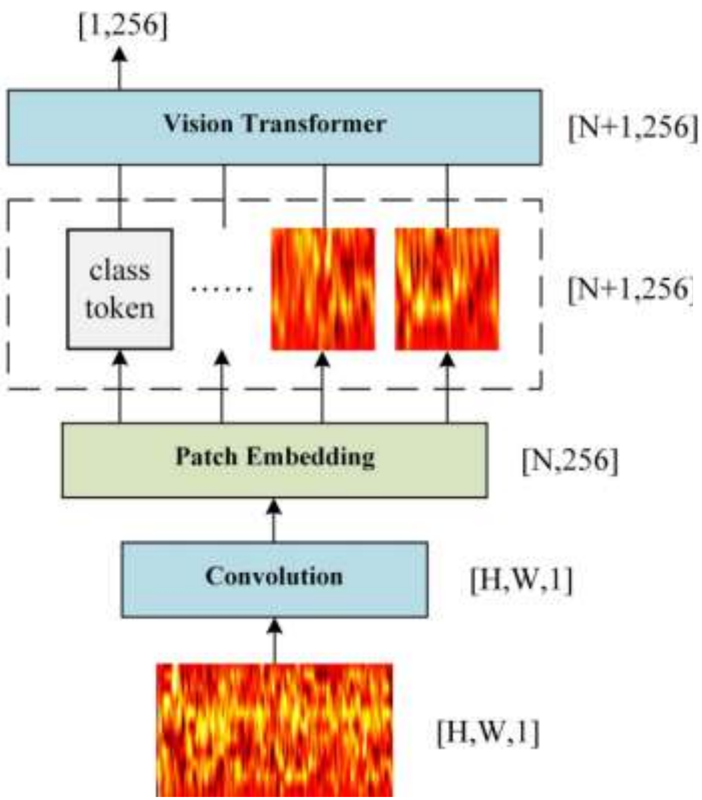
Gesture recognition

Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning

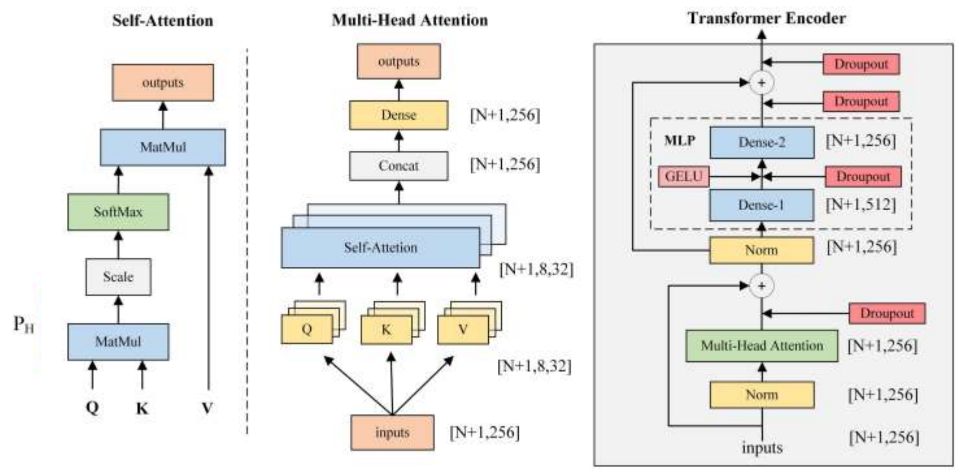




Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning

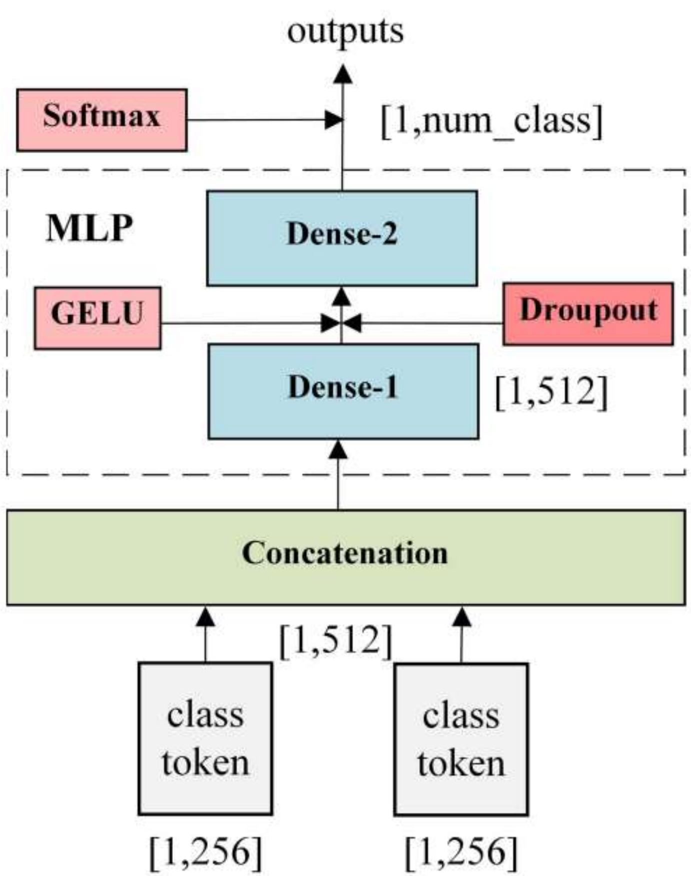


Primary Classifier



Patch Embedding
Vision Transformer

Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning

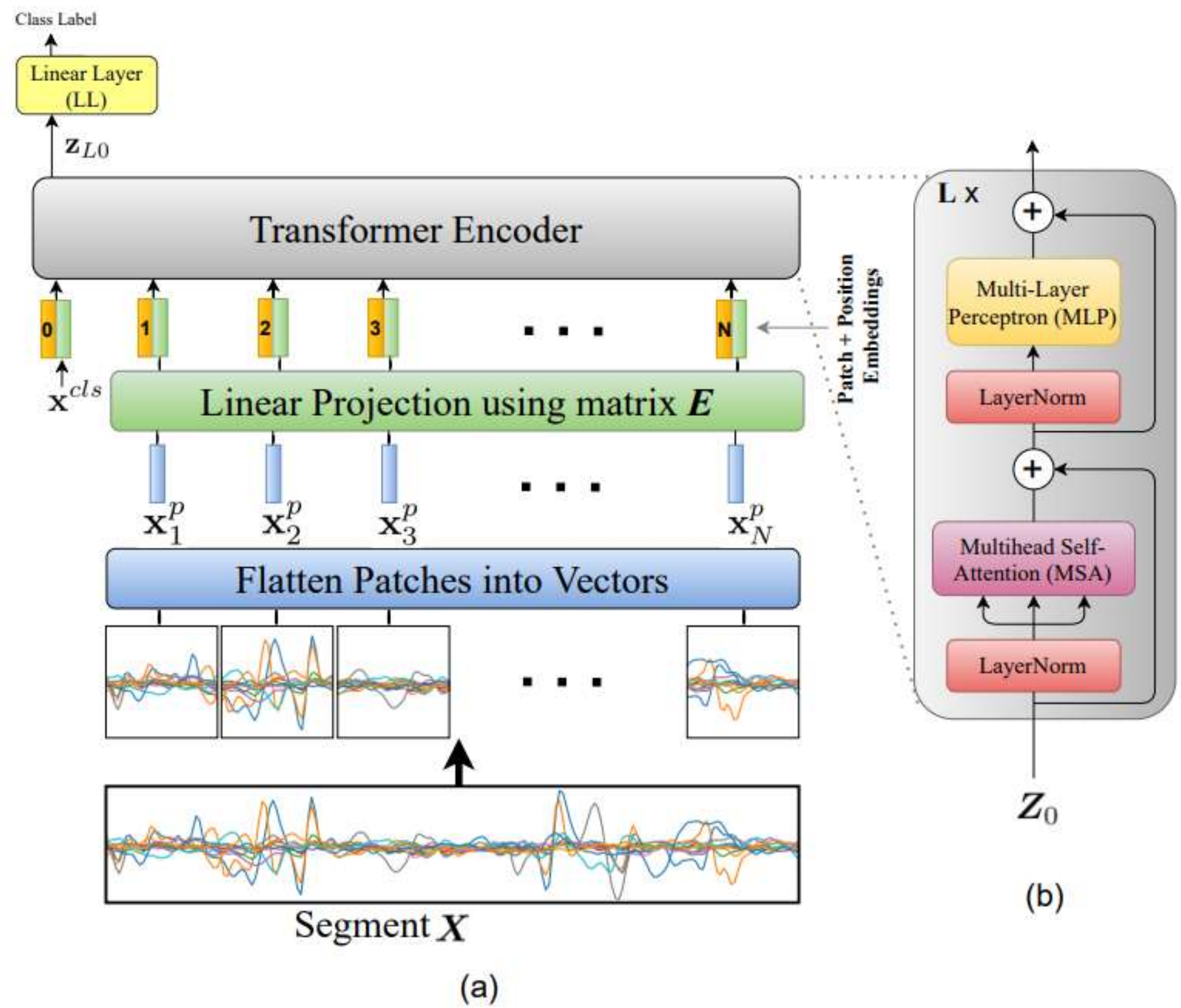


Result

Stacking Strategies	Accuracy
Time	79.58%
Frequency	78.65%
Time + Time	81.06%
Frequency + Frequency	79.27%
Time + Frequency	**81.20%**

Method	Database	Number of movements	Window length	Accuracy
Random Forest [22]	NinaPro DB2	49	200ms	75.27%
CNN [4]	NinaPro DB2	49	200ms	78.71%
CvIT	NinaPro DB2	49	200ms	**80.02%**
LSTM [8]	NinaPro DB2-E1	17	300ms	79.19%
ViT [13]	NinaPro DB2-E1	17	200ms / 300ms	82.05% / 82.93%
CvIT	NinaPro DB2-E1	17	200ms / 300ms	**83.47% / 84.09%**
LDA [2]	NinaPro DB5-A/B	12 / 17	200ms	69.49% / 61.75%
SVM [2]	NinaPro DB5-A/B	12 / 17	200ms	67.9% / 58.27%
LSTM-CNN [10]	NinaPro DB5-A/B	12 / 17	200ms	71.66% / 61.4%
CvIT	NinaPro DB5-A/B	12 / 17	200ms	**76.83% / 73.23%**

TEMGNet: Deep Transformer-based Decoding of Upper-limb sEMG for Hand Gestures Recognition

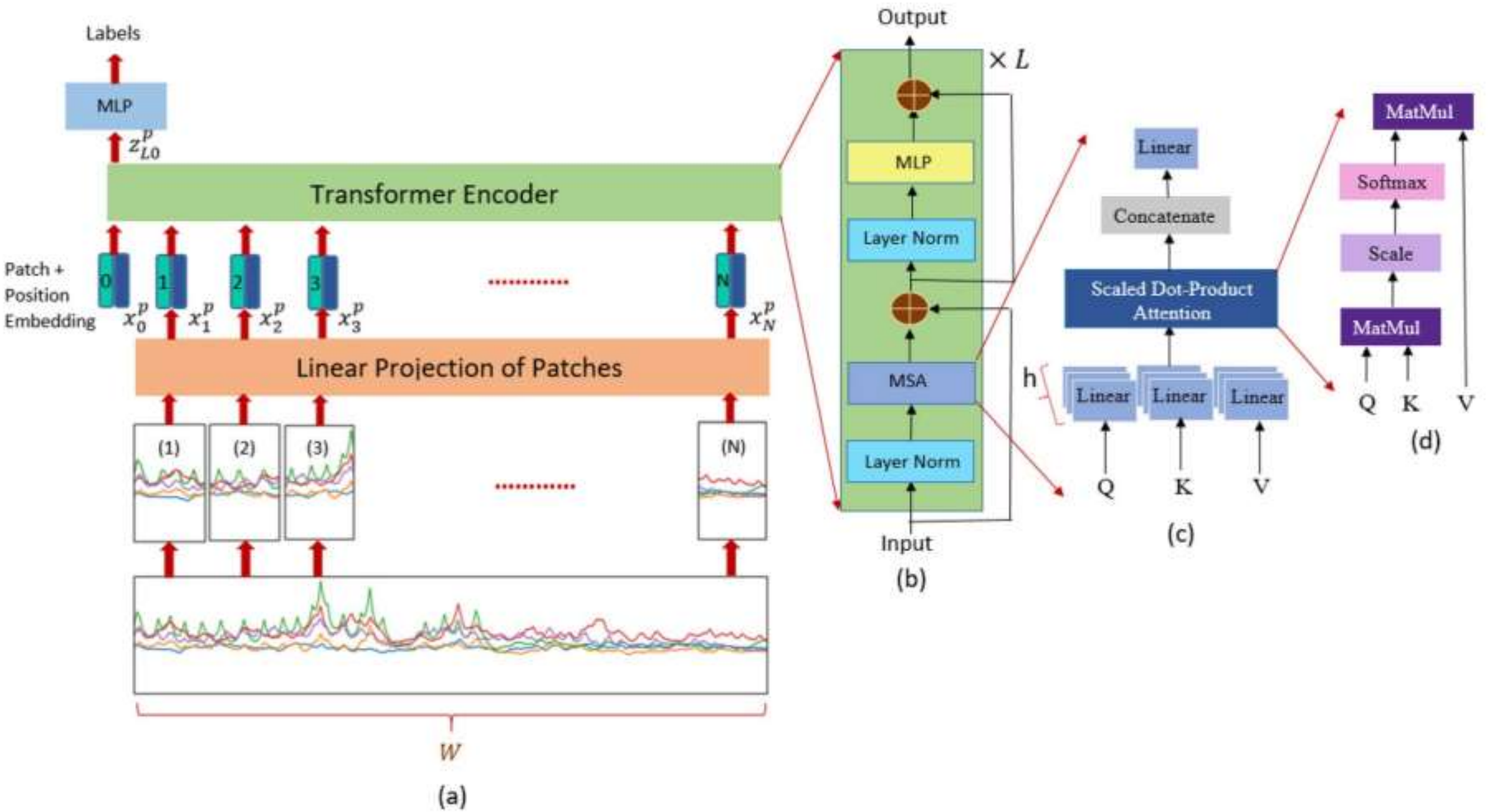


TEMGNet: Deep Transformer-based Decoding of Upper-limb sEMG for Hand Gestures Recognition

Window size	Model ID	Layers	Model dimension d	MLP size	Heads	Params
200ms	**1**	1	32	128	8	20,049
	2	2	32	128	8	32,657
	3	3	32	128	8	45,265
	4	1	64	256	8	64,625
300ms	**1**	1	32	128	8	20,593
	2	2	32	128	8	33,201
	3	3	32	128	8	45,809
	4	1	64	256	8	65,713

		200ms		300ms	
		Params	Accuracy (%)	Params	Accuracy (%)
Reference [8]	4-layer 3rd Order Dilation	–	79.0	466, 944	82.4
	4-layer 3rd Order Dilation (pure LSTM)	–	–	–	79.7
	SVM	–	26.9	–	30.7
Our Method	Model 1	**20,049**	80.39	**20,593**	80.88
	Model 4	64, 625	**82.05**	65, 713	**82.93**

Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals



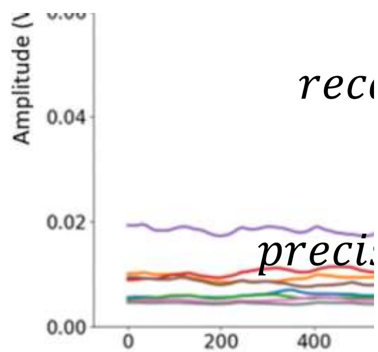
Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals

Reference	Window size (ms)	# Channels	Accuracy (%)	Train/Test Split
Ref[43]	200	128	84.6 (NA)	5-fold Cross Validation
CT-HGR-V1	250	128	91.98 (±2.22)	5-fold Cross Validation
CT-HGR-V2	250	128	92.88 (±2.10)	5-fold Cross Validation
Ref[41]	32	32	81.39 (±10.77)	NA
CT-HGR-V1	31.25	32	86.23 (±2.94)	5-fold Cross Validation
Ref[41]	256	128	96.14 (±4.67)	NA
CT-HGR-V1	250	128	91.98 (±2.22)	5-fold Cross Validation
CT-HGR-V2	250	128	92.88 (±2.10)	5-fold Cross Validation
Ref[42]	31.7	128	91.25 (±4.92)	NA
CT-HGR-V1	31.25	128	90.53 (±2.43)	5-fold Cross Validation
CT-HGR-V2	31.25	128	91.51 (±2.35)	5-fold Cross Validation
Ref[44]	32	128	94 (NA)	NA
CT-HGR-V1	31.25	128	90.53 (±2.43)	5-fold Cross Validation
CT-HGR-V2	31.25	128	91.51 (±2.35)	5-fold Cross Validation
Ref[44]	256	128	97.2 (NA)	NA
CT-HGR-V1	250	128	91.98 (±2.22)	5-fold Cross Validation
CT-HGR-V2	250	128	92.88 (±2.10)	5-fold Cross Validation

Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals

μ-law normalization

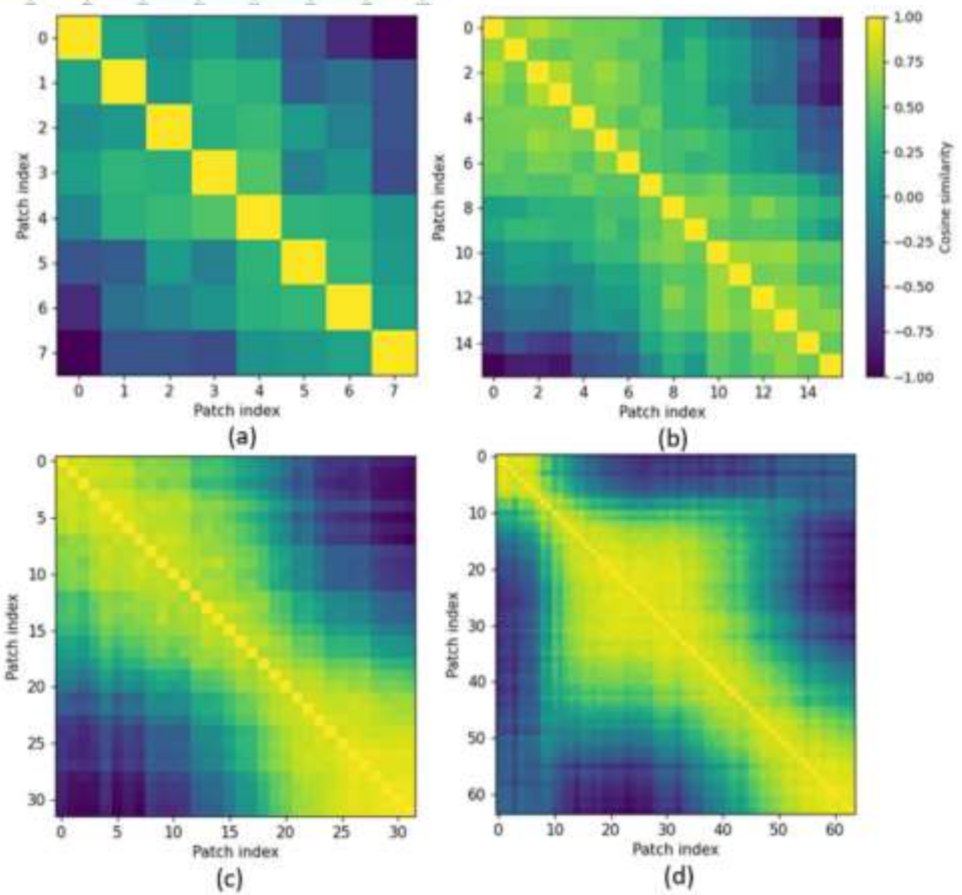
Cosine similarity



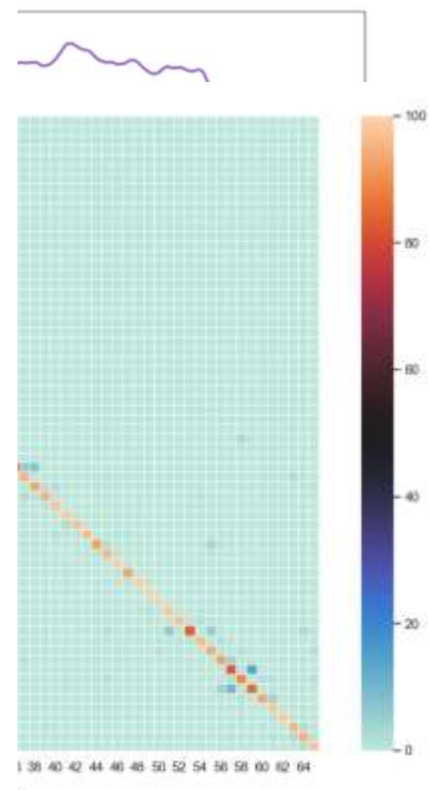
rec

$preci$

$f1score =$



(a) (b) (c) (d)

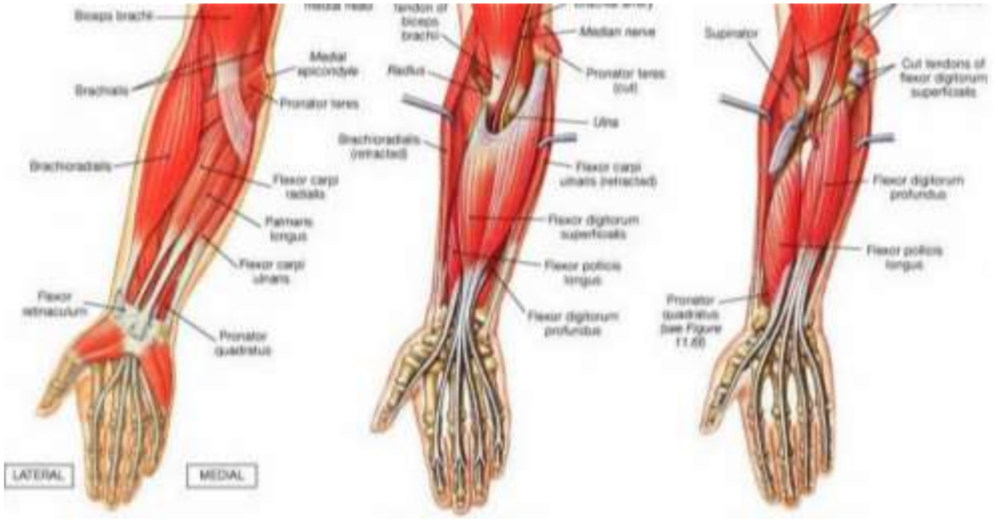


Summary

	Data set	Testing method	patch	Position embedding	optimizer
Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning	NinaPro DB2 NinaPro DB5	Each movement repeat 6 times. 1,3,4,6 training set 2,5 testing set	Overlapping	NO	Adam β1(N/A) β2(N/A) learning rate = 0.0001 (change as training) Weigh decay (N/A)
TEMGNet: Deep Transformer-based Decoding of Upper-limb sEMG for Hand Gestures Recognition	NinaPro DB2	Same as above	Non-overlapping	YES	Adam β1 = 0.9 β2 = 0.999 learning rate (N/A) Weigh decay = 0.001
Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals	(8x8)grid electrode	Each movement repeat 5 times. Randomly choose 1 for testing , the remains are training set	N/A	YES	Adam β1 = 0.9 β2 = 0.999 learning rate = 0.0001 Weigh decay = 0.001

Discussion

- Increasing the window size, and more utilizing channel can improve the accuracy, while the latter is more significant.
- Different kind of optimizer (Adam ,SDG)
- Statistic method(Confusion matrix, Wilcoxon signed-rank test, InterQuartile Range, f1-score, cosine similarity)
- Knowledge about upper-limb can be important in the primary electrode position.[5]



Reference – Overview

[1] A Low-Cost, Wireless, 3-D-Printed Custom Armband for sEMG Hand Gesture Recognition

[2] Surface-Electromyography-Based Gesture Recognition by Multi-View Deep Learning

[3]Movements Classification Through sEMG With Convolutional Vision Transformer and Stacking Ensemble Learning

[4] TEMGNet: Deep Transformer-based Decoding of Upperlimb sEMG for Hand Gestures Recognition

[5] Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals | Scientific Reports (nature.com)

[6] HYDRA-HGR: A Hybrid Transformer-Based Architecture for Fusion of Macroscopic and Microscopic Neural Drive Information

[7] TraHGR: Transformer for Hand Gesture Recognition via ElectroMyography

Undergraduate Group

TraHGR

Presenter: Miguel

Teammates: Shawn, Howard

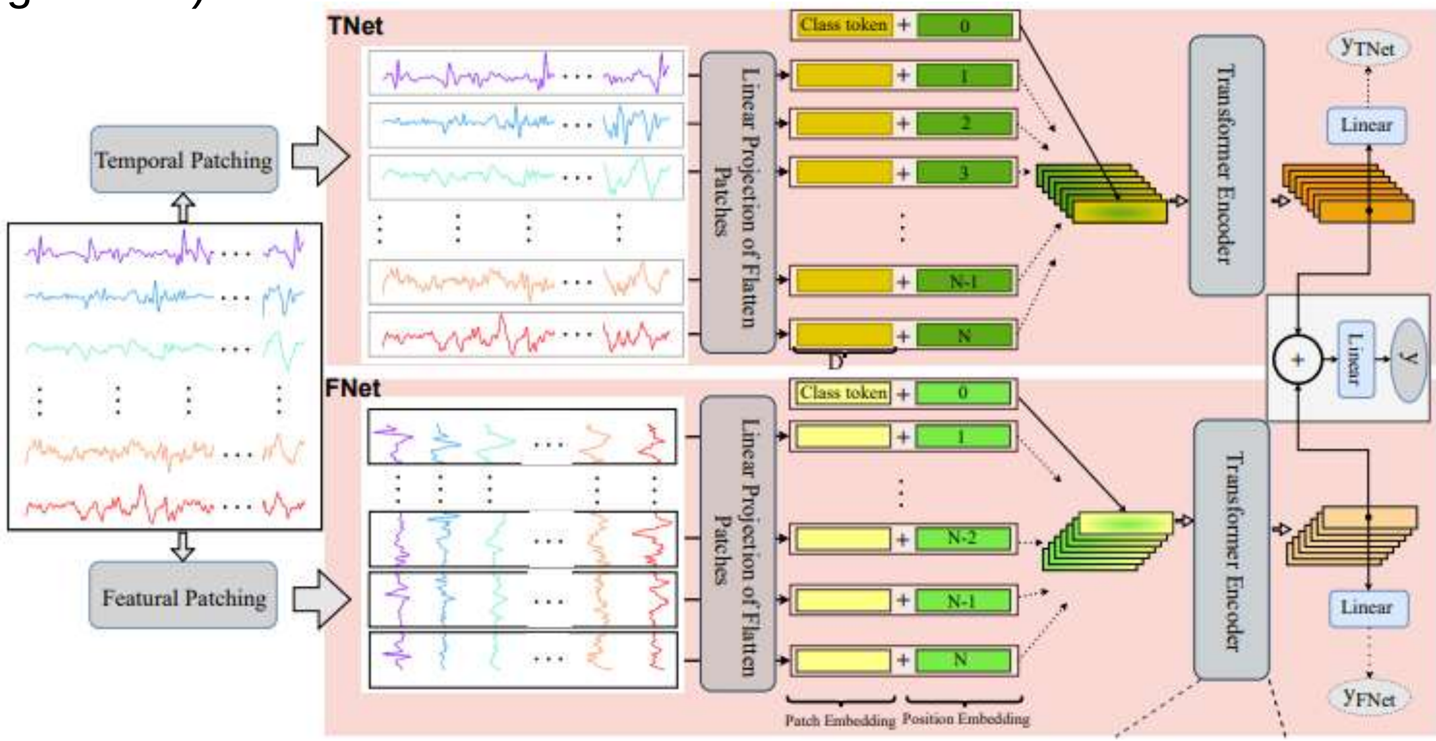
Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/4

The proposed TraHGR architecture

- Database: DB2
  - DB2-B (17 gestures)
  - DB2-C (23 gestures)
  - DB2-D (9 gestures)
- Embedded Patches
  - Temporal
  - Featural

$$\mathbf{z}_0 = [\mathbf{x}_{\text{class}}; \mathbf{x}_p^1\mathbf{E}; \mathbf{x}_p^2\mathbf{E}; \ldots; \mathbf{x}_p^N\mathbf{E}] + \mathbf{E}_{pos}$$



Transformer Encoder & Output

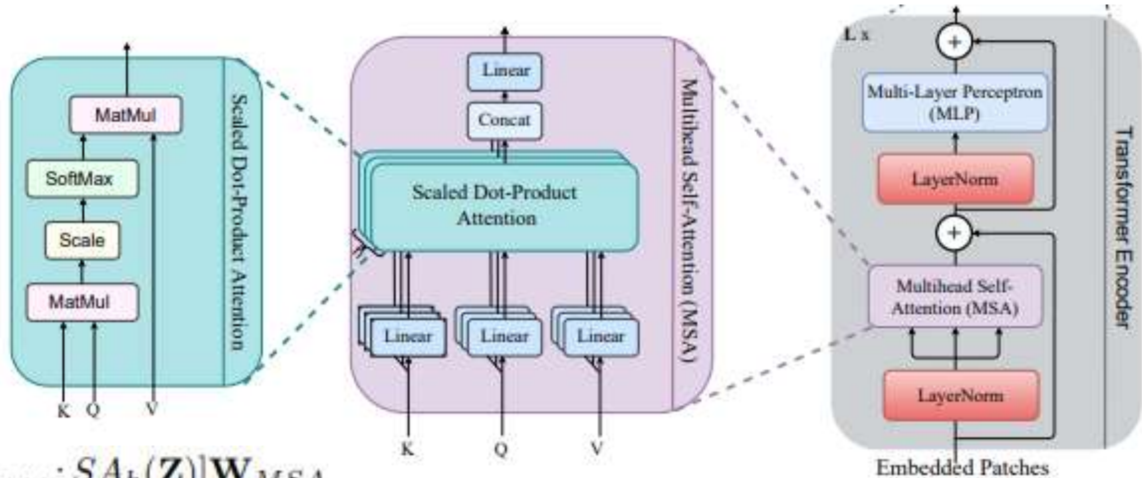
- MLP
  - Two linear layers
  - the first layer is followed by Gaussian Error Linear Unit (GELU) activation function
- MSA

$$[\mathbf{Q},\mathbf{K},\mathbf{V}] = \mathbf{Z}\mathbf{W}_{QKV}$$

$$\mathbf{P} = softmax(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{D_h}})$$

$$SA(\mathbf{Z}) = \mathbf{P}\mathbf{V}$$

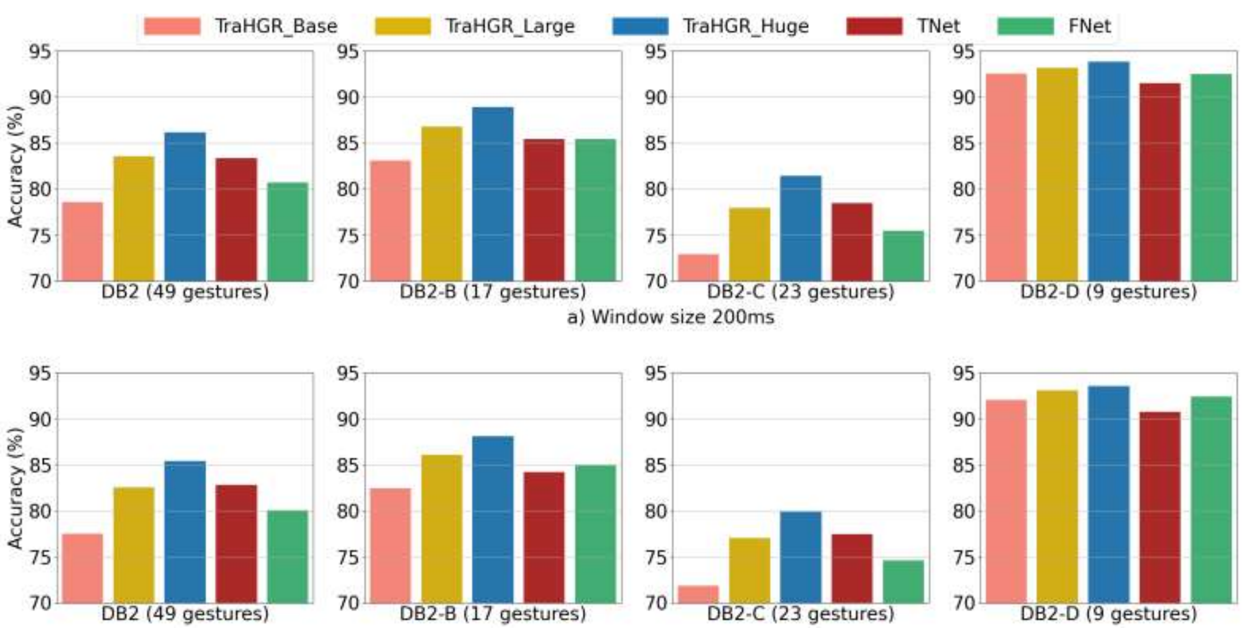
$$MSA(\mathbf{Z}) = [SA_1(\mathbf{Z}); SA_2(\mathbf{Z}); \ldots; SA_h(\mathbf{Z})]\mathbf{W}_{MSA}$$

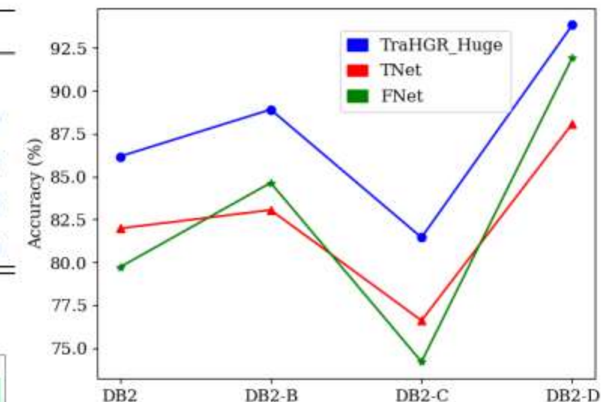


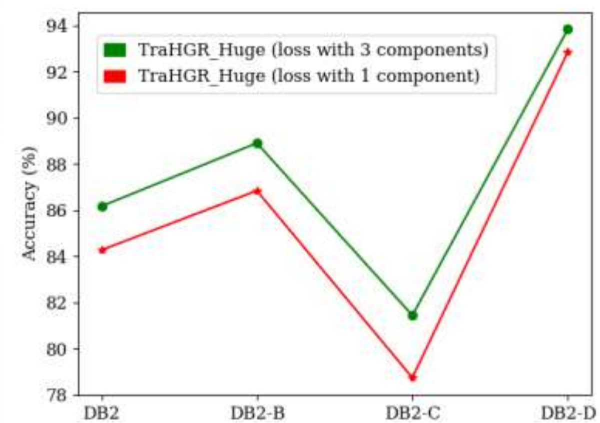
- The output of each path → $y_{path} = Linear(LayerNorm(\mathbf{Z}_L^0)_{path})$
- The output of the TraHGR → $y = Linear(LayerNorm[(\mathbf{Z}_L^0)_{\mathrm{TNet}} + (\mathbf{Z}_L^0)_{\mathrm{FNet}}])$

Experiments and Results

Model	Layers (L)	Model dimension (D)	MLP size	Number of heads (h)	Params		
					200ms	150ms	100ms
TraHGR-Base	1	32	128	4	83,731	74,259	63,603
TraHGR-large	2	64	256	4	316,051	297,107	275,795
TraHGR-Huge	1	144	720	8	846,579	803,955	756,003
TNet	1	144	720	8	472,513	431,041	384,385
FNet	1	144	720	8	366,673	365,521	364,225







$$\mathcal{L} = \mathcal{L}_{\text{TNet}} + \mathcal{L}_{\text{FNet}} + \mathcal{L}_{\text{TraHGR}}$$

$$\mathcal{L} = \mathcal{L}_{\text{TraHGR}}$$

<u>Undergraduate Group</u>

HYDRA-HGR

Presenter: Miguel

Teammates: Shawn, Howard

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/4

The proposed HYDRA-HGR architecture

- Macro Path
  - Data: one dimension in time and two dimensions in space

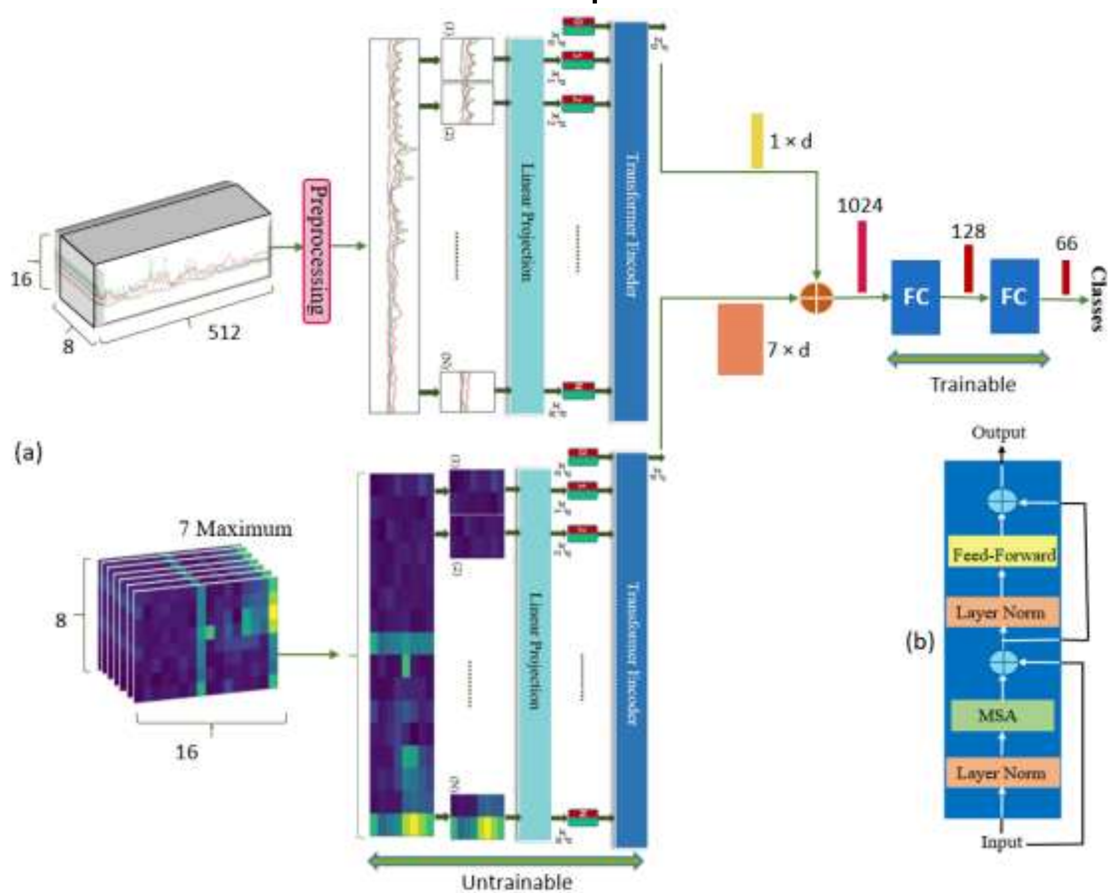
low-pass butterworth filter

↓

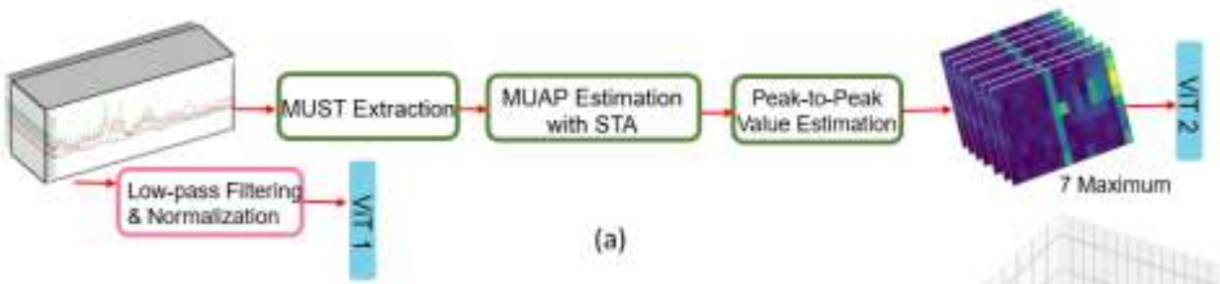
normalized through the μ-law algorithm

↓

split into windows of size 512 (250 ms) with a skip step of 256



- Micro Path
  - The basic BSS assumption $\boldsymbol{X}(t) = \Sigma_{l=0}^{L-1} \boldsymbol{H}(l)\boldsymbol{S}(t-l) + \underline{\boldsymbol{\nu}}(t)$
    - gradient Convolution Kernel Compensation (gCKC) $\hat{\boldsymbol{s}}_j(t) = \hat{\boldsymbol{c}}_{\boldsymbol{s}_j x}^T \boldsymbol{C}_{xx}^{-1} \boldsymbol{x}(t)$
    - fast Independent Component Analysis (fastICA) $\hat{\boldsymbol{s}}_j(t) = w_j^T(k)\boldsymbol{Z}(t)$
  - The Spike-Triggered Averaging (STA) technique



Experiments and Results

Model's Name	Fold1(%)	Fold2	Fold3	Fold4	Fold5	Average
Stand-alone Macro Model	79.92 (±3.39)	91.43 (±2.48)	93.84 (±2.05)	92.57 (±2.28)	88.96 (±2.83)	89.34 (±2.61)
Stand-alone Micro Model	81.53 (±3.45)	88.03 (±2.66)	89.63 (±2.39)	89.11 (±4.02)	84.92 (±2.97)	86.64 (±3.10)
The HYDRA-HGR	**89.38 (±2.88)**	**96.86 (±1.82)**	**96.82 (±1.75)**	**96.65 (±2.75)**	**94.61 (±1.90)**	**94.86 (±2.22)**

Summary

	TraHGR	HYDRA-HGR
Dataset	DB2 (sparse multi-channel sEMG)	65 isometric hand movements (HD-sEMG)
Architecture	Hybrid (two parallel paths of ViT)	
Transformer Encoder	Layer Norm MSA Layer Norm MLP	Layer Norm MSA Layer Norm Feed-Forward
Accuracy	86.18%(±4.99%) ☹	94.86%(±2.22%) ☺
Complexity	Low ☺	High ☹

Reference – Overview

[1] Soheil Zabihi, Elahe Rahimian, Amir Asif, and Arash Mohammadi , “TraHGR: Transformer for Hand Gesture Recognition via ElectroMyography”, arXiv:2203.16336v2 [eess.SP] 31 Mar 2022

[2] Mansooreh Montazerin, Elahe Rahimian, Farnoosh Naderkhani, S. Farokh Atashzar, Hamid Alinejad-Rokny, Arash Mohammadi , “HYDRA-HGR: A HYBRID TRANSFORMER-BASED ARCHITECTURE FOR FUSION OF MACROSCOPIC AND MICROSCOPIC NEURAL DRIVE INFORMATION”, ICASSP 2023