A Layered sEMG–FMG Hybrid Sensor for Hand Motion Recognition From Forearm Muscle Activities

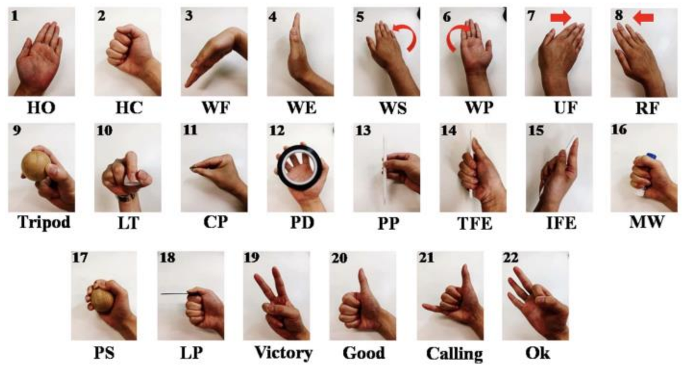
Presenter: Shawn

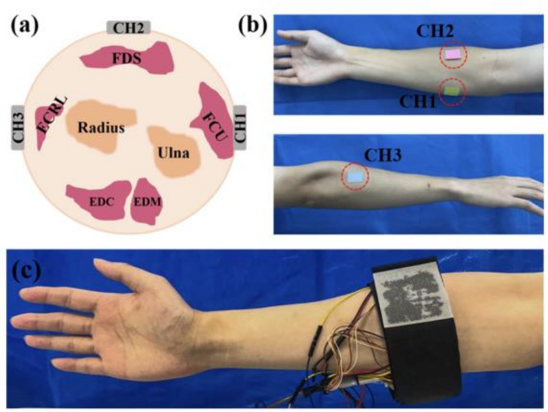
Advisor: Prof. An-Yeu (Andy) Wu

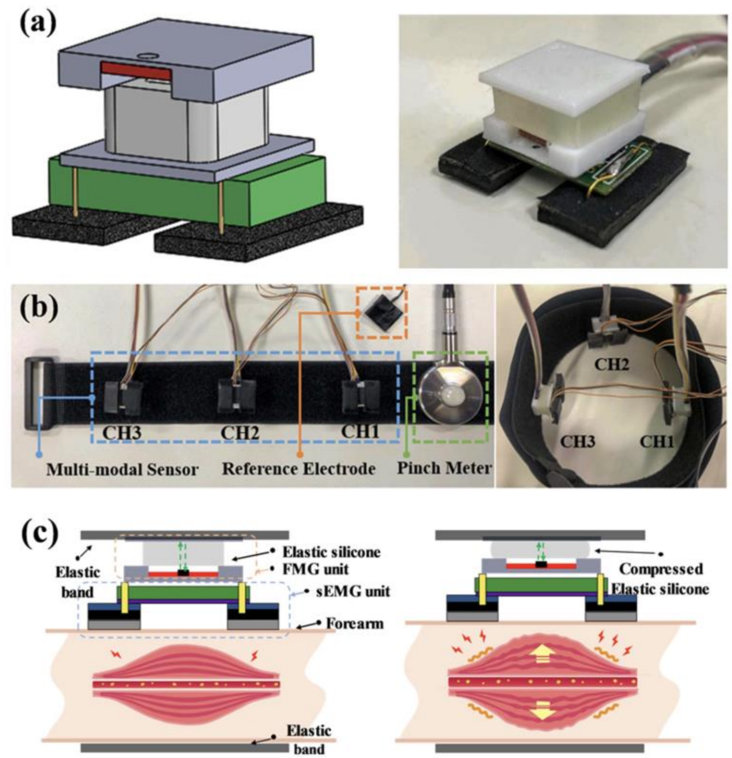
Date: 2023/10/20

Introduction

- The sEMG complicated measurement system induces both physical and mental stress in subjects, leading to classification degradation
- FMG is more stable than sEMG on prosthesis control







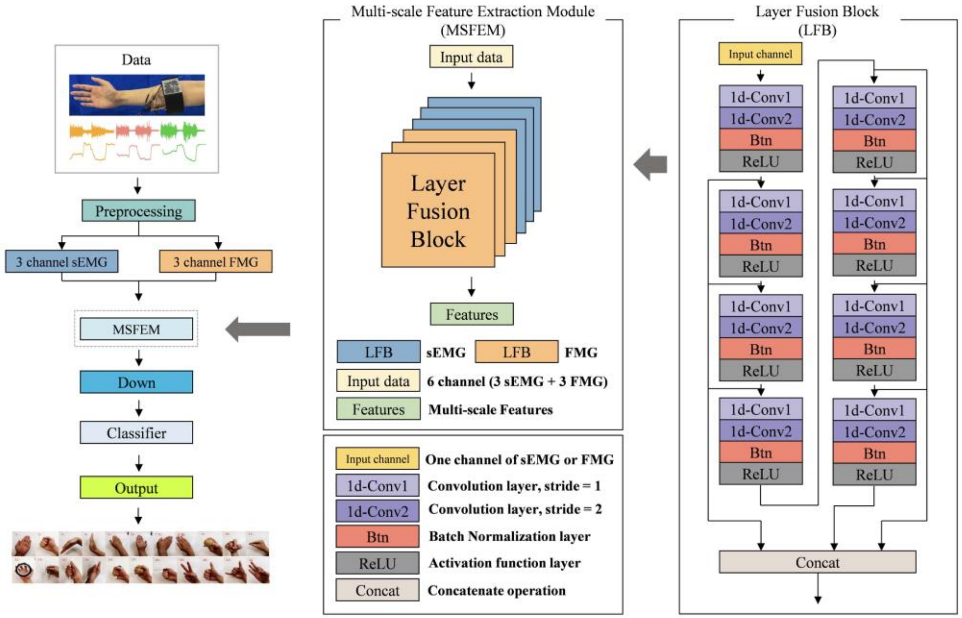
Layer Fusion CNN (LFC)

Shallow layer => low-level features
Deep layer => high-level features

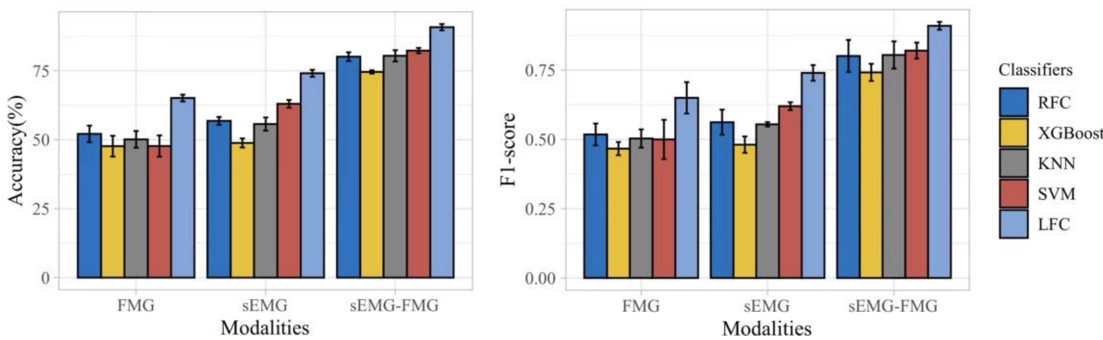
	HO	HC	WF	WE	WS	WP	UF	RF	Tripod	LT	CP	PD	PP	TFE	IFE	MW	PS	LP	Victory	Good	Calling	Ok
scale1 (CNN)	0.62	0.73	0.73	0.7	0.71	0.73	0.74	0.63	0.48	0.54	0.54	0.55	0.58	0.55	0.55	0.71	0.73	0.73	0.71	0.7	0.69	0.7
scale2 (CNN)	0.72	0.9	0.9	0.93	0.88	0.91	0.84	0.81	0.74	0.7	0.66	0.65	0.68	0.75	0.71	0.79	0.76	0.78	0.82	0.8	0.85	0.82
scale3 (CNN)	0.85	0.89	0.89	0.94	0.88	0.91	0.88	0.85	0.7	0.76	0.73	0.77	0.74	0.74	0.82	0.77	0.78	0.84	0.86	0.81	0.9	0.86
scale4 (CNN)	0.9	0.88	0.92	0.95	0.91	0.92	0.85	0.85	0.75	0.74	0.76	0.77	0.85	0.78	0.8	0.79	0.81	0.86	0.88	0.87	0.91	0.84
scale5 (CNN)	0.9	0.95	0.92	0.95	0.9	0.95	0.88	0.86	0.79	0.78	0.76	0.73	0.77	0.83	0.81	0.84	0.8	0.82	0.88	0.82	0.96	0.85
scale6 (CNN)	0.9	0.91	0.92	0.93	0.88	0.94	0.88	0.86	0.76	0.84	0.78	0.76	0.79	0.82	0.85	0.85	0.79	0.9	0.88	0.82	0.94	0.86
scale7 (CNN)	0.86	0.91	0.93	0.92	0.88	0.88	0.84	0.82	0.72	0.75	0.8	0.76	0.74	0.82	0.83	0.84	0.86	0.81	0.86	0.8	0.93	0.82
scale8 (CNN)	0.84	0.9	0.93	0.91	0.89	0.89	0.82	0.81	0.81	0.74	0.82	0.73	0.76	0.85	0.87	0.83	0.82	0.82	0.82	0.83	0.93	0.86
all scales (LFC)	0.94	0.96	0.97	0.96	0.94	0.94	0.96	0.9	0.88	0.83	0.9	0.86	0.82	0.89	0.89	0.88	0.86	0.91	0.91	0.9	0.96	0.93

High-level features are not always the optimal scale

The layer-fusion structure allows the extraction of different-scale representation from the input signals

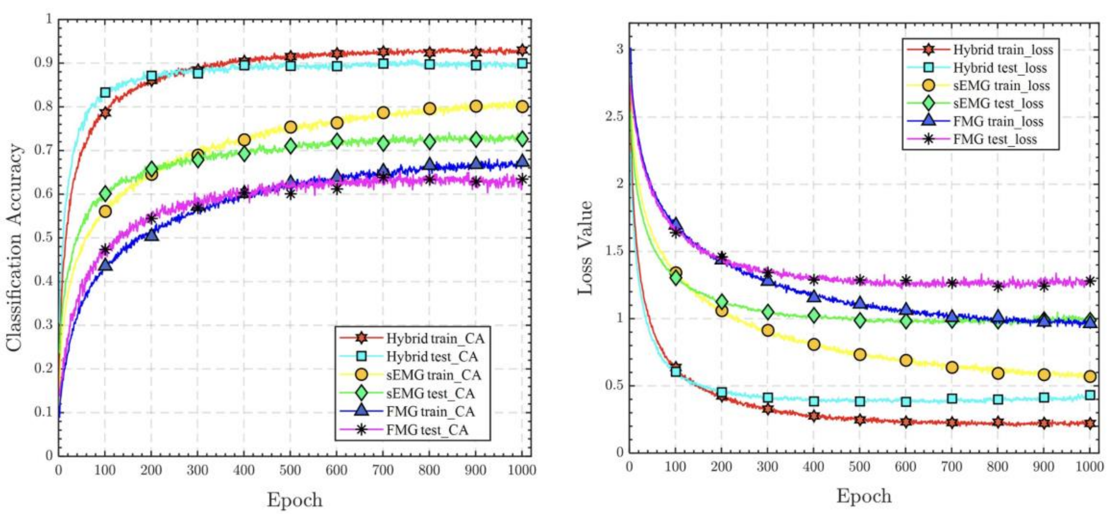


Result



Classifiers	Modalities	Mean ± SD
RFC	sEMG	61.69% ± 8.08%
	FMG	58.5% ± 5.51%
	Hybrid	81.62% ± 3.91%
XGBoost	sEMG	59.25% ± 6.22%
	FMG	57.53% ± 4.73%
	Hybrid	75.79% ± 3.5%
KNN	sEMG	59.22% ± 6.2%
	FMG	53.53% ± 7.66%
	Hybrid	81.58% ± 7.32%
SVM	sEMG	65.34% ± 4.75%
	FMG	57.57% ± 4.6%
	Hybrid	86.4% ± 4.83%
LFC	sEMG	78.41% ± 5.39%
	FMG	74.83% ± 6.22%
	Hybrid	93.06% ± 3.26%

The proposed LFC outperforms conventional ML methods



In Hybrid mode
the Mean is higher
the SD is lower

Hybrid mode performs better

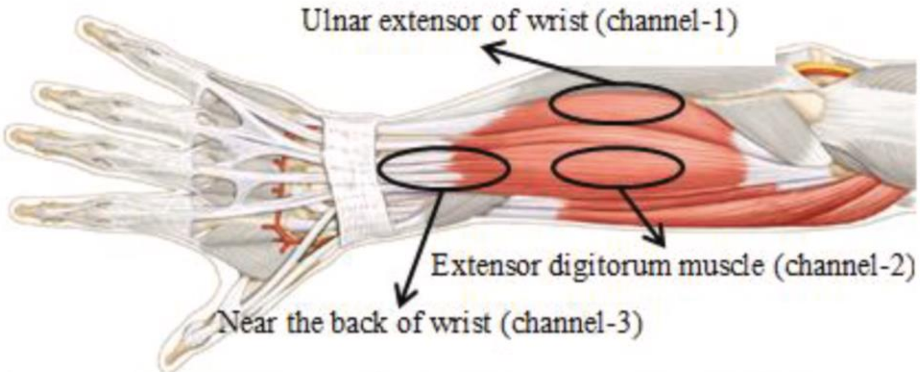
Prediction and classification of sEMG-based pinch force between different fingers

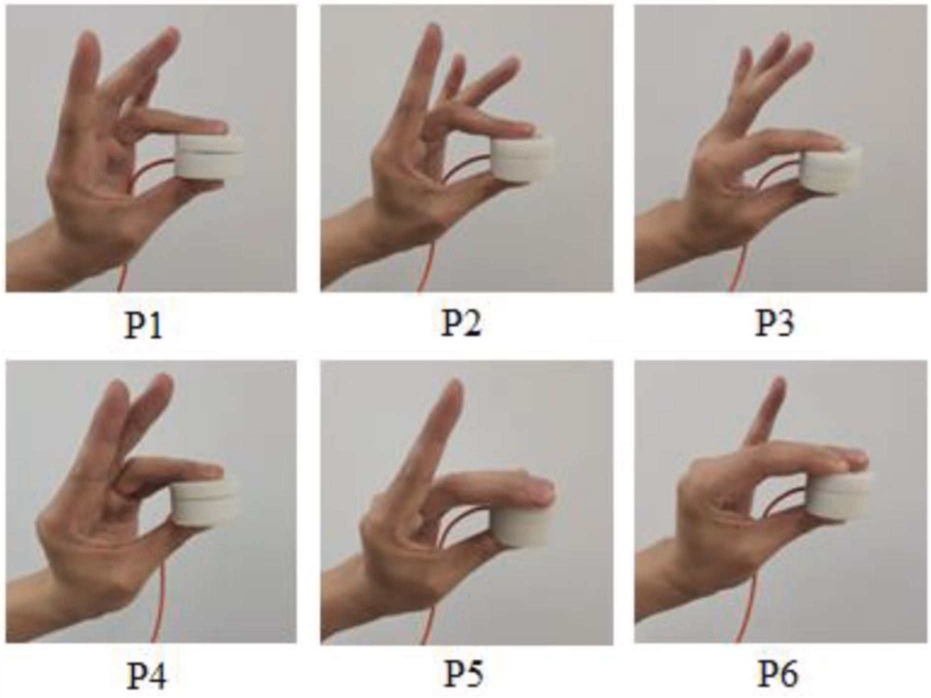
Presenter: Shawn

Advisor: Prof. An-Yeu (Andy) Wu

Date: 2023/10/20

Data Acquisition

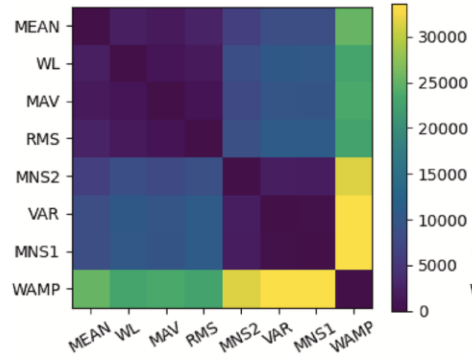




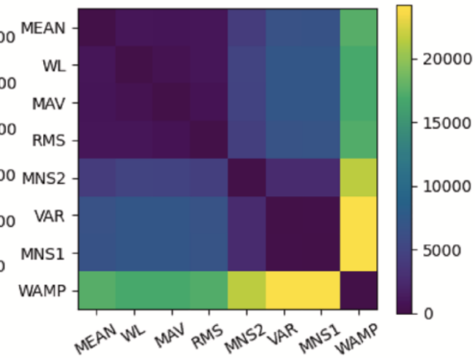
sEMG Signal Preprocessing

- Window size : 100 (ms), Step size : 50 (ms)
- Manhattan distance is used to evaluate similarity

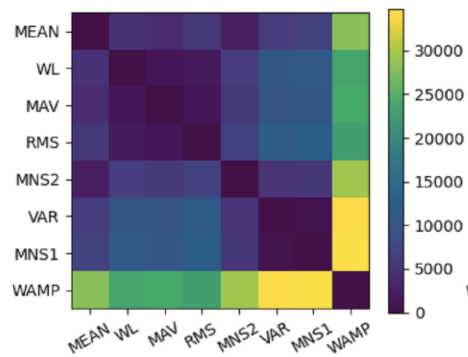
Abbreviation	Full feature name	Detailed theoretical calculation
WL	Waveform Length	$\sum_{i=1}^{N-1}\lvert x_{i+1}-x_i\rvert$
MAV	Mean Absolute Value	$\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert$
RMS	Root Mean Square	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2}$
VAR	Variance of EMG	$\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$
WAMP	Willison Amplitude	$\sum_{i=1}^{N-1}f(i),\ f(i)=\begin{cases}1, & if i \geqslant th\\ 0, & otherwise\end{cases},\ i=\lvert x_i-x_{i+1}\rvert$
MNS1	mean spectrum of 0–500 Hz	$\frac{1}{M}\sum_{i=1}^{M}P_i, 0\le P_i\le 500$
MNS2	mean spectrum of 0–350 Hz	$\frac{1}{M}\sum_{i=1}^{M}P_i, 0\le P_i\le 350$
MEAN	Mean of sEMG amplitude	$\frac{1}{N}\sum_{i=1}^{N}x_i$



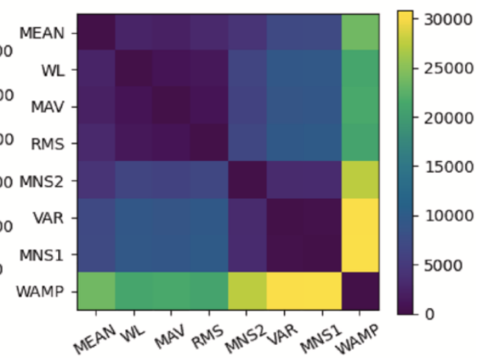
(a)



(b)



(c)

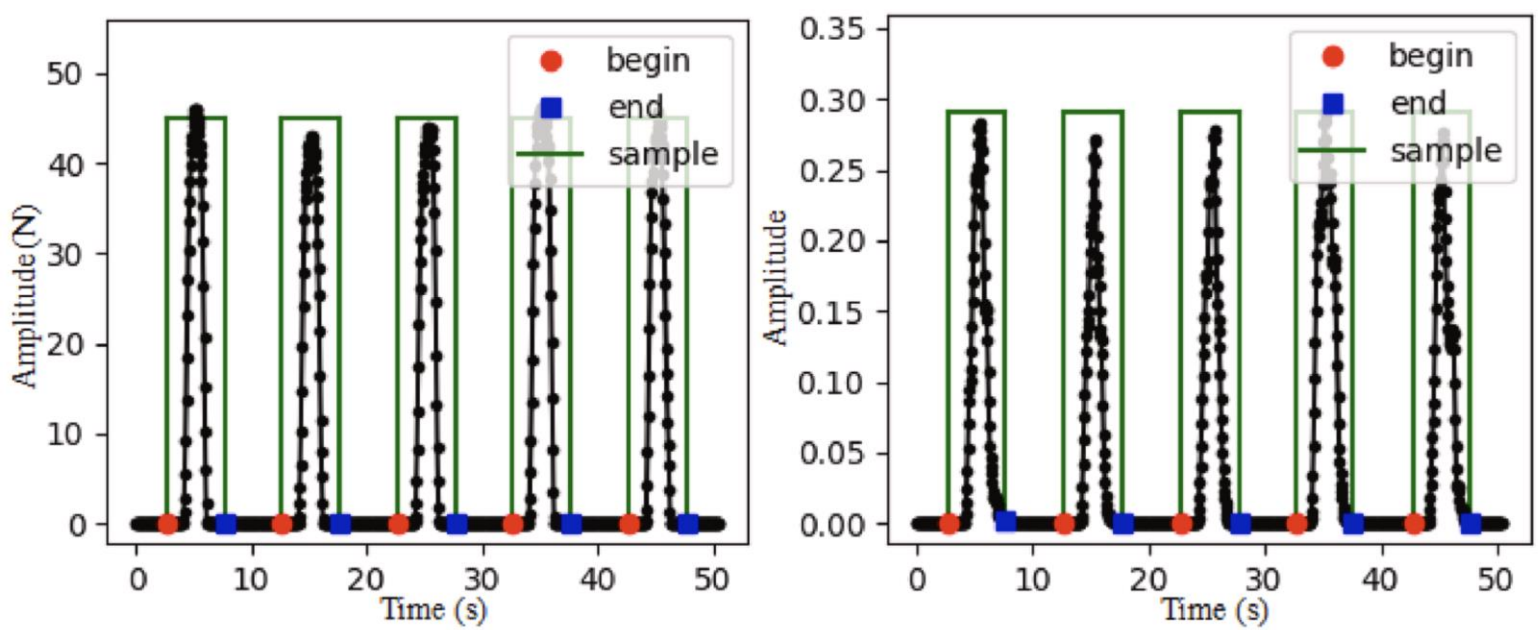


(d)

MEAN, VAR, WAMP are chosen

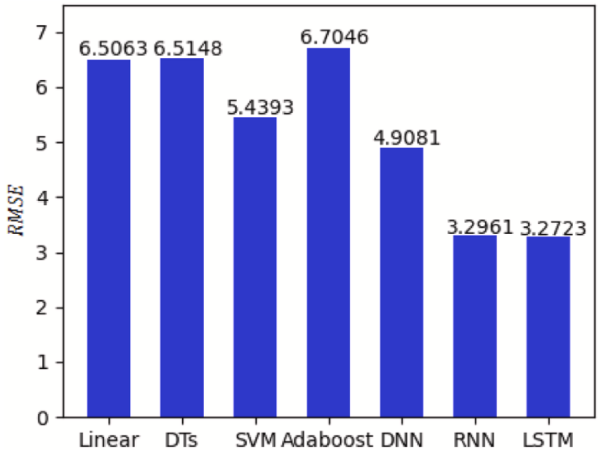
Data Augmentation

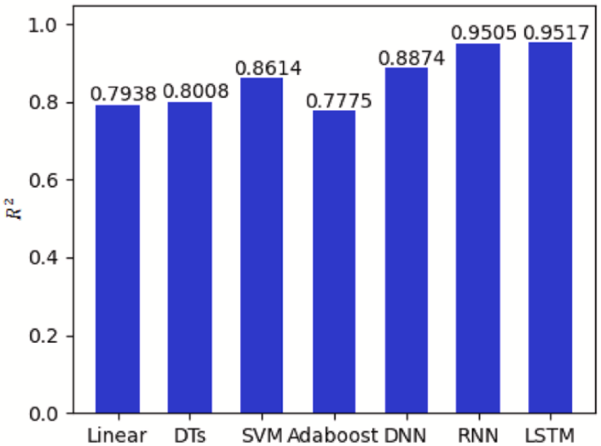
- Expand/Reduce the data length for each sample sequence to 5 (s)
- Sample 100 data points
- Normalize amplitude to [0,1]



(a) (b)

Prediction of Pinch Force

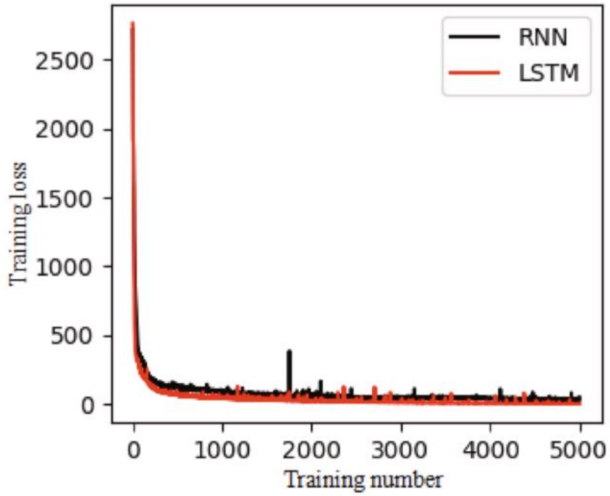




$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\bar{y}_i - y_i)^2}{n}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\bar{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

RNN and LSTM perform better



LSTM
1. Converges slightly faster
2. Is more stable
3. Needs much more time to train

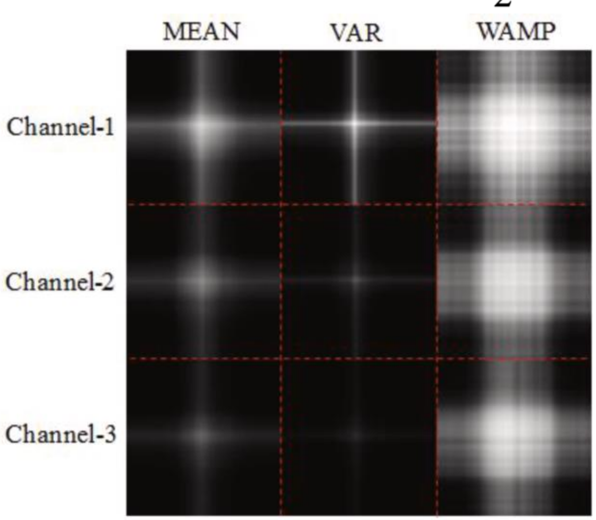
Image Transformation of sEMG signal

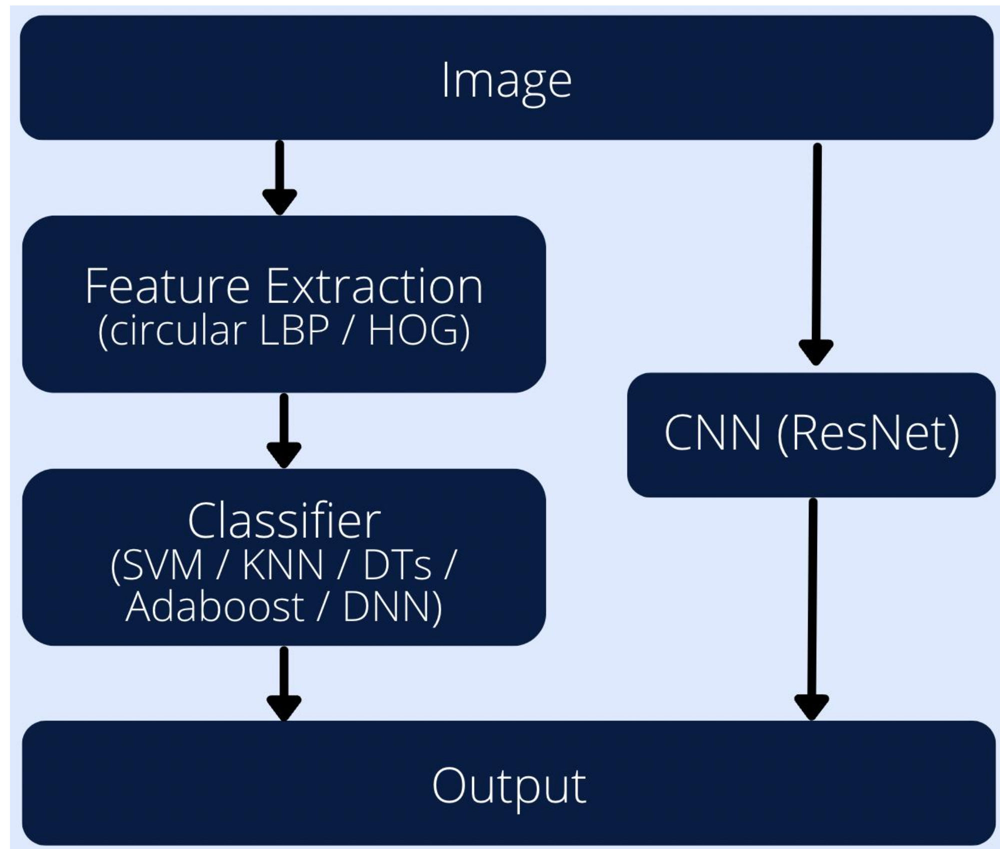
- Why ? Pinch Force → ? → Finger Movement

$$\varphi = cos^{-1}x$$

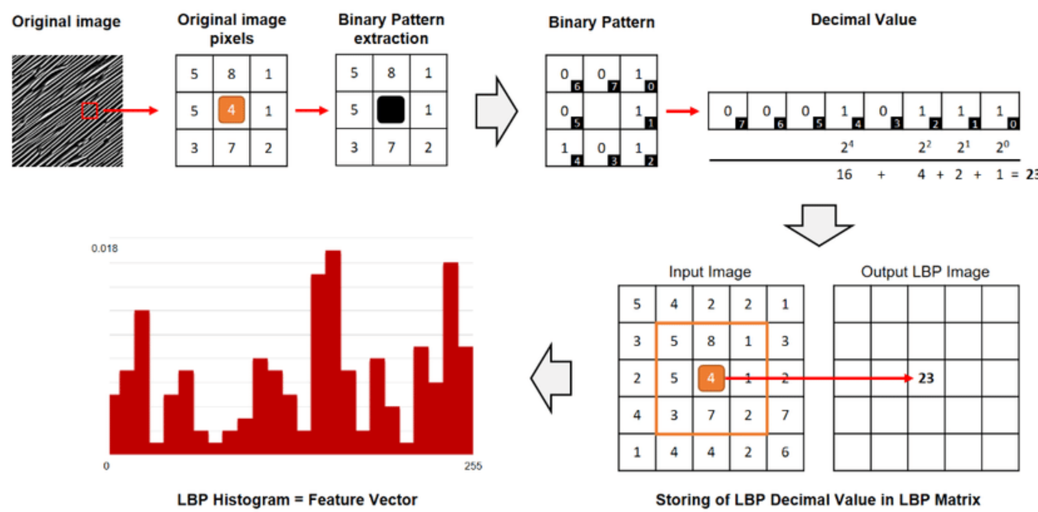
$$M = \begin{vmatrix} \cos(\varphi_1+\varphi_1) & \cos(\varphi_1+\varphi_2) & \cdots & \cos(\varphi_1+\varphi_n) \\ \cos(\varphi_2+\varphi_1) & \cos(\varphi_2+\varphi_2) & \cdots & \cos(\varphi_2+\varphi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\varphi_n+\varphi_1) & \cos(\varphi_n+\varphi_2) & \cdots & \cos(\varphi_n+\varphi_n) \end{vmatrix}$$

$$normalized\ pixel\ matrix = \frac{(M+1)}{2}$$

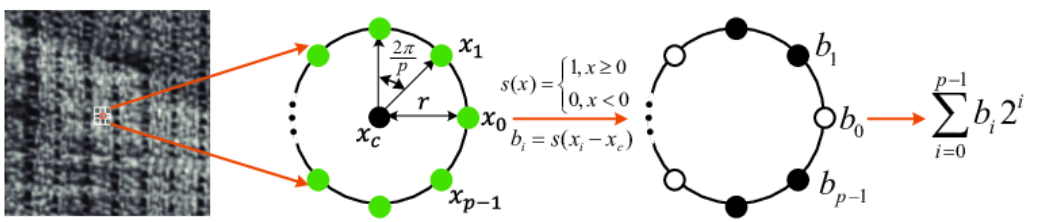


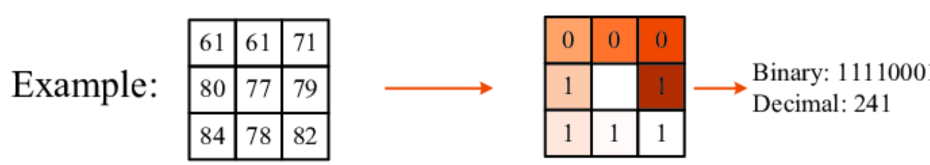


LBP vs Circular LBP



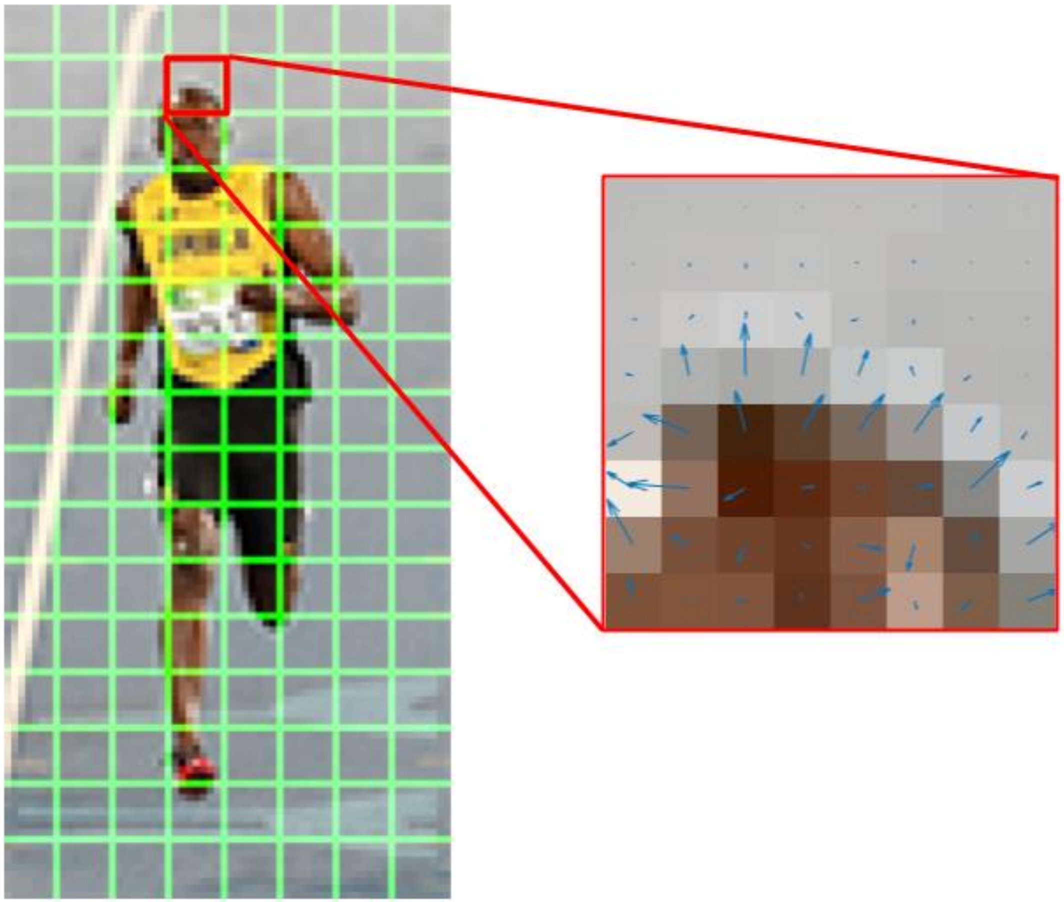
LBP captures the local patterns





Circular LBP captures the local patterns with more area

HOG



2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	21	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

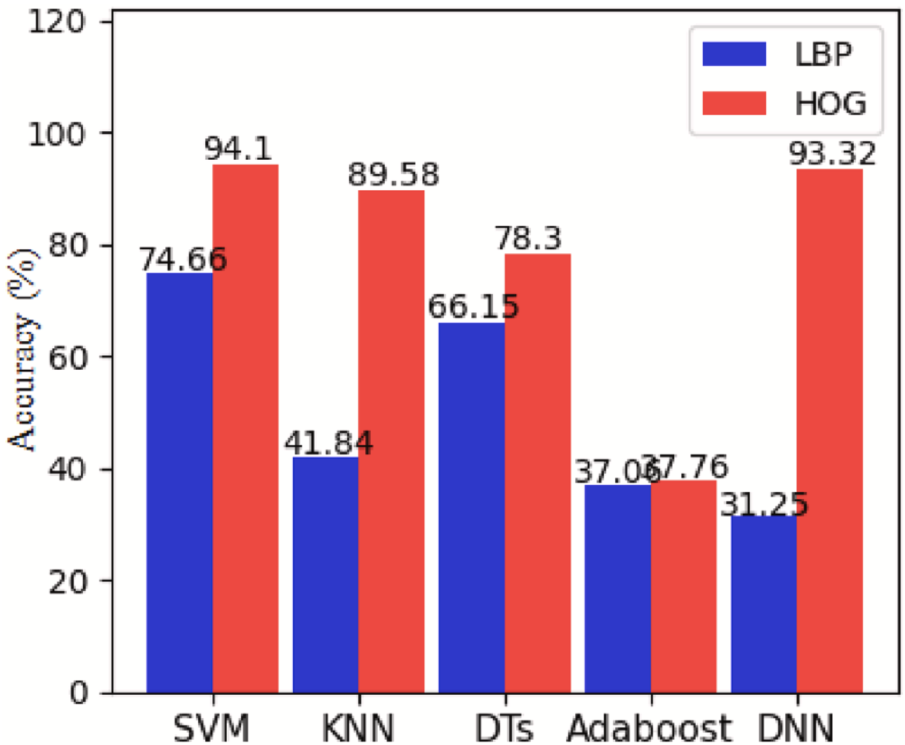
Gradient Magnitude

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

Gradient Direction

HOG is great at capturing edges and corners in images

Result



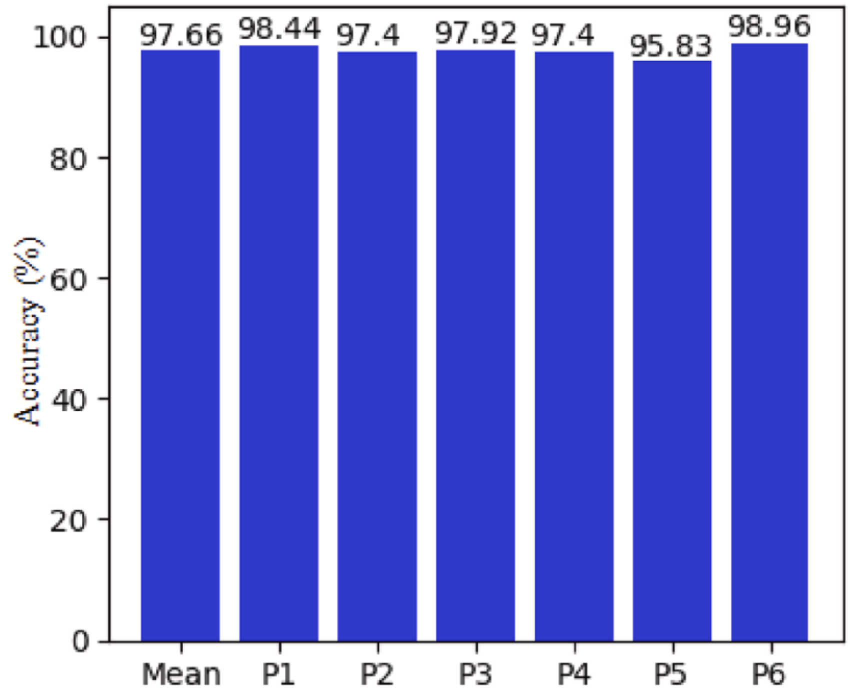


Fig. 17. The classification accuracy of CNN.

CNN
1. Outperforms the other classifiers
2. Needs large computational resources

Undergraduate Group

Paper survey

Transformer on sEMG Hand Gesture Recognition

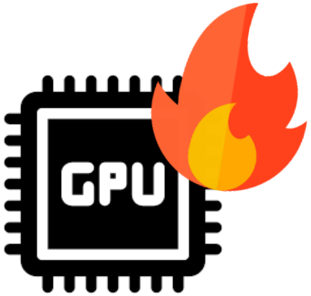
Presenter: Howard

Advisor: Prof. An-Yeu (Andy) Wu

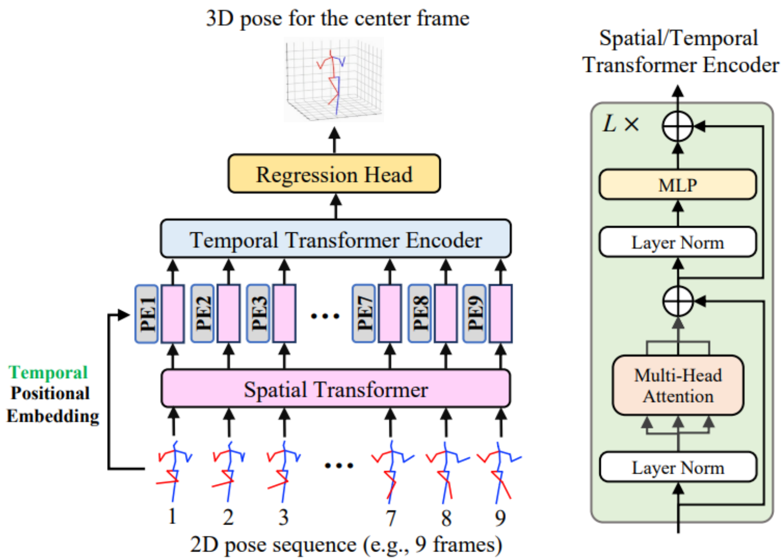
Date: 2023/10/22

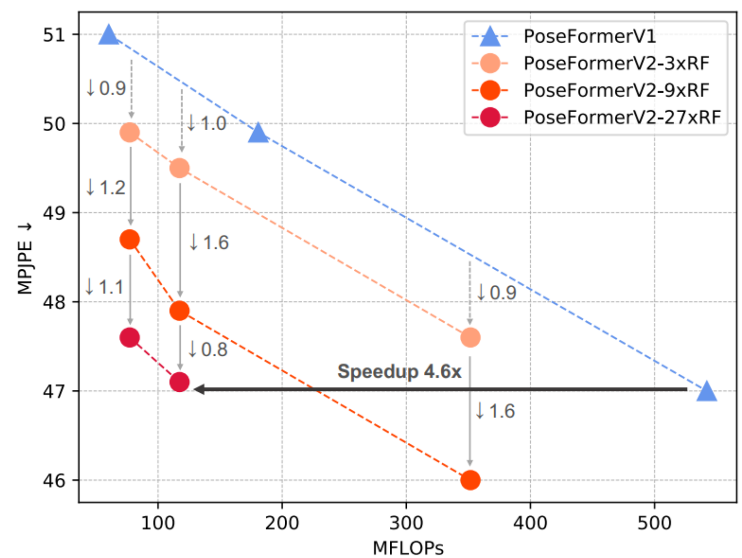
Outline

- Review of three papers
  - PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation
  - LST-EMG-Net: Long short-term transformer feature fusion network for sEMG gesture recognition
- Summary



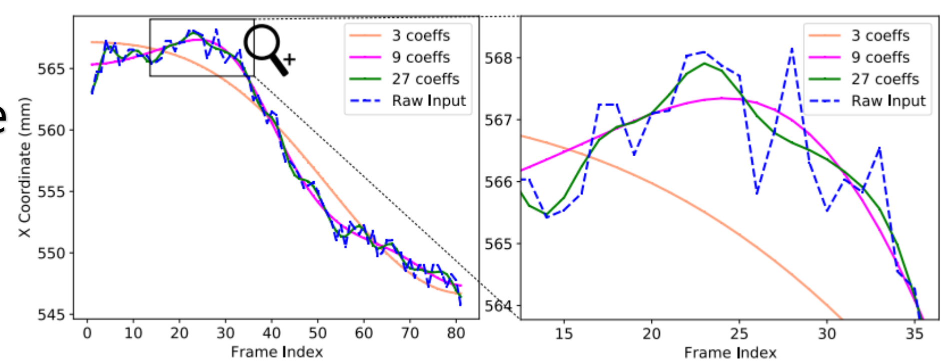
Original Poseformer



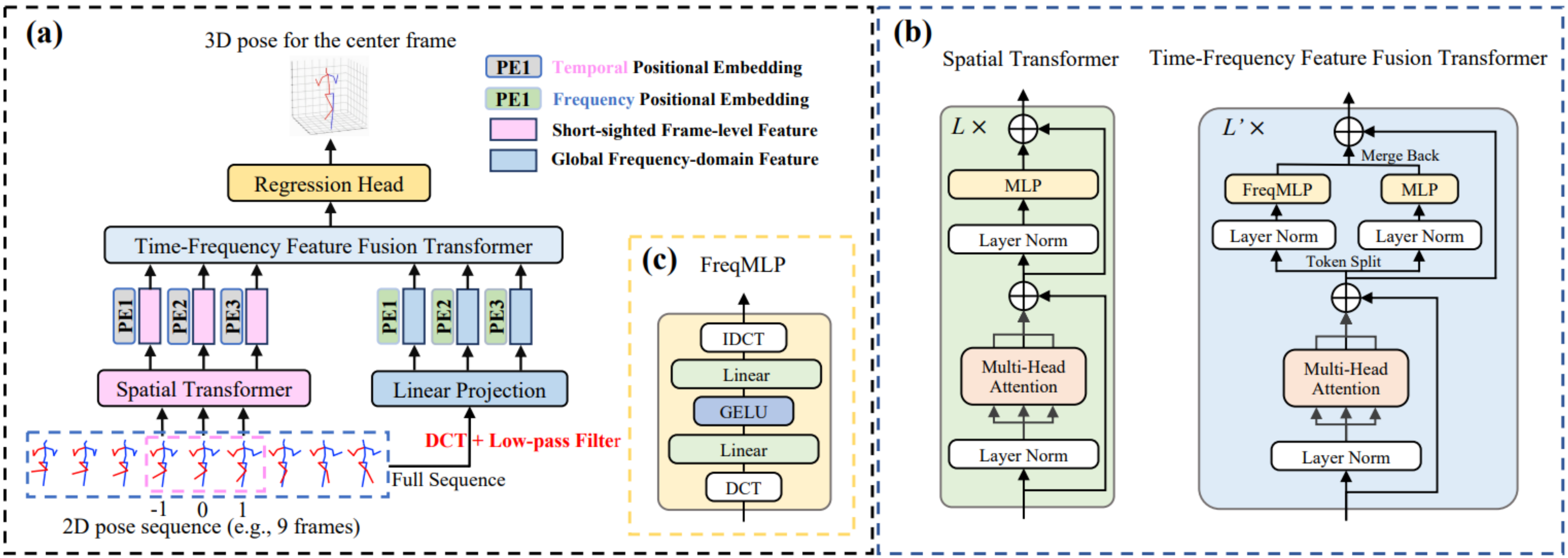


- More sequence → high accuracy
  - Less detail → avoid high freq noise

- Tradeoff between performance drop and computational cost



Method



Spatial transformer → single frame joint correlation modeling

Temporal encoder → cross-frame human motion modeling

DCT (Discrete Cosine Transform) → a portion of low-freq coefficient

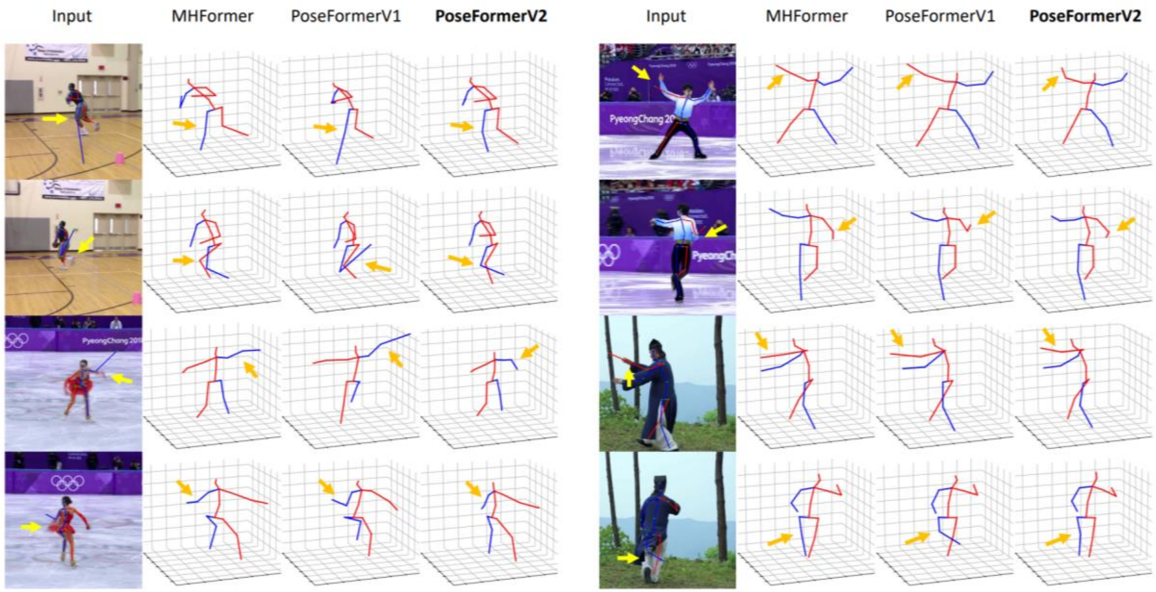
Time-Frequency Feature Fusion Transformer

$$\mathbf{z}_k^{'} = \mathrm{MSA}(\mathbf{z}_k),$$

$$\mathbf{z}_k^{Time}, \mathbf{z}_k^{Freq} = \mathbf{z}_k^{'}[:F'], \mathbf{z}_k^{'}[F':],$$

$$\mathbf{z}_{k+1} = \mathrm{Concat}(\mathrm{FreqMLP}(\mathbf{z}_k^{Time}), \mathrm{MLP}(\mathbf{z}_k^{Freq})),$$

Result

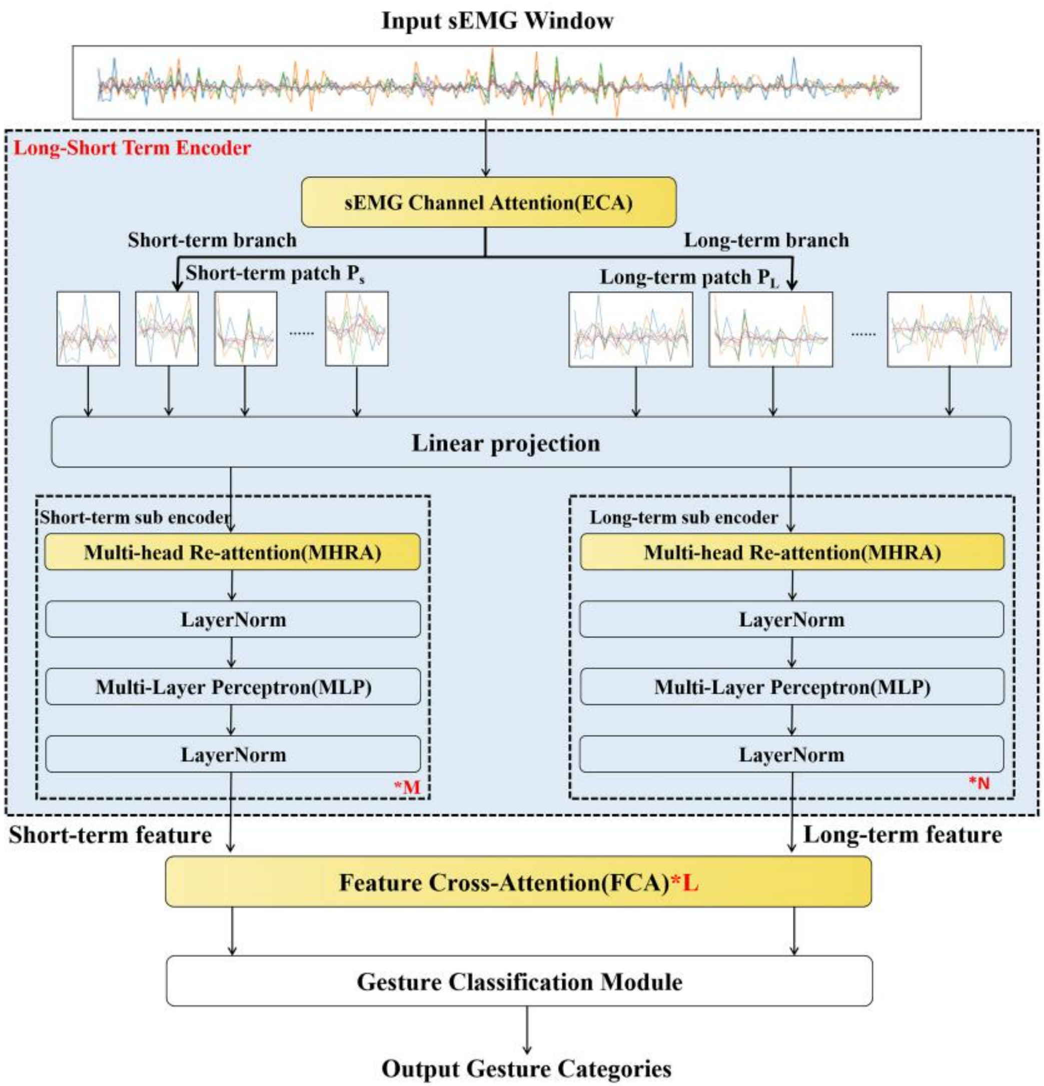


Method	f	Seq. Len.	MFLOPs	MPJPE ↓ / P-MPJPE ↓
PoseFormerV1 [41] ICCV'21	27	27	542.1	47.0/–
StridedTrans. [14] TMM'22	81	81	342.5	47.5/–
MixSTE [40](†) CVPR'22	3	3	3420	49.6/38.9
MHFormer [15] CVPR'22	9	9	342.9	47.8/–
MHFormer [15] CVPR'22	27	27	1031.8	45.9/–
P-STMO [29](*) ECCV'22	27	81	163	46.8/–
P-STMO [29](*) ECCV'22	81	81	493	45.6/–
Einfalt *et al.* [9] WACV'23	9	81	543	47.9/–
PoseFormerV2	1	9	**77.2**	49.9/38.7
PoseFormerV2	1	27	**77.2**	48.7/37.8
PoseFormerV2	1	81	**77.2**	**47.6/37.3**
PoseFormerV2	3	9	117.3	49.5/38.5
PoseFormerV2	3	27	117.3	47.9/37.4
PoseFormerV2	3	81	117.3	**47.1/37.3**
PoseFormerV2	9	27	351.7	47.6/37.1
PoseFormerV2	9	81	351.7	**46.0/36.1**
PoseFormerV2	27	243	1054.8	**45.2/35.6**

Frame Number (f)	Coefficient Number (n)	Full Length	MFLOPs	MPJPE
1	1	27	39.2	51.1
1	3	27	77.2	48.7 (2.4↓)
3	1	27	79.4	50.1 (1.0↓)
3	3	27	117.3	47.9 (3.2↓)
9	9	27	351.7	47.6 (3.5↓)

Compared to other model ,achieve less MPJPE with approximate MFLOPs

model



sEMG Channel Attention (ECA)

$$P_i = X_i + Softmax\left(\frac{Avgpooling(Q) \times Avgpooling(K)}{\sqrt{d_k}}\right)X_i$$

Linear Projection

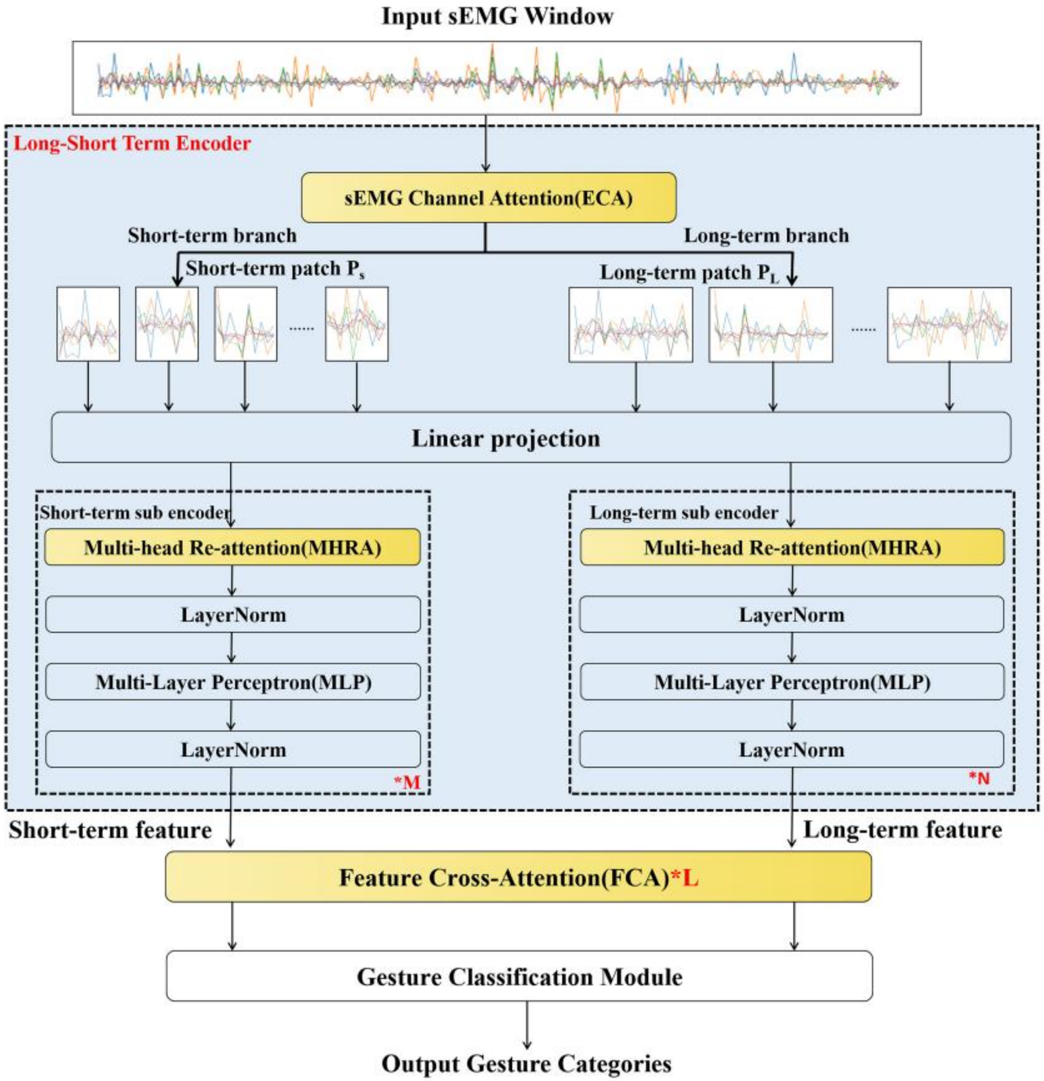
$$Z_0^S = [p_{cls}^S; p_1^S E^S; p_2^S E^S; \ldots; p_S^{N_S} E^S] + E_{pos}^S$$

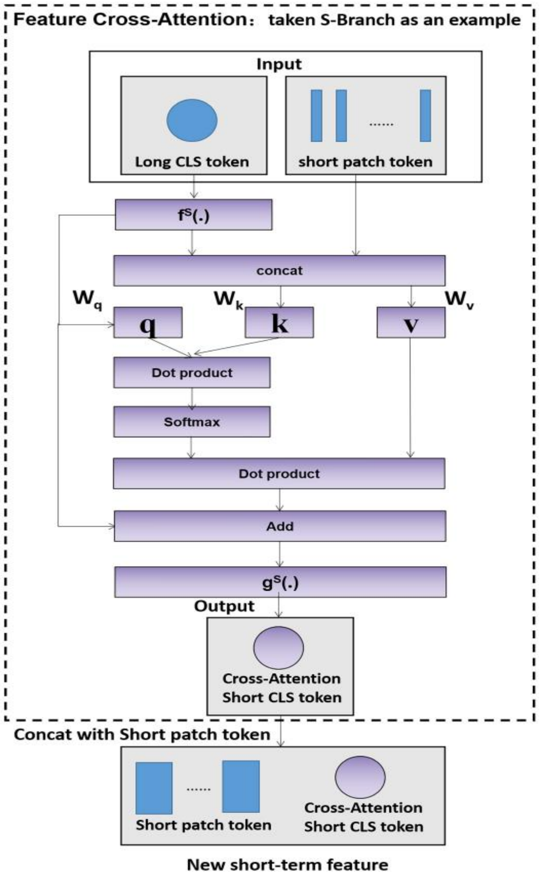
$$Z_0^L = [p_{cls}^L; p_1^L E^L; p_2^L E^L; \ldots; p_{N_L}^L E^L] + E_{pos}^L$$

Multi-head Re-attention

$$Re\text{−}Attention(Q, K, V) = \mathrm{Norm}(\theta^{\mathrm{T}}(Soft\max(QK^T/\sqrt{d_k})))V$$

Feature Cross-Attention





$$Z_M^{S'} = [z_{\text{cls}}^{S''}; z_1^S; \ldots; z_{N_S}^S]$$

$z_{\text{cl}}^{S'}$

z

$$Z_N^{L'} = [z_{\text{cls}}^{L''}; z_1^L; \ldots; z_{N_L}^L]$$

Result

Dataset	Model name	Accuracy	Inference time
DB2 exercise B	MSCNN	71.89%	5.60 ms
	BiTCN	65.79%	5.75 ms
	TEMG	78.77%	1.09 ms
	LSTEMGNet [ours]	**81.47%**	6.47 ms
DB5 exercise C	MSCNN	79.14%	7.27 ms
	BiTCN	83.75%	7.29 ms
	TEMG	68.18%	1.18 ms
	LSTEMGNet [ours]	**88.24%**	6.36 ms
CapgMyo DB-C	MSCNN	86.67%	7.78 ms
	BiTCN	98.38%	7.30 ms
	TEMG	92.90%	1.12 ms
	LSTEMGNet [ours]	**98.80%**	6.32 ms

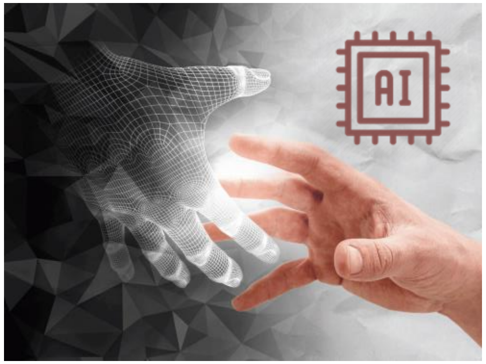
With HD-sEMG data achieve up to 98.80%?

Summary

	Accuracy	
CT-HGR (2023)	94.86	8x8 grid electrode
LSTM(2023)	98.80	CapgMyo DB-C
LSTM(2023)	81.47	DB2-B
LSTM(2023)	88.24	DB2-C
CViT(2022)	84.09	DB2-E1
TraHGR(2022)	88.91	DB2-B
TraHGR(2022)	81.44	DB2-C
TEMGNet(2021)	82.93	DB2

Reference – Overview

[1] PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation
[2] LST-EMG-Net: Long short-term transformer feature fusion network for sEMG gesture recognition



Undergraduate Group

sEMG Neural Spikes Reconstruction Using BSS

Presenter: Miguel
Teammates: Shawn, Howard
Advisor: Prof. An-Yeu (Andy) Wu
Date: 2023/10/20

Outline

- Introduction to blind source separation (BSS)
- A method based on fast Independent Component Analysis (fastICA)
- A method based on convolution kernel compensation (CKC)

Introduction to BSS

- Re-extract the motor unit action potential trains (MUAPTs)
- Reconstruct neural spikes

$$x_i(t) = \sum_{j=1}^{N} \sum_{\tau=0}^{L-1} h_{ij}(\tau) s_j(t-\tau) + \omega_i(t)$$

$$\bar{\mathbf{x}}(t) = \bar{\mathbf{H}}\bar{\mathbf{s}}(t) + \bar{\omega}(t)$$

$$\mathbf{y}(n) = \sum_{k \in Z} \mathbf{B}(k)\mathbf{x}(n-k) \,.$$

Our target : s

Use BSS to handle the inverse problem

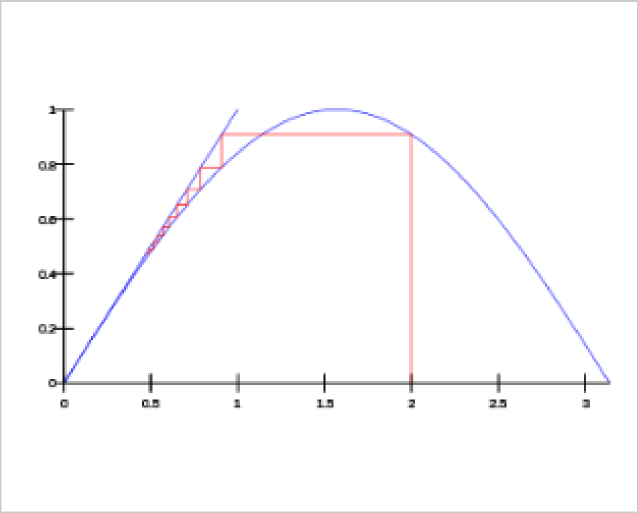
ICA-based algorithm

- Centering and whitening of the extended observation vector
- Fast-fixed point iteration
  - Gram-Schmidt orthogonalization

$$J(y) \propto (\mathbb{E}[G(y)] - \mathbb{E}[G(\nu)])^2$$

$$G(y) = \log\cosh(y)$$

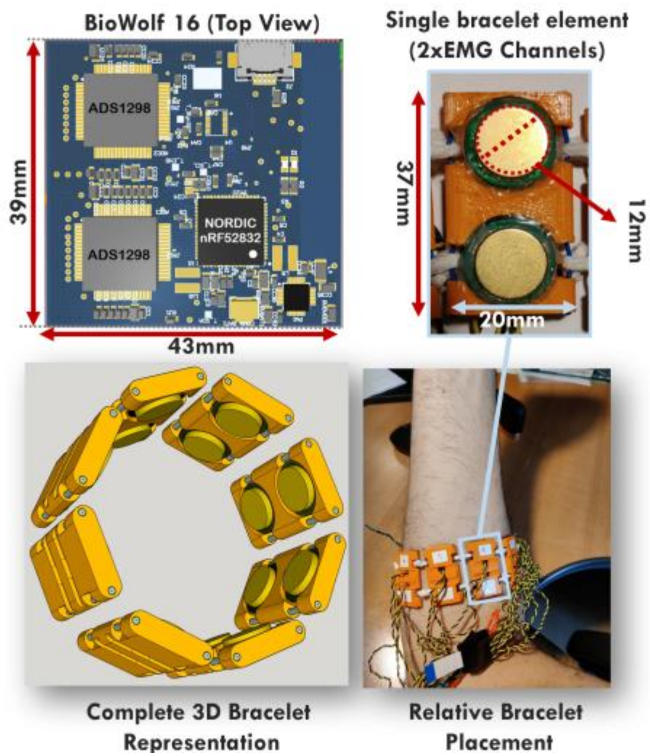
- Peak detection & classifying the detected peaks
  - K-means with K = 2
- Silhouette measure (SIL)
  - kept if SIL ≥ 0.9
- Post-processing stage : discarding
  - *inactive* MUs
  - *duplicate* MUs

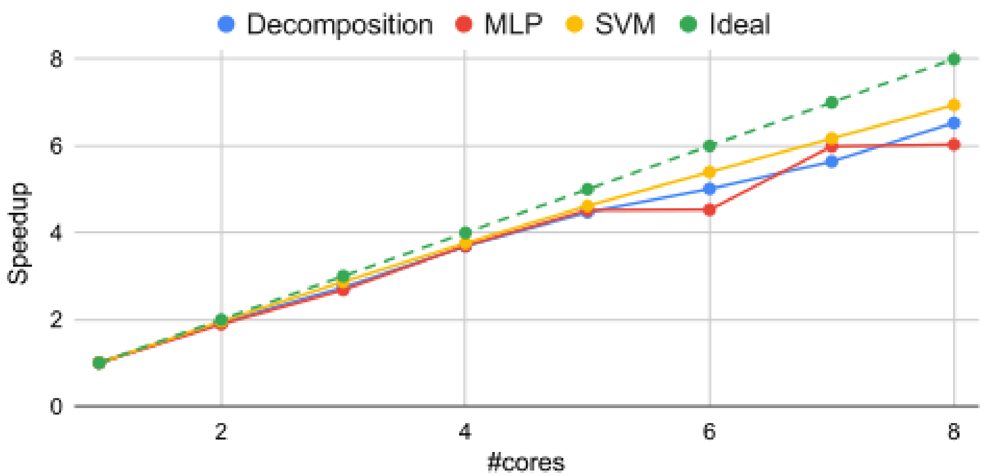


Machine Learning setup

- Training set
  - all the 200 ms slices from the first 4
  - decomposition to these windows
  - binary matrices s (19×800)
- Test set
  - the remaining 1 repetition
- Support Vector Machine (SVM)
  - 19-dimensional vectors containing # of spikes for each MU
- Multilayer Perceptron (MLP) : 3 FC layers
  - binary matrices s (19×800)
  - 4 : time dimension
  - 16 : channel dimension
  - 5 : return

Experiments and Results





High accuracy & power budgeting

STAGE		Memory (kB)	MAC	Latency Cycles	Latency Time (ms)	Energy (μJ)	Test accuracy Unbalanced	Test accuracy Balanced
Decomposition		73.0	4249k	2902k	29.02	536.87	-	-
Classification	option 1: 3-layer MLP	18.1	62k	115k	1.15	21.27	95.28%	92.48%
	option 2: linear SVM	1.0	15k	20k	0.20	3.70	95.38%	92.63%

CKC in a deflation procedure

- Kurtosis contrast function

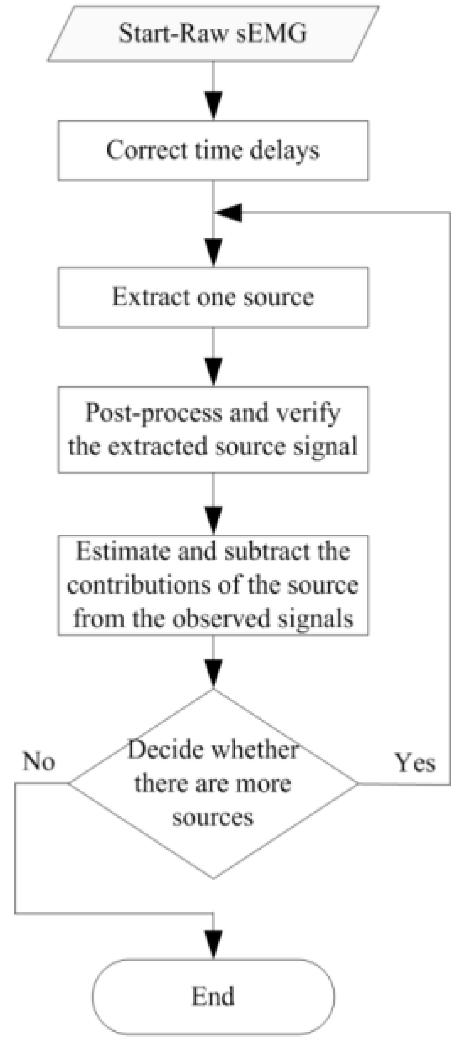
$$J(\mathbf{b}(k)) = |kurt\,(y)|^2 = \left| \frac{E\{y(n)^4\}}{[E\{y(n)^2\}]^2} - 3 \right|^2$$

- A gradient algorithm
  - maximize the contrast function

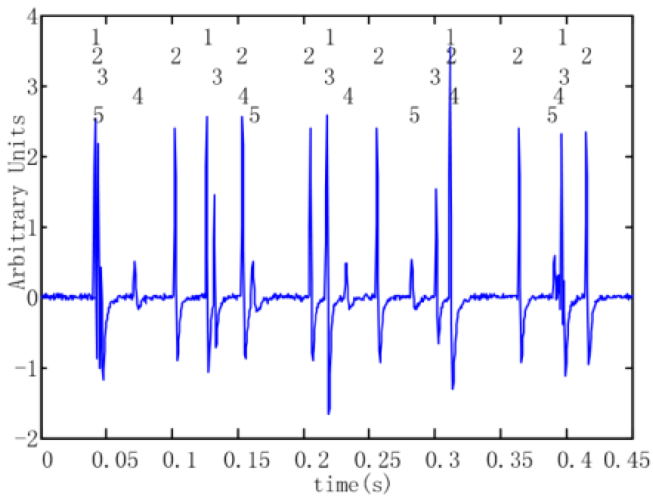
$$y^{(p)}(n) = (\mathbf{b}^{(p)} * \mathbf{x})(n)$$

$$\mathbf{b}^{(p+1)}(k) = \mathbf{b}^{(p)}(k) + \mu^{(p)}(\partial J(\mathbf{b}(k)) / \partial \mathbf{b}(k))$$

- Extract one by one
- Post-processing stage
  - peak : th=5×rms

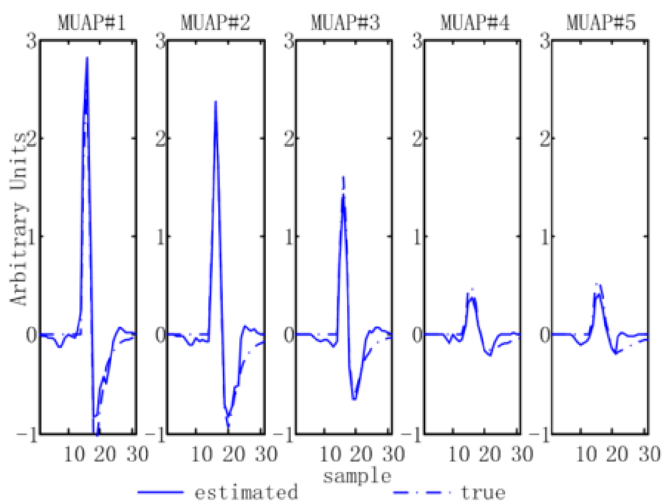


Experiments and Results



(a) Identified MUAP trains

Number of simulated MUs and SNR(dB)		Average Number of identified MUs	Average depth in muscle		Average number of fibers	
			identified	*missed*	*identified*	*missed*
5	20	4.7±0.5	4.8±0.5	7.0±0.9	192±62	146±60
	15	3.7±0.5	4.2±0.6	6.4±1.0	199±65	165±51
10	20	7.1±0.6	4.1±1.4	7.1±1.0	197±70	166±56
	15	5.2±0.8	3.7±1.2	6.3±1.5	202±65	173±66



(b) MUAP templates

Several MUs were missed

# of identified MUs decreased with increasing noise power

Active MUs with weak contribution to the recordings may be missed

Reference – Overview

[1] 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)

[2] 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

[3] https://en.wikipedia.org/wiki/Fixed-point_iteration

[4] https://zh.wikipedia.org/wiki/%E5%B3%B0%E5%BA%A6

[5] https://zh.wikipedia.org/wiki/%E4%B8%AD%E5%BF%83%E7%9F%A9