NEURAL INFORMATION PROCESSING SYSTEMS

Con4m: Context-aware Consistency Learning Framework for Segmented Time Series Classification

Junru Chen[1], Tianyu Cao[1], Jing Xu[2], Jiahe Li[1], Zhilong Chen[1], Tao Xiao[2], Yang Yang[1†]

[1]Zhejiang University [2]State Grid Power Supply Co. Ltd.

jrchen_cali@zju.edu.cn; yangya@zju.edu.cn

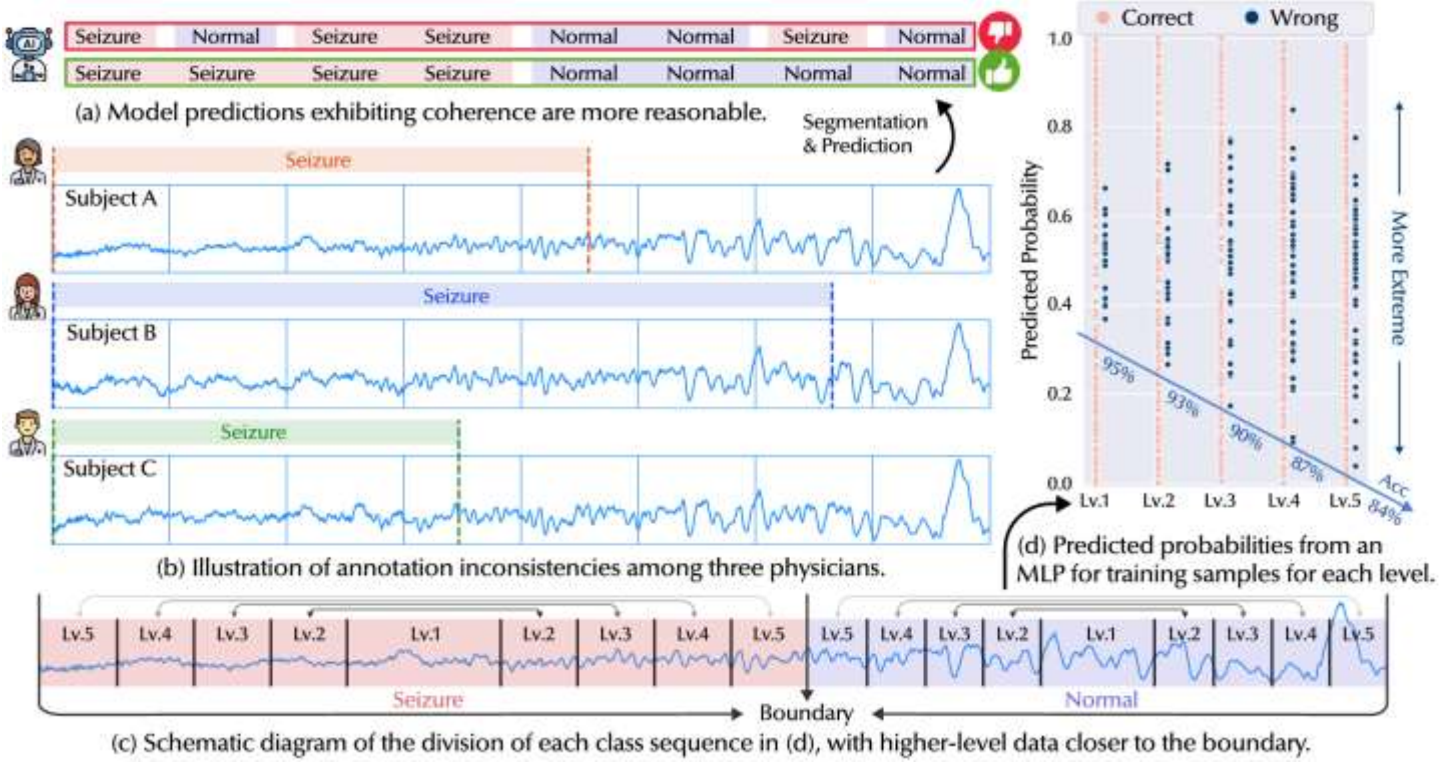
Group Homepage

浙江大學
ZHEJIANG UNIVERSITY

Contextual Inconsistency

❑ Leverage contextual information.

- A natural temporal dependency exists between consecutive classified samples.
- Not only exists at the data level but also manifests in the changes of labels.



❑ Inconsistent boundary labels.

- Manual annotations determine the start and end times for each class.
- Lacking of unified quantification standards leads to experiential differences.
- Inconsistent labels leads to unstable model training.

Con4m: Overview

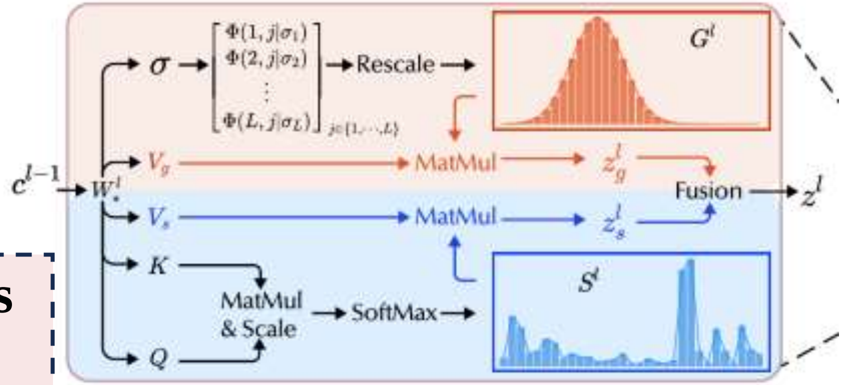
Theorem 2.1.
The more the introduced contextual instance set **enhance the discriminative power of the target instance**, the greater the benefit for the classification task.

Con4m: Overview

Theorem 2.1.
The more the introduced contextual instance set **enhance the discriminative power of the target instance**, the greater the benefit for the classification task.

Contexts at Data level:
Consecutive segments within the same state should be classified into the same class.

Continuous Encoder



Con4m: Overview

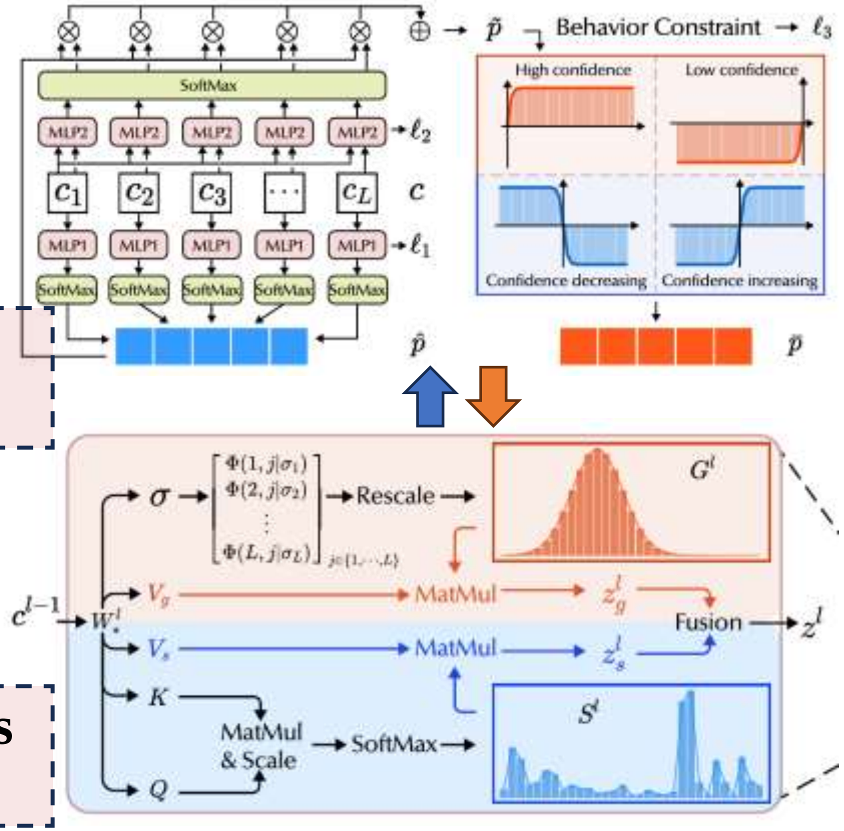
Theorem 2.1.
The more the introduced contextual instance set **enhance the discriminative power of the target instance**, the greater the benefit for the classification task.

Contexts at Label level:
The predictions for consecutive segments should exhibit a constrained monotonicity over time.

Coherent Predictor

Contexts at Data level:
Consecutive segments within the same state should be classified into the same class.

Continuous Encoder



Con4m: Overview

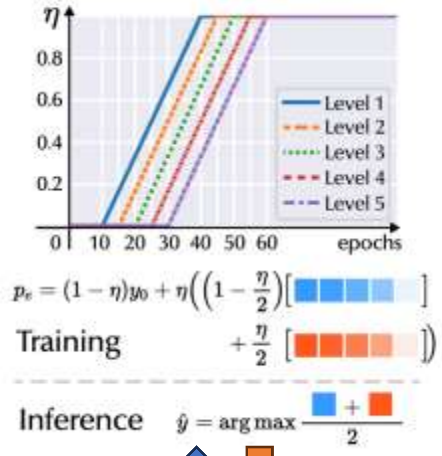
Label Consistency Training Framework:
A progressive harmonization approach for handling inconsistent training labels from easy (core) to the hard (transition) part is designed to yield a more robust model.

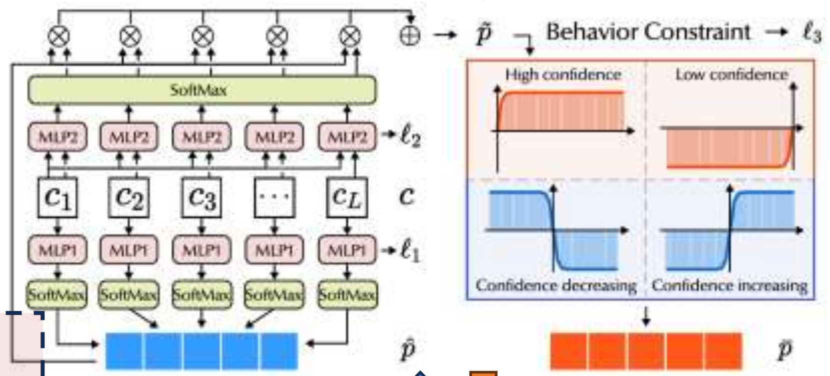
Theorem 2.1.
The more the introduced contextual instance set **enhance the discriminative power of the target instance**, the greater the benefit for the classification task.

Contexts at Label level:
The predictions for consecutive segments should exhibit a constrained monotonicity over time.

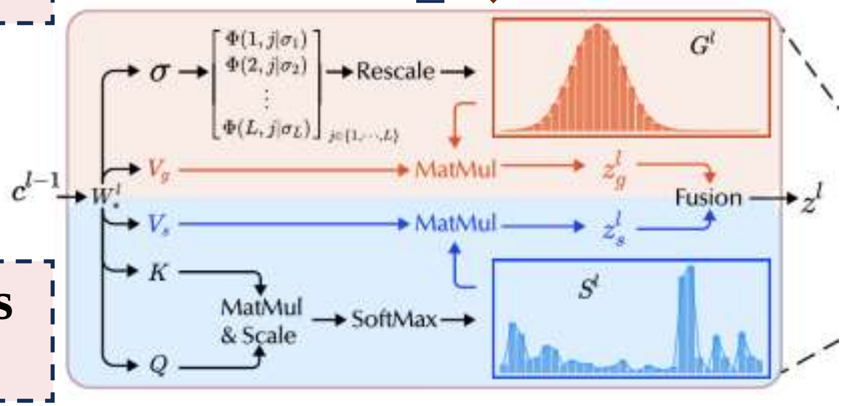
Contexts at Data level:
Consecutive segments within the same state should be classified into the same class.

Consistent Trainer



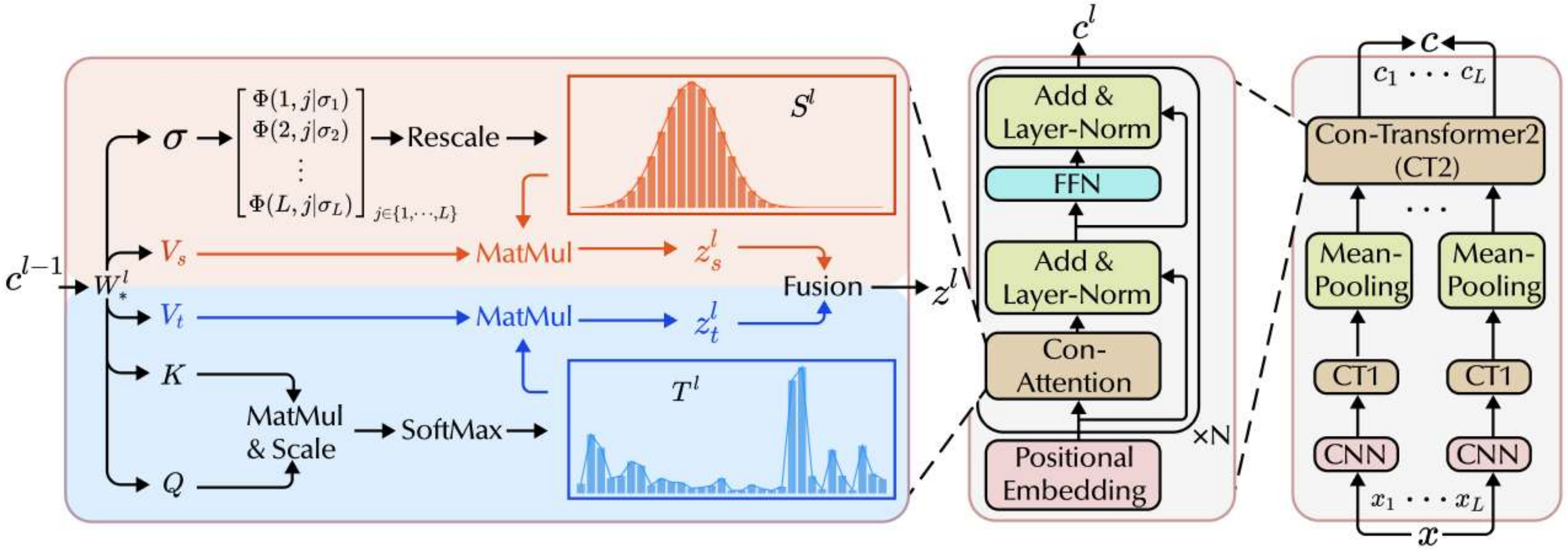


Coherent Predictor



Continuous Encoder

Con4m – Continuous Encoder



$$Q,K,V_s,V_g,\sigma = c^{l-1}W_Q^l,\ c^{l-1}W_K^l,\ c^{l-1}W_{V_s}^l,\ c^{l-1}W_{V_g}^l,\ c^{l-1}W_\sigma^l,$$

$$S^l = \mathrm{SoftMax}\left(\frac{QK^\top}{\sqrt{d}}\right), \quad G^l = \mathrm{Rescale}\left(\left[\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(-\frac{|j-i|^2}{2\sigma_i^2}\right)\right]_{i,j\in\{1,\dots,L\}}\right),$$

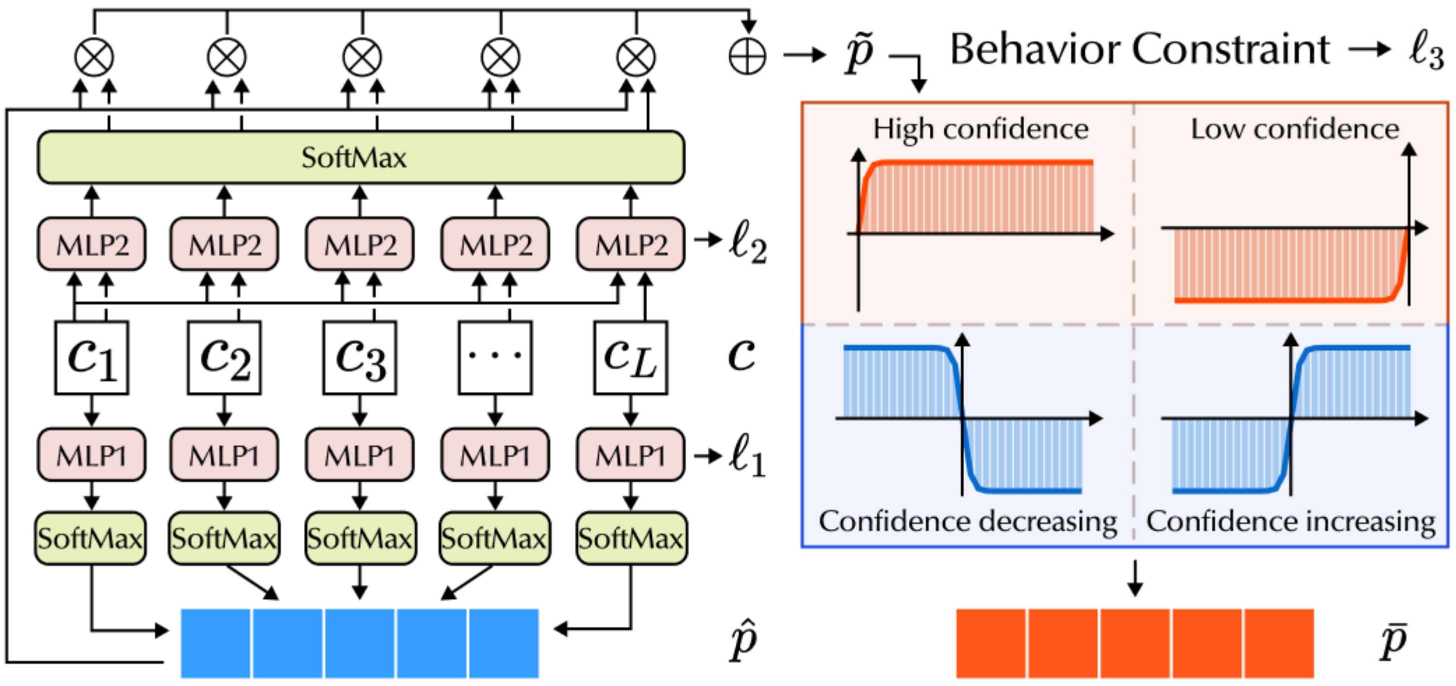
$$z_s^l = S^l V_s, \quad z_g^l = G^l V_g, \quad z^l = \mathrm{Fusion}(z_s^l, z_g^l),$$

Contexts at Data level:

Consecutive segments within the same state should be classified into the same class.

- Smoothing with a **Gaussian kernel** promotes the continuity of representations of time segments in a local temporal window.
- **Aggregating neighbor information** belonging to the same class can improve the discriminative power of the target instance.

Con4m – Coherent Predictor



$$\bar{p} = \mathrm{Tanh}(x|a,k,b,h)$$
$$= a \times \mathrm{Tanh}\,(k \times (x+b)) + h$$
$$\ell_3 = \|\mathrm{Tanh}(x|a,k,b,h) - \tilde{p}\|^2$$

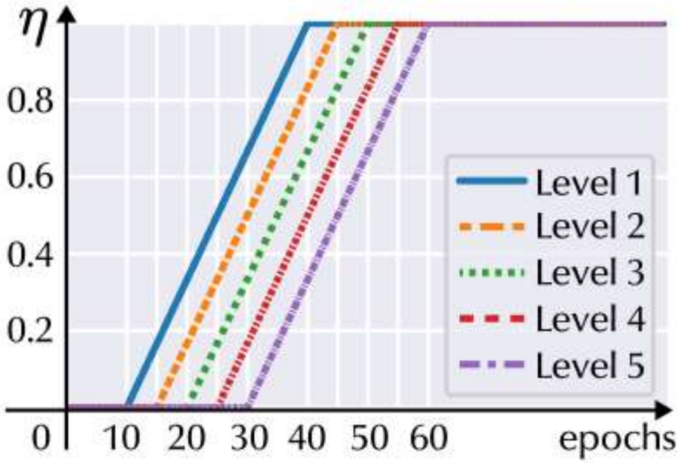
$$\hat{p} = \mathrm{SoftMax}\,(\mathrm{MLP}_1\,(c))$$
$$\hat{R} = \mathrm{SoftMax}\left([\mathrm{MLP}_2\,(c_i\|c_j)]_{i,j\in\{1,\ldots,L\}}\right)$$
$$\tilde{p} = \hat{R}_{:,:,1}\hat{p}$$
$$\ell_1 = \mathrm{CrossEntropy}(\hat{p}, y)$$
$$\ell_2 = \mathrm{CrossEntropy}(\hat{R}, \tilde{Y})$$

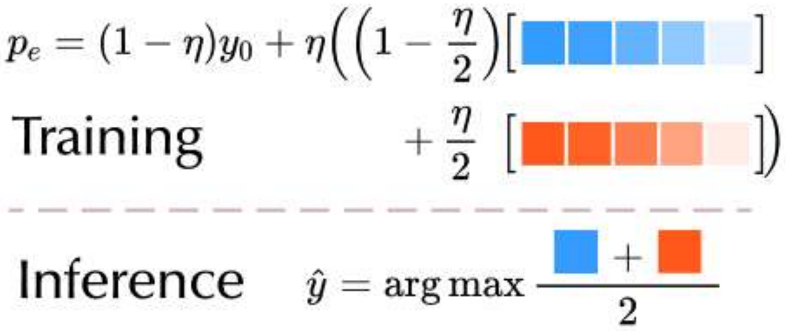
Contexts at Label level:

The predictions for consecutive segments should exhibit a constrained monotonicity over time.

- By weightedly aggregating predictions from similar time segments, the model can focus on contexts more likely to **belong to the same class**.
- Utilize contextual label information to ensure the **monotonicity of predictions** across consecutive segments through hard constraints.

Con4m – Consistent Trainer





$$\omega_e = \mathrm{Rescale}([\exp((e-m)/2)]_{m\in\{0,\dots,4\}})$$

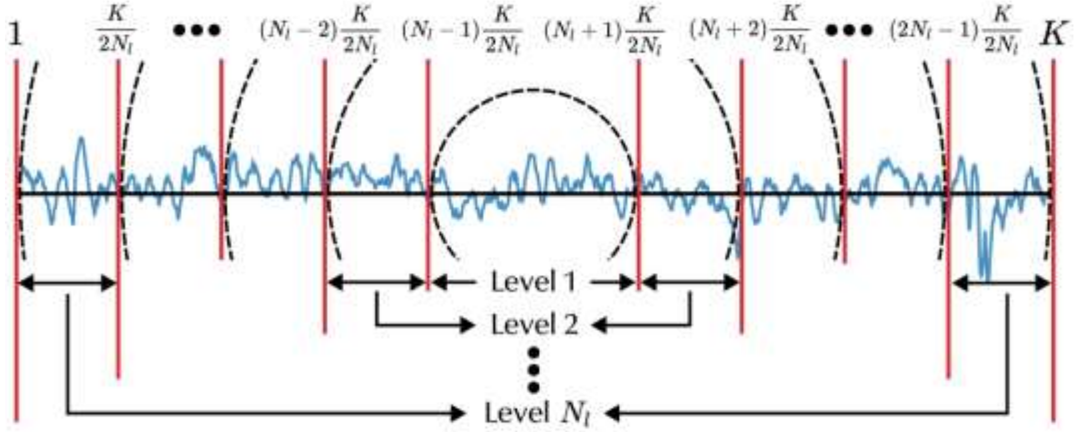
$$\hat{p}_e^5 = \omega_e \cdot [\hat{p}_{e-m}]_{m\in\{0,\dots,4\}}, \quad \bar{p}_e^5 = \omega_e \cdot [\bar{p}_{e-m}]_{m\in\{0,\dots,4\}},$$

$$p_e = (1-\eta)\, y_0 + \eta\left(\left(1-\frac{\eta}{2}\right)\hat{p}_e^5 + \frac{\eta}{2}\bar{p}_e^5\right),$$

Label Consistency Training Framework:

Although people may have differences in the fuzzy transitions between classes, they tend to reach an **agreement** on the most significant **core part** of each class.

- Adopt **curriculum learning** techniques to help the model learn instances from the easy (core) to the hard (transition) part.



- Adopt **noisy label learning** techniques to gradually change the raw labels to harmonize the inconsistency.

Model Comparison

	Model ($r\%$)	fNIRS [31] 0%	fNIRS [31] 20%	fNIRS [31] 40%
TAS	MS-TCN2 [42]	71.48	70.99	69.40
	ASFormer [68]	**71.69**	70.75	69.18
	DiffAct [45]	71.15	69.72	65.45
TSC	MiniRocket [13]	61.28	60.41	57.87
	TimesNet [64]	67.47	65.39	63.45
	PatchTST [51]	*51.79*	*55.38*	*52.67*
NLL	SIGUA [28]	67.37	65.24	63.47
	UNICON [36]	61.15	60.45	57.35
	Sel-CL [43]	63.86	62.45	61.75
TSC & NLL	SREA [9]	70.10	69.65	69.40
	Scale-T [47]	70.40	68.06	66.51
	Con4m	71.28	**71.27**	**70.04**

competitive performance

Data	Sample Frequency	# of Features	# of Classes	Subjects	Groups	Cross Validation	Total Intervals	Interval Length	Window Length	Slide Length	Total Segments
fNIRS	5.2Hz	8	2	68	4	12	4,080	38.46s	4.81s	0.96s	146,880
HHAR	50Hz	6	6	9	3	6	5,400	60s	4s	2s	156,600
Sleep	100Hz	2	5	154	3	6	6,000	40s	2.5s	1.25s	186,000
SEEG	250Hz	1	2	8	4	3	8,000	16s	1s	0.5s	248,000

Model Comparison

	Model \ r%	fNIRS [31]			HHAR [7]			Sleep [37]			SEEG
		0%	20%	40%	0%	20%	40%	0%	20%	40%	raw
TAS	MS-TCN2 [42]	71.48	70.99	69.40	69.79	66.72	62.29	60.07	59.03	56.17	61.88
	ASFormer [68]	**71.69**	70.75	69.18	62.52	60.92	60.77	59.09	55.52	53.89	56.71
	DiffAct [45]	71.15	69.72	65.45	56.76	53.86	50.63	49.12	*43.32*	*38.86*	60.62
TSC	MiniRocket [13]	61.28	60.41	57.87	70.34	63.32	59.25	62.00	61.75	58.38	62.39
	TimesNet [64]	67.47	65.39	63.45	72.07	70.19	66.76	59.50	57.72	55.73	*50.99*
	PatchTST [51]	*51.79*	*55.38*	*52.67*	*52.00*	*45.46*	*45.69*	58.40	56.16	53.05	58.45
NLL	SIGUA [28]	67.37	65.24	63.47	68.94	68.47	67.60	54.28	53.07	51.32	53.19
	UNICON [36]	61.15	60.45	57.35	62.26	61.63	58.34	62.26	61.63	58.34	60.53
	Sel-CL [43]	63.86	62.45	61.75	73.00	72.28	72.81	63.48	63.45	61.72	60.50
TSC & NLL	SREA [9]	70.10	69.65	69.40	68.64	66.02	65.67	*48.81*	48.80	45.72	55.21
	Scale-T [47]	70.40	68.06	66.51	77.77	76.71	**75.97**	63.21	63.40	60.77	67.64
	Con4m	71.28	**71.27**	**70.04**	**80.29**	**78.59**	75.52	**68.02**	**66.31**	**64.31**	**72.00**

competitive performance

3.24% ↑

7.15% ↑

6.45% ↑

Data	Sample Frequency	# of Features	# of Classes	Subjects	Groups	Cross Validation	Total Intervals	Interval Length	Window Length	Slide Length	Total Segments
fNIRS	5.2Hz	8	2	68	4	12	4,080	38.46s	4.81s	0.96s	146,880
HHAR	50Hz	6	6	9	3	6	5,400	60s	4s	2s	156,600
Sleep	100Hz	2	5	154	3	6	6,000	40s	2.5s	1.25s	186,000
SEEG	250Hz	1	2	8	4	3	8,000	16s	1s	0.5s	248,000

Model Comparison

Model		fNIRS [31] 0%	fNIRS [31] 20%	fNIRS [31] 40%	HHAR [7] 0%	HHAR [7] 20%	HHAR [7] 40%	Sleep [37] 0%	Sleep [37] 20%	Sleep [37] 40%	SEEG raw
TAS	MS-TCN2 [42]	71.48	70.99	69.40	69.79	66.72	62.29	60.07	59.03	56.17	61.88
	ASFormer [68]	**71.69**	70.75	69.18	62.52	60.92	60.77	59.09	55.52	53.89	56.71
	DiffAct [45]	71.15	69.72	65.45	56.76	53.86	50.63	49.12	*43.32*	*38.86*	60.62
TSC	MiniRocket [13]	61.28	60.41	57.87	70.34	63.32	59.25	62.00	61.75	58.38	62.39
	TimesNet [64]	67.47	65.39	63.45	72.07	70.19	66.76	59.50	57.72	55.73	*50.99*
	PatchTST [51]	*51.79*	*55.38*	*52.67*	*52.00*	*45.46*	*45.69*	58.40	56.16	53.05	58.45
NLL	SIGUA [28]	67.37	65.24	63.47	68.94	68.47	67.60	54.28	53.07	51.32	53.19
	UNICON [36]	61.15	60.45	57.35	62.26	61.63	58.34	62.26	61.63	58.34	60.53
	Sel-CL [43]	63.86	62.45	61.75	73.00	72.28	72.81	63.48	63.45	61.72	60.50
TSC & NLL	SREA [9]	70.10	69.65	69.40	68.64	66.02	65.67	*48.81*	48.80	45.72	55.21
	Scale-T [47]	70.40	68.06	66.51	77.77	76.71	**75.97**	63.21	63.40	60.77	67.64
	Con4m	71.28	**71.27**	**70.04**	**80.29**	**78.59**	75.52	**68.02**	**66.31**	**64.31**	**72.00**

competitive performance

3.01% ↓
5.23% ↓
1.92% ↓
3.34% ↓
3.22% ↓
2.37% ↓

3.24% ↑
7.15% ↑
6.45% ↑

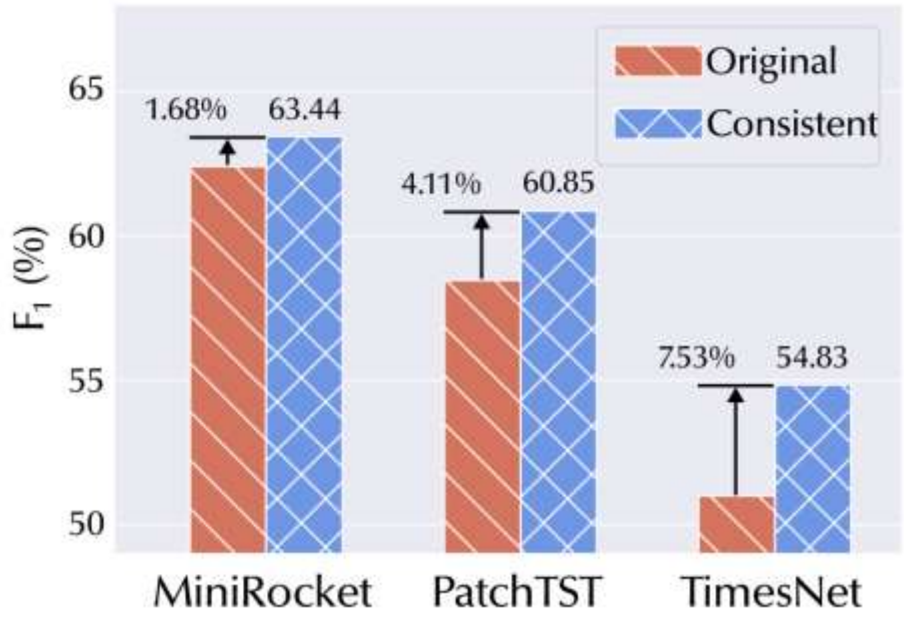
Data	Sample Frequency	# of Features	# of Classes	Subjects	Groups	Cross Validation	Total Intervals	Interval Length	Window Length	Slide Length	Total Segments
fNIRS	5.2Hz	8	2	68	4	12	4,080	38.46s	4.81s	0.96s	146,880
HHAR	50Hz	6	6	9	3	6	5,400	60s	4s	2s	156,600
Sleep	100Hz	2	5	154	3	6	6,000	40s	2.5s	1.25s	186,000
SEEG	250Hz	1	2	8	4	3	8,000	16s	1s	0.5s	248,000

Model Comparison

	Model	fNIRS [31] 0%	fNIRS [31] 20%	fNIRS [31] 40%	HHAR [7] 0%	HHAR [7] 20%	HHAR [7] 40%	Sleep [37] 0%	Sleep [37] 20%	Sleep [37] 40%	SEEG raw
TAS	MS-TCN2 [42]	71.48	70.99	69.40	69.79	66.72	62.29	60.07	59.03	56.17	61.88
	ASFormer [68]	**71.69**	70.75	69.18	62.52	60.92	60.77	59.09	55.52	53.89	56.71
	DiffAct [45]	71.15	69.72	65.45	56.76	53.86	50.63	49.12	*43.32*	*38.86*	60.62
TSC	MiniRocket [13]	61.28	60.41	57.87	70.34	63.32	59.25	62.00	61.75	58.38	62.39
	TimesNet [64]	67.47	65.39	63.45	72.07	70.19	66.76	59.50	57.72	55.73	*50.99*
	PatchTST [51]	*51.79*	*55.38*	*52.67*	*52.00*	*45.46*	*45.69*	58.40	56.16	53.05	58.45
									53.07	51.32	53.19
									61.63	58.34	60.53
									63.45	61.72	60.50
									48.80	45.72	55.21
									63.40	60.77	67.64
									66.31	**64.31**	**72.00**

competitive performance

6.45% ↑



Label Substitution Experiment

Verify the **effectiveness** of the label harmonization process on SEEG data.

lide ength	Total Segments
.96s	146,880
2s	156,600
.25s	186,000
.5s	248,000

Case Study

- Con4m demonstrates a **more coherent narrative** by constraining the prediction behavior and aligning with the contextual data information.
- Con4m **accurately identifies the consistent boundary** within the time interval spanning across two classes.

Conclusion

- We are the **first** to propose a **practical** consistency learning framework Con4m for the segmented TSC based on the raw MVD.
- By comprehensively integrating prior knowledge from the **data and label** perspectives, we guide the model to focus on **effective contextual information**.
- Based on context-aware predictions, a **progressive harmonization approach** for handling inconsistent training labels is designed to yield a more **robust** model.
- Extensive experiments on three public and one private MVD datasets demonstrate the superior performance of Con4m.

THANKS | Q&A

More relevant research of our group: ***http://yangy.org***

Contact: ***jrchen_cali@zju.edu.cn; yangya@zju.edu.cn***

Group Homepage