CHANNEL-IMPOSED FUSION: A SIMPLE YET EFFECTIVE METHOD FOR MEDICAL TIME SERIES CLASSIFICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

A simple model with good features beats a complex model with poor features. Deep learning in healthcare has increasingly embraced a model-centric paradigm, and medical time series is no exception, with increasingly complex architectures designed to capture richer temporal, cross-channel, and multi-scale dependencies. We argue that this paradigm may overlook a more fundamental bottleneck: the quality and discriminability of the input signals themselves. We therefore shift the focus from designing increasingly complex classifiers toward improving data representations before classification. To this end, we propose **CIF**, a simple, plug-and-play, and model-agnostic framework for enhancing medical time series representations. CIF can be seamlessly integrated with existing classification architectures without modifying their model designs. Extensive experiments across diverse medical time series datasets and representative classification architectures demonstrate that CIF consistently and substantially improves classification performance, providing strong evidence for the effectiveness of a data-centric alternative. We hope this work encourages the community to reconsider the core of medical time series classification: **should progress be driven primarily by data-centric strategies, model-centric design, or both?** Code: https://anonymous.4open.science/r/CIF-8F76.

1 INTRODUCTION

Medical time series (MedTS), such as electroencephalogram (EEG) and electrocardiogram (ECG) signals, are widely used for monitoring patient health and supporting the diagnosis of neurological and cardiovascular diseases Arif et al. (2024); Xiao et al. (2023); Zhu et al. (2025); Wang et al. (2024b; 2025b); Strodthoff et al. (2020). Accurate classification of these signals is essential for early anomaly detection, personalized treatment, and improved clinical decision-making Liu et al. (2024a); Tian et al. (2023). With the development of deep learning, convolutional neural networks (CNNs), such as EEGNet Lawhern et al. (2018), have demonstrated the ability to learn informative representations directly from raw signals.

In recent years, Transformer-based models Vaswani et al. (2017), inspired by the self-attention mechanism Bahdanau et al. (2014), have achieved remarkable success in time series modeling by capturing long-range dependencies and global contextual information Liu et al. (2021); Zhou et al. (2021). However, as shown in Fig. 1, increasing model complexity does not necessarily lead to consistent improvements in MedTS classification. This observation suggests that the challenge may not lie solely in temporal modeling, but also in the quality of the input representation. **(1) Domain-specific inter-channel relationships are not explicitly incorporated into input representations.** Unlike general-purpose time series, MedTS signals are generated by physiological systems with structured relationships across channels. ECG leads exhibit characteristic electrical conduction relationships Auricchio et al. (2014), while EEG electrodes exhibit structured spatial and rhythmic dependencies Palva & Palva (2014); Fries (2015). Such relationships provide valuable domain priors for constructing informative representations. However, existing time series architectures, including Autoformer Wu et al. (2021), Crossformer Zhang & Yan (2023), and Reformer Kitaev et al. (2020), primarily learn temporal and cross-channel dependencies through data-driven mechanisms. Although these models may implicitly capture domain-specific relationships, they must discover such structures

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

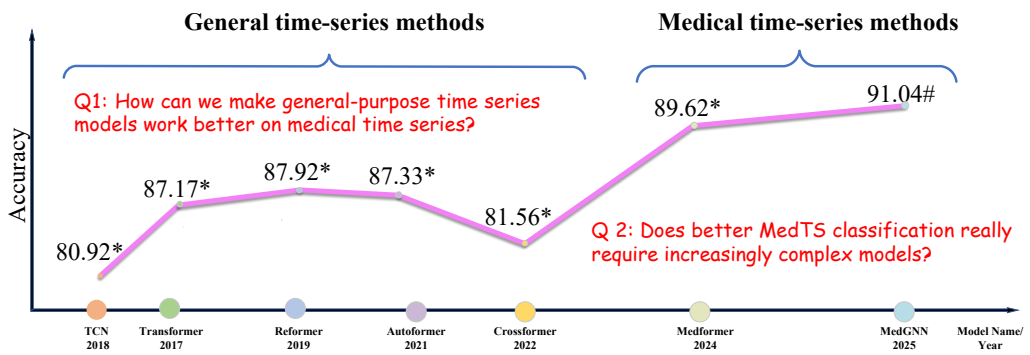


Figure 1: Results of various methods on the TDBrain dataset (EEG). $*$ denotes results reported by *Medformer* Wang et al. (2024a), while # denotes results reported by *MedGNN* Fan et al. (2025). Q1 and Q2 highlight the two main questions motivating this work.

during representation learning. This raises a natural question: **can domain-specific inter-channel relationships be explicitly incorporated into the input representation before temporal modeling? (2) Noisy observations can limit the discriminability of raw representations.** MedTS signals are often acquired under noisy conditions and may contain substantial artifacts Del Rio et al. (2011); Sraitih et al. (2022); Sharma (2017); Mohd Apandi et al. (2020); Jia et al. (2024). Such interference can obscure clinically relevant patterns and make the raw representation less discriminative. Increasing the capacity of a downstream classifier does not necessarily recover information that is weak or masked in the input. Recent MedTS-specific methods, such as MedGNN Fan et al. (2025) and Medformer Wang et al. (2024a), mainly address these challenges through architectural innovations, while leaving the underlying signal representation largely unchanged. This leads to a second question: **can selectively improving the discriminability of the input representation reduce the burden on downstream temporal modeling?**

These observations motivate a shift from a purely *model-centric* paradigm toward a *data-centric* perspective for MedTS classification. Rather than relying solely on increasingly complex architectures to discover useful structures from noisy observations, we ask whether domain knowledge can be directly exploited to construct more informative representations before temporal modeling. In this view, representation quality and model capacity are complementary factors: a stronger temporal model cannot fully compensate for an input representation that lacks discriminative information. Following this perspective, we propose *Channel-Imposed Fusion (CIF)*, a lightweight and plug-and-play representation transformation that explicitly incorporates domain-specific inter-channel relationships into the input signals. Given two channels x and y, CIF constructs an additional representation through

$$x_{\text{new}} = ax + by, \tag{1}$$

where the signs of a and b encode the expected direction of the inter-channel relationship, while their magnitudes can adapt to the data. Thus, domain knowledge determines *how* channels should interact, whereas data-driven optimization determines *how strongly* they should be combined.

CIF is motivated by two complementary signal-level hypotheses. **(1) Physiological Coupling Hypothesis.** When two ECG leads contain shared physiological components, positively combining them can reinforce information jointly expressed across the channels Platonov (2012). **(2) Common-Component Suppression Hypothesis.** In EEG recordings, ocular artifacts such as blinks can be correlated across frontal electrodes such as Fp1 and Fp2 Croft & Barry (2000). A differential combination with opposite coefficient signs can therefore suppress shared components while preserving channel-specific information. Importantly, CIF is not intended as a general-purpose denoising method or a global SNR optimization technique. Its effect is selective and representation-specific: depending on the underlying signal relationships, fusion may reinforce informative physiological components or attenuate shared interference. The ultimate objective is to construct representations that are more informative and discriminative for downstream classification.

To examine whether improved representations can reduce the need for complex temporal models, we deliberately employ a lightweight temporal backbone, termed *Hidden-layer Mixed Bidirectional Temporal Convolutional Network (HMTCN)*. This design allows us to evaluate the contribution of representation quality while minimizing the confounding effect of architectural complexity. CIF combined with HMTCN consistently outperforms sophisticated Transformer-based approaches across multiple MedTS datasets and achieves competitive, in several cases state-of-the-art, performance on general time series classification benchmarks. Furthermore, CIF can be seamlessly incorporated into existing Transformer architectures, demonstrating that data-centric representation improvement is complementary to, rather than a replacement for, advanced temporal modeling. Contributions are:

- **A Simple Data-Centric Approach.** We propose CIF, a simple, plug-and-play, and model-agnostic approach that enhances time series representations by incorporating inter-channel relationships, without modifying the underlying model architecture.
- **Better Representations, Simpler Models.** We combine CIF with a simple HMTCN and show that better feature representations can enable simple models to outperform more sophisticated Transformer-based models across multiple datasets.
- **A Broader Question for the Field.** We hope this work encourages the community to reconsider the core of medical time series classification: *should progress be driven primarily by data-centric strategies, model-centric design, or a combination of both?*

2 Related Work

Data-Centric Spatial Filtering. Spatial filtering provides a lightweight yet fundamental way to incorporate channel-level priors before temporal modeling in multichannel medical time series. Existing approaches encode different assumptions about spatial relationships, ranging from fixed structural transforms such as global referencing (CAR) Goldman (1950); Offner (1950) and local adjacency (bipolar montage Osselton (1965) and Laplacian Hjorth (1975)), to training-estimated spatial operators, including supervised class-discriminative projections (CSP Koles et al. (1990); Ramoser et al. (2000)), unsupervised statistical source-separation and subspace methods (ICA Bell & Sejnowski (1995); Makeig et al. (1995) and SSP Uusitalo & Ilmoniemi (1997)), and physics-based or subspace-reconstruction approaches (SSS Taulu & Kajola (2005)). In contrast, a fully learnable linear mixing layer removes these structural constraints but introduces $O(C^2)$ parameters and may be unnecessarily over-parameterized for small medical datasets. Our **CIF** takes an intermediate, structured approach: it uses a small number of coefficients to fuse semantically separated channel groups while preserving most of the original channels. This design explicitly introduces a *cross-region complementary prior* without relying on class labels, trial-level covariance estimation, blind source separation, artifact-subspace fitting, or a dense $C \times C$ transformation. Details in Appendix E.

Model-centric Methods for Time Series Classification. Time series classification has increasingly adopted a model-centric paradigm, with progressively sophisticated architectures designed to capture temporal, cross-channel, and multi-scale dependencies. Early approaches mainly relied on convolutional architectures, such as EEGNet Lawhern et al. (2018) and TCNs Bai et al. (2018); Lin et al. (2019), while Transformers Vaswani et al. (2017) and their variants further enabled global dependency modeling through efficient attention, frequency-domain representations, patch-based modeling, and cross-variable interactions, including Informer Zhou et al. (2021), Autoformer Wu et al. (2021), FEDformer Zhou et al. (2022), Crossformer Zhang & Yan (2023), and iTransformer Liu et al. (2024b). For medical time series, researchers have further developed modality- and domain-specific architectures, such as EEG-Conformer Song et al. (2022), Medformer Wang et al. (2024a), MedGNN Fan et al. (2025),ADformer Wang et al. (2024b) and BIOT Yang et al. (2023). One-dimensional convolutional architectures, including ResNet He et al. (2016); Wang et al. (2017), XResNet He et al. (2019), and InceptionTime Ismail Fawaz et al. (2020), have also been widely adopted. A more detailed discussion is provided in the Appendix F.

Despite these advances, existing methods predominantly pursue performance gains through increased model capacity and architectural complexity, while paying less attention to the input signals themselves. For multichannel medical time series, redundancy, noise, inter-channel interference, and modality-specific spatial structure may impose more fundamental limitations. This raises an important question: *is model capacity truly the primary bottleneck, or should we first address the quality and structure of the input signals?*

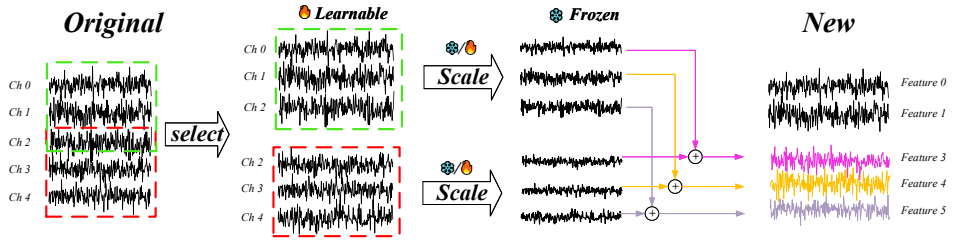


Figure 2: The implementation process of Channel-Imposed Fusion.

3 METHOD

3.1 CHANNEL-IMPOSED FUSION

As illustrated in Figure 2, we propose **Channel-Imposed Fusion (CIF)**, a lightweight and plug-and-play representation transformation that incorporates inter-channel relationships before temporal modeling. Given a multi-channel time series with N channels, CIF performs shared linear fusion between paired channels. Specifically, for two subsets $X = X_1, \ldots, X_n$ and $Y = Y_1, \ldots, Y_n$, the i-th pair is transformed as $X_i^{\text{new}} = aX_i + bY_i$, where a and b are shared fusion coefficients. The transformed channels are then concatenated with the original channels and fed into the temporal backbone. Thus, CIF introduces only two additional parameters, regardless of the number of channel pairs, while enforcing a consistent transformation across the selected channel pairs.

The signs of a and b determine the *direction* of channel interaction, while their magnitudes determine its *strength*. CIF therefore provides a simple mechanism for incorporating structured inter-channel relationships while avoiding a complex channel-dependency module. Importantly, the fusion itself is independent of any specific temporal backbone and can be applied as a preprocessing or representation transformation before temporal modeling.

Signal-level Interpretation. CIF can be interpreted through a simple signal decomposition. For two observations $x_1 = s_1 + \epsilon_1$ and $x_2 = s_2 + \epsilon_2$, their linear combination $y = ax_1 + bx_2$ yields $\text{SNR}_{\text{out}} = \text{Var}(as_1 + bs_2)/\text{Var}(a\epsilon_1 + b\epsilon_2)$. Under equal signal and noise variances, with ρ and γ denoting signal and noise correlations, respectively, this becomes $\text{SNR}_{\text{out}} = \text{SNR}_{\text{in}} \frac{a^2+b^2+2ab\rho}{a^2+b^2+2ab\gamma}$. When $\rho > \gamma$, same-sign coefficients such as $a = b$ can reinforce shared signal components (*cooperative mode*); when $\rho < \gamma$, opposite-sign coefficients such as $a = -b$ can attenuate shared interference (*differential mode*). Importantly, CIF does not aim to universally improve SNR or perform general-purpose denoising; rather, this analysis explains how structured linear combinations can selectively improve the resulting representation. Detailed derivations are provided in Appendix A.

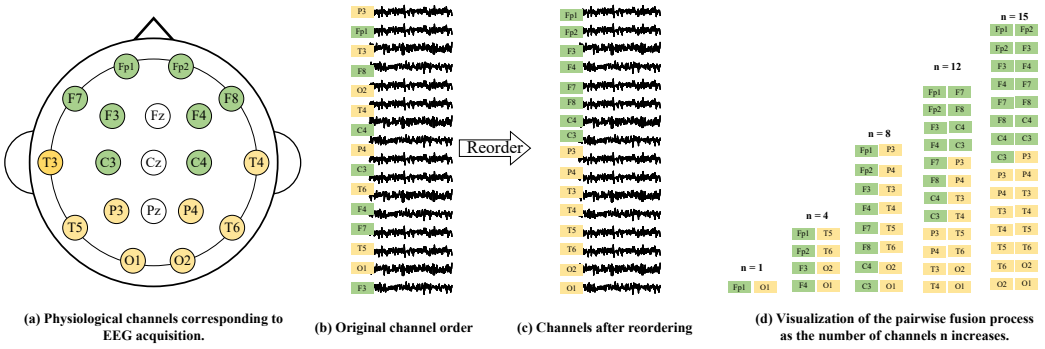


Figure 3: Illustration of the channel reordering and pairing strategy in CIF. When needed, channels are reorganized according to modality-specific spatial, anatomical, or physiological relationships, after which the first n channels are paired with the corresponding last n channels for linear fusion. Panel (d) illustrates the resulting channel-pairing patterns under different values of n.

Channel Pairing. CIF does not require a specific channel ordering. In many medical time series, channels are acquired and represented according to established measurement protocols, which already encode meaningful spatial, anatomical, or physiological relationships. CIF can therefore directly operate on the original channel organization when such structure is informative. In cases where the original ordering does not expose the desired relationships, channels can be simply reordered according to available domain knowledge before pairing. For example, EEG channels may be reorganized according to their spatial arrangement, while ECG leads may be grouped according to their measurement roles. Such reordering is not an intrinsic requirement of CIF, but rather a flexible way to make relevant channel relationships explicit when needed.

After pairing, the first n channels are combined with the corresponding n channels in the second subset, as illustrated in Figure 3(d). The parameter n controls the coverage of the transformation, whereas (a, b) determines the strength and direction of channel fusion. When a and b are learnable, they are jointly optimized with the temporal backbone in an end-to-end manner.

3.2 HMTCN Structure Design

To evaluate the contribution of improved input representations without relying on excessive model capacity, as illustrated in Figure 4, we adopt a lightweight temporal backbone, termed **Hidden-layer Mixed Bidirectional Temporal Convolutional Network (HMTCN)**. Unlike conventional TCNs that primarily model temporal dependencies through causal convolutions, HMTCN introduces bidirectional feature mixing, allowing each layer to exploit temporal information from both forward and backward directions.

CIF and HMTCN operate on complementary dimensions: CIF explicitly exploits physiological relationships along the *channel dimension*, while HMTCN extracts temporal patterns along the *temporal dimension*. The overall pipeline is therefore Raw MedTS $\xrightarrow{\text{CIF}}$ Structured Representation $\xrightarrow{\text{HMTCN}}$ Temporal Features $\xrightarrow{\text{Classifier}}$ Prediction. This design provides a relatively clean evaluation of the data-centric hypothesis, since the lightweight backbone minimizes the influence of architectural complexity. When a and b are learnable, CIF and HMTCN are jointly optimized end-to-end, allowing the channel transformation and temporal modeling to adapt simultaneously to the classification objective. Further details are provided in Appendix G and Appendix H.

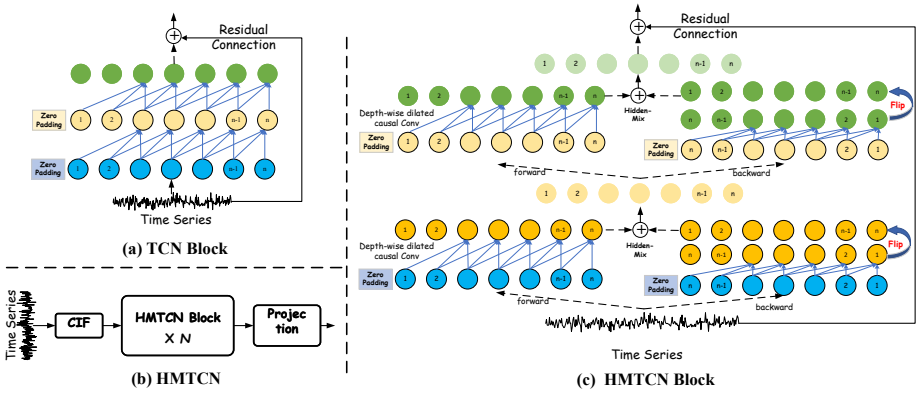


Figure 4: Architecture of HMTCN.

4 Experiments

Medical Time Series Datasets. We evaluate our method on five medical time-series datasets, including three EEG datasets (APAVA Escudero et al. (2006), TDBRAIN van Dijk et al. (2022), and ADFTD Miltiadous et al. (2023b;a)) and two ECG datasets (PTB PhysioBank (2000) and PTB-XL Wagner et al. (2020)). We follow the data preprocessing protocols of Medformer Wang et al. (2024a) and MedGNN Fan et al. (2025). Details in Appendix C.

Baselines. We compare with time series transformer methods: Autoformer Wu et al. (2021), Crossformer Zhang & Yan (2023), FEDformer Zhou et al. (2022), Informer Zhou et al. (2021), iTransformer Liu et al. (2024b), MTST Zhang et al. (2024), Nonformer Liu et al. (2022), PatchTST Nie et al. (2023), Reformer Kitaev et al. (2020), vanilla Transformer Vaswani et al. (2017), Medformer Wang et al. (2024a), MedGNN Fan et al. (2025). We further include ADformer Wang et al. (2024b), BIOT Yang et al. (2023), ResNet He et al. (2016); Wang et al. (2017), XResNet He et al. (2019), EEG-Conformer Song et al. (2022), Inception Ismail Fawaz et al. (2020), EEGInception Zhang et al. (2021), and EEGNet Lawhern et al. (2018) as additional baselines in Appendix L.

Spatial Preprocessing Baseline. We compare CIF against eight representative spatial preprocessing methods: Bipolar Osselton (1965), CAR) Goldman (1950); Offner (1950), CSP Koles et al. (1990); Ramoser et al. (2000), ICA Bell & Sejnowski (1995); Makeig et al. (1995), Laplacian Hjorth (1975), Linear Layer, SSP Uusitalo & Ilmoniemi (1997), and SSS Taulu & Kajola (2005).

Implementation. We employ six evaluation metrics: accuracy, precision (macro-averaged), recall (macro-averaged), F1 score (macro-averaged), AUROC (macro-averaged), and AUPRC (macro-averaged). The training process is conducted with five random seeds (41-45) on fixed training, validation, and test sets to compute the mean and standard deviation of the models. All experiments were conducted using a RTX 3090 GPU and implemented with PyTorch version 1.11.0. Following Medformer Wang et al. (2024a) and MedGNN Fan et al. (2025), we adopt a **Subject-Independent Split**, where each subject is exclusively assigned to the training, validation, or test set.

Table 1: Test performance (%) of spatial preprocessing methods with HMTCN.

Dataset	Metric	Bipolar	CAR	CSP	ICA	Laplacian	Linear	SSP	SSS	CIF
ADFTD	Acc	51.12 ± 0.83	53.10 ± 1.41	50.24 ± 1.09	52.29 ± 2.55	49.85 ± 1.89	53.16 ± 2.65	53.93 ± 1.66	54.15 ± 1.09	**58.56 ± 1.04**
	F1	48.88 ± 1.57	50.76 ± 1.78	49.24 ± 1.05	49.80 ± 2.57	48.00 ± 0.80	50.42 ± 4.11	51.69 ± 1.87	49.43 ± 1.26	**55.42 ± 0.92**
APAVA	Acc	86.18 ± 1.93	80.73 ± 2.66	83.48 ± 1.07	82.12 ± 2.70	**86.40 ± 1.05**	84.33 ± 1.84	77.96 ± 0.62	79.06 ± 1.44	86.30 ± 1.17
	F1	85.18 ± 2.38	79.31 ± 2.81	81.75 ± 1.36	81.09 ± 2.83	85.46 ± 0.98	83.45 ± 2.05	74.96 ± 1.29	77.63 ± 1.01	**85.71 ± 1.22**
PTB	Acc	79.01 ± 4.61	83.39 ± 2.50	76.54 ± 1.15	82.15 ± 1.84	79.52 ± 2.79	84.98 ± 2.79	79.69 ± 1.51	85.25 ± 1.63	**88.29 ± 1.62**
	F1	70.13 ± 8.58	78.39 ± 3.87	68.00 ± 1.95	76.78 ± 2.60	71.02 ± 5.18	80.83 ± 4.18	73.50 ± 2.40	81.19 ± 2.42	**85.59 ± 2.20**
TDBRAIN	Acc	77.19 ± 3.99	74.25 ± 1.68	90.38 ± 2.66	91.06 ± 2.61	78.83 ± 2.65	88.88 ± 3.33	72.10 ± 5.77	86.33 ± 1.36	**93.12 ± 1.58**
	F1	76.83 ± 4.63	74.03 ± 1.58	90.33 ± 2.69	91.05 ± 2.62	78.80 ± 2.65	88.82 ± 3.41	71.85 ± 5.82	86.25 ± 1.39	**93.12 ± 1.58**

Table 2: Test performance (%) of spatial preprocessing methods with ResNet.

Dataset	Metric	Bipolar	CAR	CSP	ICA	Laplacian	Linear	SSP	SSS	CIF
ADFTD	Acc	49.79 ± 2.46	51.88 ± 1.00	47.95 ± 3.34	49.78 ± 2.90	50.11 ± 1.34	50.30 ± 1.19	51.02 ± 1.11	50.73 ± 1.11	**53.59 ± 1.28**
	F1	48.07 ± 2.04	48.57 ± 0.55	45.47 ± 3.20	45.11 ± 2.39	47.42 ± 1.81	47.82 ± 1.42	**48.95 ± 1.28**	45.50 ± 1.32	48.25 ± 1.10
APAVA	Acc	70.59 ± 6.44	66.16 ± 4.33	75.70 ± 3.46	70.83 ± 3.57	72.23 ± 7.72	66.30 ± 5.31	70.26 ± 8.79	69.76 ± 7.94	**81.84 ± 1.63**
	F1	68.75 ± 5.83	62.91 ± 5.33	74.46 ± 3.78	68.85 ± 3.21	70.84 ± 7.72	64.86 ± 6.09	66.74 ± 8.98	65.55 ± 7.25	**80.20 ± 1.90**
PTB	Acc	77.06 ± 2.19	80.41 ± 3.07	74.95 ± 3.24	78.86 ± 3.42	75.91 ± 2.83	77.63 ± 3.03	73.60 ± 2.88	**82.81 ± 1.37**	82.19 ± 1.67
	F1	66.74 ± 4.99	73.24 ± 5.57	63.95 ± 7.86	70.92 ± 5.96	64.51 ± 5.70	68.46 ± 5.66	61.05 ± 6.17	**77.73 ± 2.03**	77.20 ± 2.43
TDBRAIN	Acc	83.54 ± 3.31	84.90 ± 2.93	80.83 ± 3.41	86.98 ± 1.21	82.54 ± 4.53	85.69 ± 2.84	83.62 ± 3.98	85.83 ± 2.89	**88.73 ± 1.06**
	F1	83.48 ± 3.31	84.77 ± 3.07	80.78 ± 3.43	86.92 ± 1.23	82.48 ± 4.57	85.63 ± 2.86	83.47 ± 4.12	85.74 ± 2.95	**88.71 ± 1.06**

4.1 Main Results

We first evaluate CIF against commonly used spatial preprocessing methods, including Bipolar, CAR, CSP, ECG Lead, ICA, Laplacian, Linear, SSP, and SSS, using both HMTCN and ResNet backbones. As shown in Tables 1 and 2, CIF consistently outperforms conventional spatial preprocessing, achieving the best Accuracy on three of four datasets with both backbones and the best F1 on all four datasets with HMTCN. These results demonstrate that physiologically informed inter-channel fusion provides more discriminative representations than fixed spatial transformations.

We then compare CIF+HMTCN with existing state-of-the-art models under the subject-independent setting. Tables 3 and 21 summarize the reported and reproduced results, respectively. CIF+HMTCN achieves the highest average performance across six metrics on four of five datasets. On PTB-XL, it obtains the best AUROC and AUPRC and ranks second in Accuracy among reported methods. Against the reproduced baselines, CIF+HMTCN ranks first in Accuracy, Precision, AUROC, and AUPRC, demonstrating that a simple temporal model equipped with better input representations can outperform substantially more complex architectures.

Table 3: **Results of Subject-Independent Setup.** The results we compare include those reported by Medformer Wang et al. (2024a) and MedGNN Fan et al. (2025). Additionally, we have reproduced all the methods in the Table 21 to ensure a **fairer comparison**.The best result is highlighted in **bold**, and the second-best is underlined.

Datasets	Models	Accuracy ↑	Precision ↑	Recall ↑	F1 score ↑	AUROC ↑	AUPRC ↑
APAVA (2-Classes) Reported	**Autoformer**	68.64±1.82	68.48±2.10	68.77±2.27	68.06±1.94	75.94±3.61	74.38±4.05
	Crossformer	73.77±1.95	79.29±4.36	68.86±1.70	68.93±1.85	72.39±3.33	72.05±3.65
	FEDformer	74.94±2.15	74.59±1.50	73.56±3.55	73.51±3.39	83.72±1.97	82.94±2.37
	Informer	73.11±4.40	75.17±6.06	69.17±4.56	69.47±5.06	70.46±4.91	70.75±5.27
	iTransformer	74.55±1.66	74.77±2.10	71.76±1.72	72.30±1.79	85.59±1.55	84.39±1.57
	MTST	71.14±1.59	79.30±0.97	65.27±2.28	64.01±3.16	68.87±2.34	71.06±1.60
	Nonformer	71.89±3.81	71.80±4.58	69.44±3.56	69.74±3.84	70.55±2.96	70.78±4.08
	PatchTST	67.03±1.65	78.76±1.28	59.91±2.02	55.97±3.10	65.65±0.28	67.99±0.76
	Reformer	78.70±2.00	82.50±3.95	75.00±1.61	75.93±1.82	73.94±1.40	76.04±1.14
	Transformer	76.30±4.72	77.64±5.95	73.09±5.01	73.75±5.38	72.50±6.60	73.23±7.60
	Medformer	78.74±0.64	81.11±0.84	75.40±0.66	76.31±0.71	83.20±0.91	83.66±0.92
	MedGNN	82.60±0.35	**87.70±0.22**	78.93±0.09	80.25±0.16	85.93±0.26	-
	HMTCN + CIF	**86.30±1.05**	86.16±1.09	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	**94.42±0.49**
TDBrain (2-Classes) Reported	**Autoformer**	87.33±3.79	88.06±3.56	87.33±3.79	87.26±3.84	93.81±2.26	93.32±2.42
	Crossformer	81.56±2.19	81.97±2.25	81.56±2.19	81.50±2.20	91.20±1.78	91.51±1.71
	FEDformer	78.13±1.98	78.52±1.91	78.13±1.98	78.04±2.01	86.56±1.86	86.48±1.99
	Informer	89.02±2.50	89.43±2.14	89.02±2.50	88.98±2.54	96.64±0.68	96.75±0.63
	iTransformer	74.67±1.06	74.71±1.06	74.67±1.06	74.65±1.06	83.37±1.14	83.73±1.27
	MTST	76.96±3.76	77.24±3.59	76.96±3.76	76.88±3.83	85.27±4.46	82.81±5.64
	Nonformer	87.88±2.48	88.86±1.84	87.88±2.48	87.78±2.56	97.05±0.68	96.99±0.68
	PatchTST	79.25±3.79	79.60±4.09	79.25±3.79	79.20±3.77	87.95±4.96	86.36±6.67
	Reformer	87.92±2.01	88.64±1.40	87.92±2.01	87.85±2.08	96.30±0.54	96.40±0.45
	Transformer	87.17±1.67	87.99±1.68	87.17±1.67	87.10±1.68	96.28±0.92	96.34±0.81
	Medformer	89.62±0.81	89.68±0.78	89.62±0.81	89.62±0.81	96.41±0.35	96.51±0.33
	MedGNN	91.04±0.09	91.15±0.12	91.04±0.20	91.04±0.08	96.74±0.04	-
	HMTCN + CIF	**93.13±1.41**	**93.33±1.37**	**93.13±1.41**	**93.12±1.42**	**98.62±0.66**	**98.68±0.63**
ADFTD (3-Classes) Reported	**Autoformer**	45.25±1.48	43.67±1.94	42.96±2.03	42.59±1.85	61.02±1.82	43.10±2.30
	Crossformer	50.45±2.31	45.57±1.63	45.88±1.82	45.50±1.70	66.45±2.03	48.33±2.05
	FEDformer	46.30±0.59	46.05±0.76	44.22±1.38	43.91±1.37	62.62±1.75	46.11±1.44
	Informer	48.45±1.96	46.54±1.68	46.06±1.84	45.74±1.38	65.87±1.27	47.60±1.30
	iTransformer	52.60±1.59	46.79±1.27	47.28±1.29	46.79±1.13	67.26±1.16	49.53±1.21
	MTST	45.60±2.03	44.70±1.33	45.05±1.30	44.31±1.74	62.50±0.81	45.16±0.85
	Nonformer	49.95±1.05	47.71±0.97	47.46±1.50	46.96±1.35	66.23±1.37	47.33±1.78
	PatchTST	44.37±0.95	42.40±1.13	42.06±1.48	41.97±1.37	60.08±1.50	42.49±1.79
	Reformer	50.78±1.17	49.64±1.49	49.89±1.67	47.94±0.69	69.17±1.58	51.73±1.94
	Transformer	50.47±2.14	49.13±1.83	48.01±1.53	48.09±1.59	67.93±1.59	48.93±2.02
	Medformer	53.27±1.54	51.02±1.57	50.71±1.55	50.65±1.51	70.93±1.19	51.21±1.32
	MedGNN	56.12±0.11	55.07±0.09	55.47±0.34	55.00±0.24	74.68±0.33	-
	HMTCN + CIF	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**
PTB (2-Classes) Reported	**Autoformer**	73.35±2.10	72.11±2.89	63.24±3.17	63.69±3.84	78.54±3.48	74.25±3.53
	Crossformer	80.17±3.79	85.04±1.83	71.25±6.29	72.75±7.19	88.55±3.45	87.31±3.25
	FEDformer	76.05±2.54	77.58±3.61	66.10±3.55	67.14±4.37	85.93±4.31	82.59±5.42
	Informer	78.69±1.68	82.87±1.02	69.19±2.90	70.84±3.47	92.09±0.53	90.02±0.60
	iTransformer	83.89±0.71	88.25±1.18	76.39±1.01	79.06±1.06	91.18±1.16	90.93±0.98
	MTST	76.59±1.90	79.88±1.90	66.31±2.95	67.38±3.71	86.86±2.75	83.75±2.84
	Nonformer	78.66±0.49	82.77±0.86	69.12±0.87	70.90±1.00	89.37±2.51	86.67±2.38
	PatchTST	74.74±1.62	76.94±1.51	63.89±2.71	64.36±3.38	88.79±0.91	83.39±0.96
	Reformer	77.96±2.13	81.72±1.61	68.20±3.35	69.65±3.88	91.13±0.74	88.42±1.30
	Transformer	77.37±1.02	81.84±0.66	67.14±1.80	68.47±2.19	90.08±1.76	87.22±1.68
	Medformer	83.50±2.01	85.19±0.94	77.11±3.39	79.18±3.31	92.81±1.48	90.32±1.54
	MedGNN	84.53±0.28	87.35±0.45	77.90±0.66	80.40±0.62	93.31±0.46	-
	HMTCN + CIF	**88.29±1.45**	**90.66±1.48**	**83.21±2.02**	**85.59±1.96**	**94.28±0.93**	**93.78±1.11**
PTB-XL (5-Classes) Reported	**Autoformer**	61.68±2.72	51.60±1.64	49.10±1.52	48.85±2.27	82.04±1.44	51.93±1.71
	Crossformer	73.30±0.14	65.06±0.35	**61.23±0.33**	62.59±0.14	90.02±0.06	67.43±0.22
	FEDformer	57.20±9.47	52.38±6.09	49.04±7.26	47.89±8.44	82.13±4.17	52.31±7.03
	Informer	71.43±0.32	62.64±0.60	59.12±0.47	60.44±0.43	88.65±0.09	64.76±0.17
	iTransformer	69.28±0.22	59.59±0.45	54.62±0.18	56.20±0.19	86.71±0.10	60.27±0.21
	MTST	72.14±0.27	63.84±0.72	60.01±0.81	61.43±0.38	88.97±0.33	65.83±0.51
	Nonformer	70.56±0.55	61.57±0.66	57.75±0.72	59.10±0.66	88.32±0.36	63.40±0.79
	PatchTST	73.23±0.25	65.70±0.64	60.82±0.76	**62.61±0.34**	89.74±0.19	67.32±0.22
	Reformer	71.72±0.43	63.12±1.02	59.20±0.75	60.69±0.18	88.80±0.24	64.72±0.47
	Transformer	70.59±0.44	61.57±0.65	57.62±0.35	59.05±0.25	88.21±0.16	63.36±0.29
	Medformer	72.87±0.23	64.14±0.42	60.60±0.46	62.02±0.37	89.66±0.13	66.39±0.22
	MedGNN	**73.87±0.18**	**66.26±0.29**	61.13±0.29	62.54±0.20	90.21±0.15	-
	HMTCN + CIF	73.73±0.30	65.41±0.67	60.70±1.08	61.89±0.91	**90.53±0.22**	**67.75±0.75**

4.1.1 Efficiency Analysis

We evaluate the model efficiency in terms of accuracy, training speed, and memory footprint using two datasets: APAVA and TDBRAIN. In Figure 5, a marker closer to the upper-left corner indicates higher accuracy and faster training speed, while a smaller marker area corresponds to lower memory

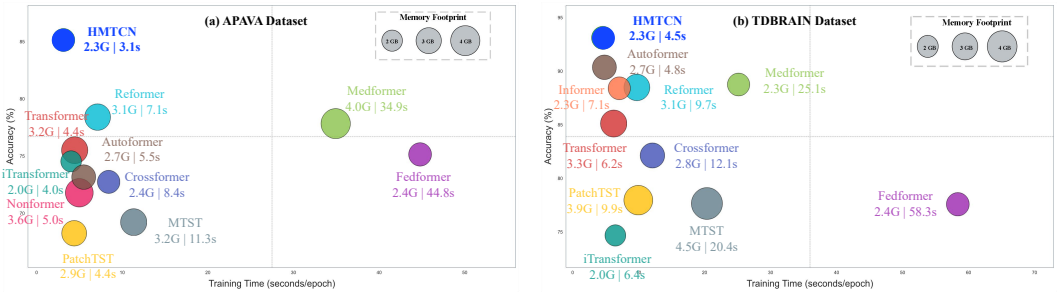


Figure 5: Effectiveness and efficiency on two datasets (subject-based).

usage. The results show that HMTCN achieves the best overall performance among all baseline methods, demonstrating its high efficiency and reliability across different application scenarios.

4.2 Ablation Study

(1) Effectiveness of CIF: Table 4 demonstrates the excellent performance of combining HMTCN with CIF, confirming the compatibility of the HMTCN with CIF. See Appendix M.1 for details. Appendix I presents ablation studies on the HMTCN architecture, the performance improvements from integrating CIF into its components, and comparisons with the vanilla TCN structure.

Table 4: Exploring the Integration of HMTCN Structure with CIF.

Datasets	APAVA		ADFTD		PTB		TDBRAIN		PTB-XL	
Metrics	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score	Accuracy	F1 Score
w/ CIF	86.30±1.05	85.71±1.09	58.56±0.93	55.42±0.82	88.29±1.45	85.59±1.96	93.13±1.41	93.12±1.42	73.73±0.30	61.89±0.91
w/o CIF	82.49±1.40	81.60±1.39	52.05±2.22	49.48±2.70	81.87±1.87	75.84±3.20	84.90±2.60	84.76±2.74	72.92±0.88	61.49±0.82
Improvement	**+3.81%**	**+4.11%**	**+6.51%**	**+5.94%**	**+6.42%**	**+9.75%**	**+8.23%**	**+8.36%**	**+0.81%**	**+0.40%**

(2) Hyperparameter Transfer and Adaptation: We evaluate the transferability of key hyperparameters (e.g., a, b, n) from HMTCN to other models. If transferred settings underperform, we further fine-tune them for adaptation. See Appendices M.2 and M.3. Figure 6 illustrates the outstanding performance of CIF when combined with other models.

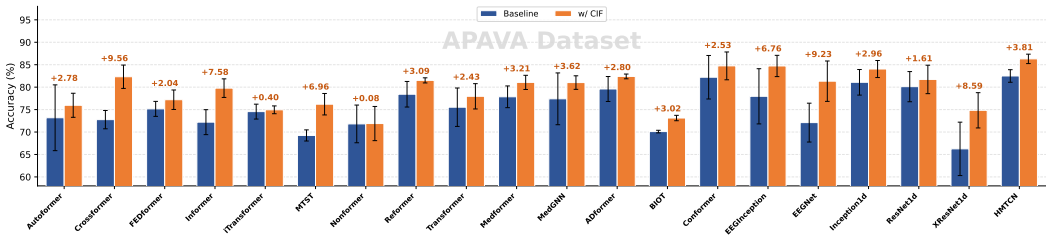


Figure 6: The improvements achieved by various baselines when combined with the CIF method.

(3) Exploring alternative fusion strategies: In Appendix J, we systematically evaluate several fusion approaches, including the fusion using only the Fp1 and Fp2 channels mentioned in the Introduction, the fusion of a reduced set of left-right symmetric channels, and our further exploration using canonical correlation analysis (CCA) for fusion. The experimental results demonstrate that CIF exhibits strong scalability, maintaining performance advantages across different fusion strategies.

(4)Results on general time series classification tasks. We follow the design of Medformer Wang et al. (2024a) and test it on two human activity recognition (HAR) datasets: FLAAP(13,123 samples, 10 classes) Kumar & Suresh (2022) and UCI-HAR(10,299 samples, 6 classes) Anguita et al. (2013). Additionally, to conduct a more comprehensive evaluation, following TimeMixer++ Wang et al.

(2025a), we used 10 multivariate datasets from the UEA Time Series Classification Archive (2018) for the assessment of classification tasks.

Table 5: Performance on the HAR and UCI-HAR non-medical time series datasets. Bold numbers indicate the best results. * denotes the results reported by Medformer.

Dataset / Metric		Crossformer * Zhang & Yan (2023)	Reformer * Kitaev et al. (2020)	Transformer * Vaswani et al. (2017)	TCN * Bai et al. (2018)	ModernTCN * Luo & Wang (2024)	Mamba * Gu & Dao (2023)	Medformer * Wang et al. (2024a)	HMTCN (This work)	HMTCN + CIF (This work)
FLAAP	Accuracy	75.84±0.52	71.65±1.27	74.96±1.25	66.48±1.66	74.80±0.96	64.87±2.78	76.44±0.64	76.08±0.81	**76.82**±1.32
(10 Classes)	F1 Score	75.52±0.66	71.14±1.45	74.49±1.39	65.29±1.74	74.35±0.85	64.14±2.70	76.25±0.65	75.54±0.94	**76.39**±1.18
UCI-HAR	Accuracy	89.74±1.08	88.44±2.02	88.86±1.65	93.08±0.95	91.44±1.01	87.78±1.10	91.65±0.74	93.72±0.73	**93.78**±0.32
(6 Classes)	F1 Score	89.70±1.10	88.34±1.98	88.80±1.67	93.19±0.88	91.47±0.98	87.72±1.10	91.61±0.75	93.69±0.76	**93.74**±0.34

As shown in Table 5, and Fig. 7, the combination of HMTCN and CIF consistently outperforms other architectures in general time series classification, achieving a 5.5% improvement over the original TCN and surpassing current SOTA methods. Although CIF was originally designed for MedTS, its integration with HMTCN significantly outperforms Transformer-based models in both medical and general time series classification tasks, demonstrating the effectiveness of our data-centric approach.

The results in Tables 4, 5, 19 and Figure 6 show that CIF achieves significant improvements in MedTS classification, while the gains on non-medical data are relatively limited. This observation further demonstrates that a data-driven perspective is particularly effective for MedTS classification with physiological characteristics.

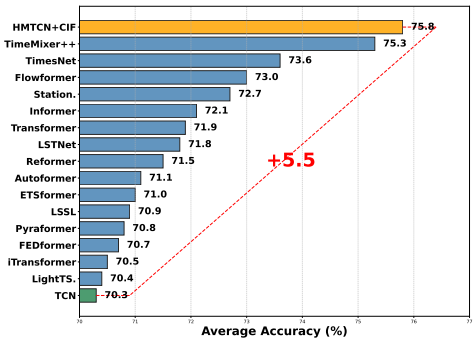


Figure 7: Average accuracy of various methods on the UEA dataset. More details in Appendix K.

(5) Additional analysis of CIF. We further examine the robustness and structural properties of CIF. The results show that the optimal fusion ratio (n) varies across datasets but generally remains around half of the available channels, indicating that its optimal value depends on the underlying signal and inter-channel structure rather than electrode density alone. We also find that the learned fusion coefficients remain stable across random seeds, while CIF retains a constrained physiological interaction topology and therefore does not degenerate into an unconstrained linear channel-mixing layer. Moreover, its channel-preserving design does not impose symmetry, allowing channel-specific and asymmetric physiological patterns to be retained. Finally, controlled experiments with a fixed HMTCN and ResNet backbone show that CIF consistently outperforms or remains competitive with representative spatial preprocessing methods, providing an apple-to-apple evaluation of its contribution. Detailed analyses of hyperparameter sensitivity, coefficient stability, and channel-wise effects are provided in Appendix O.

5 Conclusion

The time series field is rich with interesting observations, and the community stands to benefit from diverse modeling approaches as well as empirically driven studies. In this work, we contribute to this broader direction by introducing CIF, a data-centric framework for medical time series classification that shifts the focus from increasingly complex model architectures toward improving the quality and discriminability of input signals. CIF selectively enhances informative channels while preserving the original temporal structure, providing a lightweight and model-agnostic approach to improving the input representation before downstream temporal modeling. Extensive experiments across diverse EEG and ECG classification benchmarks and multiple backbone architectures demonstrate that CIF consistently improves classification performance with minimal architectural overhead. Our empirical findings reveal that improving the input representation can provide substantial benefits even without increasing the complexity of the downstream classifier, suggesting that *better features can be as important as better models*. We hope this work encourages the time series community to further explore data-centric approaches alongside model-centric innovations, particularly for medical time series where signal quality and task-relevant information are fundamental to reliable classification.

AI USE STATEMENT

In this work, we used generative AI tools solely for language polishing, including improving grammar, phrasing, clarity, and readability. The use of generative AI was limited to linguistic refinement and did not involve research ideation, methodological design, experimental design, data analysis, code generation, or the generation of scientific claims or results. All AI-assisted text was carefully reviewed and revised by the authors. We take full responsibility for the final content of this work.

REFERENCES

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, Jorge Luis Reyes-Ortiz, et al. A public domain dataset for human activity recognition using smartphones. In *Esann*, volume 3, pp. 3, 2013.

Aniqa Arif, Yihe Wang, Rui Yin, Xiang Zhang, and Ahmed Helmy. Ef-net: Mental state recognition by analyzing multimodal eeg-fnirs via cnn. *Sensors*, 24(6):1889, 2024.

Angelo Auricchio, Joost Lumens, and Frits W Prinzen. Does cardiac resynchronization therapy benefit patients with right bundle branch block: cardiac resynchronization therapy has a role in patients with right bundle branch block. *Circulation: Arrhythmia and Electrophysiology*, 7(3): 532–542, 2014.

Anthony Bagnall, Hoang Anh Dau, Jason Lines, Michael Flynn, James Large, Aaron Bostrom, Paul Southam, and Eamonn Keogh. The uea multivariate time series classification archive, 2018. *arXiv preprint arXiv:1811.00075*, 2018.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

Anthony J Bell and Terrence J Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.

Rodney J Croft and Robert J Barry. Removal of ocular artifact from the eeg: a review. *Neurophysiologie Clinique/Clinical Neurophysiology*, 30(1):5–19, 2000.

Abhimanyu Das, Weihao Kong, Andrew Leach, Shaan Mathur, Rajat Sen, and Rose Yu. Long-term forecasting with tide: Time-series dense encoder. *Trans. Mach. Learn. Res.*, 2023, 2023. URL https://openreview.net/forum?id=pCbC3aQB5W.

B Aldecoa Sanchez Del Rio, T Lopetegi, and I Romero. Assessment of different methods to estimate electrocardiogram signal quality. In *2011 Computing in Cardiology*, pp. 609–612. IEEE, 2011.

J Escudero, Daniel Abásolo, Roberto Hornero, Pedro Espino, and Miguel López. Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy. *Physiological measurement*, 27(11):1091, 2006.

Wei Fan, Jingru Fei, Dingyu Guo, Kun Yi, Xiaozhuang Song, Haolong Xiang, Hangting Ye, and Min Li. Towards multi-resolution spatiotemporal graph learning for medical time series classification. In *Proceedings of the ACM on Web Conference 2025*, pp. 5054–5064, 2025.

Jean-Yves Franceschi, Aymeric Dieuleveut, and Martin Jaggi. Unsupervised scalable representation learning for multivariate time series. *Advances in neural information processing systems*, 32, 2019.

Pascal Fries. Rhythms for cognition: communication through coherence. *Neuron*, 88(1):220–235, 2015.

Douglas Goldman. The clinical use of the "average" reference electrode in monopolar recording. *Electroencephalography and clinical neurophysiology*, 2(1-4):209–212, 1950.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023.

Albert Gu, Karan Goel, and Christopher Ré. Efficiently modeling long sequences with structured state spaces. In *ICLR*, 2022.

Alex Hanson, Koutilya Pnvr, Sanjukta Krishnagopal, and Larry Davis. Bidirectional convolutional lstm for the detection of violence in videos. In *European Conference on Computer Vision*, pp. 280–295. Springer, 2018.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. Bag of tricks for image classification with convolutional neural networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 558–567. IEEE, 2019.

Bo Hjorth. An on-line transformation of eeg scalp potentials into orthogonal source derivations. *Electroencephalography and clinical neurophysiology*, 39(5):526–530, 1975.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

Harold Hotelling. Relations between two sets of variates. In *Breakthroughs in statistics: methodology and distribution*, pp. 162–190. Springer, 1992.

Ming Hu, Jianfu Yin, Jing Wang, Yuqi Wang, Bingliang Hu, and Quan Wang. Specslice-convlstm: Medical hyperspectral image segmentation using spectral slicing and convlstm. In *International Conference on Pattern Recognition*, pp. 211–225. Springer, 2024.

Ming Hu, Jianfu Yin, Zhuangzhuang Ma, Jianheng Ma, Feiyu Zhu, Bingbing Wu, Ya Wen, Meng Wu, Cong Hu, Bingliang Hu, et al. beta-fft: Nonlinear interpolation and differentiated training strategies for semi-supervised medical image segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 30839–30849, 2025.

Hassan Ismail Fawaz, Benjamin Lucas, Germain Forestier, Charlotte Pelletier, Daniel F Schmidt, Jonathan Weber, Geoffrey I Webb, Lhassane Idoumghar, Pierre-Alain Muller, and François Petitjean. Inceptiontime: Finding alexnet for time series classification. *Data mining and knowledge discovery*, 34(6):1936–1962, 2020.

Yifan Jia, Hongyu Pei, Jiaqi Liang, Yuheng Zhou, Yanfei Yang, Yangyang Cui, and Min Xiang. Preprocessing and denoising techniques for electrocardiography and magnetocardiography: A review. *Bioengineering*, 11(11):1109, 2024.

Steven M Kay. *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

Nikita Kitaev, Lukasz Kaiser, and Anselm Levskaya. Reformer: The efficient transformer. In *International Conference on Learning Representations*, 2020.

Zoltan J Koles, Michael S Lazar, and Steven Z Zhou. Spatial patterns underlying population differences in the background eeg. *Brain topography*, 2(4):275–284, 1990.

Prabhat Kumar and S Suresh. Flaap: An open human activity recognition (har) dataset for learning and finding the associated activity patterns. *Procedia Computer Science*, 212:64–73, 2022.

Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. In *SIGIR*, 2018.

Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018.

Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.

Lei Lin, Beilei Xu, Wencheng Wu, Trevor W Richardson, and Edgar A Bernal. Medical time series classification with hierarchical attention-based temporal convolutional networks: A case study of myotonic dystrophy diagnosis. In *CVPR workshops*, volume 2, 2019.

Shizhan Liu, Hang Yu, Cong Liao, Jianguo Li, Weiyao Lin, Alex X Liu, and Schahram Dustdar. Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting. In *International conference on learning representations*, 2021.

Xi Liu, Xinming Zhang, Tao Yu, Ruochen Dang, Jian Li, Bingliang Hu, Quan Wang, and Rong Luo. Classification of self-limited epilepsy with centrotemporal spikes by classical machine learning and deep learning based on electroencephalogram data. *Brain Research*, 1830:148813, 2024a.

Yong Liu, Haixu Wu, Jianmin Wang, and Mingsheng Long. Non-stationary transformers: Exploring the stationarity in time series forecasting. *Advances in Neural Information Processing Systems*, 35: 9881–9893, 2022.

Yong Liu, Tengge Hu, Haoran Zhang, Haixu Wu, Shiyu Wang, Lintao Ma, and Mingsheng Long. itransformer: Inverted transformers are effective for time series forecasting. *International Conference on Learning Representations*, 2024b.

Donghao Luo and Xue Wang. Moderntcn: A modern pure convolution structure for general time series analysis. In *The Twelfth International Conference on Learning Representations*, 2024.

Scott Makeig, Anthony Bell, Tzyy-Ping Jung, and Terrence Sejnowski. Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, 8, 1995.

Andreas Miltiadous, Emmanouil Gionanidis, Katerina D Tzimourta, Nikolaos Giannakeas, and Alexandros T Tzallas. Dice-net: a novel convolution-transformer architecture for alzheimer detection in eeg signals. *IEEE Access*, 2023a.

Andreas Miltiadous, Katerina D Tzimourta, Theodora Afrantou, Panagiotis Ioannidis, Nikolaos Grigoriadis, Dimitrios G Tsalikakis, Pantelis Angelidis, Markos G Tsipouras, Euripidis Glavas, Nikolaos Giannakeas, et al. A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg. *Data*, 8(6):95, 2023b.

Ziti Fariha Mohd Apandi, Ryojun Ikeura, Soichiro Hayakawa, and Shigeyoshi Tsutsumi. An analysis of the effects of noisy electrocardiogram signal on heartbeat detection performance. *Bioengineering*, 7(2):53, 2020.

Guy P Nason and Rainer von Sachs. Wavelets in time-series analysis. *Philosophical transactions of the royal society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357 (1760):2511–2526, 1999.

Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, and Jayant Kalagnanam. A time series is worth 64 words: Long-term forecasting with transformers. *ICLR*, 2023.

Franklin F Offner. The eeg as potential mapping: the value of the average monopolar reference. *Electroencephalography and clinical neurophysiology*, 2(2):213–214, 1950.

B Sc JW Osselton. Acquisition of eeg data by bipolar unipolar and average reference methods: a theoretical comparison. *Electroencephalography and clinical neurophysiology*, 19(5):527–528, 1965.

Jaakko Matias Palva and Satu Palva. The correlation of the neuronal long-range temporal correlations, avalanche dynamics with the behavioral scaling laws and interindividual variability. *Criticality in Neural Systems*, pp. 105–126, 2014.

Bo Peng, Eric Alcaide, Quentin Anthony, Alon Albalak, Samuel Arcadinho, Stella Biderman, Huanqi Cao, Xin Cheng, Michael Chung, Matteo Grella, et al. Rwkv: Reinventing rnns for the transformer era. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 936–951, 2023. URL https://aclanthology.org/2023.findings-emnlp.936/.

PhysioToolkit PhysioBank. Physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

Pyotr G Platonov. P-wave morphology: underlying mechanisms and clinical implications. *Annals of Noninvasive Electrocardiology*, 17(3):161–169, 2012.

Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4): 441–446, 2000.

David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

Nand Sharma. Single-trial P300 classification using PCA with lda, QDA and neural networks. *CoRR*, abs/1712.01977, 2017. URL http://arxiv.org/abs/1712.01977.

Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2022.

Mohamed Sraitih, Younes Jabrane, and Amir Hajjam El Hassani. A robustness evaluation of machine learning algorithms for ecg myocardial infarction detection. *Journal of Clinical Medicine*, 11(17): 4935, 2022.

Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, and Wojciech Samek. Deep learning for ecg analysis: Benchmarks and insights from ptb-xl. *IEEE journal of biomedical and health informatics*, 25(5):1519–1528, 2020.

Samu Taulu and Matti Kajola. Presentation of electromagnetic multichannel data: the signal space separation method. *Journal of Applied Physics*, 97(12), 2005.

Ziwei Tian, Bingliang Hu, Yang Si, and Quan Wang. Automatic seizure detection and prediction based on brain connectivity features and a cnns meet transformers classifier. *Brain Sciences*, 13(5): 820, 2023.

Mikko A Uusitalo and Risto J Ilmoniemi. Signal-space projection method for separating meg or eeg into components. *Medical and biological engineering and computing*, 35(2):135–140, 1997.

Hanneke van Dijk, Guido van Wingen, Damiaan Denys, Sebastian Olbrich, Rosalinde van Ruth, and Martijn Arns. The two decades brainclinics research archive for insights in neurophysiology (tdbrain) database. *Scientific data*, 9(1):333, 2022.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dieter Kreiseler, Fatima I Lunze, Wojciech Samek, and Tobias Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset. *Scientific data*, 7(1):1–15, 2020.

Shiyu Wang, Jiawei Li, Xiaoming Shi, Zhou Ye, Baichuan Mo, Wenze Lin, Shengtong Ju, Zhixuan Chu, and Ming Jin. Timemixer++: A general time series pattern machine for universal predictive analysis. In *International Conference on Learning Representations*, volume 2025, pp. 16980–17016, 2025a.

Yihe Wang, Nan Huang, Taida Li, Yujun Yan, and Xiang Zhang. Medformer: A multi-granularity patching transformer for medical time-series classification. *arXiv preprint arXiv:2405.19363*, 2024a.

Yihe Wang, Nadia Mammone, Darina Petrovsky, Alexandros T Tzallas, Francesco C Morabito, and Xiang Zhang. Adformer: A multi-granularity transformer for eeg-based alzheimer's disease assessment. *arXiv preprint arXiv:2409.00032*, 2024b.

Yihe Wang, Nan Huang, Nadia Mammone, Marco Cecchi, and Xiang Zhang. Lead: An eeg foundation model for alzheimer's disease detection. *arXiv preprint arXiv:2502.01678*, 2025b.

Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International joint conference on neural networks (IJCNN)*, pp. 1578–1585. IEEE, 2017.

Gerald Woo, Chenghao Liu, Doyen Sahoo, Akshat Kumar, and Steven Hoi. Etsformer: Exponential smoothing transformers for time-series forecasting. *arXiv preprint arXiv:2202.01381*, 2022.

Haixu Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting. *Advances in Neural Information Processing Systems*, 34:22419–22430, 2021.

Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. *arXiv preprint arXiv:2210.02186*, 2022a.

Haixu Wu, Jialong Wu, Jiehui Xu, Jianmin Wang, and Mingsheng Long. Flowformer: Linearizing transformers with conservation flows. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 24226–24242. PMLR, 17–23 Jul 2022b. URL https://proceedings.mlr.press/v162/wu22m.html.

Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, and Mingsheng Long. Timesnet: Temporal 2d-variation modeling for general time series analysis. In *International Conference on Learning Representations*, 2023.

Qiao Xiao, Khuan Lee, Siti Aisah Mokhtar, Iskasymar Ismail, Ahmad Luqman bin Md Pauzi, Qiuxia Zhang, and Poh Ying Lim. Deep learning-based ecg arrhythmia classification: A systematic review. *Applied Sciences*, 13(8):4964, 2023.

Chaoqi Yang, M Westover, and Jimeng Sun. Biot: Biosignal transformer for cross-data learning in the wild. *Advances in Neural Information Processing Systems*, 36:78240–78260, 2023.

Kun Yi, Qi Zhang, Wei Fan, Longbing Cao, Shoujin Wang, Hui He, Guodong Long, Liang Hu, Qingsong Wen, and Hui Xiong. A survey on deep learning based time series analysis with frequency transformation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pp. 6206–6215, 2025.

Jianfu Yin, Nan Wang, Binliang Hu, Yao Wang, and Quan Wang. Degradation-aware deep unfolding network with transformer prior for video compressive imaging. *Signal Processing*, 227:109660, 2025.

Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.

Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. Are transformers effective for time series forecasting? *AAAI*, 2023.

Ce Zhang, Young-Keun Kim, and Azim Eskandarian. Eeg-inception: an accurate and robust end-to-end neural network for eeg-based motor imagery classification. *Journal of Neural Engineering*, 18 (4):046014, 2021.

Tianping Zhang, Yizhuo Zhang, Wei Cao, Jiang Bian, Xiaohan Yi, Shun Zheng, and Jian Li. Less is more: Fast multivariate time series forecasting with light sampling-oriented mlp structures. *arXiv preprint arXiv:2207.01186*, 2022.

Yitian Zhang, Liheng Ma, Soumyasundar Pal, Yingxue Zhang, and Mark Coates. Multi-resolution time-series transformer for long-term forecasting. In *International Conference on Artificial Intelligence and Statistics*, pp. 4222–4230. PMLR, 2024.

Yunhao Zhang and Junchi Yan. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *The Eleventh International Conference on Learning Representations*, 2023.

Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pp. 11106–11115. AAAI Press, 2021. doi: 10.1609/AAAI.V35I12.17325. URL https://doi.org/10.1609/aaai.v35i12.17325.

Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, and Rong Jin. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. In *International Conference on Machine Learning*, pp. 27268–27286. PMLR, 2022.

Feiyu Zhu, Jing Zhang, Ruochen Dang, Bingliang Hu, and Quan Wang. Mtnet: Multimodal transformer network for mild depression detection through fusion of eeg and eye tracking. *Biomedical Signal Processing and Control*, 100:106996, 2025.

APPENDIX

A Explanation of SNR Optimization via CIF. 18

B Offline Analysis of CIF-induced Trial-level Consistency 21

B.1 Channel-wise Trial-averaged SNR 21

B.2 CIF and Channel-wise SNR Gain 22

B.3 Offline Evaluation Protocol 23

B.4 Channel-wise Results on Training Splits 23

C Data Preprocessing 25

C.1 Channel Ordering 26

D Implementation Details 26

D.1 Evaluation metrics 28

E Detailed Comparison of Data-Centric Spatial Filtering 29

E.1 Unified Formulation 29

E.2 Fixed Spatial Transformations 29

E.3 Training-Estimated Spatial Projections 30

E.4 Fully Learnable Spatial Mixing 32

E.5 Channel-Imposed Fusion (CIF) 32

E.6 Computational Complexity and Parameterization 33

E.7 Modality Applicability 34

E.8 Conceptual Comparison of Spatial Priors 34

E.9 CIF versus Dense Spatial Matrices 35

E.10 Relationship to Data-Centric Modeling 36

E.11 Summary 36

F Model-centric Methods for Time Series Classification. 37

F.1 Transformer-based Time Series Methods. 37

F.2 Convolution-based Time Series Methods. 38

G HMTCN Structure Design and Theoretical Analysis 39

G.1 Dilated Convolution 39

G.2 Bidirectional Causal Convolution Structure 39

G.3 Multi-Scale Feature Learning and Residual Connections 39

H Pseudocode of CIF Method and Key Components of HMTCN 40

I Ablation Experiments of the HMTCN Structure 42

J Further exploration of physiological structures **43**

K Results on General Time Series **44**

L Reproduced Results of Existing Methods **45**

L.1 Results of Subject-Independent 45

M Ablation Study **47**

M.1 Module Study of CIF 47

M.2 Generalizability of Hyperparameters 66

M.3 Fine-tuning Transferred Parameters 70

N Limitations and Future Work **73**

O Detailed Analysis of CIF **73**

O.1 Sensitivity to the Fusion Ratio n 73

O.2 Stability and Structural Constraints of the Learned Coefficients 73

O.3 Preservation of Asymmetric Channel Information 74

O.4 Controlled Comparison with Spatial Preprocessing Methods 74

P Broader Impacts **74**

A EXPLANATION OF SNR OPTIMIZATION VIA CIF.

Overview: idealized mechanism vs. empirical proxy. This appendix serves two complementary purposes. First, we analyze an idealized two-channel linear model to explain *when* and *why* CIF fusion can improve output SNR under explicit statistical assumptions. Second, because real medical time series do not provide direct access to latent signal and noise components, we report a trial-averaged consistency proxy on the training split as an operational diagnostic. The SNR analysis below is therefore a mechanistic design principle, whereas the trial-level analysis in Section B does not claim to recover this true SNR on all datasets; it should be interpreted as a complementary data-level check rather than a replacement for downstream classification performance.

In analyzing the SNR improvement process of the CIF module, we consider two scenarios for the two-channel signals: the first scenario assumes equal signal variances and equal noise variances across the channels, while the second scenario allows the signal and noise variances to differ between the two channels.

SCENARIO 1: EQUAL SIGNAL AND NOISE VARIANCES

Consider the linear combination of the two observed signals:

$$y = ax_1 + bx_2, \tag{2}$$

where a and b are real coefficients. The observed signals are given by

$$x_1 = s_1 + \epsilon_1, \quad x_2 = s_2 + \epsilon_2, \tag{3}$$

with zero-mean signal and noise components:

$$\mathbb{E}[s_i] = \mathbb{E}[\epsilon_i] = 0, \quad i = 1, 2,$$

and mutually uncorrelated signal and noise components: $\mathrm{Cov}(s_i, \epsilon_j) = 0$.

The power of a zero-mean random signal is given by its variance:

$$P_s = \mathrm{Var}[s] = \mathbb{E}[(s - \mathbb{E}[s])^2] = \mathbb{E}[s^2]. \tag{4}$$

This is why, for zero-mean signals, the SNR can be expressed as a ratio of variances (or mean-square values) Kay (1993).

For the linear combination of signals:

$$\begin{aligned} \mathrm{Var}(as_1 + bs_2) &= a^2\mathrm{Var}(s_1) + b^2\mathrm{Var}(s_2) + 2ab\,\mathrm{Cov}(s_1, s_2) \\ &= a^2\sigma_s^2 + b^2\sigma_s^2 + 2ab(\rho\sigma_s^2) \\ &= \sigma_s^2\big(a^2 + b^2 + 2ab\rho\big), \end{aligned} \tag{5}$$

where $\rho = \mathrm{Corr}(s_1, s_2)$.

Similarly, the noise power of the linear combination is:

$$\begin{aligned} \mathrm{Var}(a\epsilon_1 + b\epsilon_2) &= a^2\mathrm{Var}(\epsilon_1) + b^2\mathrm{Var}(\epsilon_2) + 2ab\,\mathrm{Cov}(\epsilon_1, \epsilon_2) \\ &= a^2\sigma_\epsilon^2 + b^2\sigma_\epsilon^2 + 2ab(\gamma\sigma_\epsilon^2) \\ &= \sigma_\epsilon^2\big(a^2 + b^2 + 2ab\gamma\big), \end{aligned} \tag{6}$$

where $\gamma = \mathrm{Corr}(\epsilon_1, \epsilon_2)$.

Using the definition of SNR as the ratio of signal power to noise power:

$$\mathrm{SNR}_{\text{out}} = \frac{\mathrm{Var}(as_1 + bs_2)}{\mathrm{Var}(a\epsilon_1 + b\epsilon_2)} = \frac{\sigma_s^2(a^2 + b^2 + 2ab\rho)}{\sigma_\epsilon^2(a^2 + b^2 + 2ab\gamma)} = \mathrm{SNR}_{\text{in}} \cdot \frac{a^2 + b^2 + 2ab\rho}{a^2 + b^2 + 2ab\gamma}, \tag{7}$$

where $\mathrm{SNR}_{\text{in}} = \sigma_s^2/\sigma_\epsilon^2$.

> *Remark:* The zero-mean property ensures that the variance equals the mean-square value, which is why SNR can be expressed as a ratio of variances Kay (1993).

For SNR improvement relative to individual channels:

$$\frac{a^2 + b^2 + 2ab\rho}{a^2 + b^2 + 2ab\gamma} > 1 \quad \Rightarrow \quad 2ab(\rho - \gamma) > 0. \tag{8}$$

- **Difference Mode** ($ab < 0$): $\rho < \gamma$ — suppress correlated noise while possibly attenuating some correlated signal.
- **Cooperative Mode** ($ab > 0$): $\rho > \gamma$ — amplify correlated signals relative to less-correlated noise.

Optimization via the ratio $k = a/b$ (explicit choice of a, b)

Because the output SNR

$$\mathrm{SNR}_{\mathrm{out}} = \mathrm{SNR}_{\mathrm{in}} \cdot \frac{a^2 + b^2 + 2ab\rho}{a^2 + b^2 + 2ab\gamma} \tag{9}$$

is homogeneous of degree zero in (a, b) (i.e. invariant under common scaling $(a, b) \mapsto (ca, cb)$, $c \neq 0$), only the ratio $k = a/b$ matters. Assume $b \neq 0$ and set

$$k = \frac{a}{b}. \tag{10}$$

Define

$$F(k) \;= \frac{a^2 + b^2 + 2ab\rho}{a^2 + b^2 + 2ab\gamma} \;=\; \frac{k^2 + 1 + 2k\rho}{k^2 + 1 + 2k\gamma}, \tag{11}$$

so that

$$\mathrm{SNR}_{\mathrm{out}} = \mathrm{SNR}_{\mathrm{in}} \cdot F(k). \tag{12}$$

We therefore study extrema of $F(k)$ to determine favorable relative weights $a : b$.

Let

$$N(k) = k^2 + 1 + 2k\rho, \qquad D(k) = k^2 + 1 + 2k\gamma, \tag{13}$$

so $F(k) = N(k)/D(k)$. Then

$$N'(k) = 2k + 2\rho, \qquad D'(k) = 2k + 2\gamma. \tag{14}$$

Using the quotient rule,

$$F'(k) = \frac{N'(k)D(k) - N(k)D'(k)}{D(k)^2}. \tag{15}$$

The numerator simplifies and factorizes as

$$N'(k)D(k) - N(k)D'(k) = 2(\gamma - \rho)(k - 1)(k + 1), \tag{16}$$

so we obtain

$$F'(k) = \frac{2(\gamma - \rho)(k - 1)(k + 1)}{(k^2 + 1 + 2k\gamma)^2}. \tag{17}$$

Critical points and their nature

From equation 17, the critical points satisfy

$$F'(k) = 0 \iff k = -1,\ k = +1 \quad \text{or} \quad \rho = \gamma. \tag{18}$$

- If $\rho > \gamma$ (so $\gamma - \rho < 0$), then $(k - 1)(k + 1) < 0$ for $k \in (-1, 1)$ and $(k - 1)(k + 1) > 0$ for $k > 1$ or $k < -1$. Hence $F'(k) > 0$ on $(-1, 1)$ and $F'(k) < 0$ outside, so

$$k = 1 \text{ is a local maximum}, \qquad k = -1 \text{ is a local minimum}. \tag{19}$$

- If $\rho < \gamma$ (so $\gamma - \rho > 0$), the inequalities reverse: $F'(k) < 0$ on $(-1, 1)$ and $F'(k) > 0$ for $|k| > 1$, thus

$$k = -1 \text{ is a local maximum}, \qquad k = 1 \text{ is a local minimum}. \tag{20}$$

- If $\rho = \gamma$, then from equation 11 we have

$$F(k) \equiv 1, \quad \forall k \in \mathbb{R}, \tag{21}$$

so no choice of k changes the SNR.

Interpretation in terms of a, b and explicit choice

Recall from equation 10 that $k = a/b$. The critical ratios $k = \pm 1$ correspond to equal-magnitude weights:

$$k = +1 \iff a = b, \qquad k = -1 \iff a = -b. \tag{22}$$

Since $\mathrm{SNR}_{\mathrm{out}}$ in equation 9 is invariant to a common scaling of (a, b), we may fix a convenient normalization:

Fix $b = 1$. Then $k = a$ and the recommended choices are

$$\text{Cooperative mode } (\rho > \gamma): \quad (a, b) = (1, 1), \tag{23}$$

$$\text{Differential mode } (\rho < \gamma): \quad (a, b) = (-1, 1). \tag{24}$$

Fix $b = -1$. Then $k = a$ and the recommended choices are

$$\text{Cooperative mode } (\rho > \gamma): \quad (a, b) = (-1, -1), \tag{25}$$

$$\text{Differential mode } (\rho < \gamma): \quad (a, b) = (1, -1). \tag{26}$$

Degenerate and neutral cases

- If $\rho = \gamma$, then by equation 21 we have

$$\mathrm{SNR}_{\mathrm{out}} = \mathrm{SNR}_{\mathrm{in}}, \quad \forall (a, b) \neq (0, 0), \tag{27}$$

  i.e. the linear combination cannot improve SNR (except for singular noise-cancellation cases).
- The denominator in equation 11 vanishes when

$$k^2 + 1 + 2k\gamma = 0, \tag{28}$$

  which corresponds to zero output noise variance. This is a nongeneric, degenerate configuration (perfect noise cancellation).

From equation 19–equation 20 and equation 22, we conclude:

- If $\rho > \gamma$: the SNR is maximized (among equal-variance combinations) by equal-phase combining $a = b$ (cooperative mode).
- If $\rho < \gamma$: the SNR is maximized by equal-magnitude opposite-phase combining $a = -b$ (differential mode).
- If $\rho = \gamma$: $F(k) \equiv 1$ and no linear combining improves SNR, up to degenerate cases.

Remarks on the choice of (a, b) ranges. If the goal is to find the SNR extremum points, it is sufficient to consider (a, b) combinations where a and b are either of the same sign (both positive or both negative) or of opposite signs. For example, exploring (a, b) within the rectangle $[-1, 1] \times [-1, 1]$ already includes the critical ratios $k = \pm 1$ and thus captures the SNR maxima and minima. While this rectangle does not cover all possible k values (i.e., all real ratios a/b), it is enough to determine the locations of the extremal points, thanks to the homogeneity of SNR in (a, b).

SCENARIO 2: UNEQUAL SIGNAL AND NOISE VARIANCES

We now relax the equal-variance assumption and allow $\mathrm{Var}(s_i) = \sigma_{s,i}^2$ and $\mathrm{Var}(\epsilon_i) = \sigma_{\epsilon,i}^2$ for $i \in \{1, 2\}$. With $\mathrm{Cov}(s_1, s_2) = \rho\, \sigma_{s,1}\sigma_{s,2}$ and $\mathrm{Cov}(\epsilon_1, \epsilon_2) = \gamma\, \sigma_{\epsilon,1}\sigma_{\epsilon,2}$, the output SNR factor becomes

$$\frac{\mathrm{Var}(as_1 + bs_2)}{\mathrm{Var}(a\epsilon_1 + b\epsilon_2)} = \frac{a^2\sigma_{s,1}^2 + b^2\sigma_{s,2}^2 + 2ab\rho\sigma_{s,1}\sigma_{s,2}}{a^2\sigma_{\epsilon,1}^2 + b^2\sigma_{\epsilon,2}^2 + 2ab\gamma\sigma_{\epsilon,1}\sigma_{\epsilon,2}}. \tag{29}$$

When $\sigma_{s,1} = \sigma_{s,2} = \sigma_s$ and $\sigma_{\epsilon,1} = \sigma_{\epsilon,2} = \sigma_\epsilon$, Eq. equation 29 reduces to Scenario 1. Thus, Scenario 1 is not a separate mechanism, but a tractable special case that yields closed-form extrema.

In the general case, the optimal (a, b) depends on both correlation structure (ρ, γ) and channel-wise variance ratios. This motivates treating (a, b) in CIF as dataset-adaptive parameters (fixed by domain priors, selected by validation, or learned end-to-end), rather than restricting the method to the symmetric choices $a = \pm b$ derived in Scenario 1. Importantly, the qualitative rule remains: fusion is beneficial when the chosen combination emphasizes correlated signal structure relative to correlated noise structure under the channel variance profile of the dataset.

While the idealized analysis above characterizes CIF under explicit latent signal/noise assumptions, real MedTS recordings do not provide direct access to (s_i, ϵ_i) or to $(\rho, \gamma, \sigma_s, \sigma_\epsilon)$. Section B therefore adopts a trial-averaged consistency proxy as a practical surrogate for examining whether CIF increases reproducible structure in the modified channels on the training split.

B Offline Analysis of CIF-induced Trial-level Consistency

To characterize the effect of CIF on the consistency of multichannel responses across trials, we perform an offline analysis using a *trial-averaged SNR* estimator. This measure is used as a **proxy for trial-level consistency**, rather than as a direct measurement of physiological signal-to-noise ratio.

Scope and limitations. This section should be interpreted as an *operational consistency analysis* under real-data constraints, not as a claim that CIF optimizes a physiologically exact SNR on all MedTS settings. The idealized model above assumes access to latent signal and noise components; in practice, these components are unobservable, so we estimate a trial-averaged consistency proxy from repeated segments in the training split. This estimator is most appropriate when trials contain approximately phase-locked or repeatable structure (e.g., event-related responses). For spontaneous, non-phase-locked resting-state EEG or unaligned ECG segments, averaging independent trials can attenuate oscillatory components through destructive interference. In such cases, $\mathrm{SNR}_{\mathrm{trial}}^{(c)}$ should be read as the strength of *trial-reproducible temporal structure* relative to *cross-trial variability*, which may include phase-invariant artifacts, baseline wander, or other repeatable nuisance components. Accordingly, $\Delta\mathrm{SNR}^{(c)}$ is used only as a data-level diagnostic complement to downstream classification, not as a standalone proof of physiological SNR improvement.

The estimator follows the conventional trial-averaging principle when repeatable structure is present: components that are consistently present across trials are reinforced by averaging, whereas trial-specific variability is attenuated. For non-phase-locked spontaneous activity, however, the cross-trial average primarily captures reproducible structure rather than the full neural signal, and oscillatory components may be partially cancelled by averaging. Specifically, the cross-trial average is treated as the trial-consistent component, while the variance across trials is used to characterize inter-trial variability.

Since different EEG channels may exhibit distinct spatial and temporal characteristics, pooling multiple channels into a single global SNR can obscure channel-specific effects. Therefore, we perform the analysis **independently for each channel modified by CIF**, and report the corresponding SNR gain on a per-channel basis.

B.1 Channel-wise Trial-averaged SNR

Let a data split contain B EEG trials, where the b-th trial is denoted by $\mathbf{x}_b \in \mathbb{R}^{T\times C}$, with T temporal samples and C channels. For an evaluated signal tensor $\mathbf{X} = \{\mathbf{x}_b\}_{b=1}^{B}$, the cross-trial mean response of channel c is defined as

$$\bar{x}_{t,c} = \frac{1}{B}\sum_{b=1}^{B} x_{b,t,c}, \qquad t = 1, \ldots, T, \quad c = 1, \ldots, C. \tag{30}$$

For each channel c, we estimate the trial-consistent signal power as the temporal variance of the cross-trial mean:

$$P_{\mathrm{s}}^{(c)} = \mathrm{Var}_t\left(\bar{x}_{t,c}\right), \tag{31}$$

and characterize inter-trial variability by the average variance across trials over all temporal locations:

$$P_{\mathrm{n}}^{(c)} = \frac{1}{T}\sum_{t=1}^{T} \mathrm{Var}_b\left(x_{b,t,c}\right). \tag{32}$$

Here, $\mathrm{Var}_b(\cdot)$ denotes the population variance along the trial dimension. Equivalently,

$$P_{\mathrm{n}}^{(c)} = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{B}\sum_{b=1}^{B} x_{b,t,c}^2 - \left(\frac{1}{B}\sum_{b=1}^{B} x_{b,t,c}\right)^2\right]. \tag{33}$$

The channel-wise trial-averaged SNR is then defined in decibels as

$$\mathrm{SNR}_{\mathrm{trial}}^{(c)} = 10\log_{10}\left(\frac{P_{\mathrm{s}}^{(c)}}{P_{\mathrm{n}}^{(c)}}\right). \tag{34}$$

Thus, a larger $\mathrm{SNR}_{\mathrm{trial}}^{(c)}$ indicates that the channel exhibits a stronger trial-reproducible component relative to its inter-trial variability. This metric should be interpreted as a consistency-oriented SNR proxy rather than a strict physiological signal/noise decomposition, especially on resting-state EEG datasets where phase-unlocked oscillations may not survive trial averaging.

B.2 CIF and Channel-wise SNR Gain

CIF (Channel Interpolation Fusion) performs a linear fusion between two groups of EEG channels. Given an input tensor $\mathbf{X}$, the first and last n channels are denoted as

$$\mathbf{X}_{\mathrm{front}} = \mathbf{X}[:,:,:n], \qquad \mathbf{X}_{\mathrm{back}} = \mathbf{X}[:,:,-n:]. \tag{35}$$

The fused representation is computed as

$$\mathbf{X}_{\mathrm{fused}} = a\mathbf{X}_{\mathrm{front}} + b\mathbf{X}_{\mathrm{back}}, \tag{36}$$

where $a, b \in \mathbb{R}$ are dataset-specific fusion coefficients. Depending on the sign of t, the fused representation replaces either the first n channels ($t > 0$) or the last n channels ($t < 0$).

Let $\mathcal{C}_{\mathrm{chg}}$ denote the set of channels modified by CIF, with

$$|\mathcal{C}_{\mathrm{chg}}| = n. \tag{37}$$

Let $\mathbf{X}_{\mathrm{before}}$ denote the original values of the modified channels before fusion, and let $\mathbf{X}'$ denote the corresponding CIF-transformed channels. For each modified channel $c \in \mathcal{C}_{\mathrm{chg}}$, the CIF-induced SNR gain is defined as

$$\Delta\mathrm{SNR}^{(c)} = \mathrm{SNR}_{\mathrm{trial}}^{(c)}\left(\mathbf{X}'\right) - \mathrm{SNR}_{\mathrm{trial}}^{(c)}\left(\mathbf{X}_{\mathrm{before}}\right). \tag{38}$$

A positive $\Delta\mathrm{SNR}^{(c)}$ indicates that CIF increases the relative strength of the trial-reproducible component in channel c with respect to its inter-trial variability under the consistency proxy defined above. Conversely, a negative value indicates a reduction in the corresponding trial-level reproducibility.

Importantly, the gain is computed independently for each modified channel rather than by pooling channels. This design preserves the spatial specificity of CIF and allows its effect to be examined at the same channel resolution at which the preprocessing operation is applied.

B.3 Offline Evaluation Protocol

To estimate the effect of CIF without relying on training logs or learned model representations, we perform the analysis directly on the complete training split. The procedure consists of the following steps:

1. **Load the training data.** All B trials in the training split are collected to form $\mathbf{X} \in \mathbb{R}^{B\times T\times C}$.
2. **Apply CIF.** The same CIF configuration (t, n, a, b) used by the model is applied to obtain $\mathbf{X}'$. The original values of the modified channels are retained as $\mathbf{X}_{\text{before}}$.
3. **Compute channel-wise SNR.** For every $c \in \mathcal{C}_{\text{chg}}$, the trial-averaged SNR is independently computed for both $\mathbf{X}_{\text{before}}$ and $\mathbf{X}'$ according to Eq. equation 34.
4. **Compute channel-wise gain.** The difference between the post-CIF and pre-CIF SNR values is calculated according to Eq. equation 38.

For memory efficiency, the first- and second-order moments across trials are accumulated in a streaming manner:

$$\mathbb{E}_b[x] = \frac{1}{B}\sum_{b=1}^{B} x_b, \qquad \mathbb{E}_b[x^2] = \frac{1}{B}\sum_{b=1}^{B} x_b^2. \tag{39}$$

The inter-trial variance is subsequently obtained as

$$\mathrm{Var}_b(x) = \mathbb{E}_b[x^2] - \mathbb{E}_b[x]^2. \tag{40}$$

This streaming computation is equivalent to directly evaluating Eqs. equation 30–equation 33, while avoiding the need to simultaneously store all training trials in memory. When $B < 2$, the inter-trial variance cannot be reliably estimated, and the corresponding channel-wise SNR is treated as undefined.

For the channel indices reported in the following tables, we use local indices within $\mathcal{C}_{\text{chg}}$. Specifically, when $t > 0$, local indices $0, \ldots, n-1$ correspond to global channels $0, \ldots, n-1$; when $t < 0$, local index i corresponds to global channel $C - n + i$.

B.4 Channel-wise Results on Training Splits

The following tables report the channel-wise trial-averaged SNR before and after CIF, together with the corresponding gain, for every modified channel on the training split. All values are reported in dB.

Table 6: Channel-wise trial-averaged SNR on the APAVA training split ($t = 1$, $n = 9$, $a = -0.80$, $b = -0.60$).

Channel Index	Pre-CIF $\mathrm{SNR}^{(c)}_{\text{trial}}$	Post-CIF $\mathrm{SNR}^{(c)}_{\text{trial}}$	$\Delta\mathrm{SNR}^{(c)}$
0	-31.65	-32.61	-0.96
1	-30.99	-31.26	-0.27
2	-32.67	-32.88	-0.21
3	-34.76	-33.49	$+1.26$
4	-34.98	-34.62	$+0.36$
5	-35.53	-34.53	$+1.00$
6	-31.10	-31.31	-0.21
7	-34.76	-32.68	$+2.07$
8	-36.29	-36.20	$+0.09$

Across the four datasets, the channel-wise results demonstrate that the effect of CIF is inherently **channel-dependent**. For example, on PTB, channel 2 achieves a gain of $+11.32$ dB, whereas channel

Table 7: Channel-wise trial-averaged SNR on the ADFTD training split ($t = 1$, $n = 10$, $a = -0.19$, $b = -0.27$).

Channel Index	Pre-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	Post-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	$\Delta\mathrm{SNR}^{(c)}$
0	-48.14	-50.53	-2.40
1	-49.45	-48.91	$+0.54$
2	-49.57	-49.94	-0.38
3	-50.69	-50.42	$+0.28$
4	-49.89	-49.80	$+0.09$
5	-50.91	-51.21	-0.30
6	-51.01	-51.10	-0.08
7	-51.22	-50.98	$+0.23$
8	-50.44	-50.93	-0.49
9	-51.78	-51.54	$+0.24$

Table 8: Channel-wise trial-averaged SNR on the PTB training split ($t = 1$, $n = 8$, $a = +0.21$, $b = -0.50$).

Channel Index	Pre-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	Post-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	$\Delta\mathrm{SNR}^{(c)}$
0	-5.62	-2.60	$+3.02$
1	-9.17	-6.71	$+2.46$
2	-19.23	-7.91	$+11.32$
3	-6.39	-3.67	$+2.73$
4	-11.10	-7.02	$+4.09$
5	-14.28	-4.50	$+9.78$
6	-4.44	-8.03	-3.58
7	-3.73	-3.13	$+0.60$

Table 9: Channel-wise trial-averaged SNR on the TDBRAIN training split ($t = -1$, $n = 25$, $a = +1.00$, $b = +1.10$). Local index i corresponds to global channel $C - n + i$.

Channel Index	Pre-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	Post-CIF $\mathrm{SNR}^{(c)}_{\mathrm{trial}}$	$\Delta\mathrm{SNR}^{(c)}$
0	-35.74	-35.16	$+0.58$
1	-37.08	-36.56	$+0.52$
2	-33.88	-37.81	-3.93
3	-37.82	-31.95	$+5.87$
4	-37.77	-36.74	$+1.02$
5	-33.91	-32.74	$+1.17$
6	-31.71	-31.43	$+0.27$
7	-34.56	-37.10	-2.54
8	-36.25	-35.80	$+0.45$
9	-36.99	-38.02	-1.02
10	-36.85	-34.66	$+2.19$
11	-36.24	-37.04	-0.81
12	-36.43	-37.35	-0.93
13	-36.36	-37.31	-0.95
14	-33.66	-31.83	$+1.83$
15	-36.46	-35.40	$+1.07$
16	-35.25	-37.18	-1.93
17	-35.89	-36.52	-0.63
18	-36.69	-35.97	$+0.73$
19	-26.19	-28.44	-2.25
20	-36.93	-35.48	$+1.45$
21	-35.57	-34.62	$+0.95$
22	-34.22	-33.81	$+0.41$
23	-31.16	-33.12	-1.97
24	-27.30	-30.82	-3.52

6 exhibits a decrease of -3.58 dB. Similarly, APAVA shows a gain of $+2.07$ dB on channel 7 but a decrease of -0.96 dB on channel 0.

These results highlight the importance of evaluating CIF at the channel level. Rather than assuming a uniform effect across channels, the proposed analysis reveals the **spatially heterogeneous effect** of channel fusion. The resulting $\Delta\mathrm{SNR}^{(c)}$ provides an approximate data-level estimate of whether CIF enhances trial-reproducible structure in each modified channel, while serving as a complementary diagnostic measure rather than a replacement for downstream classification performance. On resting-state EEG datasets (APAVA, TDBRAIN, and ADFTD), positive $\Delta\mathrm{SNR}^{(c)}$ should therefore not be over-interpreted as universal physiological denoising; it indicates increased cross-trial reproducibility under the chosen segmentation protocol.

Table 10: **Comparison of HMTCN, HMTCN + CIF, and HMTCN + CIF (select) under Subject-Independent Setup.** CIF (select) applies channel fusion only to channels whose trial-averaged consistency proxy increases on the training set; all other channels remain unchanged.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
APAVA (2-Classes)	HMTCN	82.49±1.40	82.38±1.79	81.20±1.32	81.60±1.39	91.10±1.63	91.30±1.71
	HMTCN + CIF	**86.30±1.05**	**86.16±1.09**	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	**94.42±0.49**
	HMTCN + CIF (select)	83.61±2.28	83.34±2.71	82.72±1.93	82.93±2.21	92.48±1.56	92.76±1.50
ADFTD (3-Classes)	HMTCN	52.05±2.22	50.45±3.00	50.40±2.55	49.48±2.70	69.43±2.84	50.99±3.15
	HMTCN + CIF	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**
	HMTCN + CIF (select)	53.08±2.02	50.12±2.21	50.02±2.25	49.61±2.50	70.28±1.61	52.59±1.61
PTB (2-Classes)	HMTCN	81.87±1.87	86.50±1.24	73.49±2.90	75.84±3.20	94.20±0.29	93.04±0.45
	HMTCN + CIF	**88.29±1.45**	**90.66±1.48**	**83.21±2.02**	**85.59±1.96**	**94.28±0.93**	**93.78±1.11**
	HMTCN + CIF (select)	82.16±2.74	86.63±1.64	73.98±4.34	76.28±4.60	93.76±1.48	92.42±2.17
TDBrain (2-Classes)	HMTCN	84.90±2.60	86.02±2.00	84.90±2.60	84.76±2.74	93.94±1.92	94.20±1.85
	HMTCN + CIF	**93.13±1.41**	**93.33±1.37**	**93.13±1.41**	**93.12±1.42**	**98.62±0.66**	**98.68±0.63**
	HMTCN + CIF (select)	91.25±1.44	91.30±1.48	91.25±1.44	91.25±1.43	97.68±0.98	97.78±0.94

To further investigate the relationship between CIF and channel-wise consistency-proxy changes, we analyze the effect of global CIF on individual channels. We observe that global CIF increases the trial-averaged consistency proxy of some channels while decreasing that of others, indicating that CIF does not uniformly improve this metric across all channels. To further examine whether the channels with higher proxy values are also beneficial for classification, we introduce **CIF (select)**, which applies channel fusion only to channels whose trial-averaged SNR proxy increases after CIF, while leaving the remaining channels unchanged. As shown in Table 10, CIF (select) consistently improves upon the HMTCN baseline across all four datasets, suggesting an exploratory association between consistency-proxy gains and classification performance. However, CIF (select) remains inferior to global CIF overall. This gap indicates that the effectiveness of CIF cannot be fully explained by the trial-averaged proxy alone, as channels without proxy gains may still contain complementary task-relevant information. In addition, global CIF incorporates information from all available channel pairs, providing a broader utilization of inter-channel information. These results therefore support a partial association between channel-wise consistency improvement and classification gains, while also highlighting that the benefit of CIF extends beyond this proxy alone.

C Data Preprocessing

(1) **APAVA** Escudero et al. (2006) is an EEG dataset where each sample is assigned a binary label indicating whether the subject has Alzheimer's disease. (2) **TDBRAIN** van Dijk et al. (2022) is an EEG dataset with a binary label assigned to each sample, indicating whether the subject has Parkinson's disease. (3) **ADFTD** Miltiadous et al. (2023b;a) is an EEG dataset with a three-class label for each sample, categorizing the subject as Healthy, having Frontotemporal Dementia, or Alzheimer's disease. (4) **PTB** PhysioBank (2000) is an ECG dataset where each sample is labeled with a binary indicator of Myocardial Infarction. (5) **PTB-XL** Wagner et al. (2020) is an ECG dataset with a five-class label for each sample, representing various heart conditions. The data processing methodology is the same as that of Medformer Wang et al. (2024a) and MedGNN Fan et al. (2025).

Table 11: **The information of processed datasets.** The table shows the number of subjects, samples, classes, channels, sampling rate, sample timestamps, modality of MedTS, and file size. Here, **#-Timestamps** indicates the number of timestamps per sample.All data processing procedures follow those of Medformer Wang et al. (2024a).

Datasets	#-Subject	#-Sample	#-Class	#-Channel	#-Timestamps	Sampling Rate	Modality	File Size
APAVA	23	5,967	2	16	256	256Hz	EEG	186MB
ADFTD	88	69,752	3	19	256	256Hz	EEG	2.52GB
TDBrain	72	6,240	2	33	256	256Hz	EEG	571MB
PTB	198	64,356	2	15	300	250Hz	ECG	2.15GB
PTB-XL	17,596	191,400	5	12	250	250Hz	ECG	4.28GB

C.1 Channel Ordering

For all datasets, we preserve the channel ordering adopted by the data preprocessing pipeline provided by Medformer Wang et al. (2024a), ensuring consistency with the corresponding benchmark and model implementation. The channel orders used in our experiments are as follows:

- **APAVA (EEG):** C3, C4, F3, F4, F7, F8, Fp1, Fp2, O1, O2, P3, P4, T3, T4, T5, T6.
- **ADFTD (EEG):** Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2.
- **TDBrain (EEG):** Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7/T5, P3, Pz, P4, P8/T6, O1, Oz, O2, VPVA, VNVB, HOHL, HNHR, Erbs, OrbOcc, MASS.
- **PTB (ECG):** I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6, Vx, Vy, Vz.
- **PTB-XL (ECG):** I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, V6.

Maintaining a consistent channel ordering is particularly important for CIF, as its channel-wise fusion operation is defined with respect to the fixed channel positions. Therefore, all channel indices and fusion directions reported in our experiments directly correspond to the channel ordering specified above.

Table 12: Datasets and mapping details of UEA dataset (Bagnall et al., 2018).

Dataset	Sample Numbers(train set,test set)	Variable Number	Series Length
EthanolConcentration	(261, 263)	3	1751
FaceDetection	(5890, 3524)	144	62
Handwriting	(150, 850)	3	152
Heartbeat	(204, 205)	61	405
JapaneseVowels	(270, 370)	12	29
PEMSSF	(267, 173)	963	144
SelfRegulationSCP1	(268, 293)	6	896
SelfRegulationSCP2	(200, 180)	7	1152
SpokenArabicDigits	(6599, 2199)	13	93
UWaveGestureLibrary	(120, 320)	3	315

D Implementation Details

We implement our method and all the baselines based on the Time-Series-Library project[1] from Tsinghua University Wu et al. (2023), which integrates all methods under the same framework and training techniques to ensure a relatively fair comparison. The 12 baseline time series transformer methods are Autoformer Wu et al. (2021), Crossformer Zhang & Yan (2023), FEDformer Zhou et al. (2022), Informer Zhou et al. (2021), iTransformer Liu et al. (2024b), MTST Zhang et al. (2024),

[1] https://github.com/thuml/Time-Series-Library

Nonformer Liu et al. (2022), PatchTST Nie et al. (2023), Reformer Kitaev et al. (2020), vanilla Transformer Vaswani et al. (2017), Medformer Wang et al. (2024a), and MedGNN Fan et al. (2025).

Autoformer Autoformer Wu et al. (2021) employs an auto-correlation mechanism to replace self-attention for time series forecasting. Additionally, they use a time series decomposition block to separate the time series into trend-cyclical and seasonal components for improved learning. The raw source code is available at `https://github.com/thuml/Autoformer`.

Crossformer Crossformer Zhang & Yan (2023) designs a single-channel patching approach for token embedding. They utilize two-stage self-attention to leverage both temporal features and channel correlations. A router mechanism is proposed to reduce time and space complexity during the cross-dimension stage. The raw code is available at `https://github.com/Thinklab-SJTU/Crossformer`.

FEDformer FEDformer Zhou et al. (2022) leverages frequency domain information using the Fourier transform. They introduce frequency-enhanced blocks and frequency-enhanced attention, which are computed in the frequency domain. A novel time series decomposition method replaces the layer norm module in the transformer architecture to improve learning. The raw code is available at `https://github.com/MAZiqing/FEDformer`.

Informer Informer Zhou et al. (2021) is the first paper to employ a one-forward procedure instead of an autoregressive method in time series forecasting tasks. They introduce ProbSparse self-attention to reduce complexity and memory usage. The raw code is available at `https://github.com/zhouhaoyi/Informer2020`.

iTransformer iTransformer Liu et al. (2024b) questions the conventional approach of embedding attention tokens in time series forecasting tasks and proposes an inverted approach by embedding the whole series of channels into a token. They also invert the dimension of other transformer modules, such as the layer norm and feed-forward networks. The raw code is available at `https://github.com/thuml/iTransformer`.

MTST MTST Zhang et al. (2024) uses the same token embedding method as Crossformer and PatchTST. It highlights the importance of different patching lengths in forecasting tasks and designs a method that can take different sizes of patch tokens as input simultaneously. The raw code is available at `https://github.com/networkslab/MTST`.

Nonformer Nonformer Liu et al. (2022) analyzes the impact of non-stationarity in time series forecasting tasks and its significant effect on results. They design a de-stationary attention module and incorporate normalization and denormalization steps before and after training to alleviate the over-stationarization problem. The raw code is available at `https://github.com/thuml/Nonstationary_Transformers`.

PatchTST PatchTST Nie et al. (2023) embeds a sequence of single-channel timestamps as a patch token to replace the attention token used in the vanilla transformer. This approach enlarges the receptive field and enhances forecasting ability. The raw code is available at `https://github.com/yuqinie98/PatchTST`.

Reformer Reformer Kitaev et al. (2020) replaces dot-product attention with locality-sensitive hashing. They also use a reversible residual layer instead of standard residuals. The raw code is available at `https://github.com/lucidrains/reformer-pytorch`.

Transformer Transformer Vaswani et al. (2017), commonly known as the vanilla transformer, is introduced in the well-known paper "Attention is All You Need." It can also be applied to time series by embedding each timestamp of all channels as an attention token. The PyTorch version of the code is available at `https://github.com/jadore801120/attention-is-all-you-need-pytorch`.

Medformer Medformer Wang et al. (2024a) uses cross-channel patch embedding to model spatiotemporal dependencies.The raw code is available at `https://github.com/DL4mHealth/Medformer`

MedGNN MedGNN Fan et al. (2025) employs multi-resolution spatiotemporal graph learning to extract dynamic features across multiple time scales. The raw code is available at `https://github.com/aikunyi/MedGNN`.

D.1 EVALUATION METRICS

For all methods, the optimizer used is Adam, with a learning rate of 1e-4. The batch size is set to {32,32,128,128,128} for the datasets APAVA, TDBrain, ADFD, PTB, and PTB-XL, respectively. Training is conducted for 100 epochs, with early stopping triggered after 10 epochs without improvement in the F1 score on the validation set. We save the model with the best F1 score on the validation set and evaluate it on the test set. We employ six evaluation metrics: accuracy, precision (macro-averaged), recall (macro-averaged), F1 score (macro-averaged), AUROC (macro-averaged), and AUPRC (macro-averaged). Both subject-dependent and subject-independent setups are implemented for different datasets. Each experiment is run with 5 random seeds (41-45) and fixed training, validation, and test sets to compute the average results and standard deviations.

To comprehensively and fairly evaluate the performance of each model in the classification task, we select five evaluation metrics: Accuracy, Precision, Recall, F1 score, and AUROC. The definitions and specific calculation formulas for each metric are presented below:

Accuracy measures the proportion of correct predictions out of the total number of predictions. It's calculated as:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}. \tag{41}$$

This metric is useful when the classes are balanced but may be misleading in cases of class imbalance.

Precision focuses on the quality of positive predictions and measures the proportion of correctly predicted positive instances out of all instances predicted as positive. It's especially useful when false positives need to be minimized. The formula is:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}. \tag{42}$$

Recall measures the proportion of actual positive instances that were correctly identified. It's important when false negatives are costly. The formula is:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}. \tag{43}$$

It shows how well the model captures all relevant instances.

The F1 score is the harmonic mean of precision and recall, balancing the two when one is more important than the other. It's particularly useful when dealing with imbalanced datasets, as it accounts for both false positives and false negatives. The formula is:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{44}$$

It gives a single metric that reflects both precision and recall performance.

The Area Under the Receiver Operating Characteristic Curve (AUROC) measures the ability of a model to distinguish between classes, defined as

$$\text{AUROC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}), \tag{45}$$

where

$$\text{TPR} = \frac{TP}{TP + FN}, \qquad \text{FPR} = \frac{FP}{FP + TN}.$$

The Area Under the Precision–Recall Curve (AUPRC) summarizes the trade-off between precision and recall across different thresholds, defined as

$$\text{AUPRC} = \int_0^1 \text{Precision}(\text{Recall})\, d(\text{Recall}), \tag{46}$$

where

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad \text{Recall} = \frac{TP}{TP + FN}.$$

E Detailed Comparison of Data-Centric Spatial Filtering

This appendix provides a detailed comparison between **CIF (Channel-Imposed Fusion)** and representative spatial preprocessing strategies for multichannel time series. We first establish a unified formulation and then analyze each method from four perspectives: operational scope, mathematical form, computational complexity, and modality applicability. We further summarize their parameterization and spatial priors, highlighting the design space occupied by CIF relative to fixed spatial transforms, statistically estimated projections, and fully learnable channel mixing.

E.1 Unified Formulation

Consider a multichannel time series $\mathbf{X} \in \mathbb{R}^{T\times C}$, where T denotes the temporal length and C denotes the number of channels. Let $\mathbf{x}_t \in \mathbb{R}^C$ represent the channel vector at time step t. Most spatial preprocessing methods can be expressed in the general form

$$\widetilde{\mathbf{x}}_t = f(\mathbf{x}_t; \boldsymbol{\theta}), \qquad t = 1, \ldots, T, \tag{47}$$

where $f(\cdot)$ denotes a spatial transformation and $\boldsymbol{\theta}$ contains either fixed structural parameters, statistics estimated from the training data, or learnable parameters.

From an operational perspective, the considered methods can be divided into two groups. First, *full-channel transformations* operate on all channels, including CAR, bipolar montage, Laplacian filtering, CSP, ICA, SSP, SSS, and fully learnable linear mixing. Second, *partial channel transformations* modify only a selected subset of channels while preserving the remaining raw channels; CIF belongs to this category.

The parameters of these transformations also have different origins. For fixed spatial operators such as CAR, bipolar montage, and Laplacian, no training data are required. For training-estimated spatial operators—including supervised class-discriminative projections (CSP), unsupervised statistical source-separation and subspace methods (ICA and SSP), and physics-based or subspace-reconstruction approaches (SSS)—the transformation is estimated from training signals and then fixed during downstream temporal modeling. In contrast, a learnable linear layer and CIF coefficients can be jointly optimized with the temporal classifier.

The nine spatial operators considered in this study are therefore organized into four conceptual classes:

1. fixed per-channel or local operators;
2. training-data-estimated spatial projections;
3. low-parameter structured channel fusion;
4. fully learnable dense channel mixing.

This organization allows the comparison to focus not only on the specific signal-processing techniques, but also on the amount and type of spatial prior introduced before temporal modeling.

E.2 Fixed Spatial Transformations

E.2.1 Common Average Reference (CAR)

The common average reference (CAR) removes the instantaneous mean across all channels. For a channel vector $\mathbf{x}_t$, CAR is defined as

$$\bar{x}_t = \frac{1}{C}\sum_{c=1}^{C} x_{t,c}, \tag{48}$$

and

$$\widetilde{\mathbf{x}}_t = \mathbf{x}_t - \bar{x}_t\mathbf{1}. \tag{49}$$

Equivalently,

$$\widetilde{\mathbf{x}}_t = \left(\mathbf{I} - \frac{1}{C}\mathbf{1}\mathbf{1}^\top\right)\mathbf{x}_t. \tag{50}$$

CAR therefore imposes a global common-mode assumption: information shared approximately equally across channels is treated as a reference component and removed. The transformation requires no fitted parameters and has linear inference complexity $O(C)$ per time step.

CAR is particularly common in EEG and can also be applied to other multichannel signals, including ECG or generic multivariate time series, provided that a common-reference interpretation is meaningful. Importantly, CAR does not model specific functional regions or inter-channel complementary relationships; its spatial prior is the global channel average.

E.2.2 BIPOLAR MONTAGE

A bipolar transformation constructs local channel differences. For an ordered channel sequence, a simple adjacent-channel formulation is

$$\widetilde{x}_{t,c} = x_{t,c+1} - x_{t,c}, \qquad c = 1, \ldots, C-1. \tag{51}$$

When a fixed number of output channels is required, the final output can repeat the last available difference or follow the implementation-specific boundary convention.

Bipolar transformation emphasizes local spatial gradients and suppresses components shared by neighboring channels. Its inference complexity is $O(C)$ per time step and it contains no fitted parameters.

The effectiveness of this operation depends strongly on the channel ordering. For EEG, bipolar montages are normally constructed according to a meaningful electrode arrangement rather than arbitrary channel indices. The same formulation can be applied to ECG or other ordered sensor arrays, but its interpretation is only justified when neighboring channels have a meaningful spatial or functional relationship.

E.2.3 LAPLACIAN FILTERING

Laplacian filtering extends the local-difference idea by comparing each channel with its neighborhood. For a path-graph implementation, the interior channels can be written as

$$\widetilde{x}_{t,c} = 2x_{t,c} - x_{t,c-1} - x_{t,c+1}, \tag{52}$$

with an appropriate boundary convention for the first and last channels.

This operation acts as a spatial high-pass filter: activity that varies locally across channels is emphasized, whereas broad spatially synchronous activity is attenuated. Similar to bipolar filtering, the inference complexity is $O(C)$ and no training data are required.

The formulation above should be understood as an *adjacency-based/path-graph Laplacian implementation*. For EEG, a canonical surface Laplacian normally uses electrode geometry or a neighborhood graph derived from sensor locations. Thus, when channel indices are used directly, the validity of the spatial prior depends on whether the adopted channel ordering preserves meaningful adjacency.

E.3 TRAINING-ESTIMATED SPATIAL PROJECTIONS

Several classical spatial filtering methods estimate a channel-space transformation from training data. Although their objectives differ, they can be represented using a channel-space matrix

$$\widetilde{\mathbf{x}}_t = \mathbf{W}^\top \mathbf{x}_t, \qquad \mathbf{W} \in \mathbb{R}^{C \times C}, \tag{53}$$

or its equivalent row-vector form

$$\widetilde{\mathbf{x}}_t = \mathbf{x}_t^\top \mathbf{W}. \tag{54}$$

The key distinction from fixed transforms is that $\mathbf{W}$ is estimated from training signals rather than predefined analytically.

E.3.1 COMMON SPATIAL PATTERN (CSP)

CSP seeks spatial projections that maximize the variance contrast between two classes. Given class-specific covariance matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, a generalized eigenvalue problem can be formulated as

$$\mathbf{\Sigma}_1 \mathbf{w} = \lambda \left(\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2 \right) \mathbf{w}. \tag{55}$$

The resulting eigenvectors form a spatial projection matrix $\mathbf{W}_{\mathrm{CSP}}$, and the transformed signal is

$$\widetilde{\mathbf{x}}_t = \mathbf{W}_{\mathrm{CSP}}^{\top}\mathbf{x}_t. \tag{56}$$

CSP is inherently supervised because its covariance estimates depend on class-specific trials. In multiclass settings, CSP is commonly extended through one-versus-rest or related pairwise decompositions.

The dominant computational cost arises during covariance estimation and eigendecomposition. For N training trials with average length $\bar{T}$, covariance accumulation is approximately $O(NC^2\bar{T})$, followed by a matrix decomposition whose cost depends on C. Once fitted, applying the resulting dense projection requires $O(C^2)$ operations per time step.

CSP is particularly effective in EEG and brain-computer interface settings, where discriminative spatial covariance structure is often strong. However, its reliance on trial-level statistics and class labels makes it fundamentally different from label-free fixed spatial transformations and from CIF.

E.3.2 Independent Component Analysis (ICA)

ICA assumes that the observed multichannel signal is a mixture of statistically independent latent sources:

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t, \tag{57}$$

where $\mathbf{A}$ denotes the mixing matrix and $\mathbf{s}_t$ contains latent source components. An estimated unmixing matrix provides

$$\mathbf{s}_t = \mathbf{A}^{-1}\mathbf{x}_t. \tag{58}$$

After identifying components considered undesirable, for example artifact-related components, the reconstructed signal can be written schematically as

$$\widetilde{\mathbf{x}}_t = \mathbf{A}_{\bar{\mathcal{E}}}\mathbf{s}_{\bar{\mathcal{E}},t}, \tag{59}$$

where $\mathcal{E}$ denotes the excluded component set.

Thus, after fitting and component selection, ICA can be viewed as an effective channel-space reconstruction operator. Its computational cost is dominated by source separation and the associated iterative optimization, while applying the resulting transformation to a new sample generally requires dense matrix operations of approximately $O(C^2)$ per time step.

ICA is widely used in EEG preprocessing, particularly for separating neural activity from stereotyped artifacts. Although the underlying mathematical formulation is not restricted to EEG, meaningful ICA-based artifact rejection generally depends on adequate training data and a domain-specific criterion for identifying undesirable components.

E.3.3 Signal Space Projection (SSP)

SSP removes signal components associated with an estimated artifact subspace. Given artifact directions $\mathbf{v}_1, \ldots, \mathbf{v}_K$, the corresponding projection can be written as

$$\mathbf{P} = \mathbf{I} - \sum_{k=1}^{K} \frac{\mathbf{v}_k\mathbf{v}_k^{\top}}{\mathbf{v}_k^{\top}\mathbf{v}_k}, \tag{60}$$

when the projection directions are mutually orthogonal. More generally, the artifact subspace can be orthonormalized before constructing the projection.

The filtered signal is

$$\widetilde{\mathbf{x}}_t = \mathbf{P}\mathbf{x}_t. \tag{61}$$

The main idea is to identify a low-dimensional artifact subspace and project the observations onto its orthogonal complement. Once the projection matrix is estimated, applying it costs approximately $O(C^2)$ per time step.

SSP originates primarily from EEG/MEG artifact suppression workflows. The exact construction of the projection vectors is implementation dependent. In particular, an array-only PCA-based projection should be regarded as an SSP-like subspace projection rather than as a complete reproduction of every modality-specific SSP workflow.

E.3.4 SIGNAL SPACE SEPARATION (SSS)

Classical Signal Space Separation (SSS) is a physics-based method developed primarily for MEG, where sensor measurements are decomposed according to spatial field components generated inside and outside the sensor array. Its canonical formulation therefore depends on the geometry and physical characteristics of the MEG sensor system.

For an array-only implementation without explicit sensor geometry, a covariance-subspace analogue can instead be constructed. Let $\mathbf{R}$ denote the estimated channel covariance matrix and let $\mathbf{U}_r$ contain its leading r eigenvectors:

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top, \qquad \mathbf{U}_r = [\mathbf{u}_1, \ldots, \mathbf{u}_r]. \tag{62}$$

A retained-subspace projection can then be defined as

$$\mathbf{P}_{\text{in}} = \mathbf{U}_r\mathbf{U}_r^\top, \tag{63}$$

leading to

$$\widetilde{\mathbf{x}}_t = \mathbf{P}_{\text{in}}\mathbf{x}_t. \tag{64}$$

This operation should be interpreted as an *SSS-inspired covariance-subspace reconstruction* when sensor geometry and Maxwell-equation-based field bases are not explicitly used. Its purpose in the present comparison is to represent a training-estimated low-dimensional spatial subspace rather than to claim a canonical MEG SSS implementation.

The fitted projection is dense and therefore requires approximately $O(C^2)$ operations per time step during inference. The subspace itself is estimated from training signals and does not require class labels.

E.4 FULLY LEARNABLE SPATIAL MIXING

E.4.1 LEARNABLE LINEAR MIXING

A fully learnable spatial transformation removes predefined structural constraints and directly learns a dense channel-mixing matrix:

$$\widetilde{\mathbf{x}}_t = \mathbf{W}\mathbf{x}_t, \qquad \mathbf{W} \in \mathbb{R}^{C\times C}. \tag{65}$$

The matrix contains

$$C^2 \tag{66}$$

learnable parameters and is optimized jointly with the downstream temporal classifier. Its inference complexity is $O(C^2)$ per time step.

The main advantage of this formulation is flexibility: any pair of channels can potentially interact and the model does not require a predefined spatial graph, reference scheme, or functional partition. However, this flexibility also substantially increases the number of degrees of freedom. For medical time-series datasets with limited training samples, learning C^2 unrestricted channel interactions may introduce unnecessary estimation burden and weaken the inductive bias provided by known channel structure.

Consequently, learnable linear mixing represents the opposite end of the design spectrum from fixed spatial preprocessing: it maximizes adaptability but imposes almost no explicit spatial constraint.

E.5 CHANNEL-IMPOSED FUSION (CIF)

CIF is designed as an intermediate spatial operation between fixed handcrafted transformations and fully dense learnable channel mixing. Rather than transforming every channel through a $C \times C$ matrix, CIF first partitions the channel sequence into semantically meaningful regions.

Let the first and last n channels denote two selected channel groups:

$$\mathbf{x}_t^{(1)} = \mathbf{x}_t[1:n], \qquad \mathbf{x}_t^{(2)} = \mathbf{x}_t[C-n+1:C]. \tag{67}$$

CIF constructs a cross-region fused representation as

$$\mathbf{u}_t = a\mathbf{x}_t^{(1)} + b\mathbf{x}_t^{(2)}, \tag{68}$$

where a and b are scalar fusion coefficients.

Depending on the configuration, the fused representation replaces one of the two selected channel blocks. If the first block is replaced,

$$\widetilde{\mathbf{x}}_t = [\mathbf{u}_t; \mathbf{x}_t[n+1:C]]\,. \tag{69}$$

Alternatively, if the last block is replaced,

$$\widetilde{\mathbf{x}}_t = [\mathbf{x}_t[1:C-n]; \mathbf{u}_t]\,. \tag{70}$$

The key difference from dense channel mixing is that CIF modifies only a small portion of the representation while preserving the remaining raw channels. In particular, the middle channels remain untouched, and only the selected cross-region interaction is introduced explicitly.

The spatial operation itself requires only two scalar coefficients, a and b. Thus, when $n \ll C$, its inference complexity is approximately

$$O(n), \tag{71}$$

rather than $O(C^2)$ for dense channel mixing.

The coefficients a and b can either be treated as dataset-level hyperparameters or jointly optimized with the downstream temporal network. Importantly, CIF does not require class labels, trial-level covariance estimation, source separation, or a dense spatial projection matrix.

E.5.1 Cross-Region Complementary Prior

The spatial prior introduced by CIF differs fundamentally from the priors used by classical spatial filters.

CAR assumes that a global common component should be removed. Bipolar and Laplacian filtering emphasize local spatial differences. CSP searches for class-discriminative spatial projections. ICA separates statistically independent latent sources. SSP removes an estimated artifact subspace. SSS-type methods reconstruct signals within an estimated spatial subspace. A fully learnable linear layer does not impose an explicit spatial prior.

In contrast, CIF assumes that two semantically separated channel groups may contain complementary information and explicitly introduces their interaction through a low-dimensional fusion:

$$\mathbf{u}_t = a\mathbf{x}_t^{(1)} + b\mathbf{x}_t^{(2)}. \tag{72}$$

This can be regarded as a *cross-region complementary prior*. The prior does not require the two groups to be adjacent in channel index, nor does it require their relationship to be estimated from training covariance statistics. Instead, the partition is specified according to domain knowledge, such as functional brain regions, lead groups, or semantically related sensor clusters.

Therefore, CIF is not intended to replace the statistical objectives of CSP or ICA, nor the artifact-removal objectives of SSP and SSS. Instead, it occupies a different point in the spatial-design space:

$$\text{fixed local/global rules} \quad \longleftrightarrow \quad \text{structured low-parameter fusion} \quad \longleftrightarrow \quad \text{dense learnable mixing}. \tag{73}$$

E.6 Computational Complexity and Parameterization

Table 13 summarizes the approximate fitting and inference complexity of the considered spatial transformations.

The comparison reveals three distinct parameter regimes. Fixed methods such as CAR, bipolar, and Laplacian require no learned spatial parameters. Training-estimated spatial operators such as CSP, ICA, SSP, and SSS-like projection may produce dense $C \times C$ transformations, but these

Table 13: Computational complexity and parameterization of spatial filtering methods. C denotes the number of channels, n denotes the CIF fusion block size, N denotes the number of training trials, and $\bar{T}$ denotes the average trial length. Complexity values represent the dominant order of computation and may vary with the specific implementation. CSP is the only baseline requiring class labels during spatial fitting.

Method	Fitting	Inference / step	Free parameters	Labels
CAR	None	$O(C)$	0	No
Bipolar	None	$O(C)$	0	No
Laplacian	None	$O(C)$	0	No
CSP	$\sim O(NC^2\bar{T})$	$O(C^2)$	$0^\dagger$	Yes
ICA	Iterative source separation	$O(C^2)$	$0^\dagger$	No
SSP	Covariance/subspace estimation	$O(C^2)$	$0^\dagger$	No
SSS‡	Covariance/subspace estimation	$O(C^2)$	$0^\dagger$	No
Linear	Joint training	$O(C^2)$	C^2	No
CIF	None*	$O(n)$	2	No

† Parameters are estimated during preprocessing and frozen during downstream temporal-network training; they are therefore not counted as jointly learnable model parameters.
* CIF itself requires no statistical fitting. The coefficients a, b may either be fixed as hyperparameters or jointly optimized with the temporal classifier.
‡ Here SSS refers to the array-only covariance-subspace analogue described in Section E.3.4, not the full geometry-dependent Maxwell-equation-based MEG SSS implementation.

transformations are fitted once and frozen; among them, only CSP requires class labels during spatial fitting. At the other extreme, a learnable linear layer introduces C^2 parameters and allows all channel interactions to be optimized jointly with the temporal model.

CIF uses only two scalar coefficients for the spatial fusion itself. When $n \ll C$, this leads to substantially lower inference complexity than dense channel mixing:

$$O(n) \ll O(C^2). \tag{74}$$

More importantly, the reduction is not achieved merely by removing parameters. CIF restricts the spatial interaction pattern to a predefined, interpretable cross-region structure.

E.7 Modality Applicability

The considered methods also differ substantially in their dependence on modality-specific assumptions. Table 14 summarizes their applicability.

The modality comparison highlights an important distinction between *modality-agnostic* and *structure-dependent* spatial processing. CAR and linear mixing can be applied broadly without a specific clinical measurement convention. Bipolar and Laplacian transformations are also broadly implementable, but their interpretation requires meaningful channel adjacency or ordering. CSP, ICA, SSP, and SSS are more strongly associated with EEG/MEG processing because their objectives depend on trial statistics, source separation, artifact subspaces, or sensor-space structure.

CIF is likewise not universally applicable without assumptions. Instead, it requires a meaningful partition of channels into two semantically related groups. Such a partition can be obtained from functional brain regions, physiological lead groups, or other domain-specific sensor organizations. Thus, the key requirement of CIF is not a particular sensing modality but the availability of a reasonable semantic channel partition.

E.8 Conceptual Comparison of Spatial Priors

Table 15 summarizes the conceptual role of each spatial operation.

The methods can consequently be viewed as imposing increasingly flexible spatial interaction patterns. CAR imposes a single global reference, while bipolar and Laplacian filters encode local relationships. CSP and related statistical projections infer spatial directions from training data. A fully learnable linear layer removes these restrictions entirely.

Table 14: Modality scope and structural assumptions of representative spatial preprocessing and channel transformation methods. ✓ indicates direct applicability, △ indicates applicability under additional assumptions, and × indicates that the method is not directly defined for the corresponding setting.

Method	ECG	EEG/MEG	ECG+EEG	Generic TS	Key assumption
CSP	✗	✓	✗	△	Trials and class labels
ICA	△	✓	△	△	Source-separation structure
SSP	△	✓	△	△	Artifact subspace structure
SSS†	✗	✓†	✗	△	Modality-specific spatial structure
CAR	✓	✓	✓	✓	Global common-mode assumption
Bipolar	✓	✓	✓	△	Meaningful channel ordering
Laplacian	✓	✓	✓	△	Local adjacency or geometry
Linear	✓	✓	✓	✓	Unconstrained channel mixing
CIF	✓	✓	✓	✓	Meaningful modality-specific channel grouping

† SSS is primarily defined for MEG and relies on modality-specific spatial structure.

Table 15: Conceptual comparison of representative spatial filtering strategies.

Method	Spatial Scope	Primary Objective	Spatial Prior
CAR	Full channel	Remove common-mode activity	Global channel average
Bipolar	Adjacent channel pairs	Local difference filtering	Channel-order adjacency
Laplacian	Local neighborhood	Spatial high-pass filtering	Local adjacency or graph structure
CSP	Full channel	Maximize class-specific variance contrast	Class-dependent covariance structure
ICA	Full channel	Source separation and artifact rejection	Statistical independence
SSP	Full channel	Remove artifact-related subspaces	Estimated artifact directions
SSS	Full channel	Subspace-based signal reconstruction	Sensor-space signal subspace
Linear	Full channel	Learn arbitrary channel interactions	No explicit structural constraint
CIF	Two selected channel blocks	Cross-region information fusion	Semantically defined functional partition

CIF differs from these approaches by explicitly selecting a *small structured interaction pattern*. Instead of requiring every channel to interact with every other channel, only two selected regions are coupled through two coefficients. This design provides a middle ground between overly restrictive fixed filtering and unconstrained dense channel mixing.

E.9 CIF versus Dense Spatial Matrices

A more direct comparison can be made by considering how each method modifies the channel representation.

The distinction is particularly clear when CIF is compared with a dense matrix transformation. A general spatial matrix $\mathbf{W} \in \mathbb{R}^{C\times C}$ permits

$$\widetilde{x}_{t,i} = \sum_{j=1}^{C} W_{ij} x_{t,j}, \tag{75}$$

Table 16: Comparison between CIF, statistically estimated spatial matrices, fully learnable linear mixing, and fixed spatial transforms.

Property	**CIF**	**CSP/ICA/SSP/SSS**	**Linear**	**Fixed filters**
Channel operation	Partial replacement	Full projection	Full mixing	Full/local transform
Spatial prior	Semantic partition	Training statistics	None explicit	Reference/adjacency
Uses labels	No	CSP: Yes	No	No
Parameter scale	$O(1)$	$O(C^2)$ matrix	$O(C^2)$	$O(1)$
Inference	$O(n)$	$O(C^2)$	$O(C^2)$	$O(C)$
Raw channels retained	Yes	Generally no	No	Method dependent
Cross-region fusion	Explicit	Implicit/statistical	Implicit	Usually no
Jointly learnable	Optional	No	Yes	No

meaning that every output channel can depend on every input channel. Although this provides maximum flexibility, it also introduces C^2 degrees of freedom.

CIF instead restricts the transformation to

$$\mathbf{u}_t = a\mathbf{x}_t^{(1)} + b\mathbf{x}_t^{(2)}, \tag{76}$$

where only two semantically selected regions are explicitly coupled. The number of learnable spatial coefficients therefore remains constant with respect to C.

This distinction is important for medical time-series classification, where the number of training samples can be relatively limited compared with the dimensionality of the temporal observations. CIF does not attempt to learn an unrestricted spatial covariance structure. Instead, it injects a small amount of domain-informed structure before the temporal model and leaves the remaining channels available to the downstream classifier.

E.10 Relationship to Data-Centric Modeling

The comparison above suggests a broader data-centric interpretation of spatial preprocessing. Instead of increasing the complexity of the temporal backbone, spatial filtering modifies the representation presented to the backbone by introducing an inductive bias at the input level.

Let a temporal classifier be represented by

$$\hat{y} = g_\phi(\widetilde{\mathbf{X}}), \tag{77}$$

where g_ϕ denotes the temporal model and $\widetilde{\mathbf{X}}$ is the spatially transformed input. The overall system can therefore be written as

$$\hat{y} = g_\phi\left(f_\theta(\mathbf{X})\right). \tag{78}$$

Different choices of f_θ correspond to different assumptions about what constitutes a useful spatial representation before temporal modeling. A more complex temporal model does not necessarily compensate for an inadequate spatial representation, particularly when cross-channel dependencies are difficult for the temporal backbone to discover efficiently.

CIF follows a data-centric design principle: rather than substantially increasing the capacity of g_ϕ, it modifies the input representation using a small and interpretable spatial operation. The temporal backbone can consequently remain simple, while the spatial preprocessing provides an explicit inductive bias for cross-region information interaction.

This design should not be interpreted as claiming that spatial preprocessing is universally superior to architectural modeling. Rather, it demonstrates that part of the performance gap between medical time-series classifiers can arise from how channel information is organized before temporal modeling.

E.11 Summary

In summary, existing spatial preprocessing methods encode substantially different assumptions. CAR removes a global common component, bipolar and Laplacian filters exploit local channel relationships,

CSP uses class-dependent covariance structure, ICA performs statistical source separation, and SSP and SSS-like methods construct low-dimensional spatial subspaces. Fully learnable linear mixing provides maximum flexibility but requires C^2 parameters and dense channel interaction.

CIF occupies an intermediate position in this design space. It does not attempt to estimate a full channel covariance structure or learn an unconstrained $C \times C$ transformation. Instead, it introduces a small number of coefficients to fuse two semantically selected channel groups while preserving most of the original representation. This results in an explicit cross-region complementary prior with $O(n)$ inference complexity and only $O(1)$ spatial parameters.

Therefore, CIF should be viewed not as a replacement for classical signal-processing objectives such as artifact removal or source separation, but as a lightweight data-centric mechanism for injecting structured cross-channel interactions before temporal modeling. The resulting design occupies a distinct point between handcrafted spatial filtering and fully learnable channel mixing, providing a simple and interpretable alternative for multichannel medical time-series classification.

F Model-centric Methods for Time Series Classification.

Time series classification has increasingly adopted a model-centric paradigm, where architectural innovations aim to learn increasingly rich temporal, cross-channel, and multi-scale representations from raw or minimally processed signals. Existing approaches can be broadly categorized into Transformer-based and convolution-based architectures, with different designs emphasizing global dependency modeling, local temporal patterns, cross-variable interactions, or multi-scale feature extraction.

F.1 Transformer-based Time Series Methods.

Transformer-based models have become an important family of model-centric approaches for time series analysis by leveraging self-attention to capture long-range dependencies. The vanilla Transformer Vaswani et al. (2017) introduced multi-head self-attention and positional encoding, enabling global interactions among temporal observations. Building upon this framework, a variety of architectures have been developed to improve computational efficiency or introduce inductive biases for temporal data. Informer Zhou et al. (2021) employs ProbSparse attention to focus on dominant queries and reduce the computational cost of long-sequence modeling. Reformer Kitaev et al. (2020) incorporates locality-sensitive hashing to approximate attention with reduced complexity, making Transformer-based modeling more scalable to long sequences. Autoformer Wu et al. (2021) introduces an Auto-Correlation mechanism that aggregates periodic sub-sequences to explicitly capture inherent periodicity. FEDformer Zhou et al. (2022) performs seasonal–trend decomposition in the frequency domain and applies attention over a compact set of Fourier coefficients, providing an efficient way to model frequency-specific temporal dependencies. Crossformer Zhang & Yan (2023) adopts dimension-segment-wise embedding and two-stage attention to capture dependencies across both temporal segments and variables, thereby modeling multivariate correlations. iTransformer Liu et al. (2024b) reverses the conventional tokenization strategy by treating each variable as a token and applying attention across variables, allowing the model to directly learn inter-variable dependencies. PatchTST Nie et al. (2023) segments long time series into local patches and treats them as tokens, enabling the model to capture temporal patterns at different receptive-field scales while reducing the effective sequence length.

For medical time series, Transformer-based architectures have been further adapted to incorporate the heterogeneous characteristics and domain-specific structures of biomedical signals. EEG-Conformer Song et al. (2022) combines convolutional feature extraction with self-attention to jointly model local temporal patterns and global dependencies in EEG signals. BIOT Yang et al. (2023) extends Transformer-based representation learning to heterogeneous biomedical signals by converting multi-modal physiological measurements into a unified token representation. Medformer Wang et al. (2024a) introduces multi-granularity patch embeddings and cross-channel attention to accommodate the diverse temporal resolutions and heterogeneous channel characteristics of medical time series. ADformer Wang et al. (2024b) further develops Transformer-based modeling for Alzheimer's disease analysis by incorporating temporal and channel-wise information to capture discriminative patterns in neurological signals. MedGNN Fan et al. (2025) explicitly incorporates graph-based modeling into

medical time series analysis, using medical variables as nodes to represent structured relationships among physiological measurements while combining temporal and frequency-domain information. Collectively, these approaches demonstrate a continued trend toward increasing architectural complexity to capture richer temporal, cross-channel, and domain-specific dependencies.

F.2 Convolution-based Time Series Methods.

Convolutional architectures provide an alternative model-centric paradigm that emphasizes local temporal structure and hierarchical feature extraction. EEGNet Lawhern et al. (2018) employs depthwise and separable convolutions to construct a compact architecture for EEG decoding, efficiently capturing temporal and spatial patterns while maintaining a small number of parameters. Temporal Convolutional Networks (TCNs) Bai et al. (2018) use causal and dilated convolutions to enlarge the temporal receptive field without substantially increasing computational cost, enabling effective modeling of long-range temporal dependencies. TCN-based architectures have subsequently been adapted to medical time series classification, where their hierarchical temporal receptive fields provide an efficient alternative to recurrent and attention-based models Lin et al. (2019).

Residual convolutional architectures have also been widely applied to one-dimensional time series. ResNet He et al. (2016) introduces residual connections that facilitate the optimization of deep networks, while its one-dimensional adaptations extend this principle to temporal signals Wang et al. (2017). XResNet He et al. (2019) further improves the training behavior and empirical performance of ResNet through a collection of architectural refinements. InceptionTime Ismail Fawaz et al. (2020) adapts the multi-branch Inception architecture to the temporal domain by applying convolutional kernels with different receptive-field sizes in parallel, allowing the model to capture temporal patterns at multiple scales. These convolutional models provide strong and computationally efficient baselines for time series classification, with their inductive biases favoring local and hierarchical temporal representations.

Despite the architectural differences among these approaches, they share a common underlying philosophy: improving classification performance primarily through increasingly expressive model architectures. Transformer-based methods introduce more sophisticated mechanisms for global dependency modeling, variable interaction, frequency-domain analysis, and multi-scale tokenization, while convolutional approaches expand receptive fields, network depth, and multi-scale feature extraction. Such model-centric designs have substantially advanced time series classification. However, they generally assume that the downstream model can sufficiently resolve the noise, redundancy, inter-channel interference, and heterogeneous structures contained in the input signals. For medical time series, where signals are often noisy, highly correlated, and strongly dependent on acquisition modality and channel configuration, this assumption may not always hold. This motivates an alternative perspective that places greater emphasis on improving the quality and discriminability of the input representation before applying increasingly complex classifiers.

G HMTCN STRUCTURE DESIGN AND THEORETICAL ANALYSIS

Modeling short-term and long-term dependencies in time series data is challenging. Traditional CNNs excel at capturing local features but have limited receptive fields, hindering long-range dependency learning. Transformer-based methods effectively model long-term dependencies, but their complex design lacks interpretability, which is a key issue in medical time-series classification.To address these limitations, TCNs use causal convolutions for explicit temporal modeling and dilated convolutions to expand the receptive field, overcoming the constraints of traditional CNNs. Building on the advantages of TCN, we propose the HMTCN, which combines the benefits of dilated convolutions, bidirectional causal convolution, and residual connections. This approach allows for better capture of temporal dependencies while preserving causality.

G.1 DILATED CONVOLUTION

Dilated convolution expands the receptive field without significantly increasing computational cost Yu & Koltun (2015). For a 1D input sequence $x = [x_1, x_2, \ldots, x_T]$, its output is defined as $y(t) = \sum_{i=0}^{k-1} x(t + i \cdot d) \cdot w(i)$, where t is the current time step, k is the kernel size, d is the dilation factor, and $w(i)$ is the weight at the i-th position in the kernel. Increasing d effectively enlarges the receptive field, enabling the network to capture longer-term temporal dependencies. When stacking multiple dilated convolutional layers, the receptive field grows progressively. For the l-th layer, the receptive field r_l can be expressed as $r_l = k + (k-1)\sum_{j=1}^{l-1} d_j$, where d_j is the dilation factor of the j-th layer. By gradually increasing d_j, the network captures temporal dependencies across both global and local scales, offering an effective way to model long-term dependencies in time series.

G.2 BIDIRECTIONAL CAUSAL CONVOLUTION STRUCTURE

In addition to dilated convolutions, HMTCN introduces a *bidirectional causal convolution structure*, inspired by prior bidirectional temporal modeling approaches Hanson et al. (2018); Hu et al. (2024); Yin et al. (2025). Unlike traditional TCNs that use only forward causal convolutions, our architecture applies causal convolutions in both forward and backward directions, enabling the model to capture dependencies from both past and future contexts while strictly preserving causality. The *forward causal convolution* processes the input sequence $x(t)$ in chronological order, producing output $y_{\text{forward}}(t) = \sum_{i=0}^{k-1} x(t - i \cdot d) \cdot w_{\text{forward}}(i)$, which depends only on current and past inputs. For the *backward causal convolution*, we first reverse the input sequence as $x_{\text{flip}}(t) = x(T - t)$, and then apply a causal convolution over this flipped sequence. This ensures that the model captures future-directed dependencies without introducing information leakage. The output is given by $y_{\text{backward}}(t) = \sum_{i=0}^{k-1} x(T - (t - i \cdot d)) \cdot w_{\text{backward}}(i)$. These two operations are implemented using separate convolutional layers (convforward and convbackward), and their outputs are summed to form the final bidirectional result: $y_{\text{bi}}(t) = y_{\text{forward}}(t) + \text{flip}(y_{\text{backward}}(t))$. By integrating both directions under strict causality constraints, HMTCN achieves superior temporal dependency modeling compared to unidirectional causal approaches.

G.3 MULTI-SCALE FEATURE LEARNING AND RESIDUAL CONNECTIONS

To further improve the model's capacity to capture dependencies at different temporal scales, HMTCN incorporates *multi-scale feature learning* and *residual connections*. Multi-scale Feature Learning: In HMTCN, we employ a hierarchy of dilation factors that decrease layer by layer to capture temporal dependencies at multiple scales. Lower layers use larger dilation factors to expand the receptive field, aggregating long-range information and smoothing short-term noise in highly redundant medical time series; higher layers use smaller dilation factors to focus on local dependencies and capture fine-grained features. This coarse-to-fine, global-to-local design enables the network to extract broad patterns in its initial layers and refine precise details in its later layers, thereby enhancing adaptability across a wide range of time series tasks. Residual connections: Residual connections He et al. (2016) are introduced between the dilated convolutional layers to facilitate the efficient flow of information through the network. The residual connection is defined as $y = F(x) + x$, where $F(x)$ is the convolutional output, and x is the input. This design alleviates the vanishing gradient problem and improves the overall stability of the network during training.

H PSEUDOCODE OF CIF METHOD AND KEY COMPONENTS OF HMTCN

Algorithm 1 CIF Module (Channel-Imposed Fusion with Physiological Ordering)

Require: Input $x_{enc} \in \mathbb{R}^{B \times T \times C}$, hyperparameters t, n, a, b ▷ n must satisfy $1 \leq n \leq C$ (i.e., cannot exceed total number of channels)

Step 1: Physiological Channel Reordering (optional) ▷ Only applied if channels have physiology-consistent structure (e.g., ECG, EEG). For general/unstructured data, skip this step.
if PhysiologyStructured(x_{enc}) **then**
    $x_{enc} \leftarrow$ PhysiologyReorder(x_{enc})
end if

Step 2: Split into front and back segments
if $n > C$ **then**
    Error: n cannot exceed total number of channels C
end if
$front \leftarrow x_{enc}[:, :, : n]$ ▷ Front segment: first n channels
$back \leftarrow x_{enc}[:, :, -n :]$ ▷ Back segment: last n channels

Step 3: Fusion
$x_{new} \leftarrow$ Clone(x_{enc})
$added \leftarrow front \cdot a + back \cdot b$
if $t > 0$ **then**
    $x_{new}[:, :, : n] \leftarrow added$
else
    $x_{new}[:, :, -n :] \leftarrow added$
end if
return x_{new}

Algorithm 2 Model with CCA-based Feature Interaction

Require: Encoded input $x_{\text{enc}} \in \mathbb{R}^{B \times L \times D}$, task configs
Initialize encoder and task-specific heads
procedure FORWARD(x_{enc})
    $n \leftarrow$ number of CCA components
    $X_{\text{front}} \leftarrow$ first n features of x_{enc}
    $X_{\text{back}} \leftarrow$ last n features of x_{enc}
    $x_{\text{enc_new}} \leftarrow x_{\text{enc}}$ (clone)
    $X_{\text{front_flat}} \leftarrow$ reshape X_{front} to $[B \cdot L, n]$
    $X_{\text{back_flat}} \leftarrow$ reshape X_{back} to $[B \cdot L, n]$
    Without gradient:
        $\text{CCA}_1 \leftarrow \texttt{TorchCCA}(n)$
        $(X_c^1, Y_c^1) \leftarrow \text{CCA}_1$.fit_transform($X_{\text{front_flat}}, X_{\text{back_flat}}$)
        corr_scores $\leftarrow$ correlation coefficients for each canonical component
        $\text{CCA}_2 \leftarrow \texttt{TorchCCA}(n)$
        $(X_c^2, Y_c^2) \leftarrow \text{CCA}_2$.fit_transform($X_{\text{back_flat}}, X_{\text{front_flat}}$)
        corr_scores2 $\leftarrow$ correlation coefficients for each canonical component
    $x_{\text{enc_new}}[:, :, : n] \leftarrow X_{\text{front}} *$ corr_scores $+ X_{\text{back}}$
    $x_{\text{enc_new}}[:, :, -n :] \leftarrow X_{\text{front}} + X_{\text{back}} *$ corr_scores2
    return $x_{\text{enc_new}}$
end procedure

Algorithm 3 TorchCCA: PyTorch-style Canonical Correlation Analysis (CCA)

Require: Input datasets $X \in \mathbb{R}^{N \times D_1}, Y \in \mathbb{R}^{N \times D_2}$
Require: Number of components k, scaling flag `scale`, regularization `reg`
Initialize `TorchCCA`$(k, \texttt{scale}, \texttt{reg})$
procedure FIT(X, Y)
    $X, Y \leftarrow$ CENTER_AND_SCALE(X, Y)
    $N \leftarrow$ number of rows in X
    $C_{xx} \leftarrow \frac{X^\top X}{N-1} + \texttt{reg} \cdot I$
    $C_{yy} \leftarrow \frac{Y^\top Y}{N-1} + \texttt{reg} \cdot I$
    $C_{xy} \leftarrow \frac{X^\top Y}{N-1}$
    $C_{xx}^{-1/2} \leftarrow (\text{Cholesky}(C_{xx})^{-1})^\top$
    $C_{yy}^{-1/2} \leftarrow (\text{Cholesky}(C_{yy})^{-1})^\top$
    $T \leftarrow C_{xx}^{-1/2} C_{xy} C_{yy}^{-1/2\top}$
    $(U, S, V^\top) \leftarrow \text{SVD}(T)$
    `x_weights` $\leftarrow$ first k columns of U
    `y_weights` $\leftarrow$ first k columns of V
    `corr_scores` $\leftarrow$ first k singular values S
end procedure
procedure TRANSFORM(X, Y)
    $X, Y \leftarrow$ CENTER_AND_SCALE(X, Y)
    $X_c \leftarrow X \cdot$ `x_weights`
    $Y_c \leftarrow Y \cdot$ `y_weights`
    return X_c, Y_c
end procedure
procedure FIT_TRANSFORM(X, Y)
    FIT(X, Y)
    return TRANSFORM(X, Y)
end procedure
procedure CENTER_AND_SCALE(X, Y)
    $X \leftarrow X - \text{mean}(X)$
    $Y \leftarrow Y - \text{mean}(Y)$
    if scale **then**
        $X \leftarrow X / \text{std}(X)$
        $Y \leftarrow Y / \text{std}(Y)$
    end if
    return X, Y
end procedure

Algorithm 4 BidirectionalCausalConv

Require: Input $x \in \mathbb{R}^{B \times C \times T}$, kernel size k, dilations d_f, d_b
Compute $p_f \leftarrow (k-1) \cdot d_f$
Compute $p_b \leftarrow (k-1) \cdot d_b$
$x_f \leftarrow \text{PadLeft}(x, p_f)$
$x_b \leftarrow \text{PadLeft}(\text{Flip}(x), p_b)$
$y_f \leftarrow \text{Conv1D}(x_f, \text{dilation} = d_f)$
$y_b \leftarrow \text{Flip}(\text{Conv1D}(x_b, \text{dilation} = d_b))$
return $y_f + y_b$

Algorithm 5 BidirectionalDilatedConvBlock

Require: Input x, channels C_{in}, C_{out}, kernel size k, dilation d
1: **if** $C_{in} \neq C_{out}$ or final layer **then**
2: $\quad res \leftarrow \text{Conv1D}(x, \text{kernel} = 1)$
3: **else**
4: $\quad res \leftarrow x$
5: **end if**
6: $x \leftarrow \text{GELU}(x)$
7: $x \leftarrow \text{BidirectionalCausalConv}(x, k, d, d)$
8: $x \leftarrow \text{GELU}(x)$
9: $x \leftarrow \text{BidirectionalCausalConv}(x, k, d, d)$
10: **return** $x + res$

I Ablation Experiments of the HMTCN Structure

Table 17: The ablation experiments of the HMTCN structure, where "Forward" indicates using only the forward part, and "Backward" indicates using only the backward part.

Datasets	Models	CIF	Forward	Backward	Accuracy ↑	Precision ↑	Recall ↑	F1 score ↑	AUROC ↑	AUPRC ↑
APAVA	**HMTCN**		✓		82.31±2.34	83.29±2.50	80.39±2.65	81.02±2.63	91.50±1.80	91.66±1.82
(2-Classes)	**HMTCN**			✓	79.45±3.51	80.69±3.04	77.14±4.47	77.58±4.75	87.95±3.82	88.41±3.74
	HMTCN		✓	✓	82.49±1.40	82.38±1.79	81.20±1.32	81.60±1.39	91.10±1.63	91.30±1.71
APAVA	**HMTCN**	✓	✓		80.43±5.60	80.46±5.23	79.56±5.98	79.50±5.97	89.23±4.44	89.62±4.25
(2-Classes)	**HMTCN**	✓		✓	79.39±3.44	79.49±3.76	78.09±2.94	78.35±3.33	87.62±3.09	88.11±2.91
	HMTCN	✓	✓	✓	85.16±1.55	84.76±1.62	85.33±1.27	84.82±1.49	94.06±1.07	94.21±0.99
ADFTD	**HMTCN**		✓		53.32±1.35	52.01±1.54	51.46±2.19	51.21±1.99	70.78±1.78	53.16±2.20
(3-Classes)	**HMTCN**			✓	52.80±1.18	50.16±0.77	49.23±1.22	49.24±1.02	68.65±0.71	49.95±0.97
	HMTCN		✓	✓	52.05±2.22	50.45±3.00	50.40±2.55	49.48±2.70	69.43±2.84	50.99±3.15
ADFTD	**HMTCN**	✓	✓		56.06±0.47	53.21±1.03	53.54±1.36	52.82±1.33	72.93±0.88	55.71±1.03
(3-Classes)	**HMTCN**	✓		✓	56.54±1.33	54.28±0.96	54.63±1.06	53.91±1.11	73.46±1.17	56.12±1.61
	HMTCN	✓	✓	✓	58.56±0.93	55.65±0.81	55.86±0.79	55.42±0.82	76.07±0.59	59.75±0.67
TDBrain	**HMTCN**		✓		87.23±2.87	87.75±2.48	87.23±2.87	87.17±2.93	95.55±1.69	95.73±1.60
(2-Classes)	**HMTCN**			✓	86.92±3.46	87.41±3.17	86.92±3.46	86.86±3.51	95.28±1.78	95.42±1.70
	HMTCN		✓	✓	84.90±2.60	86.02±2.00	84.90±2.60	84.76±2.74	93.94±1.92	94.20±1.85
TDBrain	**HMTCN**	✓	✓		93.29±1.73	93.34±1.73	93.29±1.73	93.29±1.73	98.50±0.63	98.56±0.60
(2-Classes)	**HMTCN**	✓		✓	93.69±1.52	93.83±1.42	93.69±1.52	93.68±1.53	98.56±0.67	98.59±0.64
	HMTCN	✓	✓	✓	93.13±1.41	93.33±1.37	93.13±1.41	93.12±1.42	98.62±0.66	98.68±0.63
PTB	**HMTCN**		✓		82.56±1.74	86.16±1.51	74.91±2.88	77.24±2.92	95.69±0.64	94.56±0.76
(2-Classes)	**HMTCN**			✓	81.07±4.24	85.36±2.71	72.50±6.59	74.33±6.71	92.83±2.38	91.28±2.79
	HMTCN		✓	✓	81.87±1.87	86.50±1.24	73.49±2.90	75.84±3.20	94.20±0.29	93.04±0.45
PTB	**HMTCN**	✓	✓		87.33±1.41	90.26±1.24	81.64±2.04	84.19±1.97	96.21±1.30	95.67±1.52
(2-Classes)	**HMTCN**	✓		✓	84.35±2.28	87.42±2.07	77.54±3.30	79.98±3.41	91.25±1.92	90.42±2.28
	HMTCN	✓	✓	✓	88.29±1.45	90.66±1.48	83.21±2.02	85.59±1.96	94.28±0.93	93.78±1.11
FLAAP	**HMTCN**		✓		70.81±2.31	72.58±1.33	69.81±2.79	70.07±2.24	95.89±0.28	76.90±1.21
(10-Classes)	**HMTCN**			✓	70.29±2.04	72.77±2.09	68.86±2.39	69.56±1.98	95.61±0.26	76.56±1.56
	HMTCN		✓	✓	76.08±0.81	76.05±0.83	75.95±0.84	75.54±0.94	96.49±0.10	81.19±0.65
FLAAP	**HMTCN**	✓	✓		72.30±1.50	72.98±1.61	71.65±1.35	71.54±1.50	95.92±0.55	77.87±2.27
(10-Classes)	**HMTCN**	✓		✓	72.81±1.04	74.05±0.80	71.86±1.25	72.12±1.17	96.20±0.21	79.22±1.25
	HMTCN	✓	✓	✓	76.82±1.32	77.38±0.85	76.52±1.24	76.39±1.18	96.48±0.06	81.77±0.81
UCI-HAR	**HMTCN**		✓		91.94±0.98	92.36±0.90	92.02±0.96	91.98±0.93	99.30±0.08	97.31±0.47
(6-Classes)	**HMTCN**			✓	93.03±0.62	93.28±0.63	93.12±0.60	93.05±0.62	99.36±0.19	97.72±0.46
	HMTCN		✓	✓	93.72±0.73	94.02±0.72	93.75±0.70	93.69±0.76	99.60±0.09	98.31±0.40
UCI-HAR	**HMTCN**	✓	✓		92.18±0.45	92.42±0.47	92.21±0.44	92.14±0.44	99.17±0.11	97.04±0.15
(6-Classes)	**HMTCN**	✓		✓	92.62±0.90	92.88±0.86	92.68±0.88	92.63±0.89	99.25±0.18	97.15±0.57
	HMTCN	✓	✓	✓	93.78±0.32	94.08±0.26	93.79±0.32	93.74±0.34	99.34±0.19	97.60±0.46

As described in Section 2 of the Methods, HMTCN is constructed by adding a backward branch to the vanilla TCN architecture; therefore, using only the Forward part is essentially equivalent to the vanilla TCN structure.

From the table 17, it can be observed that when both the Forward and Backward parts of the HMTCN structure are used simultaneously, the performance drops significantly compared to using only one of them individually. We speculate that this is mainly due to the presence of substantial noise within

medical time-series data. When both parts of the structure are applied at the same time, it is akin to capturing noise from two different directions simultaneously. Instead of enhancing the representation, this leads to noise accumulation, which ultimately results in degraded performance.

However, after processing the data with CIF to improve the signal-to-noise ratio, the combination of the Forward and Backward parts of the HMTCN structure eventually outperforms the use of either part alone. This result strongly demonstrates the feature-capturing capability of the HMTCN when both directions are utilized together. It indicates that once noise interference is effectively reduced, the bidirectional structure of HMTCN can better leverage its strengths, thereby improving overall performance.

Further observations show that using only the Forward part of HMTCN outperforms the Backward part. This is closely related to the inherent unidirectionality of medical time-series signals such as EEG and ECG, where information typically propagates forward in time (e.g., neural signal transmission in EEG or atrial-to-ventricular activation in ECG). Such characteristics enable the Forward structure to capture key features and temporal evolution more effectively, yielding better performance. This finding not only deepens the understanding of medical signal processing but also provides insights for optimizing HMTCN in related applications.

To evaluate the performance of our method on general time series, we follow the design of Medformer Wang et al. (2024a) and test it on two human activity recognition (HAR) datasets: FLAAP(13,123 samples, 10 classes) Kumar & Suresh (2022) and UCI-HAR(10,299 samples, 6 classes) Anguita et al. (2013).

Additionally, on the non-medical datasets FLAAP and UCI-HAR, we observed that integrating the bidirectional structure significantly improves performance. This indicates that in high-SNR scenarios, bidirectional modeling can more effectively capture both forward and backward feature information, enhancing overall model performance. In contrast, CIF provides relatively limited gains on these high-SNR datasets. This observation further highlights the design advantage of CIF: it is specifically tailored for low-SNR medical time series, explicitly fusing inter-channel physiological information to enhance signal quality and discriminative power, while its marginal benefit is smaller for low-noise non-medical data. Overall, these findings not only reveal the differential adaptability of model architectures under varying data characteristics but also underscore the unique value of CIF in complex medical scenarios.

J FURTHER EXPLORATION OF PHYSIOLOGICAL STRUCTURES

The parameters (a, b, t, n) in CIF are explicit hyperparameters that can be directly set and adjusted based on experience. For example, Figure. 8 shows the corresponding locations of EEG channels on the human brain, we have adjusted the AFAVA dataset, which comprises 16 channels: `C3`, `C4`, `F3`, `F4`, `F7`, `F8`, `Fp1`, `Fp2`, `O1`, `O2`, `P3`, `P4`, `T3`, `T4`, `T5`, and `T6`. For the first six channels, we performed pairwise fusion as follows:

$$C3_{\text{new}} = a \cdot C3 + b \cdot C4,$$
$$F3_{\text{new}} = a \cdot F3 + b \cdot F4,$$
$$F7_{\text{new}} = a \cdot F7 + b \cdot F8.$$

Here, the `C3`, `F3`, and `F7` channels correspond to `C4`, `F4`, and `F8`, respectively, exhibiting left–right physiological symmetry, and they also belong to the same functional region. We refer to this type of fusion as *Left–Right Physiological Symmetry Fusion* (LR-PSF). In contrast, the channel fusion strategy introduced earlier in this paper, controlled by the hyperparameter n, reflects fusion across different functional regions, and is therefore termed *Functional Region Fusion* (FRF).

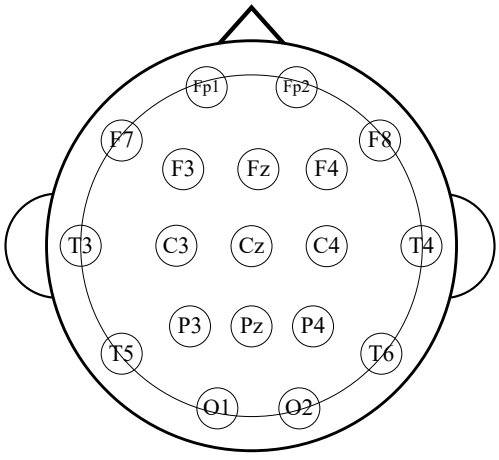


Figure 8: Physiological Placement Diagram of EEG Channels.

Table 18: **Results of Subject-Independent Setup. APAVA Dataset**

Datasets	Models	Accuracy ↑	Precision ↑	Recall ↑	F1 score ↑	AUROC ↑	AUPRC ↑
APAVA (2-Classes)	HMTCN	82.49±1.40	82.38±1.79	81.20±1.32	81.60±1.39	91.10±1.63	91.30±1.71
	HMTCN + CIF (-Fp1 + Fp2)	82.70±1.90	83.34±1.96	80.73±2.14	81.47±2.13	91.24±2.25	91.62±2.08
	HMTCN + CIF (FRF)	86.30±1.05	86.16±1.09	85.47±1.12	85.71±1.09	94.26±0.54	94.42±0.49
	HMTCN + CIF (LR-PSF)	86.23±2.09	85.82±2.14	**86.04±2.06**	85.83±2.12	94.59±1.08	94.64±1.08
	HMTCN + CIF (CCA)	**87.91**±1.60	**89.81**±0.93	85.85±2.12	**86.93**±1.91	**96.97**±0.52	**96.91**±0.49

Building on these two types of physiologically motivated fusion, we further introduce a data-driven fusion mechanism based on Canonical Correlation Analysis (CCA) Hotelling (1992) (For detailed pseudocode, see Code 2 3.). Specifically, for each input sample we first select the first n channels and the last n channels, denoted as

$$\mathbf{X}^{\text{front}} \in \mathbb{R}^{B\times L\times n}, \qquad \mathbf{X}^{\text{back}} \in \mathbb{R}^{B\times L\times n},$$

where B is the batch size, L is the sequence length, and n is the number of paired channels. We then reshape them into two matrices

$$\mathbf{F} \in \mathbb{R}^{(BL)\times n}, \qquad \mathbf{B} \in \mathbb{R}^{(BL)\times n},$$

by stacking all temporal positions and samples along the first dimension.

We apply a CCA module to these two matrices and obtain their canonical projections

$$\mathbf{F}_{\text{c}}, \mathbf{B}_{\text{c}} \in \mathbb{R}^{(BL)\times n}.$$

For each canonical component $i \in \{1, \ldots, n\}$, we compute the Pearson correlation coefficient between the corresponding canonical variables $\mathbf{F}_{\text{c}}^{(i)}$ and $\mathbf{B}_{\text{c}}^{(i)}$, yielding a correlation score vector

$$\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)^\top \in \mathbb{R}^n.$$

These scores quantify the statistical dependence between the paired channels and are used as adaptive, channel-wise fusion weights.

Concretely, we broadcast $\boldsymbol{\rho}$ back to the original tensor shape and construct two fused channel groups:

$$\widetilde{\mathbf{X}}^{\text{front}} = \mathbf{X}^{\text{front}} \odot \boldsymbol{\rho} + \mathbf{X}^{\text{back}}, \qquad \widetilde{\mathbf{X}}^{\text{back}} = \mathbf{X}^{\text{front}} + \mathbf{X}^{\text{back}} \odot \boldsymbol{\rho}',$$

where $\odot$ denotes element-wise (channel-wise) multiplication and $\boldsymbol{\rho}'$ is another correlation score vector obtained by swapping the roles of $\mathbf{F}$ and $\mathbf{B}$ in CCA. The fused tensors $\widetilde{\mathbf{X}}^{\text{front}}$ and $\widetilde{\mathbf{X}}^{\text{back}}$ are then written back to the first and last n channels of the encoder input, respectively.

In this way, the proposed CCA-based fusion adaptively strengthens or attenuates each paired channel according to its learned cross-channel correlation, providing a controllable and data-driven mechanism that complements the physiologically defined LR-PSF and FRF fusion strategies.

We consider that a full theoretical and experimental treatment of CCA would require a more extensive discussion. Therefore, here we present it solely as an exploratory method and do not include its results in the main text.

We also examined the relationship between the `Fp1` and `Fp2` channels, as introduced in the Introduction. For this purpose, we conducted experiments using the following fusion approach:

$$\texttt{Fp1}_{\text{new}} = -1 \cdot \texttt{Fp1} + 1 \cdot \texttt{Fp2},$$

which we refer to as CIF ($-\texttt{Fp1} + \texttt{Fp2}$).

The results in the table 18 reveal that explicit fusion leveraging the prior knowledge of channels can more effectively integrate channel features, thereby yielding more accurate classification outcomes. Many previous methods, especially various general time series models, are unable to incorporate such medical prior knowledge in a "controllable" manner.

K Results on General Time Series

We compared our method with various approaches on the general time series UEA dataset. The results of these methods were provided by TimeMixer++ Wang et al. (2025a). Experimental results show that CIF can enhance the performance of HMTCN on general time series classification tasks. Moreover, the combination of HMTCN with CIF achieves SOTA performance.

Table 19: Full results for the classification task. $*$. in the Transformers indicates the name of $*$former. We report the classification accuracy (%) as the result.

Datasets / Models	RNN			TCN	Transformers										MLP			CNN			
	LSTM (1997)	LSTNet (2018)	LSSL (2022)	TCN (2019)	Trans. (2017)	Re. (2020)	In. (2021)	Pyra. (2021)	Auto. (2021)	Station. (2022)	FED. (2022)	ETS. (2022)	Flow. (2022b)	iTrans. (2024b)	DLinear (2023)	LightTS. (2022)	TiDE (2023)	TimesNet (2022a)	Time Mixer++ (2025a)	**HMTCN (Ours)**	**HMTCN +CIF (Ours)**
EthanolConcentration	32.3	39.9	31.1	28.9	32.7	31.9	31.6	30.8	31.6	32.7	28.1	31.2	33.8	28.1	32.6	29.7	27.1	35.7	39.9	31.9	32.3
FaceDetection	57.7	65.7	66.7	52.8	67.3	68.6	67.0	65.7	68.4	68.0	66.0	66.3	67.6	66.3	68.0	67.5	65.3	68.6	71.8	66.8	67.2
Handwriting	15.2	25.8	24.6	53.3	32.0	27.4	32.8	29.4	36.7	31.6	28.0	32.5	33.8	24.2	27.0	26.1	23.2	32.1	26.5	49.5	51.2
Heartbeat	72.2	77.1	72.7	75.6	76.1	77.1	80.5	75.6	74.6	73.7	73.7	71.2	77.6	75.6	75.1	75.1	74.6	78.0	79.1	74.6	77.5
JapaneseVowels	79.7	98.1	98.4	98.9	98.7	97.8	98.9	98.4	96.2	99.2	98.4	95.9	98.9	96.6	96.2	96.2	95.6	98.4	97.9	97.8	98.3
PEMS-SF	39.9	86.7	86.1	68.8	82.1	82.7	81.5	83.2	82.7	87.3	80.9	86.0	83.8	87.9	75.1	88.4	86.9	89.6	91.0	82.6	86.1
SelfRegulationSCP1	68.9	84.0	90.8	84.6	92.2	90.4	90.1	88.1	84.0	89.4	88.7	89.6	92.5	90.2	87.3	89.8	89.2	91.8	93.1	89.7	91.1
SelfRegulationSCP2	46.6	52.8	52.2	55.6	53.9	56.7	53.3	53.3	50.6	57.2	54.4	55.0	56.1	54.4	50.5	51.1	53.4	57.2	65.6	61.6	62.2
SpokenArabicDigits	31.9	100.0	100.0	95.6	98.4	97.0	100.0	99.6	100.0	100.0	100.0	100.0	98.8	96.0	81.4	100.0	95.0	99.0	99.8	99.5	99.6
UWaveGestureLibrary	41.2	87.8	85.9	88.4	85.6	85.6	85.6	83.4	85.9	87.5	85.3	85.0	86.6	85.9	82.1	80.3	84.9	85.3	88.2	92.1	92.8
Average Accuracy	48.6	71.8	70.9	70.3	71.9	71.5	72.1	70.8	71.1	72.7	70.7	71.0	73.0	70.5	67.5	70.4	69.5	73.6	75.3	74.6	**75.8**

Table 20: **Comparison of HMTCN+CIF with baseline methods without CIF under Subject-Independent Setup.** Baseline methods are evaluated without CIF, while our proposed HMTCN is evaluated with CIF.

Datasets	**Models**	**Accuracy** ↑	**Precision** ↑	**Recall** ↑	**F1 score** ↑	**AUROC** ↑	**AUPRC** ↑
APAVA (2-Classes)	**ADformer**	79.59±3.11	81.75±4.03	76.51±3.30	77.42±3.50	84.70±6.97	85.41±6.71
	BIOT	70.10±0.33	73.17±1.69	65.10±0.58	64.60±0.91	68.58±1.35	69.15±1.78
	Conformer	82.21±5.43	**86.69±2.77**	78.71±6.78	79.59±7.39	90.70±2.83	91.08±2.72
	EEGInception	77.96±6.88	83.49±5.91	74.41±8.23	74.29±10.30	86.30±2.19	86.95±2.11
	ResNet1d	80.11±3.76	80.69±4.79	78.44±3.20	78.92±3.52	83.07±6.75	83.99±6.71
	XResNet1d	66.25±6.65	66.58±6.64	63.76±6.75	63.01±7.35	70.66±7.28	70.55±7.66
	Inception1d	81.08±3.18	82.63±5.24	78.78±2.54	79.58±2.97	85.65±5.78	86.53±5.43
	HMTCN + CIF	**86.30±1.05**	86.16±1.09	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	**94.42±0.49**
TDBrain (2-Classes)	**ADformer**	84.19±1.49	85.83±0.49	84.19±1.49	83.99±1.69	91.72±0.31	91.61±0.44
	BIOT	81.15±5.92	81.52±6.35	81.15±5.92	81.11±5.89	90.03±4.65	90.29±4.46
	Conformer	90.38±1.86	90.48±1.78	90.38±1.86	90.37±1.87	96.87±1.07	96.79±1.20
	EEGInception	87.42±4.79	88.55±3.77	87.42±4.79	87.29±4.93	96.52±2.05	96.63±1.98
	ResNet1d	85.40±2.57	87.42±1.70	85.40±2.57	85.18±2.69	96.11±1.42	96.24±1.36
	XResNet1d	87.88±3.85	88.21±3.54	87.88±3.85	87.84±3.90	95.38±2.82	95.55±2.72
	Inception1d	88.71±4.24	89.57±3.67	88.71±4.24	88.63±4.31	96.70±2.08	96.81±2.01
	HMTCN + CIF	**93.13±1.41**	**93.33±1.37**	**93.13±1.41**	**93.12±1.42**	**98.62±0.66**	**98.68±0.63**
ADFTD (3-Classes)	**ADformer**	52.46±3.56	51.11±3.52	50.80±3.21	50.52±3.47	72.27±2.73	54.37±3.46
	BIOT	42.49±0.71	42.17±0.70	41.44±0.92	41.66±0.81	58.61±0.90	42.76±1.85
	Conformer	49.37±2.62	48.64±1.49	48.88±1.80	48.03±2.15	68.00±2.42	49.38±1.58
	EEGInception	50.75±1.67	47.55±1.45	47.36±1.42	46.31±1.95	67.52±1.13	48.55±1.29
	ResNet1d	48.75±3.30	47.12±3.17	47.23±3.49	46.88±3.38	65.99±3.43	47.88±3.19
	XResNet1d	49.89±1.76	46.69±1.95	46.93±1.65	46.45±1.75	65.93±2.06	47.86±2.17
	Inception1d	49.85±1.93	48.40±1.46	48.33±1.33	47.22±1.17	68.06±1.30	49.62±1.08
	HMTCN + CIF	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**
PTB (2-Classes)	**ADformer**	77.61±2.17	84.81±1.25	66.80±3.57	67.90±4.63	89.64±1.00	88.93±0.90
	BIOT	77.32±2.28	81.37±2.63	67.13±3.35	68.42±4.16	91.74±1.09	87.57±1.62
	Conformer	77.27±2.00	81.33±1.27	67.22±3.69	68.39±4.39	89.13±2.36	86.44±2.68
	EEGInception	81.75±1.95	84.07±3.03	74.28±2.29	76.51±2.55	91.60±1.49	89.98±2.07
	ResNet1d	76.02±1.53	79.36±1.59	65.50±2.57	66.41±3.27	87.96±1.85	85.09±2.03
	XResNet1d	80.46±2.53	85.58±2.00	71.34±3.83	73.42±4.26	90.19±2.38	89.32±2.78
	Inception1d	79.91±2.65	84.92±2.01	70.76±4.51	72.56±5.13	89.25±3.87	88.15±2.93
	HMTCN + CIF	**88.29±1.45**	**90.66±1.48**	**83.21±2.02**	**85.59±1.96**	**94.28±0.93**	**93.78±1.11**

L Reproduced Results of Existing Methods

L.1 Results of Subject-Independent

Table 3 presents the results of various methods reported under the subject-independent setup, while Table 21 shows the results of our reproduction of these methods. Our approach achieves the highest scores on six metrics across four out of the five datasets. For the PTB-XL dataset, compared to the results reported by other methods, our method achieves the highest AUROC, with Accuracy ranking second.

Table 21: **Results of Subject-Independent Setup.** The training, validation, and test sets are distributed based on subjects according to a predetermined ratio/IDs. Results of the APAVA, TDBrain, ADFTD, PTB, and PTB-XL datasets under this setup are presented here. Unfortunately, MedGNN has only released the training parameters for the APAVA and ADFTD datasets.

Datasets	Models	Accuracy ↑	Precision ↑	Recall ↑	F1 score ↑	AUROC ↑	AUPRC ↑
APAVA (2-Classes) Reproduced	**Autoformer**	73.18±7.33	73.87±6.72	73.01±6.10	72.40±7.03	81.64±7.24	81.10±7.75
	Crossformer	72.76±2.04	79.64±2.45	67.41±2.62	66.88±3.61	71.81±4.06	71.64±3.74
	FEDformer	75.16±1.67	74.98±0.69	73.34±2.97	73.50±2.90	83.89±1.54	83.27±1.62
	Informer	72.20±2.78	73.92±4.80	68.48±2.51	68.74±2.70	70.14±3.43	70.84±3.80
	iTransformer	74.55±1.66	74.77±2.10	71.76±1.72	72.30±1.79	85.59±1.55	84.39±1.57
	MTST	69.24±1.24	75.87±2.80	63.28±1.81	61.62±2.75	66.09±3.27	68.08±2.93
	Nonformer	71.81±4.20	71.31±4.40	70.15±3.38	70.38±3.74	71.54±2.73	72.79±2.50
	PatchTST	68.27±2.11	78.56±1.88	61.53±2.60	58.52±4.07	64.61±2.18	67.14±2.06
	Reformer	78.42±2.85	80.89±4.52	75.20±2.28	76.09±2.54	75.48±2.79	77.52±2.64
	Transformer	75.53±4.28	76.90±5.05	72.14±4.87	72.64±5.44	72.30±6.04	73.04±7.15
	Medformer	77.85±2.42	80.31±3.21	74.38±2.49	75.21±2.67	80.85±3.80	81.62±3.24
	MedGNN	77.40±5.77	82.77±4.46	73.24±7.06	73.29±9.01	81.31±2.94	82.80±2.91
	HMTCN + CIF	**86.30±1.05**	**86.16±1.09**	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	**94.42±0.49**
TDBrain (2-Classes) Reproduced	**Autoformer**	90.38±3.03	91.16±2.42	90.38±3.03	90.31±3.09	95.83±2.14	95.43±2.31
	Crossformer	82.15±2.60	82.81±2.11	82.15±2.60	82.04±2.70	91.20±2.23	91.47±2.16
	FEDformer	77.60±1.23	78.25±1.52	77.60±1.23	77.48±1.19	86.31±1.23	86.48±1.36
	Informer	88.42±2.99	89.01±2.45	88.42±2.99	88.36±3.05	96.54±0.90	96.66±0.85
	iTransformer	74.69±1.02	74.76±1.04	74.69±1.02	74.67±1.02	83.35±1.24	83.65±1.41
	MTST	77.67±3.58	78.97±4.37	77.67±3.58	77.45±3.55	86.47±4.84	84.99±6.43
	Nonformer	88.10±2.39	88.76±1.74	88.10±2.39	88.04±2.47	96.56±0.91	96.36±1.21
	PatchTST	77.98±2.64	79.30±3.73	77.98±2.64	77.76±2.65	86.67±4.03	84.93±5.47
	Reformer	88.50±2.30	89.01±1.80	88.50±2.30	88.45±2.35	96.10±0.63	96.19±0.55
	Transformer	85.13±1.86	86.39±1.56	85.13±1.86	84.99±1.93	95.61±1.05	95.63±0.91
	Medformer	88.77±1.24	88.91±1.11	88.77±1.24	88.76±1.25	96.38±0.34	96.44±0.30
	MedGNN	-	-	-	-	-	-
	HMTCN + CIF	**93.13±1.41**	**93.33±1.37**	**93.13±1.41**	**93.12±1.42**	**98.62±0.66**	**98.68±0.63**
ADFTD (3-Classes) Reproduced	**Autoformer**	46.90±2.89	45.59±2.37	44.91±2.23	44.34±2.52	63.49±2.44	45.63±2.29
	Crossformer	50.18±1.97	45.97±1.84	46.30±1.73	45.90±1.84	66.68±1.67	48.65±1.89
	FEDformer	45.75±0.78	45.71±1.29	44.27±1.28	43.51±1.00	62.64±1.64	45.88±1.35
	Informer	48.42±1.99	46.94±1.60	46.41±0.99	45.76±0.43	65.99±1.14	47.49±1.07
	iTransformer	52.85±1.36	46.97±1.05	47.31±1.03	46.84±0.78	67.46±0.96	49.90±0.89
	MTST	45.77±1.70	44.39±1.73	43.70±1.82	43.36±1.98	61.38±1.57	44.01±1.60
	Nonformer	50.81±1.06	48.71±1.40	48.55±1.47	48.36±1.38	66.95±1.54	48.08±1.82
	PatchTST	43.32±0.53	41.95±0.38	41.45±1.26	40.75±1.62	60.21±0.29	42.49±0.57
	Reformer	51.28±2.60	49.68±2.75	49.64±2.02	48.45±2.06	69.20±2.53	51.74±3.24
	Transformer	50.53±0.94	49.31±0.87	48.57±1.23	48.42±1.28	67.98±0.90	49.07±1.35
	Medformer	53.70±1.18	51.51±1.32	50.49±1.48	50.35±1.53	70.48±1.17	50.91±1.13
	MedGNN	50.22±3.21	48.65±3.72	47.50±4.57	47.33±4.40	67.18±4.39	48.84±4.11
	HMTCN + CIF	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**
PTB (2-Classes) Reproduced	**Autoformer**	71.99±2.74	69.60±3.85	61.50±4.23	61.43±5.07	74.29±1.89	70.26±2.00
	Crossformer	78.06±3.44	81.53±3.13	68.62±5.63	69.76±6.53	88.31±2.07	85.81±2.43
	FEDformer	74.54±2.27	77.99±4.10	63.14±3.29	63.28±4.36	84.63±4.27	80.91±5.55
	Informer	79.59±0.65	83.33±0.77	70.58±0.95	72.58±1.08	92.77±0.48	90.89±0.57
	iTransformer	83.43±1.19	88.06±1.47	75.64±1.55	78.29±1.70	91.38±1.41	91.08±1.30
	MTST	75.53±2.45	78.72±1.87	64.78±4.06	65.30±4.81	87.76±4.09	83.60±3.92
	Nonformer	78.93±1.46	82.48±1.53	69.68±2.15	71.50±2.56	90.54±0.59	87.78±1.46
	PatchTST	75.28±2.44	77.05±2.44	64.86±4.05	65.41±5.28	88.11±2.59	82.65±2.87
	Reformer	78.11±1.65	82.70±0.80	68.17±2.68	69.68±3.34	90.77±1.56	88.14±1.20
	Transformer	76.43±1.98	81.25±1.15	65.64±3.29	66.44±4.39	90.21±1.24	87.28±1.49
	Medformer	80.99±0.75	83.01±0.72	73.35±1.16	75.47±1.21	93.10±1.18	90.69±1.04
	MedGNN	-	-	-	-	-	-
	HMTCN + CIF	**88.29±1.45**	**90.66±1.48**	**83.21±2.02**	**85.59±1.96**	**94.28±0.93**	**93.78±1.11**
PTB-XL (5-Classes) Reproduced	**Autoformer**	60.37±1.72	49.83±0.50	49.17±0.72	48.43±0.47	81.35±0.56	50.83±0.72
	Crossformer	73.15±0.17	64.62±0.31	**61.28±0.45**	**62.49±0.34**	89.94±0.15	67.14±0.29
	FEDformer	56.01±10.27	51.61±5.58	48.77±7.64	47.23±8.76	82.01±4.20	52.23±7.14
	Informer	71.28±0.27	62.28±0.53	59.18±0.54	60.38±0.38	88.57±0.09	64.57±0.18
	iTransformer	69.18±0.33	59.44±0.48	54.84±0.25	56.44±0.23	86.65±0.17	60.27±0.44
	MTST	72.14±0.29	64.01±0.61	59.31±0.27	61.07±0.32	88.82±0.19	65.52±0.52
	Nonformer	70.51±0.79	61.42±1.03	58.15±0.57	59.31±0.63	88.20±0.43	63.33±0.77
	PatchTST	73.13±0.20	65.38±0.54	60.61±0.56	62.39±0.27	89.66±0.16	66.97±0.24
	Reformer	71.24±0.46	61.78±0.79	59.67±0.99	60.51±0.65	88.67±0.21	64.29±0.29
	Transformer	70.46±0.42	61.56±0.58	57.86±0.65	59.16±0.46	88.18±0.14	63.25±0.27
	Medformer	72.94±0.15	64.39±0.32	59.98±0.49	61.61±0.33	89.67±0.10	66.25±0.25
	MedGNN	-	-	-	-	-	-
	HMTCN + CIF	**73.73±0.30**	**65.41±0.67**	60.70±1.08	61.89±0.91	**90.53±0.22**	**67.75±0.75**

In comparison with our reproduced results, our method ranks first in Accuracy, AUROC, and AUPRC, and second in Precision and Recall. Additionally, it is noteworthy that under the subject-independent setup, the F1 score for ADFTD is 55.42%, significantly lower than the 99.59% F1 score achieved under the subject-dependent setup. This comparison highlights the challenge of the subject-independent setup, which better simulates real-world scenarios.

M Ablation Study

M.1 Module Study of CIF

We conduct experiments by integrating our Channel-Imposed Fusion principle with other methods.

In the CIF structure, there are four different hyperparameters. First, t is the switch for forward and backward feature fusion. When $t = 1$, it replaces the forward features with the fused features; when $t = -1$, it replaces the backward features with the fused features. The second hyperparameter, n, determines the number of selected channels. Additionally, there are two scaling factors associated with the channels: a, the scaling factor for the first n channels, and b, the scaling factor for the remaining n channels. Both a and b can either be learnable or fixed parameters. We use 🔥 to represent learnable parameters, and ❄ to represent non-learnable parameters.

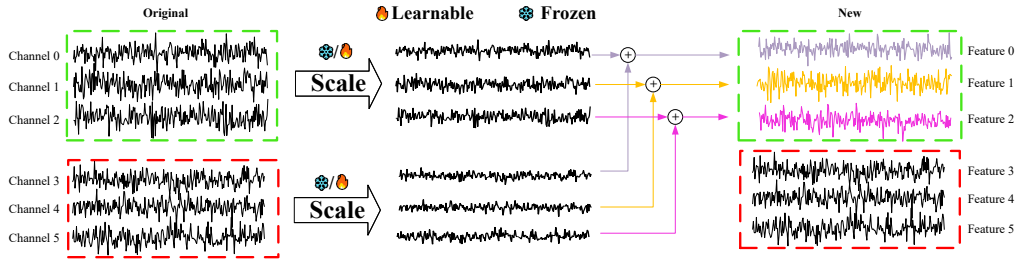


Figure 9: The case of $X_i^{new} = X_{fused}$ is marked by setting $t = 1$.

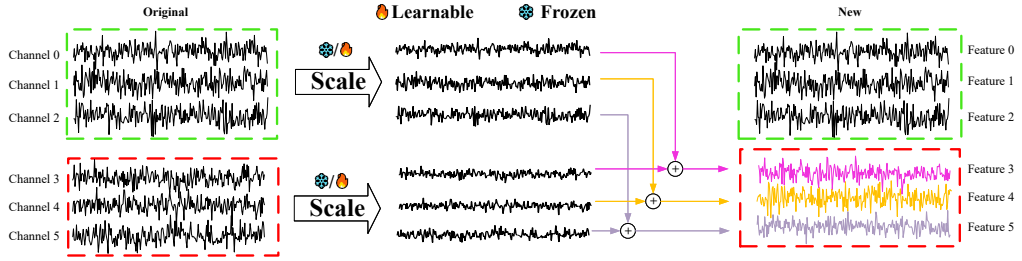


Figure 10: The case of $X_j^{new} = X_{fused}$ is marked by setting $t = -1$.

Following the theoretical analysis of the parameters a and b in Appendix A, we perform an evenly spaced grid search over the range $[-1, 1] \times [-1, 1]$ and report the overall performance trends as these hyperparameters vary.

The ablation study(Table 22 23 24) reveals that the performance of the CIF architecture is highly sensitive to its hyperparameters, with optimal configurations varying significantly across datasets. For the APAVA dataset (16 channels), the best results (86.30% accuracy, 94.26% AUROC) are achieved using backward fusion ($t = -1$) on 9 selected channels ($n = 9$), where the first 7 channels apply fixed reinforcement weights ($a = -0.8$, learnable) and the rest use fixed suppression weights ($b = -0.6$, learnable). Conversely, the ADFTD dataset (19 channels) performs best (58.56% accuracy, 76.07% AUROC) with forward fusion ($t = 1$) over 10 selected channels ($n = 10$) using dynamically learned weights for negative suppression ($a = -0.19$) and positive enhancement ($b = 0.27$). These differences, also observed in PTB (APAVA: 78–85% accuracy vs. ADFTD: 50–58%), emphasize the critical role of dataset-specific tuning. In particular, the number of selected channels (n) and

fusion direction (t) are pivotal, while adaptive weighting (a/b) further improves robustness. Overall, the results highlight the necessity of careful hyperparameter optimization—especially in channel selection and fusion strategy—for maximizing CIF performance across diverse medical time series tasks.

Overall, the number of selected channels n and the fusion direction t are the key factors influencing performance, while the adaptive weights a/b further enhance the robustness and generalization ability of the model. Taken together, these results indicate that, for diverse medical time-series tasks, fully leveraging the advantages of CIF requires careful and systematic optimization of hyperparameters such as channel selection and fusion strategy.

Table 22: Ablation study of CIF structure hyperparameters on ADFTD dataset (19 Channels). We use 🔥 to represent learnable parameters, and ❄ to represent non-learnable parameters.

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
-	-	-	-	52.05±2.22	50.45±3.00	50.40±2.55	49.48±2.70	69.43±2.84	50.99±3.15
1	10	-0.20 ❄	-0.30 ❄	57.33±0.95	54.06±1.47	54.44±1.33	53.73±1.47	74.82±1.36	58.24±1.78
1	10	-0.25 ❄	-0.40 ❄	57.08±1.46	53.31±0.98	53.24±1.14	52.87±0.98	73.77±1.28	56.86±1.47
1	11	-0.20 ❄	-0.30 ❄	56.68±1.30	53.04±2.00	53.45±2.03	52.84±2.30	73.81±1.55	57.34±1.46
1	10	-0.19 🔥	-0.27 🔥	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**
1	7	0.20 🔥	-0.60 🔥	56.36±0.58	53.91±1.10	54.34±1.03	53.78±0.91	74.19±0.69	56.62±1.28
1	9	-0.20 🔥	-0.60 🔥	55.20±1.50	53.05±1.14	53.21±1.06	52.76±1.02	73.09±1.13	55.50±1.71
1	10	-0.20 🔥	-0.20 🔥	57.11±0.92	55.04±1.43	55.53±1.57	54.58±1.29	75.13±1.33	58.47±1.43
1	11	-0.20 🔥	-0.30 🔥	57.83±1.16	54.32±1.22	54.51±1.18	54.01±1.21	74.74±1.12	57.96±1.08
1	11	-0.18 🔥	-0.25 🔥	57.00±1.32	53.00±2.28	53.17±1.94	52.32±2.19	73.97±1.89	57.00±1.93
1	7	-1.00 ❄	-1.00 ❄	51.29±1.14	49.72±1.46	49.95±1.65	49.35±1.44	69.13±0.82	51.01±1.19
1	7	-1.00 🔥	-1.00 🔥	50.46±1.57	48.50±1.19	48.68±1.29	48.20±1.38	68.18±1.08	49.73±1.24
1	7	-1.00 ❄	-0.80 ❄	50.62±2.38	49.34±2.17	49.49±2.18	48.97±2.34	68.55±1.78	50.52±2.40
1	7	-1.00 🔥	-0.80 🔥	50.51±2.57	49.04±2.31	49.26±2.60	48.73±2.66	68.32±2.13	50.38±2.53
1	7	-1.00 ❄	-0.60 ❄	51.51±1.56	50.15±1.36	50.35±1.03	49.80±1.24	69.03±0.92	51.09±1.48
1	7	-1.00 🔥	-0.60 🔥	51.09±2.05	49.84±1.97	50.24±2.16	49.33±1.84	69.04±1.83	51.15±2.55
1	7	-1.00 ❄	-0.40 ❄	51.17±1.27	48.47±1.03	48.45±1.36	47.83±1.38	68.20±0.79	49.55±1.06
1	7	-1.00 ❄	-0.40 ❄	53.59±1.01	52.16±2.37	52.16±3.11	51.01±3.24	70.77±2.29	53.17±3.59
1	7	-1.00 🔥	-0.40 🔥	51.27±0.91	48.14±0.91	48.28±1.15	47.66±1.18	67.95±1.05	48.92±1.71
1	7	-1.00 🔥	-0.40 🔥	53.03±1.70	52.41±2.70	51.95±3.09	50.35±3.43	70.69±2.34	52.99±3.58
1	7	-1.00 ❄	-0.20 ❄	51.11±0.61	48.09±0.51	47.88±0.59	47.40±0.59	67.73±0.67	48.74±0.89
1	7	-1.00 ❄	-0.20 ❄	54.26±0.58	52.53±1.78	51.58±2.01	51.31±2.32	70.87±1.43	53.25±2.30
1	7	-1.00 🔥	-0.20 🔥	50.21±2.53	47.75±2.54	47.84±2.62	47.38±2.51	67.34±2.67	48.72±3.07
1	7	-1.00 🔥	-0.20 🔥	54.91±1.15	53.99±1.35	53.54±2.01	53.16±1.56	72.41±1.23	55.67±1.70
1	8	-1.00 ❄	-1.00 ❄	51.22±0.94	49.68±1.50	49.67±2.01	49.18±1.59	68.92±1.00	51.02±1.05
1	8	-1.00 🔥	-1.00 🔥	52.00±1.48	50.10±1.12	49.87±1.43	49.56±1.17	69.38±0.97	51.20±1.26
1	8	-1.00 ❄	-0.80 ❄	52.20±1.50	49.88±1.71	49.90±1.85	49.69±1.75	69.44±1.36	51.39±1.69
1	8	-1.00 🔥	-0.80 🔥	51.19±2.06	47.57±1.30	47.61±1.79	47.18±1.62	67.52±1.78	48.94±1.69
1	8	-1.00 ❄	-0.60 ❄	51.67±1.42	47.79±0.92	47.97±1.22	47.49±1.21	67.89±1.31	49.35±1.62
1	8	-1.00 🔥	-0.60 🔥	51.15±1.78	47.52±0.70	47.29±1.10	46.62±0.71	67.69±1.50	48.90±1.31
1	8	-1.00 ❄	-0.40 ❄	52.58±1.40	49.21±2.48	49.35±1.80	48.50±2.27	68.95±1.57	50.26±2.08
1	8	-1.00 ❄	-0.40 ❄	52.40±0.54	50.11±2.26	49.92±2.13	48.87±2.64	69.13±1.85	50.62±2.47
1	8	-1.00 🔥	-0.40 🔥	52.61±1.20	49.38±2.01	49.07±1.57	48.32±1.97	68.92±1.50	50.01±1.66
1	8	-1.00 🔥	-0.40 🔥	52.47±1.46	49.68±1.96	49.59±1.70	48.30±2.21	68.60±1.70	49.58±2.21
1	8	-1.00 ❄	-0.20 ❄	51.83±1.52	48.65±2.57	48.79±2.05	48.16±2.45	67.83±1.86	49.21±2.17
1	8	-1.00 ❄	-0.20 ❄	52.14±1.35	50.70±2.37	50.25±2.35	49.75±2.07	69.64±1.73	51.02±2.65
1	8	-1.00 🔥	-0.20 🔥	52.70±1.14	50.10±1.76	49.66±1.42	48.99±1.71	69.22±1.06	50.20±1.09
1	8	-1.00 🔥	-0.20 🔥	52.64±0.73	49.69±2.78	50.02±1.94	48.95±2.76	69.15±1.81	50.69±2.16
1	9	-1.00 ❄	-1.00 ❄	53.35±1.53	49.43±0.75	49.65±0.54	49.05±1.03	69.05±0.60	51.18±0.44
1	9	-1.00 🔥	-1.00 🔥	53.42±2.39	50.34±1.98	50.23±1.85	49.47±2.21	69.43±1.40	51.76±1.87

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	9	-1.00 ❄	-0.80 ❄	52.60±1.52	50.27±0.83	50.10±0.85	49.49±0.70	69.19±0.43	51.16±0.76
1	9	-1.00 🔥	-0.80 🔥	52.41±1.31	49.15±0.81	49.14±1.22	48.13±1.43	68.53±0.45	50.22±0.82
1	9	-1.00 ❄	-0.60 ❄	52.18±0.91	49.49±1.11	49.40±1.37	48.90±0.85	68.75±1.16	50.44±1.61
1	9	-1.00 🔥	-0.60 🔥	51.92±1.60	49.03±0.83	48.73±0.78	48.13±0.88	68.31±0.40	49.95±0.46
1	9	-1.00 ❄	-0.40 ❄	51.34±0.70	49.72±0.34	49.90±0.42	49.32±0.42	68.67±0.31	50.21±0.90
1	9	-1.00 ❄	-0.40 ❄	51.36±2.78	49.89±1.37	49.48±1.07	48.83±1.52	69.11±1.57	50.22±1.88
1	9	-1.00 🔥	-0.40 🔥	52.24±1.09	49.96±0.92	49.42±0.99	49.01±0.85	68.50±0.91	49.71±1.22
1	9	-1.00 🔥	-0.40 🔥	51.23±2.51	50.34±0.95	49.68±0.70	48.94±1.37	69.81±1.24	51.01±1.47
1	9	-1.00 ❄	-0.20 ❄	51.30±0.83	49.86±0.96	49.61±0.99	49.00±0.83	68.90±0.99	50.08±1.43
1	9	-1.00 ❄	-0.20 ❄	50.63±2.05	49.20±1.98	48.76±1.75	48.08±1.47	68.55±1.89	49.46±2.35
1	9	-1.00 🔥	-0.20 🔥	53.22±2.54	51.76±2.90	50.85±3.02	50.22±3.38	69.92±2.50	51.51±3.36
1	9	-1.00 🔥	-0.20 🔥	51.03±2.23	49.86±1.36	49.59±1.59	49.03±1.66	69.13±1.44	50.35±2.08
1	10	-1.00 ❄	-1.00 ❄	51.76±1.87	49.07±2.33	48.94±2.07	48.35±2.23	68.73±1.56	50.36±2.20
1	10	-1.00 🔥	-1.00 🔥	49.94±1.17	47.35±2.22	47.48±1.98	46.89±2.20	67.17±0.98	48.61±1.84
1	10	-1.00 ❄	-0.80 ❄	52.94±2.14	50.68±2.19	50.17±2.00	49.79±1.94	69.46±1.79	51.11±2.61
1	10	-1.00 🔥	-0.80 🔥	52.13±2.00	49.49±2.73	49.51±2.54	48.81±2.72	68.65±1.77	50.42±2.08
1	10	-1.00 ❄	-0.60 ❄	51.87±2.46	49.68±2.77	49.99±2.84	49.24±2.95	69.18±2.01	51.03±2.77
1	10	-1.00 🔥	-0.60 🔥	51.86±1.69	49.10±1.32	49.76±1.38	48.69±1.49	69.01±1.02	50.41±1.35
1	10	-1.00 ❄	-0.40 ❄	52.66±1.03	47.97±1.50	48.38±0.97	47.34±1.41	68.12±1.15	49.01±1.38
1	10	-1.00 ❄	-0.40 ❄	52.51±2.82	51.49±3.07	51.10±3.57	50.33±3.43	70.03±3.22	51.80±4.02
1	10	-1.00 🔥	-0.40 🔥	52.41±1.03	48.62±1.93	49.14±1.60	47.99±1.75	68.41±1.13	49.39±1.37
1	10	-1.00 🔥	-0.40 🔥	51.49±1.49	50.82±1.60	50.21±1.68	49.51±1.59	69.39±1.74	50.83±2.24
1	10	-1.00 ❄	-0.20 ❄	52.93±0.91	50.09±2.50	49.23±1.64	48.63±1.98	69.06±1.84	50.34±2.30
1	10	-1.00 ❄	-0.20 ❄	54.29±2.40	52.85±2.43	51.05±1.91	50.86±2.26	71.26±2.20	53.12±2.65
1	10	-1.00 🔥	-0.20 🔥	53.17±2.38	49.78±3.85	50.04±3.26	49.42±3.77	69.42±3.45	50.81±4.85
1	10	-1.00 🔥	-0.20 🔥	52.73±3.06	51.96±2.78	51.07±2.84	50.37±3.15	70.70±2.84	52.46±3.47
1	11	-1.00 ❄	-1.00 ❄	52.58±0.58	48.82±1.37	49.03±1.32	48.54±1.36	68.70±0.98	50.46±1.21
1	11	-1.00 🔥	-1.00 🔥	51.62±1.79	48.69±2.60	48.58±2.66	48.48±2.49	68.51±1.74	50.69±2.33
1	11	-1.00 ❄	-0.80 ❄	52.30±1.30	49.28±1.37	48.77±0.61	48.05±1.12	68.63±1.10	50.61±1.37
1	11	-1.00 🔥	-0.80 🔥	53.35±2.23	49.09±2.26	49.08±2.12	47.91±2.25	68.62±1.60	50.38±1.64
1	11	-1.00 ❄	-0.60 ❄	51.63±1.71	49.23±1.82	48.99±1.43	48.42±1.59	69.21±1.40	51.41±1.71
1	11	-1.00 🔥	-0.60 🔥	52.79±2.04	48.88±3.26	49.16±2.06	48.29±2.78	68.38±2.51	50.14±2.64
1	11	-1.00 ❄	-0.40 ❄	51.21±2.03	48.89±1.47	48.68±0.96	47.98±1.63	68.61±1.26	50.42±1.04
1	11	-1.00 🔥	-0.40 🔥	52.76±1.32	48.29±3.35	48.87±2.77	47.83±3.61	68.48±2.72	50.29±3.06
1	11	-1.00 ❄	-0.20 ❄	52.17±2.07	49.95±1.95	49.34±1.79	48.76±2.17	69.64±1.89	51.36±1.79
1	11	-1.00 ❄	-0.20 ❄	52.94±1.57	50.48±1.48	50.62±1.19	49.65±1.34	69.49±1.11	50.60±1.77
1	11	-1.00 🔥	-0.20 🔥	52.86±1.18	49.30±2.91	49.10±2.10	48.35±2.69	69.00±1.91	50.58±2.21
1	11	-1.00 🔥	-0.20 🔥	52.79±1.55	50.06±1.70	50.33±1.44	49.25±1.58	69.63±1.11	50.74±1.63
1	12	-1.00 ❄	-1.00 ❄	52.23±0.70	49.02±1.30	48.85±1.39	48.42±1.79	69.42±1.28	51.53±1.57
1	12	-1.00 🔥	-1.00 🔥	52.87±0.91	48.84±1.15	48.80±1.09	48.17±1.61	69.83±1.51	51.81±1.39
1	12	-1.00 ❄	-0.80 ❄	51.77±1.81	49.15±1.61	48.85±1.60	48.10±1.99	69.28±1.63	51.18±1.62
1	12	-1.00 🔥	-0.80 🔥	51.30±1.90	48.55±2.06	48.64±2.13	48.15±2.30	68.67±1.85	50.49±2.12
1	12	-1.00 ❄	-0.60 ❄	51.09±2.09	48.78±1.64	48.58±1.66	48.01±1.97	68.73±1.96	50.50±2.11
1	12	-1.00 🔥	-0.60 🔥	50.75±2.02	48.16±2.49	47.98±2.48	47.38±2.42	68.05±2.19	49.60±2.52
1	12	-1.00 ❄	-0.40 ❄	50.24±1.48	47.38±1.71	47.13±1.68	46.79±1.63	67.87±1.12	49.14±1.11
1	12	-1.00 🔥	-0.40 🔥	49.54±2.95	47.32±2.66	47.10±2.16	46.00±2.48	67.31±1.79	48.73±2.04
1	12	-1.00 ❄	-0.20 ❄	50.92±1.56	48.17±1.85	48.04±1.63	47.63±1.68	68.12±1.42	49.36±1.84
1	12	-1.00 ❄	-0.20 ❄	51.41±2.32	49.46±2.48	49.58±2.61	48.42±2.26	68.68±2.27	49.71±2.82
1	12	-1.00 🔥	-0.20 🔥	51.05±1.58	48.59±2.39	48.22±1.78	47.63±1.73	68.61±1.48	49.90±1.68
1	12	-1.00 🔥	-0.20 🔥	52.52±0.88	49.16±2.35	49.51±2.45	48.47±1.92	69.29±2.02	50.12±2.34

Table 23: Ablation study of CIF structure hyperparameters on PTB dataset (15 Channels). We use 🔥 to represent learnable parameters, and ❄ to represent non-learnable parameters.

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
-	-	-	-	81.87±1.87	86.50±1.24	73.49±2.90	75.84±3.20	94.20±0.29	93.04±0.45
1	8	0.21 ❄	-0.50 ❄	**88.29±1.45**	**90.66±1.48**	**83.21±2.02**	**85.59±1.96**	94.28±0.93	**93.78±1.11**
1	8	0.20 ❄	-0.50 ❄	84.45±2.15	88.34±1.15	77.40±3.52	79.89±3.51	93.15±0.70	91.94±0.67
1	9	0.22 ❄	-0.50 ❄	85.14±0.22	87.42±1.08	79.05±0.52	81.42±0.35	93.82±1.19	92.73±1.27
1	10	0.22 ❄	-0.50 ❄	85.40±2.08	88.17±1.96	79.17±3.12	81.57±3.07	**94.57±1.13**	93.53±1.44
1	11	0.22 ❄	-0.50 ❄	84.47±0.84	88.02±0.59	77.50±1.26	80.10±1.28	92.77±0.89	92.02±0.78
1	8	0.23 ❄	-0.50 ❄	83.96±5.41	87.13±3.96	76.89±8.22	78.66±9.51	92.49±2.91	91.14±3.79
1	8	0.25 ❄	-0.50 ❄	85.56±2.09	88.96±1.77	79.03±3.03	81.61±3.14	93.61±1.11	93.07±1.03
-1	7	-1.00 ❄	-1.00 ❄	82.57±0.88	86.88±0.94	74.58±1.26	77.09±1.35	93.75±1.30	92.74±1.65
-1	7	-1.00 🔥	-1.00 🔥	82.26±3.27	86.11±2.38	74.28±4.93	76.48±5.44	92.94±2.37	91.53±2.75
-1	7	-1.00 ❄	-0.90 ❄	81.49±3.97	86.49±2.89	72.81±6.00	74.78±6.87	93.41±1.80	92.62±2.04
-1	7	-1.00 🔥	-0.90 🔥	78.44±4.44	84.94±2.03	68.22±7.05	69.04±8.81	92.58±2.17	91.48±2.14
-1	7	-1.00 ❄	-0.80 ❄	79.94±3.18	85.35±2.61	70.48±4.79	72.26±6.02	93.43±1.49	92.30±1.61
-1	7	-1.00 🔥	-0.80 🔥	80.57±4.70	85.89±3.20	71.38±7.18	72.86±9.26	92.43±0.59	91.48±1.19
-1	7	-1.00 ❄	-0.70 ❄	81.01±2.74	86.90±0.99	71.93±4.34	73.99±5.25	92.44±2.72	91.70±2.53
-1	7	-1.00 🔥	-0.70 🔥	83.64±1.03	87.21±0.54	76.36±1.71	78.88±1.65	93.50±0.94	92.83±0.73
-1	7	-1.00 ❄	-0.60 ❄	80.22±1.47	86.35±1.48	70.72±2.19	72.82±2.57	92.61±1.47	91.70±1.77
-1	7	-1.00 🔥	-0.60 🔥	81.73±2.77	86.55±1.78	73.24±4.32	75.45±4.82	90.80±1.65	90.14±1.62
-1	8	-1.00 ❄	-1.00 ❄	82.23±3.14	86.39±1.88	74.18±4.86	76.40±5.01	94.31±1.14	93.25±1.17
-1	8	-1.00 🔥	-1.00 🔥	80.07±1.79	84.28±2.52	71.10±2.50	73.16±2.85	93.63±0.95	92.04±1.42
-1	8	-1.00 ❄	-0.90 ❄	81.36±2.30	86.49±1.89	72.62±3.49	74.87±3.81	92.14±2.56	91.19±2.68
-1	8	-1.00 🔥	-0.90 🔥	79.96±1.46	84.20±1.84	70.92±2.02	72.98±2.32	92.00±1.97	90.66±2.01
-1	8	-1.00 ❄	-0.80 ❄	80.34±1.93	86.14±1.65	71.00±3.00	73.07±3.40	91.44±2.64	91.01±2.55
-1	8	-1.00 🔥	-0.80 🔥	77.08±4.30	84.12±3.23	66.14±6.66	66.44±9.24	89.68±3.65	88.73±3.53
-1	8	-1.00 ❄	-0.70 ❄	79.17±2.48	85.39±1.26	69.35±4.22	70.91±5.05	92.43±1.29	91.69±0.98
-1	8	-1.00 🔥	-0.70 🔥	80.03±4.82	85.46±2.00	70.75±7.74	71.98±8.94	91.05±2.98	90.05±3.18
-1	9	-1.00 ❄	-1.00 ❄	81.23±2.67	86.60±1.43	72.39±4.23	74.52±4.64	92.20±1.22	91.55±1.57
-1	9	-1.00 🔥	-1.00 🔥	80.05±4.78	86.50±2.17	70.45±7.40	71.73±9.27	92.73±0.93	91.87±1.14
-1	9	-1.00 ❄	-0.90 ❄	80.89±4.06	87.74±1.78	71.49±6.25	73.24±7.85	93.83±2.39	93.33±2.64
-1	9	-1.00 🔥	-0.90 🔥	82.34±3.22	87.35±2.06	74.12±5.14	76.30±5.62	94.05±2.17	93.54±2.08
-1	9	-1.00 ❄	-0.80 ❄	84.06±3.49	88.91±1.67	76.49±5.44	78.88±5.80	92.73±1.43	92.39±1.26
-1	9	-1.00 🔥	-0.80 🔥	82.50±2.14	88.19±1.12	74.03±3.33	76.51±3.67	94.33±0.84	93.83±1.05
-1	9	-1.00 ❄	-0.70 ❄	81.44±4.04	87.88±1.77	72.37±6.23	74.28±7.32	92.35±1.77	91.86±1.50
-1	9	-1.00 🔥	-0.70 🔥	80.07±3.15	87.48±1.01	70.24±4.97	71.96±6.03	92.83±2.18	92.43±2.24
-1	10	-1.00 ❄	-1.00 ❄	79.43±3.39	85.03±2.12	69.86±5.42	71.34±6.75	90.25±3.82	89.29±4.14
-1	10	-1.00 🔥	-1.00 🔥	82.20±1.03	86.48±1.03	74.07±1.47	76.53±1.61	92.05±2.25	91.29±2.26
-1	10	-1.00 ❄	-0.90 ❄	83.08±3.17	87.22±1.54	75.41±5.02	77.64±5.69	90.27±3.83	89.89±3.83
-1	10	-1.00 🔥	-0.90 🔥	80.74±3.10	86.14±1.21	71.72±4.94	73.63±5.85	92.44±1.50	91.53±1.55
-1	10	-1.00 ❄	-0.80 ❄	78.96±3.54	86.02±2.07	68.75±5.52	70.07±6.74	90.23±3.34	89.36±3.48
-1	10	-1.00 🔥	-0.80 🔥	79.20±3.45	85.98±0.63	69.31±5.83	70.58±6.97	93.24±1.60	92.07±1.78
-1	10	-1.00 ❄	-0.70 ❄	82.37±1.89	87.63±0.88	74.00±3.03	76.44±3.21	93.05±1.90	92.39±1.83
-1	10	-1.00 🔥	-0.70 🔥	77.99±4.51	85.22±0.87	67.50±7.35	67.94±9.84	91.94±3.85	90.81±3.67
-1	11	-1.00 ❄	-1.00 ❄	81.63±3.29	86.31±2.32	73.11±5.00	75.22±5.90	91.76±2.02	90.82±2.04
-1	11	-1.00 🔥	-1.00 🔥	79.09±3.82	86.30±1.79	68.91±5.95	70.15±7.70	91.35±3.78	90.46±3.63
-1	11	-1.00 ❄	-0.90 ❄	79.68±2.84	86.22±1.10	69.89±4.48	71.60±5.35	92.11±1.09	91.33±1.27
-1	11	-1.00 🔥	-0.90 🔥	78.44±3.40	86.20±1.44	67.84±5.36	68.96±6.63	90.75±5.55	90.19±5.31
-1	11	-1.00 ❄	-0.80 ❄	80.05±2.78	87.14±1.09	70.25±4.35	72.08±5.08	92.86±2.82	92.27±2.86
-1	11	-1.00 🔥	-0.80 🔥	82.18±0.99	86.96±1.56	73.88±1.47	76.36±1.56	91.48±3.29	90.85±3.31
-1	11	-1.00 ❄	-0.70 ❄	80.82±2.07	85.72±1.44	71.94±3.23	74.08±3.62	89.25±2.75	88.51±2.83
-1	11	-1.00 🔥	-0.70 🔥	78.20±2.54	84.96±1.72	67.72±3.95	69.00±5.18	90.29±0.65	89.52±0.96

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
-1	12	-1.00 ❄	-1.00 ❄	78.27±3.04	86.34±1.19	67.54±4.80	68.65±6.04	91.52±2.68	90.92±2.48
-1	12	-1.00 🔥	-1.00 🔥	82.45±1.76	87.85±0.79	74.10±3.00	76.55±3.05	92.78±1.02	92.30±0.82
-1	12	-1.00 ❄	-0.90 ❄	79.49±1.84	85.59±2.16	69.67±2.57	71.57±3.17	91.80±1.78	90.84±1.96
-1	12	-1.00 🔥	-0.90 🔥	82.03±1.90	88.37±0.83	73.22±3.06	75.64±3.42	93.21±1.67	92.67±1.74
-1	12	-1.00 ❄	-0.80 ❄	80.02±2.82	86.84±2.62	70.21±4.03	72.15±4.65	91.42±0.95	90.99±1.27
-1	12	-1.00 🔥	-0.80 🔥	81.67±3.24	87.40±2.30	72.82±4.86	75.00±5.75	92.56±1.50	91.81±1.68
-1	12	-1.00 ❄	-0.70 ❄	80.32±4.95	87.36±2.58	70.62±7.57	72.03±8.91	93.41±3.07	92.78±3.34
-1	12	-1.00 🔥	-0.70 🔥	79.00±1.99	87.52±0.83	68.43±3.05	70.02±3.74	92.27±2.22	92.16±1.82
1	7	-1.00 ❄	-1.00 ❄	85.02±1.35	88.50±1.01	78.27±1.99	80.88±1.99	95.79±0.74	95.11±0.72
1	7	-1.00 🔥	-1.00 🔥	84.81±1.84	88.77±1.20	77.86±3.06	80.43±2.94	96.33±1.03	95.64±1.19
1	7	-1.00 ❄	-0.90 ❄	83.37±2.41	87.74±1.09	75.71±3.86	78.15±4.02	94.66±1.34	93.86±1.44
1	7	-1.00 🔥	-0.90 🔥	83.70±1.62	87.88±1.13	76.19±2.45	78.77±2.47	95.73±0.89	95.00±0.68
1	7	-1.00 ❄	-0.80 ❄	84.10±1.64	88.14±1.54	76.78±2.28	79.40±2.41	95.46±1.68	94.73±1.86
1	7	-1.00 🔥	-0.80 🔥	83.59±2.52	88.30±1.38	75.88±3.99	78.38±4.01	93.55±1.14	93.01±1.42
1	7	-1.00 ❄	-0.70 ❄	82.68±2.56	87.55±1.12	74.56±4.03	76.94±4.37	95.07±1.61	94.34±1.52
1	7	-1.00 🔥	-0.70 🔥	86.59±1.96	89.57±1.34	80.67±3.11	83.14±2.92	95.86±0.99	95.33±1.08
1	7	-1.00 ❄	-0.60 ❄	83.42±3.53	87.76±2.07	75.78±5.45	78.08±5.59	95.23±1.66	94.35±1.82
1	7	-1.00 🔥	-0.60 🔥	84.09±2.80	87.38±1.49	77.12±4.43	79.45±4.54	93.88±1.22	93.14±1.40
1	8	-1.00 ❄	-1.00 ❄	82.84±2.54	84.33±2.58	76.30±3.90	78.32±3.72	94.66±1.53	92.85±1.97
1	8	-1.00 🔥	-1.00 🔥	81.17±1.65	84.79±1.21	72.91±2.68	75.10±2.91	94.65±1.08	92.31±2.18
1	8	-1.00 ❄	-0.90 ❄	83.41±0.89	86.62±1.18	76.22±1.40	78.67±1.40	94.87±1.00	93.35±1.07
1	8	-1.00 🔥	-0.90 🔥	81.23±2.07	84.90±1.55	73.00±3.35	75.14±3.64	94.30±1.54	92.69±2.04
1	8	-1.00 ❄	-0.80 ❄	82.75±2.09	85.86±1.82	75.41±3.40	77.66±3.50	94.49±1.21	92.78±1.69
1	8	-1.00 🔥	-0.80 🔥	83.49±2.10	85.88±2.27	76.66±2.93	78.97±3.02	95.08±2.27	93.62±2.66
1	8	-1.00 ❄	-0.70 ❄	82.14±2.38	85.38±1.66	74.44±3.79	76.63±4.11	94.74±1.66	93.38±1.67
1	8	-1.00 🔥	-0.70 🔥	83.04±1.70	85.52±1.90	76.05±2.52	78.33±2.58	94.25±1.34	92.44±1.50
1	8	-1.00 ❄	-0.60 ❄	83.87±2.61	86.85±2.41	76.95±4.01	79.26±4.08	95.20±2.86	93.94±3.31
1	8	-1.00 🔥	-0.60 🔥	84.24±2.01	87.41±1.67	77.34±2.97	79.81±3.00	94.88±0.96	93.38±1.20
1	9	-1.00 ❄	-1.00 ❄	82.58±3.05	86.28±1.69	74.89±4.91	77.07±4.91	93.56±2.06	92.48±1.79
1	9	-1.00 🔥	-1.00 🔥	81.97±2.26	85.64±1.61	73.99±3.46	76.26±3.77	93.68±1.04	92.21±1.27
1	9	-1.00 ❄	-0.90 ❄	81.24±2.88	84.95±2.58	72.97±4.37	75.05±4.74	94.07±2.15	92.79±2.34
1	9	-1.00 🔥	-0.90 🔥	83.23±1.06	86.49±2.05	75.95±1.17	78.42±1.32	93.78±1.86	92.82±2.05
1	9	-1.00 ❄	-0.80 ❄	81.84±2.15	86.93±1.44	73.30±3.27	75.65±3.53	93.56±0.88	92.60±1.15
1	9	-1.00 🔥	-0.80 🔥	82.90±1.57	86.86±1.99	75.18±2.19	77.67±2.29	94.64±0.85	93.78±0.94
1	9	-1.00 ❄	-0.70 ❄	83.52±2.03	86.44±2.85	76.45±2.50	78.89±2.76	93.94±1.29	92.76±1.71
1	9	-1.00 🔥	-0.70 🔥	79.94±3.23	84.83±2.00	70.69±5.00	72.45±5.82	93.59±1.56	92.47±1.42
1	9	-1.00 ❄	-0.60 ❄	81.79±3.50	87.48±1.74	73.12±5.52	75.20±6.04	94.39±1.57	93.35±2.09
1	9	-1.00 🔥	-0.60 🔥	82.20±1.47	87.39±2.01	73.78±2.18	76.26±2.31	94.56±3.17	93.59±3.93
1	10	-1.00 ❄	-1.00 ❄	82.13±3.26	87.02±1.61	73.82±5.11	75.97±5.88	93.67±1.96	92.80±2.01
1	10	-1.00 🔥	-1.00 🔥	82.69±4.49	86.75±2.85	74.85±7.00	76.75±7.98	94.19±0.91	93.34±0.96
1	10	-1.00 ❄	-0.90 ❄	81.80±2.20	86.99±2.63	73.17±3.06	75.57±3.48	92.92±1.69	92.09±2.04
1	10	-1.00 🔥	-0.90 🔥	86.15±1.41	88.87±0.92	80.15±2.16	82.62±2.05	95.97±0.65	95.32±0.57
1	10	-1.00 ❄	-0.80 ❄	82.52±2.43	87.71±1.90	74.19±3.65	76.64±4.01	93.67±3.19	92.93±3.20
1	10	-1.00 🔥	-0.80 🔥	82.92±1.74	87.86±1.61	74.83±2.52	77.40±2.76	93.42±3.18	92.76±2.99
1	10	-1.00 ❄	-0.70 ❄	84.52±1.47	88.52±0.74	77.38±2.32	80.01±2.38	95.45±1.12	94.74±1.04
1	10	-1.00 🔥	-0.70 🔥	83.00±1.47	87.12±0.99	75.26±2.31	77.76±2.42	94.24±1.35	93.54±1.19
1	11	-1.00 ❄	-1.00 ❄	83.24±2.22	87.34±1.29	75.60±3.43	78.05±3.68	94.73±1.95	93.82±1.85
1	11	-1.00 🔥	-1.00 🔥	82.60±0.88	87.11±0.95	74.55±1.37	77.08±1.43	92.63±1.52	91.79±1.51
1	11	-1.00 ❄	-0.90 ❄	80.82±2.45	85.77±1.59	71.90±3.72	74.02±4.24	93.53±1.82	92.37±1.58
1	11	-1.00 🔥	-0.90 🔥	81.41±3.01	86.76±2.08	72.68±4.66	74.78±5.44	94.61±1.16	93.24±1.68
1	11	-1.00 ❄	-0.80 ❄	82.73±2.62	87.66±2.10	74.59±3.93	77.01±4.30	94.59±0.65	93.61±0.69
1	11	-1.00 🔥	-0.80 🔥	77.98±2.41	85.65±1.32	67.28±4.08	68.40±4.92	92.95±3.27	92.12±3.26
1	11	-1.00 ❄	-0.70 ❄	83.28±3.98	87.98±2.44	75.47±6.08	77.72±6.22	94.63±1.63	93.23±2.43

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	11	-1.00 🔥	-0.70 🔥	81.95±3.19	87.27±1.44	73.43±5.07	75.57±5.85	94.49±1.31	93.62±1.35
1	12	-1.00 ❄	-1.00 ❄	82.56±2.11	85.78±2.15	75.05±3.29	77.34±3.22	92.70±3.46	91.46±4.17
1	12	-1.00 🔥	-1.00 🔥	79.36±2.38	85.74±2.69	69.50±3.66	71.23±4.45	89.66±2.53	88.86±2.93
1	12	-1.00 ❄	-0.90 ❄	79.23±3.52	86.61±1.48	69.12±5.66	70.47±6.76	92.23±2.15	91.33±2.28
1	12	-1.00 🔥	-0.90 🔥	82.68±2.77	88.63±1.29	74.20±4.25	76.64±4.70	94.50±2.22	94.05±2.22
1	12	-1.00 ❄	-0.80 ❄	80.16±4.61	84.27±5.05	71.06±6.59	72.75±7.90	92.14±2.38	90.81±3.40
1	12	-1.00 🔥	-0.80 🔥	81.78±3.53	86.61±2.21	73.41±5.56	75.44±6.14	92.98±2.44	92.10±2.42
1	12	-1.00 ❄	-0.70 ❄	81.35±2.77	87.19±2.53	72.42±4.21	74.62±4.55	94.66±1.02	93.97±1.37
1	12	-1.00 🔥	-0.70 🔥	85.25±2.74	88.51±2.37	78.74±4.12	81.17±3.99	95.07±1.05	94.48±1.36

Table 24: Ablation study of CIF structure hyperparameters on APAVA dataset (16 Channels). We use 🔥 to represent learnable parameters, and ❄ to represent non-learnable parameters.

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
-	-	-	-	82.49±1.40	82.38±1.79	81.20±1.32	81.60±1.39	91.10±1.63	91.30±1.71
-1	5	1 ❄	1 ❄	81.48±1.96	81.99±2.38	79.52±2.00	80.21±2.07	90.93±1.74	91.17±1.77
-1	6	1 🔥	1 🔥	84.00±2.37	84.33±2.61	82.37±2.44	83.02±2.50	92.67±2.57	92.87±2.55
-1	7	1 🔥	1 🔥	81.77±2.31	82.67±3.24	79.78±2.04	80.49±2.26	89.92±2.60	90.23±2.53
1	9	1 ❄	1 ❄	85.16±1.55	84.76±1.62	85.33±1.27	84.82±1.49	94.06±1.07	94.21±0.99
-1	10	1 🔥	1 🔥	82.60±1.82	82.30±2.04	81.92±2.29	81.92±1.95	91.94±1.55	92.16±1.50
-1	11	1 ❄	2 ❄	78.98±0.97	78.57±1.09	77.60±0.93	77.93±0.97	87.29±0.93	87.35±0.96
1	6	-1.00 ❄	-1.00 ❄	81.20±1.28	81.50±1.16	79.51±1.84	80.01±1.67	90.03±1.00	90.28±1.03
1	6	-1.00 🔥	-1.00 🔥	81.24±1.16	81.57±0.95	79.56±1.81	80.05±1.62	90.29±0.90	90.61±0.88
1	7	-1.00 ❄	-1.00 ❄	81.33±2.24	81.36±2.52	79.71±2.22	80.24±2.31	89.57±1.98	90.02±1.89
1	7	-1.00 ❄	-1.00 ❄	81.33±2.24	81.36±2.52	79.71±2.22	80.24±2.31	89.57±1.98	90.02±1.89
1	7	-1.00 🔥	-1.00 🔥	81.62±2.58	81.63±2.95	80.09±2.48	80.60±2.62	89.94±2.18	90.33±2.07
1	7	-1.00 ❄	-0.90 ❄	81.45±2.29	81.48±2.56	79.83±2.31	80.37±2.39	89.73±1.95	90.17±1.85
1	7	-1.00 🔥	-0.90 🔥	81.58±2.27	81.72±2.66	79.88±2.21	80.47±2.33	89.67±2.01	90.10±1.93
1	7	-1.00 ❄	-0.80 ❄	82.28±1.80	82.97±1.82	80.30±2.14	81.01±2.09	90.47±1.67	90.83±1.58
1	7	-1.00 🔥	-0.80 🔥	81.80±1.95	82.45±2.33	80.01±2.01	80.59±2.07	89.90±1.61	90.35±1.53
1	7	-1.00 ❄	-0.70 ❄	82.17±2.06	82.87±2.44	80.41±2.08	80.99±2.16	90.51±1.76	90.91±1.64
1	7	-1.00 🔥	-0.70 🔥	82.05±1.11	83.13±1.64	79.89±1.42	80.66±1.33	90.47±1.20	90.86±1.11
1	7	-1.00 ❄	-0.60 ❄	82.38±1.55	83.21±2.12	80.41±1.74	81.12±1.69	90.96±1.20	91.26±1.16
1	7	-1.00 🔥	-0.60 🔥	82.19±1.49	83.02±2.15	80.23±1.61	80.93±1.60	90.81±1.26	91.15±1.18
1	7	-1.00 ❄	-0.50 ❄	82.94±0.93	83.86±1.07	80.95±1.45	81.68±1.27	91.39±0.77	91.69±0.71
1	7	-1.00 🔥	-0.50 🔥	82.52±1.10	83.35±1.63	80.60±1.40	81.28±1.30	91.17±0.98	91.49±0.92
1	7	-1.00 ❄	-0.40 ❄	82.96±0.89	83.73±1.10	81.07±1.32	81.77±1.16	91.79±0.60	92.03±0.60
1	7	-1.00 🔥	-0.40 🔥	83.02±0.79	83.82±1.14	81.08±1.05	81.82±0.95	91.85±0.60	92.09±0.60
1	7	-1.00 ❄	-0.30 ❄	83.00±1.71	83.51±2.29	81.31±1.43	81.95±1.61	91.78±1.64	92.02±1.63
1	7	-1.00 🔥	-0.30 🔥	82.78±1.68	83.49±2.38	81.10±1.04	81.70±1.43	91.98±1.35	92.19±1.36
1	7	-1.00 ❄	-0.20 ❄	83.13±1.72	83.47±2.29	81.71±1.60	82.18±1.67	92.09±1.65	92.25±1.65
1	7	-1.00 🔥	-0.20 🔥	83.16±1.93	83.61±2.47	81.55±1.79	82.14±1.91	92.13±1.74	92.30±1.72
1	7	-1.00 ❄	-0.10 ❄	84.39±0.90	84.68±0.94	82.92±1.22	83.48±1.06	93.08±0.92	93.20±0.98
1	7	-1.00 🔥	-0.10 🔥	84.56±0.97	84.89±0.69	83.05±1.42	83.64±1.20	93.19±0.94	93.30±1.01
1	7	-1.00 ❄	0.10 ❄	84.25±1.46	84.47±1.37	82.77±1.79	83.33±1.64	92.77±1.08	92.90±1.13
1	7	-1.00 🔥	0.10 🔥	84.39±1.55	84.56±1.49	82.96±1.83	83.50±1.71	92.84±1.18	92.95±1.22
1	7	-1.00 ❄	0.20 ❄	84.02±1.48	84.15±1.52	82.56±1.61	83.12±1.60	92.51±1.21	92.66±1.20
1	7	-1.00 🔥	0.20 🔥	83.37±1.81	83.65±2.04	81.85±1.83	82.40±1.87	92.11±1.31	92.27±1.26
1	7	-1.00 ❄	0.30 ❄	82.95±2.19	82.86±2.64	81.70±1.98	82.11±2.18	91.84±1.67	92.01±1.62
1	7	-1.00 🔥	0.30 🔥	82.66±1.88	82.62±2.43	81.40±1.56	81.79±1.81	91.34±1.88	91.43±1.94
1	7	-1.00 ❄	0.40 ❄	82.33±1.91	82.35±2.29	80.91±1.87	81.38±1.97	91.20±1.98	91.32±2.00
1	7	-1.00 🔥	0.40 🔥	83.12±1.75	83.32±2.07	81.57±1.79	82.13±1.84	91.91±1.65	92.08±1.59
1	7	-1.00 ❄	0.50 ❄	82.07±1.77	82.32±2.31	80.33±1.61	80.96±1.78	90.99±1.65	91.21±1.59
1	7	-1.00 🔥	0.50 🔥	82.35±2.42	82.57±2.49	80.63±2.64	81.24±2.63	91.04±1.86	91.27±1.79
1	7	-1.00 ❄	0.60 ❄	82.03±2.11	82.19±2.65	80.39±2.01	80.97±2.16	90.74±1.94	90.95±1.87
1	7	-1.00 🔥	0.60 🔥	82.08±2.01	82.22±2.48	80.48±1.97	81.04±2.08	90.80±1.92	91.00±1.86
1	7	-1.00 ❄	0.70 ❄	81.73±2.22	81.66±2.39	80.24±2.36	80.72±2.38	90.43±1.90	90.63±1.82
1	7	-1.00 🔥	0.70 🔥	81.79±2.20	81.87±2.48	80.13±2.25	80.70±2.32	90.50±1.90	90.71±1.84
1	7	-1.00 ❄	0.80 ❄	81.37±1.94	81.69±2.61	79.56±1.75	80.19±1.94	90.28±2.07	90.50±2.00
1	7	-1.00 🔥	0.80 🔥	81.83±2.37	81.99±2.63	80.10±2.44	80.71±2.52	90.31±1.97	90.54±1.91
1	7	-1.00 ❄	0.90 ❄	81.72±1.73	82.11±2.28	79.84±1.62	80.52±1.76	90.29±1.72	90.49±1.67
1	7	-1.00 🔥	0.90 🔥	82.07±2.30	82.18±2.41	80.39±2.48	80.98±2.48	90.38±1.84	90.59±1.78
1	7	-1.00 ❄	1.00 ❄	81.29±1.73	81.08±1.65	79.84±2.05	80.27±1.93	89.88±1.38	90.09±1.36
1	7	-1.00 🔥	1.00 🔥	81.31±2.34	82.15±2.11	79.34±3.26	79.89±2.98	89.94±1.88	90.14±1.84

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	7	-0.90 ❄	-1.00 ❄	81.34±2.29	81.35±2.56	79.74±2.33	80.26±2.41	89.54±2.01	89.96±1.97
1	7	-0.90 🔥	-1.00 🔥	81.30±2.17	81.41±2.48	79.59±2.20	80.16±2.28	89.53±1.99	89.96±1.95
1	7	-0.90 ❄	-0.90 ❄	81.58±2.33	81.58±2.82	80.05±2.21	80.56±2.36	89.96±2.13	90.33±2.04
1	7	-0.90 🔥	-0.90 🔥	81.89±2.19	82.12±2.15	80.15±2.54	80.73±2.47	90.10±2.14	90.41±2.09
1	7	-0.90 ❄	-0.80 ❄	82.00±1.62	82.67±1.83	80.05±1.93	80.73±1.86	90.12±1.47	90.49±1.41
1	7	-0.90 🔥	-0.80 🔥	81.75±1.77	82.20±2.33	80.00±1.77	80.58±1.82	89.88±1.42	90.30±1.36
1	7	-0.90 ❄	-0.70 ❄	81.31±1.74	81.92±2.21	79.29±1.77	79.99±1.84	89.40±1.63	89.88±1.52
1	7	-0.90 🔥	-0.70 🔥	81.87±2.08	82.28±2.50	80.18±2.09	80.74±2.16	90.02±1.74	90.44±1.63
1	7	-0.90 ❄	-0.60 ❄	81.98±1.47	82.62±1.97	80.09±1.50	80.75±1.54	90.37±1.07	90.73±1.03
1	7	-0.90 🔥	-0.60 🔥	81.93±1.35	82.56±2.14	80.08±0.93	80.73±1.17	90.09±1.10	90.49±1.07
1	7	-0.90 ❄	-0.50 ❄	81.72±1.50	82.60±2.15	79.70±1.32	80.40±1.44	90.11±1.68	90.46±1.61
1	7	-0.90 🔥	-0.50 🔥	81.33±1.40	81.65±2.17	79.66±1.26	80.19±1.32	90.08±1.36	90.45±1.21
1	7	-0.90 ❄	-0.40 ❄	82.35±0.75	82.80±1.60	80.67±0.25	81.25±0.47	91.06±0.76	91.35±0.71
1	7	-0.90 🔥	-0.40 🔥	82.45±0.46	82.95±1.18	80.68±0.54	81.31±0.41	91.13±0.59	91.40±0.55
1	7	-0.90 ❄	-0.30 ❄	81.38±1.83	81.35±2.47	80.09±1.34	80.46±1.64	90.22±1.37	90.48±1.47
1	7	-0.90 🔥	-0.30 🔥	81.31±1.88	81.24±2.53	80.15±1.24	80.44±1.63	90.21±1.42	90.47±1.52
1	7	-0.90 ❄	-0.20 ❄	82.01±1.85	81.92±2.42	80.77±1.62	81.13±1.79	90.86±1.71	91.07±1.77
1	7	-0.90 🔥	-0.20 🔥	82.42±1.76	82.42±2.22	80.98±1.67	81.47±1.77	91.10±1.67	91.28±1.72
1	7	-0.90 ❄	-0.10 ❄	83.94±0.79	84.11±0.90	82.51±1.01	83.04±0.89	92.50±0.78	92.65±0.83
1	7	-0.90 🔥	-0.10 🔥	82.92±1.67	83.22±1.78	81.44±2.00	81.92±1.87	91.66±1.60	91.82±1.62
1	7	-0.90 ❄	0.10 ❄	83.61±1.65	83.89±1.99	82.02±1.54	82.64±1.67	92.12±1.49	92.28±1.51
1	7	-0.90 🔥	0.10 🔥	83.75±2.23	84.27±2.21	81.95±2.53	82.66±2.49	92.21±2.01	92.36±1.99
1	7	-0.90 ❄	0.20 ❄	83.37±1.46	83.92±2.01	81.60±1.40	82.29±1.47	92.18±1.65	92.33±1.64
1	7	-0.90 🔥	0.20 🔥	83.51±1.25	84.15±1.64	81.66±1.33	82.40±1.32	92.33±1.47	92.47±1.48
1	7	-0.90 ❄	0.30 ❄	82.99±1.52	83.25±1.74	81.45±1.87	81.98±1.70	92.01±1.52	92.15±1.50
1	7	-0.90 🔥	0.30 🔥	81.62±1.47	81.77±2.19	80.25±1.19	80.64±1.34	90.76±1.81	90.98±1.74
1	7	-0.90 ❄	0.40 ❄	81.50±1.62	81.95±2.46	79.67±1.41	80.30±1.57	90.76±1.91	90.97±1.86
1	7	-0.90 🔥	0.40 🔥	81.59±1.74	81.90±2.32	79.81±1.63	80.43±1.76	90.73±1.84	90.94±1.79
1	7	-0.90 ❄	0.50 ❄	81.68±1.91	81.84±2.40	80.16±2.12	80.62±2.04	90.85±1.90	91.04±1.84
1	7	-0.90 🔥	0.50 🔥	81.71±1.71	82.05±2.40	79.99±1.72	80.56±1.76	90.82±1.98	91.01±1.93
1	7	-0.90 ❄	0.60 ❄	81.45±2.07	81.58±2.65	79.96±2.13	80.41±2.16	90.34±1.90	90.54±1.86
1	7	-0.90 🔥	0.60 🔥	81.52±1.84	81.68±2.44	80.00±1.89	80.47±1.90	90.34±1.84	90.55±1.79
1	7	-0.90 ❄	0.70 ❄	81.44±2.60	81.43±2.92	79.84±2.66	80.37±2.75	89.97±2.15	90.19±2.10
1	7	-0.90 🔥	0.70 🔥	81.02±1.46	81.49±2.75	79.19±1.13	79.80±1.29	90.21±2.15	90.38±2.09
1	7	-0.90 ❄	0.80 ❄	81.05±1.88	81.28±2.59	79.29±1.71	79.88±1.87	89.82±2.03	89.96±2.01
1	7	-0.90 🔥	0.80 🔥	81.41±2.61	81.44±2.87	79.75±2.70	80.31±2.77	89.82±2.10	90.00±2.05
1	7	-0.90 ❄	0.90 ❄	81.23±1.87	81.02±1.80	79.77±2.18	80.21±2.08	89.72±1.71	89.91±1.67
1	7	-0.90 🔥	0.90 🔥	81.59±2.37	81.53±2.33	80.01±2.69	80.52±2.60	89.95±1.92	90.17±1.87
1	7	-0.90 ❄	1.00 ❄	81.26±2.06	81.47±2.19	79.39±2.20	80.03±2.23	89.73±1.76	89.91±1.74
1	7	-0.90 🔥	1.00 🔥	80.78±1.60	80.71±1.59	79.13±1.81	79.65±1.76	89.51±1.56	89.63±1.48
1	7	-0.80 ❄	-1.00 ❄	81.24±2.14	81.34±2.24	79.59±2.40	80.11±2.36	89.28±1.74	89.65±1.80
1	7	-0.80 🔥	-1.00 🔥	81.24±1.98	81.67±2.43	79.34±2.05	79.98±2.11	89.23±1.77	89.63±1.83
1	7	-0.80 ❄	-0.90 ❄	81.44±2.14	81.48±2.46	79.88±2.12	80.38±2.22	89.56±1.82	89.93±1.83
1	7	-0.80 🔥	-0.90 🔥	81.12±1.93	81.32±2.23	79.44±1.93	79.97±2.01	89.23±1.52	89.66±1.56
1	7	-0.80 ❄	-0.80 ❄	81.22±1.66	81.38±2.25	79.53±1.40	80.09±1.59	89.32±1.30	89.77±1.30
1	7	-0.80 🔥	-0.80 🔥	80.96±1.29	81.22±2.05	79.24±1.08	79.79±1.18	88.96±1.01	89.44±1.05
1	7	-0.80 ❄	-0.70 ❄	81.73±1.63	82.03±2.08	80.00±1.63	80.59±1.66	89.71±1.26	90.08±1.25
1	7	-0.80 🔥	-0.70 🔥	81.93±1.59	82.45±2.15	80.08±1.61	80.73±1.64	89.96±1.34	90.30±1.33
1	7	-0.80 ❄	-0.60 ❄	81.73±1.42	81.72±1.72	80.22±1.34	80.71±1.42	89.55±1.25	89.95±1.21
1	7	-0.80 🔥	-0.60 🔥	81.52±1.45	81.87±2.22	79.83±1.16	80.39±1.31	89.60±1.05	90.00±1.04
1	7	-0.80 ❄	-0.50 ❄	81.40±1.46	81.51±1.70	79.76±1.51	80.29±1.52	89.54±1.56	89.88±1.47
1	7	-0.80 🔥	-0.50 🔥	80.99±1.82	81.09±1.84	79.30±2.07	79.83±2.03	89.14±1.94	89.44±1.89
1	7	-0.80 ❄	-0.40 ❄	81.09±1.65	81.21±1.75	79.40±1.80	79.94±1.79	89.26±1.86	89.49±1.78

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	7	-0.80 🔥	-0.40 🔥	80.98±1.50	81.04±1.59	79.32±1.68	79.84±1.65	89.21±1.65	89.48±1.66
1	7	-0.80 ❄	-0.30 ❄	82.03±1.41	82.48±2.02	80.30±1.15	80.90±1.31	90.40±1.25	90.68±1.21
1	7	-0.80 🔥	-0.30 🔥	82.08±1.29	82.60±1.95	80.30±0.89	80.93±1.12	90.39±1.17	90.69±1.14
1	7	-0.80 ❄	-0.20 ❄	81.93±1.35	82.02±1.85	80.40±0.99	80.91±1.21	90.31±1.25	90.60±1.27
1	7	-0.80 🔥	-0.20 🔥	81.33±1.84	81.22±2.32	79.94±1.59	80.36±1.77	89.96±1.68	90.25±1.72
1	7	-0.80 ❄	-0.10 ❄	82.59±1.87	82.58±2.35	81.24±1.57	81.68±1.79	91.01±1.85	91.20±1.87
1	7	-0.80 🔥	-0.10 🔥	81.73±1.76	81.47±2.19	80.55±1.54	80.88±1.72	90.39±1.72	90.60±1.76
1	7	-0.80 ❄	0.10 ❄	82.89±1.77	83.56±2.02	81.01±1.90	81.72±1.91	91.67±1.74	91.83±1.74
1	7	-0.80 🔥	0.10 🔥	82.25±1.79	82.78±2.06	80.35±1.89	81.04±1.91	91.08±1.69	91.28±1.69
1	7	-0.80 ❄	0.20 ❄	82.03±2.50	82.65±2.68	80.14±2.73	80.78±2.76	90.83±2.46	91.04±2.40
1	7	-0.80 🔥	0.20 🔥	82.70±1.52	83.16±2.03	80.95±1.45	81.59±1.52	91.58±1.60	91.73±1.62
1	7	-0.80 ❄	0.30 ❄	82.07±1.68	82.51±2.06	80.25±1.76	80.89±1.80	90.43±2.29	90.65±2.22
1	7	-0.80 🔥	0.30 🔥	81.27±1.65	81.43±2.11	79.82±1.66	80.23±1.70	90.11±2.22	90.33±2.14
1	7	-0.80 ❄	0.40 ❄	81.23±1.54	81.37±2.17	79.65±1.41	80.15±1.51	90.19±1.80	90.43±1.75
1	7	-0.80 🔥	0.40 🔥	82.29±1.25	82.70±1.85	80.50±1.06	81.16±1.18	90.92±1.67	91.11±1.64
1	7	-0.80 ❄	0.50 ❄	80.92±1.92	81.08±2.68	79.51±1.75	79.89±1.85	89.95±2.06	90.15±2.03
1	7	-0.80 🔥	0.50 🔥	80.94±1.93	80.81±2.53	79.72±1.83	80.01±1.91	89.87±2.00	90.07±1.99
1	7	-0.80 ❄	0.60 ❄	80.92±1.66	80.55±1.65	79.75±1.98	80.02±1.85	89.67±1.66	89.82±1.68
1	7	-0.80 🔥	0.60 🔥	80.99±1.90	80.75±2.03	79.70±2.09	80.04±2.04	89.65±1.59	89.87±1.57
1	8	-1.00 ❄	-1.00 ❄	84.12±1.05	84.18±1.59	82.90±0.95	83.32±1.01	92.68±0.93	92.78±0.89
1	8	-1.00 🔥	-1.00 🔥	84.91±0.96	85.15±1.53	83.57±0.62	84.09±0.85	93.08±0.98	93.16±0.87
1	8	-1.00 ❄	-0.90 ❄	84.32±1.62	84.10±1.90	83.36±1.39	83.64±1.57	92.63±1.28	92.79±1.21
1	8	-1.00 🔥	-0.90 🔥	84.89±1.03	85.03±1.45	83.63±0.85	84.11±0.98	92.88±1.08	93.06±1.05
1	8	-1.00 ❄	-0.80 ❄	85.27±0.85	85.49±1.00	84.10±1.06	84.51±0.90	93.30±0.76	93.44±0.74
1	8	-1.00 🔥	-0.80 🔥	85.10±0.87	85.61±1.22	83.72±1.26	84.24±1.03	93.38±0.90	93.52±0.86
1	8	-1.00 ❄	-0.70 ❄	84.91±0.91	85.51±1.65	83.42±0.99	84.00±0.91	93.67±1.08	93.80±1.07
1	8	-1.00 🔥	-0.70 🔥	84.56±0.64	85.52±1.80	82.89±0.74	83.55±0.52	93.70±1.03	93.81±1.03
1	8	-1.00 ❄	-0.60 ❄	84.40±0.69	85.27±1.55	82.70±0.97	83.37±0.77	93.56±1.25	93.70±1.19
1	8	-1.00 🔥	-0.60 🔥	84.98±0.66	85.65±1.36	83.45±1.01	84.05±0.77	93.82±1.08	93.95±1.03
1	8	-1.00 ❄	-0.50 ❄	85.21±0.92	85.91±1.28	83.61±1.16	84.28±1.02	93.70±1.11	93.83±1.07
1	8	-1.00 🔥	-0.50 🔥	85.34±0.93	85.35±0.76	84.18±1.28	84.60±1.09	93.49±0.94	93.63±0.90
1	8	-1.00 ❄	-0.40 ❄	84.98±0.90	85.77±1.28	83.22±1.02	83.97±0.97	93.80±0.85	93.92±0.84
1	8	-1.00 🔥	-0.40 🔥	84.47±1.51	85.36±1.72	82.54±1.64	83.37±1.65	93.27±1.55	93.45±1.46
1	8	-1.00 ❄	-0.30 ❄	84.25±1.35	85.00±1.36	82.38±1.61	83.16±1.53	93.02±1.34	93.17±1.26
1	8	-1.00 🔥	-0.30 🔥	84.56±0.86	85.30±0.89	82.75±1.15	83.51±1.02	93.59±0.72	93.72±0.69
1	8	-1.00 ❄	-0.20 ❄	83.35±2.14	83.95±2.93	81.69±1.57	82.34±1.96	92.57±2.06	92.73±1.95
1	8	-1.00 🔥	-0.20 🔥	83.87±1.29	84.91±1.72	81.79±1.22	82.67±1.33	93.04±1.41	93.18±1.31
1	8	-1.00 ❄	-0.10 ❄	83.47±2.14	83.80±2.53	81.93±2.24	82.49±2.21	92.22±1.91	92.39±1.80
1	8	-1.00 🔥	-0.10 🔥	83.58±2.04	83.92±2.43	81.99±2.02	82.60±2.09	92.37±1.95	92.53±1.84
1	8	-1.00 ❄	0.10 ❄	83.59±1.46	83.74±1.51	82.09±1.67	82.65±1.60	92.21±1.25	92.36±1.22
1	8	-1.00 🔥	0.10 🔥	82.91±2.16	83.13±2.21	81.23±2.38	81.85±2.36	91.17±2.32	91.37±2.17
1	8	-1.00 ❄	0.20 ❄	81.31±2.64	81.25±2.71	79.83±2.89	80.27±2.87	90.09±2.07	90.27±2.07
1	8	-1.00 🔥	0.20 🔥	80.77±2.03	80.77±2.39	79.25±2.13	79.69±2.16	89.65±2.12	89.73±2.13
1	8	-1.00 ❄	0.30 ❄	81.29±2.45	81.41±2.65	79.66±2.56	80.17±2.60	89.46±2.36	89.71±2.24
1	8	-1.00 🔥	0.30 🔥	81.30±2.50	81.36±2.75	79.72±2.57	80.21±2.63	89.43±2.52	89.69±2.40
1	8	-1.00 ❄	0.40 ❄	80.31±2.06	80.88±2.68	78.15±1.95	78.87±2.09	89.10±2.26	89.35±2.17
1	8	-1.00 🔥	0.40 🔥	80.17±2.20	80.59±2.73	78.06±2.13	78.76±2.28	88.60±2.62	88.90±2.50
1	8	-1.00 ❄	0.50 ❄	79.04±2.24	79.87±3.14	76.60±2.07	77.34±2.24	87.44±2.78	87.79±2.67
1	8	-1.00 🔥	0.50 🔥	79.40±2.11	79.94±2.83	77.16±1.89	77.86±2.08	88.00±2.29	88.35±2.21
1	8	-1.00 ❄	0.60 ❄	78.32±2.40	78.88±2.95	76.20±2.09	76.76±2.29	86.79±2.20	87.16±2.11
1	8	-1.00 🔥	0.60 🔥	78.53±2.37	79.03±2.94	76.44±2.10	77.01±2.29	87.11±2.10	87.48±2.06
1	8	-1.00 ❄	0.70 ❄	78.16±2.00	78.58±2.40	75.89±1.95	76.51±2.07	86.46±2.07	86.89±2.01
1	8	-1.00 🔥	0.70 🔥	77.83±1.94	78.63±2.46	75.29±1.89	75.96±2.01	85.92±2.25	86.34±2.19

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	8	-1.00 ❄	0.80 ❄	77.23±1.36	77.84±1.99	74.72±1.32	75.35±1.38	85.34±1.85	85.69±1.77
1	8	-1.00 🔥	0.80 🔥	77.23±1.63	78.03±2.45	74.63±1.34	75.29±1.49	85.38±2.03	85.82±2.03
1	8	-1.00 ❄	0.90 ❄	76.79±1.74	77.49±2.37	74.23±1.77	74.82±1.83	84.53±2.01	85.01±2.00
1	8	-1.00 🔥	0.90 🔥	76.52±1.27	76.86±1.96	74.11±1.03	74.68±1.13	84.70±1.39	85.11±1.30
1	8	-1.00 ❄	1.00 ❄	77.04±1.46	77.68±2.25	74.54±0.97	75.16±1.16	84.76±1.19	85.27±1.21
1	8	-1.00 🔥	1.00 🔥	76.90±1.48	77.73±2.41	74.32±0.83	74.94±1.05	84.52±1.16	85.02±1.15
1	8	-0.90 ❄	-1.00 ❄	83.35±1.73	83.28±1.97	82.35±1.80	82.59±1.79	92.15±0.89	92.26±0.86
1	8	-0.90 🔥	-1.00 🔥	83.07±1.89	82.95±2.03	82.08±2.08	82.30±2.01	91.80±1.12	91.91±1.09
1	8	-0.90 ❄	-0.90 ❄	83.58±1.84	83.36±2.19	82.69±1.73	82.88±1.84	92.20±1.08	92.31±1.04
1	8	-0.90 🔥	-0.90 🔥	84.43±1.40	84.28±1.65	83.42±1.39	83.72±1.43	92.62±0.87	92.67±0.85
1	8	-0.90 ❄	-0.80 ❄	83.86±2.07	83.70±2.40	82.92±1.89	83.16±2.04	92.31±1.21	92.44±1.21
1	8	-0.90 🔥	-0.80 🔥	83.70±1.81	83.76±2.48	82.63±1.49	82.94±1.69	92.52±1.24	92.64±1.27
1	8	-0.90 ❄	-0.70 ❄	84.36±0.98	84.79±1.91	82.93±0.80	83.47±0.86	92.78±1.16	92.93±1.16
1	8	-0.90 🔥	-0.70 🔥	83.97±1.05	84.46±2.13	82.56±0.81	83.06±0.88	92.77±1.07	92.92±1.07
1	8	-0.90 ❄	-0.60 ❄	84.74±1.29	84.95±1.72	83.49±1.01	83.94±1.21	93.28±0.90	93.38±0.89
1	8	-0.90 🔥	-0.60 🔥	84.58±1.37	85.19±2.20	83.14±0.87	83.70±1.18	93.33±1.18	93.43±1.16
1	8	-0.90 ❄	-0.50 ❄	84.64±1.22	84.72±1.40	83.44±1.35	83.86±1.29	93.01±1.12	93.17±1.07
1	8	-0.90 🔥	-0.50 🔥	84.72±1.08	85.00±1.39	83.29±1.05	83.86±1.10	93.01±1.17	93.17±1.12
1	8	-0.90 ❄	-0.40 ❄	84.75±1.32	84.76±1.42	83.51±1.45	83.97±1.41	93.01±1.02	93.14±1.02
1	8	-0.90 🔥	-0.40 🔥	84.57±0.89	84.90±1.15	83.05±0.96	83.66±0.94	93.08±0.92	93.20±0.93
1	8	-0.90 ❄	-0.30 ❄	84.92±1.28	85.22±1.41	83.47±1.46	84.05±1.40	93.35±1.01	93.45±1.02
1	8	-0.90 🔥	-0.30 🔥	84.25±1.82	84.32±2.00	82.91±1.89	83.41±1.90	92.82±1.33	92.92±1.37
1	8	-0.90 ❄	-0.20 ❄	84.44±1.35	84.82±1.20	82.87±1.72	83.49±1.56	92.75±1.31	92.89±1.26
1	8	-0.90 🔥	-0.20 🔥	83.97±1.12	84.59±1.12	82.14±1.29	82.89±1.26	92.55±1.22	92.71±1.14
1	8	-0.90 ❄	-0.10 ❄	83.90±0.94	84.18±0.92	82.35±1.26	82.94±1.10	92.53±0.92	92.69±0.88
1	8	-0.90 🔥	-0.10 🔥	83.61±0.93	84.06±1.27	81.91±1.05	82.57±1.00	92.44±1.04	92.60±1.03
1	8	-0.90 ❄	0.10 ❄	82.14±2.78	82.44±3.33	80.65±2.35	81.15±2.65	91.02±2.35	91.26±2.25
1	8	-0.90 🔥	0.10 🔥	82.38±1.98	82.59±2.42	80.78±1.74	81.35±1.94	91.09±2.04	91.31±1.97
1	8	-0.90 ❄	0.20 ❄	82.18±2.34	82.46±2.89	80.66±2.11	81.16±2.29	90.44±2.32	90.70±2.24
1	8	-0.90 🔥	0.20 🔥	81.24±2.99	81.32±3.43	79.81±2.77	80.24±2.98	89.57±2.76	89.77±2.69
1	8	-0.90 ❄	0.30 ❄	80.75±2.19	80.96±2.76	79.06±2.09	79.59±2.20	88.97±2.63	89.23±2.54
1	8	-0.90 🔥	0.30 🔥	80.66±2.40	80.85±2.84	78.97±2.38	79.49±2.46	89.08±2.65	89.35±2.56
1	8	-0.90 ❄	0.40 ❄	79.64±2.51	79.80±3.08	77.79±2.29	78.35±2.48	88.31±2.32	88.62±2.25
1	8	-0.90 🔥	0.40 🔥	79.76±2.79	79.68±3.28	78.15±2.63	78.62±2.80	88.16±2.37	88.48±2.32
1	8	-0.90 ❄	0.50 ❄	78.55±2.61	78.46±3.01	76.71±2.53	77.22±2.67	86.46±2.46	86.85±2.44
1	8	-0.90 🔥	0.50 🔥	78.46±2.48	78.51±2.87	76.57±2.47	77.08±2.57	86.58±2.23	86.96±2.21
1	8	-0.90 ❄	0.60 ❄	77.90±1.95	77.77±2.17	76.02±2.07	76.51±2.11	85.68±2.56	86.09±2.56
1	8	-0.90 🔥	0.60 🔥	78.36±1.87	78.18±2.15	76.55±1.86	77.05±1.93	86.32±1.81	86.69±1.86
1	8	-0.90 ❄	0.70 ❄	78.03±1.81	78.65±2.39	75.67±1.81	76.29±1.89	85.96±2.14	86.36±2.11
1	8	-0.90 🔥	0.70 🔥	77.76±1.79	78.52±2.30	75.11±1.76	75.83±1.87	85.64±2.09	86.08±2.07
1	8	-0.90 ❄	0.80 ❄	77.64±2.23	77.93±2.83	75.36±2.06	75.98±2.20	85.53±1.89	85.99±1.91
1	8	-0.90 🔥	0.80 🔥	77.41±1.95	78.00±2.28	74.82±2.10	75.49±2.15	85.18±2.41	85.53±2.47
1	8	-0.90 ❄	0.90 ❄	76.72±2.13	77.53±2.91	74.22±1.95	74.77±2.11	84.26±2.17	84.69±2.23
1	8	-0.90 🔥	0.90 🔥	76.70±1.98	77.20±2.64	74.36±1.82	74.89±1.95	84.27±2.06	84.68±2.15
1	8	-0.90 ❄	1.00 ❄	77.01±1.66	77.79±2.43	74.34±1.23	75.02±1.41	84.23±1.65	84.73±1.64
1	8	-0.90 🔥	1.00 🔥	77.09±1.36	77.87±2.13	74.47±1.04	75.12±1.18	84.27±1.44	84.70±1.52
1	8	-0.80 ❄	-1.00 ❄	83.51±1.91	83.37±1.61	82.43±2.47	82.70±2.24	91.74±1.31	91.81±1.33
1	8	-0.80 🔥	-1.00 🔥	83.33±1.98	83.22±1.60	82.31±2.64	82.51±2.38	91.59±1.46	91.68±1.46
1	8	-0.80 ❄	-0.90 ❄	83.44±1.35	83.84±2.08	82.11±1.44	82.52±1.37	92.12±0.96	92.23±1.04
1	8	-0.80 🔥	-0.90 🔥	83.72±1.48	83.55±1.62	82.97±1.74	83.05±1.58	92.21±0.94	92.26±0.98
1	8	-0.80 ❄	-0.80 ❄	83.86±1.36	84.20±2.17	82.59±1.33	83.00±1.30	92.50±1.00	92.58±1.04
1	8	-0.80 🔥	-0.80 🔥	83.87±1.40	83.95±1.92	82.76±1.41	83.08±1.38	92.48±1.02	92.57±1.05
1	8	-0.80 ❄	-0.70 ❄	83.89±1.05	84.03±2.00	82.76±0.86	83.10±0.91	92.50±1.13	92.62±1.15

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	8	-0.80 🔥	-0.70 🔥	83.72±1.05	83.94±2.00	82.58±0.99	82.91±0.93	92.52±1.11	92.63±1.14
1	8	-0.80 ❄	-0.60 ❄	84.30±1.36	84.49±2.21	83.14±0.97	83.53±1.21	92.69±1.25	92.78±1.28
1	8	-0.80 🔥	-0.60 🔥	84.29±1.12	84.50±1.95	83.16±0.81	83.52±0.95	92.73±1.18	92.81±1.21
1	8	-0.80 ❄	-0.50 ❄	84.33±1.52	84.47±2.04	83.07±1.26	83.53±1.45	92.61±1.20	92.72±1.16
1	8	-0.80 🔥	-0.50 🔥	84.28±1.48	84.19±1.77	83.14±1.37	83.52±1.48	92.58±1.10	92.71±1.06
1	8	-0.80 ❄	-0.40 ❄	84.01±1.26	83.98±1.67	82.81±1.01	83.22±1.19	92.53±0.98	92.69±0.96
1	8	-0.80 🔥	-0.40 🔥	83.86±1.77	83.75±2.12	82.70±1.70	83.08±1.79	92.17±1.43	92.34±1.40
1	8	-0.80 ❄	-0.30 ❄	83.84±1.98	83.97±2.53	82.59±2.03	83.01±2.02	92.42±1.58	92.54±1.61
1	8	-0.80 🔥	-0.30 🔥	83.98±1.79	84.35±1.64	82.36±2.16	82.98±2.04	92.34±1.49	92.50±1.45
1	8	-0.80 ❄	-0.20 ❄	83.82±1.36	83.99±1.36	82.28±1.54	82.87±1.49	92.19±1.17	92.35±1.16
1	8	-0.80 🔥	-0.20 🔥	83.40±1.55	83.71±1.40	81.74±1.88	82.36±1.78	91.81±1.25	91.99±1.18
1	8	-0.80 ❄	-0.10 ❄	83.61±1.09	83.97±1.26	81.94±1.17	82.59±1.17	92.05±1.05	92.19±1.09
1	8	-0.80 🔥	-0.10 🔥	83.49±1.10	83.75±1.18	81.86±1.18	82.49±1.18	92.00±1.16	92.16±1.18
1	8	-0.80 ❄	0.10 ❄	80.88±2.92	81.02±3.31	79.41±2.37	79.85±2.74	89.83±2.36	90.14±2.27
1	8	-0.80 🔥	0.10 🔥	81.72±3.09	81.92±3.52	80.36±2.59	80.77±2.93	90.46±2.33	90.72±2.27
1	8	-0.80 ❄	0.20 ❄	81.36±1.99	81.66±2.51	79.67±1.84	80.22±1.97	89.72±2.24	89.99±2.21
1	8	-0.80 🔥	0.20 🔥	81.06±2.13	81.59±2.58	79.18±2.12	79.79±2.19	89.30±2.43	89.60±2.33
1	8	-0.80 ❄	0.30 ❄	79.82±2.62	80.09±3.48	78.20±2.10	78.66±2.42	88.38±2.57	88.69±2.51
1	8	-0.80 🔥	0.30 🔥	80.84±2.17	81.23±2.76	79.13±1.84	79.66±2.06	89.12±2.10	89.41±2.09
1	8	-0.80 ❄	0.40 ❄	79.54±2.09	79.91±2.97	77.54±1.65	78.17±1.91	87.71±2.30	88.05±2.27
1	8	-0.80 🔥	0.40 🔥	79.50±2.29	79.84±3.09	77.50±1.94	78.12±2.17	87.79±2.39	88.12±2.34
1	8	-0.80 ❄	0.50 ❄	78.56±2.26	78.97±3.48	76.50±1.65	77.11±1.97	86.58±2.73	86.88±2.66
1	8	-0.80 🔥	0.50 🔥	78.18±2.26	78.50±3.41	76.17±1.67	76.74±1.97	86.59±2.73	86.97±2.71
1	8	-0.80 ❄	0.60 ❄	77.86±1.91	77.81±2.33	75.84±1.89	76.39±1.99	85.38±2.65	85.79±2.67
1	8	-0.80 🔥	0.60 🔥	78.25±1.87	78.47±2.91	76.18±1.47	76.77±1.69	86.13±2.27	86.51±2.29
1	9	-1.00 ❄	-1.00 ❄	86.16±1.34	85.62±1.38	86.00±1.37	85.77±1.37	94.32±1.02	94.47±0.99
1	9	-1.00 🔥	-1.00 🔥	86.07±1.14	85.54±1.21	86.20±1.51	85.73±1.24	94.53±1.28	94.66±1.22
1	9	-1.00 ❄	-0.90 ❄	85.48±2.00	85.21±1.97	85.39±1.57	85.08±1.92	94.20±1.28	94.35±1.23
1	9	-1.00 🔥	-0.90 🔥	85.70±2.10	85.35±2.03	85.65±1.80	85.32±2.06	94.25±1.28	94.40±1.24
1	9	-1.00 ❄	-0.80 ❄	85.81±1.73	85.47±1.72	85.78±1.38	85.44±1.67	94.35±1.09	94.48±1.05
1	9	-1.00 🔥	-0.80 🔥	85.37±2.62	85.34±2.22	85.34±1.87	84.97±2.49	94.12±1.18	94.27±1.15
1	9	-1.00 ❄	-0.70 ❄	82.91±4.09	83.81±3.21	83.49±2.51	82.57±3.81	93.75±1.11	93.92±1.09
1	9	-1.00 🔥	-0.70 🔥	84.01±3.68	84.45±2.76	84.01±2.30	83.57±3.44	93.89±1.01	94.04±0.99
1	9	-1.00 ❄	-0.60 ❄	85.31±1.43	84.94±1.55	85.29±1.13	84.93±1.35	94.10±0.93	94.26±0.87
1	9	-1.00 🔥	-0.60 🔥	86.26±0.76	85.84±0.76	86.17±0.49	85.88±0.67	94.49±0.44	94.62±0.42
1	9	-1.00 ❄	-0.50 ❄	86.54±1.55	86.33±1.88	86.10±1.11	86.08±1.46	94.43±0.70	94.57±0.67
1	9	-1.00 🔥	-0.50 🔥	86.39±1.06	86.04±1.20	86.14±0.66	85.97±0.96	94.53±0.50	94.66±0.49
1	9	-1.00 ❄	-0.40 ❄	85.74±1.00	85.57±1.09	85.18±1.26	85.20±1.07	94.29±0.59	94.43±0.57
1	9	-1.00 🔥	-0.40 🔥	85.77±1.17	85.60±1.22	85.19±1.35	85.23±1.24	94.19±0.58	94.34±0.54
1	9	-1.00 ❄	-0.30 ❄	85.56±1.36	85.57±1.26	84.83±1.99	84.93±1.62	93.99±0.82	94.14±0.78
1	9	-1.00 🔥	-0.30 🔥	85.34±1.66	85.45±1.64	84.50±2.06	84.67±1.87	93.87±0.82	94.03±0.76
1	9	-1.00 ❄	-0.20 ❄	85.31±0.85	85.05±1.04	84.51±0.82	84.70±0.84	93.58±0.71	93.75±0.67
1	9	-1.00 🔥	-0.20 🔥	85.52±1.39	85.49±1.78	84.50±1.14	84.85±1.34	93.65±0.88	93.81±0.82
1	9	-1.00 ❄	-0.10 ❄	84.65±0.96	84.43±1.24	83.79±0.92	84.00±0.94	93.34±0.74	93.51±0.72
1	9	-1.00 🔥	-0.10 🔥	85.03±1.42	85.30±2.07	83.69±1.03	84.23±1.32	93.62±0.93	93.77±0.88
1	9	-1.00 ❄	0.10 ❄	83.35±0.99	84.57±1.70	81.21±0.99	82.07±1.03	92.63±1.25	92.79±1.15
1	9	-1.00 🔥	0.10 🔥	83.65±1.32	84.36±1.72	81.81±1.37	82.54±1.38	92.57±1.34	92.74±1.25
1	9	-1.00 ❄	0.20 ❄	83.90±2.14	84.40±2.06	82.13±2.45	82.83±2.39	92.65±1.78	92.81±1.66
1	9	-1.00 🔥	0.20 🔥	83.41±1.93	83.88±1.94	81.62±2.15	82.32±2.14	92.21±1.67	92.38±1.52
1	9	-1.00 ❄	0.30 ❄	83.16±2.07	83.70±1.92	81.31±2.47	82.01±2.38	92.20±1.86	92.34±1.73
1	9	-1.00 🔥	0.30 🔥	83.10±1.87	83.90±1.95	81.06±2.07	81.87±2.08	92.02±1.84	92.18±1.72
1	9	-1.00 ❄	0.40 ❄	82.84±1.76	83.51±1.33	80.89±2.31	81.60±2.16	91.59±1.89	91.73±1.73
1	9	-1.00 🔥	0.40 🔥	81.84±1.34	82.72±1.52	79.61±1.45	80.44±1.49	90.87±1.86	91.04±1.66

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	9	-1.00 ❄	0.50 ❄	81.91±2.08	81.99±2.27	80.27±2.18	80.84±2.20	91.05±2.02	91.20±1.78
1	9	-1.00 🔥	0.50 🔥	81.50±1.66	81.65±1.87	79.77±1.79	80.35±1.77	90.67±1.72	90.79±1.47
1	9	-1.00 ❄	0.60 ❄	81.38±1.09	81.89±1.47	79.36±1.06	80.08±1.11	90.67±1.42	90.73±1.33
1	9	-1.00 🔥	0.60 🔥	81.44±0.93	82.07±1.00	79.32±1.10	80.08±1.07	90.78±1.29	90.84±1.22
1	9	-1.00 ❄	0.70 ❄	81.09±0.90	82.04±0.97	78.79±1.28	79.57±1.15	90.33±1.26	90.35±1.10
1	9	-1.00 🔥	0.70 🔥	81.23±0.71	81.92±0.61	79.07±1.05	79.82±0.95	90.42±1.15	90.46±1.03
1	9	-1.00 ❄	0.80 ❄	80.52±0.98	81.47±1.32	78.10±1.03	78.92±1.06	89.87±1.24	89.80±1.24
1	9	-1.00 🔥	0.80 🔥	80.17±1.14	81.05±1.37	77.76±1.33	78.54±1.32	89.68±1.13	89.65±1.09
1	9	-1.00 ❄	0.90 ❄	79.92±0.64	80.92±0.83	77.36±0.64	78.19±0.68	89.39±1.07	89.30±1.05
1	9	-1.00 🔥	0.90 🔥	79.99±0.84	80.70±1.14	77.66±0.97	78.41±0.95	89.50±1.11	89.37±1.07
1	9	-1.00 ❄	1.00 ❄	79.27±0.64	79.94±1.08	76.90±0.67	77.62±0.66	88.76±1.02	88.54±1.14
1	9	-1.00 🔥	1.00 🔥	78.91±1.49	80.03±1.67	76.16±1.62	76.97±1.69	88.21±2.36	88.14±2.21
1	9	-0.90 ❄	-1.00 ❄	85.58±2.16	85.25±1.89	85.67±1.94	85.22±2.14	93.97±1.34	94.13±1.32
1	9	-0.90 🔥	-1.00 🔥	85.76±2.20	85.40±2.06	85.78±2.04	85.39±2.20	94.03±1.44	94.18±1.41
1	9	-0.90 ❄	-0.90 ❄	85.81±1.11	85.23±1.13	85.83±1.19	85.46±1.15	94.12±0.98	94.29±0.94
1	9	-0.90 🔥	-0.90 🔥	85.65±1.63	85.48±1.80	85.39±1.02	85.20±1.49	94.14±0.90	94.29±0.87
1	9	-0.90 ❄	-0.80 ❄	85.39±2.84	85.26±2.48	85.48±1.93	85.04±2.68	94.04±1.05	94.18±1.05
1	9	-0.90 🔥	-0.80 🔥	85.26±3.24	85.15±2.81	85.33±2.33	84.90±3.09	94.01±1.18	94.15±1.18
1	9	-0.90 ❄	-0.70 ❄	85.23±2.83	85.34±2.47	85.06±2.08	84.78±2.67	94.23±1.11	94.38±1.07
1	9	-0.90 🔥	-0.70 🔥	85.23±2.88	85.30±2.44	85.09±2.07	84.79±2.71	94.23±1.16	94.38±1.14
1	9	-0.90 ❄	-0.60 ❄	84.00±3.17	84.01±2.61	84.68±2.39	83.76±3.04	93.89±1.28	94.04±1.22
1	9	-0.90 🔥	-0.60 🔥	85.74±2.99	85.56±2.43	85.91±2.20	85.42±2.86	94.44±1.14	94.58±1.10
1	9	-0.90 ❄	-0.50 ❄	86.01±0.83	85.89±1.17	85.53±0.97	85.50±0.82	94.34±0.45	94.49±0.42
1	9	-0.90 🔥	-0.50 🔥	84.85±2.47	85.08±2.36	84.56±1.95	84.34±2.34	94.14±0.89	94.29±0.85
1	9	-0.90 ❄	-0.40 ❄	85.45±1.78	85.68±2.13	84.67±2.06	84.80±1.87	93.88±1.21	94.04±1.16
1	9	-0.90 🔥	-0.40 🔥	85.46±1.62	85.56±1.93	84.80±1.90	84.86±1.70	93.92±1.26	94.08±1.22
1	9	-0.90 ❄	-0.30 ❄	85.59±1.58	85.46±1.68	85.02±1.71	85.04±1.63	93.95±0.86	94.11±0.81
1	9	-0.90 🔥	-0.30 🔥	85.14±1.96	84.89±2.06	84.51±2.26	84.56±2.08	93.45±1.52	93.62±1.45
1	9	-0.90 ❄	-0.20 ❄	85.00±1.14	85.29±1.28	84.03±1.91	84.25±1.49	93.49±0.63	93.66±0.59
1	9	-0.90 🔥	-0.20 🔥	84.40±2.40	84.20±2.47	83.33±2.64	83.66±2.57	92.83±1.83	93.07±1.77
1	9	-0.90 ❄	-0.10 ❄	84.89±0.93	85.20±1.54	83.66±1.03	84.09±0.95	93.41±0.79	93.58±0.75
1	9	-0.90 🔥	-0.10 🔥	84.79±1.06	84.88±1.53	83.67±0.82	84.05±0.97	93.44±0.72	93.60±0.69
1	9	-0.90 ❄	0.10 ❄	84.21±1.85	84.73±2.37	82.51±1.67	83.22±1.85	92.75±1.64	92.94±1.55
1	9	-0.90 🔥	0.10 🔥	84.04±1.91	84.66±2.38	82.26±1.84	83.00±1.96	92.82±1.64	93.01±1.55
1	9	-0.90 ❄	0.20 ❄	83.82±1.99	84.27±2.17	82.10±2.16	82.78±2.16	92.62±1.74	92.80±1.64
1	9	-0.90 🔥	0.20 🔥	83.93±2.12	84.38±2.25	82.22±2.30	82.90±2.30	92.63±1.80	92.82±1.69
1	9	-0.90 ❄	0.30 ❄	83.06±1.92	83.52±1.95	81.26±2.20	81.94±2.16	92.13±1.86	92.29±1.73
1	9	-0.90 🔥	0.30 🔥	83.91±1.34	84.43±1.38	82.20±1.62	82.87±1.52	92.59±1.48	92.71±1.35
1	9	-0.90 ❄	0.40 ❄	82.99±1.07	83.76±1.19	81.01±1.36	81.77±1.24	91.72±1.28	91.82±1.25
1	9	-0.90 🔥	0.40 🔥	83.45±1.14	84.40±1.42	81.39±1.19	82.23±1.21	92.28±1.39	92.41±1.32
1	9	-0.90 ❄	0.50 ❄	82.73±1.34	83.39±1.40	80.72±1.49	81.49±1.48	91.71±1.39	91.77±1.37
1	9	-0.90 🔥	0.50 🔥	82.71±1.31	83.34±1.55	80.75±1.39	81.51±1.40	91.80±1.40	91.85±1.34
1	9	-0.90 ❄	0.60 ❄	82.03±1.18	82.84±1.20	79.86±1.37	80.66±1.34	91.20±1.10	91.20±1.06
1	9	-0.90 🔥	0.60 🔥	81.97±1.26	82.81±1.29	79.79±1.48	80.59±1.45	91.15±1.22	91.12±1.18
1	9	-0.90 ❄	0.70 ❄	81.58±1.13	82.53±1.21	79.32±1.44	80.12±1.35	90.72±1.24	90.70±1.19
1	9	-0.90 🔥	0.70 🔥	80.92±0.64	81.94±0.93	78.50±0.63	79.35±0.66	90.35±1.14	90.30±1.10
1	9	-0.90 ❄	0.80 ❄	79.69±1.67	80.50±2.38	77.32±1.40	78.08±1.57	88.77±1.94	88.78±1.85
1	9	-0.90 🔥	0.80 🔥	80.03±1.27	80.91±1.87	77.60±1.08	78.40±1.21	89.27±1.66	89.27±1.60
1	9	-0.90 ❄	0.90 ❄	79.80±0.84	81.00±1.11	77.13±0.84	77.99±0.89	89.43±1.03	89.35±1.03
1	9	-0.90 🔥	0.90 🔥	79.82±0.94	81.20±1.20	77.06±0.97	77.94±1.02	89.43±1.15	89.24±1.17
1	9	-0.90 ❄	1.00 ❄	79.36±0.74	80.15±1.22	76.89±0.67	77.66±0.71	88.77±1.28	88.57±1.26
1	9	-0.90 🔥	1.00 🔥	79.23±0.91	80.23±0.95	76.59±1.04	77.40±1.07	88.79±1.17	88.64±1.18
1	9	-0.80 ❄	-1.00 ❄	85.37±1.64	85.90±0.60	84.34±2.90	84.54±2.30	93.82±1.30	94.00±1.21

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	9	-0.80 🔥	-1.00 🔥	84.46±1.74	85.26±1.23	83.67±2.36	83.67±2.05	93.72±1.05	93.92±1.00
1	9	-0.80 ❄	-0.90 ❄	85.65±1.36	86.08±0.56	84.55±2.41	84.86±1.86	93.96±0.79	94.13±0.76
1	9	-0.80 🔥	-0.90 🔥	85.95±1.70	86.43±0.92	84.76±2.68	85.14±2.19	94.10±0.89	94.27±0.86
1	9	-0.80 ❄	-0.80 ❄	85.95±1.20	86.05±0.99	85.26±1.92	85.34±1.51	94.28±0.77	**94.44±0.72**
1	9	-0.80 🔥	-0.80 🔥	86.00±1.20	86.01±1.14	85.32±1.74	85.40±1.42	94.20±0.76	94.36±0.72
1	9	-0.80 ❄	-0.70 ❄	85.19±2.09	85.57±2.04	84.32±1.74	84.50±2.02	94.10±0.86	94.26±0.83
1	9	-0.80 🔥	-0.70 🔥	85.41±2.39	85.68±2.21	84.70±2.05	84.79±2.33	93.98±0.91	94.14±0.87
1	9	-0.80 ❄	-0.60 ❄	86.23±1.17	**86.35±1.35**	85.14±1.24	85.55±1.22	94.09±0.81	94.27±0.75
1	9	-0.80 🔥	-0.60 🔥	**86.30±1.05**	86.16±1.09	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	94.42±0.49
1	9	-0.80 ❄	-0.50 ❄	85.67±2.74	85.40±2.42	85.61±2.03	85.30±2.60	94.18±1.24	94.34±1.20
1	9	-0.80 🔥	-0.50 🔥	85.67±1.39	85.34±1.59	85.20±1.02	85.18±1.30	94.07±0.66	94.24±0.63
1	9	-0.80 ❄	-0.40 ❄	86.08±1.04	85.69±1.09	85.51±1.17	85.57±1.09	93.97±0.94	94.14±0.89
1	9	-0.80 🔥	-0.40 🔥	86.18±0.77	85.81±0.86	85.74±0.91	85.70±0.79	94.14±0.70	94.30±0.65
1	9	-0.80 ❄	-0.30 ❄	85.49±1.89	85.45±2.18	84.68±2.02	84.87±1.94	93.73±1.45	93.87±1.37
1	9	-0.80 🔥	-0.30 🔥	85.02±0.76	84.90±0.99	84.06±0.84	84.34±0.76	93.43±0.63	93.63±0.60
1	9	-0.80 ❄	-0.20 ❄	84.84±2.03	85.20±2.56	83.38±1.87	83.98±2.02	93.22±1.64	93.44±1.57
1	9	-0.80 🔥	-0.20 🔥	84.28±1.67	84.97±2.20	82.55±1.54	83.26±1.65	93.04±1.47	93.26±1.41
1	9	-0.80 ❄	-0.10 ❄	84.07±1.50	84.56±1.98	82.37±1.38	83.07±1.51	92.70±1.15	92.96±1.05
1	9	-0.80 🔥	-0.10 🔥	83.83±2.37	84.58±3.02	81.99±2.32	82.75±2.44	92.33±1.96	92.60±1.88
1	9	-0.80 ❄	0.10 ❄	84.21±1.34	85.39±1.96	82.14±1.24	83.03±1.33	92.98±1.31	93.20±1.20
1	9	-0.80 🔥	0.10 🔥	84.29±1.48	85.49±1.95	82.23±1.52	83.11±1.57	92.95±1.33	93.17±1.23
1	9	-0.80 ❄	0.20 ❄	84.32±1.36	84.94±1.39	82.54±1.57	83.28±1.51	92.76±1.43	92.97±1.32
1	9	-0.80 🔥	0.20 🔥	83.83±1.93	84.13±2.52	82.35±1.71	82.91±1.88	92.37±1.75	92.57±1.66
1	9	-0.80 ❄	0.30 ❄	83.48±1.29	84.23±1.50	81.54±1.39	82.32±1.39	92.28±1.43	92.47±1.33
1	9	-0.80 🔥	0.30 🔥	83.44±1.09	84.45±1.17	81.33±1.28	82.19±1.23	92.35±1.32	92.53±1.23
1	9	-0.80 ❄	0.40 ❄	82.66±1.24	83.38±1.47	80.65±1.39	81.41±1.36	91.36±1.59	91.50±1.43
1	9	-0.80 🔥	0.40 🔥	83.05±1.19	83.81±1.18	81.04±1.42	81.82±1.36	91.86±1.36	91.94±1.24
1	9	-0.80 ❄	0.50 ❄	82.63±1.22	83.53±1.38	80.47±1.32	81.31±1.34	91.70±1.26	91.70±1.33
1	9	-0.80 🔥	0.50 🔥	82.35±1.09	83.13±1.13	80.25±1.30	81.04±1.25	91.58±1.31	91.57±1.33
1	9	-0.80 ❄	0.60 ❄	81.38±1.59	82.08±1.55	79.21±1.89	79.97±1.84	90.25±2.10	90.37±1.94
1	9	-0.80 🔥	0.60 🔥	81.75±0.95	82.79±1.28	79.46±1.20	80.29±1.12	90.97±1.32	91.01±1.29
1	10	-1.00 ❄	-1.00 ❄	82.39±1.78	82.06±1.69	81.98±1.15	81.83±1.61	91.79±0.89	91.99±0.84
1	10	-1.00 🔥	-1.00 🔥	82.10±1.92	81.76±1.63	81.89±1.15	81.60±1.73	91.81±0.77	92.01±0.74
1	10	-1.00 ❄	-0.90 ❄	83.10±2.42	83.35±2.48	82.35±1.46	82.41±2.15	92.36±1.11	92.55±1.08
1	10	-1.00 🔥	-0.90 🔥	82.92±1.43	83.61±1.93	81.32±1.96	81.82±1.76	91.78±1.19	92.04±1.07
1	10	-1.00 ❄	-0.80 ❄	84.00±1.41	84.69±0.61	82.39±2.37	82.93±2.02	92.73±1.12	92.93±1.00
1	10	-1.00 🔥	-0.80 🔥	83.13±1.40	83.55±1.35	81.69±2.11	82.12±1.85	91.99±0.97	92.24±0.85
1	10	-1.00 ❄	-0.70 ❄	83.02±1.22	83.67±1.74	81.51±1.96	81.96±1.56	91.92±1.09	92.20±0.97
1	10	-1.00 🔥	-0.70 🔥	83.93±1.30	84.13±1.04	82.68±2.01	83.05±1.67	92.64±1.28	92.86±1.15
1	10	-1.00 ❄	-0.60 ❄	83.09±1.47	83.24±1.73	81.98±1.92	82.24±1.66	92.03±1.14	92.31±0.99
1	10	-1.00 🔥	-0.60 🔥	83.31±1.34	83.63±1.34	82.04±2.18	82.38±1.72	91.86±1.11	92.16±0.95
1	10	-1.00 ❄	-0.50 ❄	83.63±1.38	83.31±1.76	83.30±1.35	83.12±1.32	92.94±1.13	93.05±1.07
1	10	-1.00 🔥	-0.50 🔥	83.84±1.54	83.50±1.90	83.36±1.48	83.29±1.50	92.92±1.14	93.04±1.07
1	10	-1.00 ❄	-0.40 ❄	83.44±1.22	83.16±1.60	82.60±1.28	82.76±1.23	92.10±0.94	92.32±0.86
1	10	-1.00 🔥	-0.40 🔥	83.68±1.47	83.26±1.69	82.90±1.40	83.04±1.48	92.41±1.30	92.60±1.17
1	10	-1.00 ❄	-0.30 ❄	84.14±1.54	84.13±1.65	83.02±1.94	83.35±1.70	92.59±1.25	92.85±1.10
1	10	-1.00 🔥	-0.30 🔥	84.14±1.51	84.26±1.76	82.97±1.83	83.33±1.63	92.54±1.17	92.80±1.02
1	10	-1.00 ❄	-0.20 ❄	84.49±1.39	84.53±1.51	83.19±1.53	83.67±1.48	92.77±1.18	92.96±0.97
1	10	-1.00 🔥	-0.20 🔥	84.99±1.11	84.68±1.17	84.10±1.17	84.35±1.17	93.07±1.01	93.22±0.88
1	10	-1.00 ❄	-0.10 ❄	85.12±1.28	84.99±1.30	84.03±1.45	84.40±1.38	93.20±1.06	93.40±0.90
1	10	-1.00 🔥	-0.10 🔥	85.07±1.18	84.90±1.22	84.03±1.31	84.38±1.26	93.20±0.98	93.41±0.84
1	10	-1.00 ❄	0.10 ❄	85.20±1.27	85.52±1.67	83.79±1.30	84.36±1.31	93.83±0.98	93.95±0.84
1	10	-1.00 🔥	0.10 🔥	85.39±0.95	85.67±1.26	84.03±1.05	84.58±1.01	93.85±0.93	93.96±0.81

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	10	-1.00 ❄	0.20 ❄	84.61±2.08	84.76±2.42	83.28±2.06	83.79±2.13	93.05±1.55	93.15±1.51
1	10	-1.00 🔥	0.20 🔥	83.96±1.32	84.31±1.63	82.40±1.41	83.00±1.38	92.88±1.26	92.97±1.16
1	10	-1.00 ❄	0.30 ❄	82.95±1.59	84.10±1.90	80.95±2.19	81.65±2.02	92.40±1.45	92.51±1.41
1	10	-1.00 🔥	0.30 🔥	82.36±1.83	83.87±2.10	80.25±2.70	80.91±2.54	92.27±1.35	92.38±1.32
1	10	-1.00 ❄	0.40 ❄	82.33±1.92	84.13±2.36	79.84±2.27	80.75±2.23	91.95±1.82	92.11±1.74
1	10	-1.00 🔥	0.40 🔥	82.08±1.93	84.03±2.22	79.57±2.57	80.41±2.50	91.87±1.80	92.03±1.74
1	10	-1.00 ❄	0.50 ❄	81.20±2.56	83.46±2.24	78.47±3.40	79.26±3.51	91.14±2.18	91.32±2.13
1	10	-1.00 🔥	0.50 🔥	82.10±2.55	84.46±1.10	79.38±3.61	80.21±3.66	92.01±1.75	92.19±1.77
1	10	-1.00 ❄	0.60 ❄	82.56±2.79	84.15±3.28	80.11±2.92	81.05±3.03	91.53±3.05	91.68±2.95
1	10	-1.00 🔥	0.60 🔥	82.43±2.81	83.64±3.29	80.12±2.90	81.01±3.01	90.84±3.48	91.09±3.34
1	10	-1.00 ❄	0.70 ❄	82.85±1.77	84.55±1.60	80.32±2.11	81.30±2.12	91.51±2.11	91.75±2.01
1	10	-1.00 🔥	0.70 🔥	82.01±2.88	83.43±3.38	79.59±2.98	80.49±3.10	90.63±3.37	90.85±3.24
1	10	-1.00 ❄	0.80 ❄	82.50±1.48	84.59±1.19	79.75±1.79	80.80±1.81	90.85±2.07	91.15±1.98
1	10	-1.00 🔥	0.80 🔥	82.63±1.73	84.52±1.40	79.98±2.10	80.99±2.11	90.85±2.43	91.08±2.44
1	10	-1.00 ❄	0.90 ❄	81.82±2.02	83.59±2.26	79.11±2.16	80.11±2.26	90.07±3.10	90.44±2.88
1	10	-1.00 🔥	0.90 🔥	81.59±1.80	83.13±2.47	79.02±1.71	79.96±1.84	89.65±3.19	90.07±2.93
1	10	-1.00 ❄	1.00 ❄	81.13±1.69	82.88±1.79	78.36±1.88	79.32±1.95	89.14±2.90	89.58±2.66
1	10	-1.00 🔥	1.00 🔥	81.10±1.70	82.79±1.98	78.36±1.83	79.32±1.90	89.02±3.03	89.47±2.76
1	10	-0.90 ❄	-1.00 ❄	82.59±2.68	82.56±2.82	81.84±1.90	81.91±2.44	91.85±1.39	92.04±1.33
1	10	-0.90 🔥	-1.00 🔥	83.28±2.87	83.28±2.41	82.93±1.80	82.75±2.62	92.44±0.97	92.59±1.00
1	10	-0.90 ❄	-0.90 ❄	83.91±2.02	83.97±2.62	83.04±1.41	83.23±1.82	92.80±1.18	92.94±1.19
1	10	-0.90 🔥	-0.90 🔥	83.90±1.48	84.63±1.54	82.33±2.23	82.86±1.93	92.59±1.29	92.79±1.19
1	10	-0.90 ❄	-0.80 ❄	83.69±1.20	84.52±1.09	82.10±2.17	82.61±1.76	92.59±0.84	92.78±0.76
1	10	-0.90 🔥	-0.80 🔥	83.72±1.01	83.93±0.61	82.53±1.91	82.84±1.51	92.38±1.01	92.60±0.91
1	10	-0.90 ❄	-0.70 ❄	83.52±1.21	83.79±0.44	82.42±2.31	82.64±1.83	92.53±1.04	92.74±0.93
1	10	-0.90 🔥	-0.70 🔥	83.02±2.15	83.25±2.19	82.83±1.28	82.49±1.88	92.67±0.95	92.83±0.95
1	10	-0.90 ❄	-0.60 ❄	84.05±1.60	84.14±2.22	83.16±1.74	83.34±1.63	92.85±1.20	93.01±1.16
1	10	-0.90 🔥	-0.60 🔥	83.55±1.59	83.72±2.20	82.50±1.81	82.75±1.65	92.31±1.35	92.57±1.21
1	10	-0.90 ❄	-0.50 ❄	83.41±1.23	83.06±1.42	82.63±1.44	82.74±1.31	92.32±0.92	92.53±0.92
1	10	-0.90 🔥	-0.50 🔥	84.28±1.78	84.12±2.11	83.44±2.07	83.60±1.90	92.76±1.27	92.91±1.20
1	10	-0.90 ❄	-0.40 ❄	83.73±1.30	84.51±1.75	82.02±1.61	82.66±1.45	92.44±1.13	92.73±1.01
1	10	-0.90 🔥	-0.40 🔥	83.58±1.75	83.79±2.16	82.18±1.77	82.67±1.79	92.26±1.36	92.52±1.26
1	10	-0.90 ❄	-0.30 ❄	83.02±1.73	83.49±1.89	81.56±2.52	81.98±2.15	92.01±1.34	92.28±1.26
1	10	-0.90 🔥	-0.30 🔥	83.30±2.09	83.70±2.11	81.87±2.82	82.29±2.48	92.21±1.43	92.46±1.34
1	10	-0.90 ❄	-0.20 ❄	84.79±1.35	84.91±1.27	83.52±1.79	83.98±1.56	92.90±1.31	93.16±1.15
1	10	-0.90 🔥	-0.20 🔥	84.93±1.34	85.25±1.30	83.45±1.62	84.05±1.50	92.89±1.25	93.15±1.08
1	10	-0.90 ❄	-0.10 ❄	84.75±1.72	84.97±1.26	83.42±2.34	83.88±2.08	93.04±1.28	93.26±1.20
1	10	-0.90 🔥	-0.10 🔥	85.28±0.92	85.48±0.80	83.97±1.31	84.48±1.10	93.47±0.87	93.67±0.76
1	10	-0.90 ❄	0.10 ❄	85.17±1.27	85.83±1.30	83.45±1.48	84.20±1.42	93.55±1.07	93.71±0.95
1	10	-0.90 🔥	0.10 🔥	85.26±1.33	86.12±1.10	83.45±1.76	84.23±1.59	93.62±1.02	93.79±0.90
1	10	-0.90 ❄	0.20 ❄	84.40±1.92	85.22±2.03	82.53±2.18	83.31±2.12	93.24±1.44	93.38±1.36
1	10	-0.90 🔥	0.20 🔥	84.64±1.99	85.41±2.03	82.81±2.26	83.58±2.21	93.42±1.53	93.54±1.45
1	10	-0.90 ❄	0.30 ❄	83.79±2.29	84.16±2.57	82.12±2.35	82.79±2.41	92.42±2.07	92.59±2.04
1	10	-0.90 🔥	0.30 🔥	84.09±1.92	84.88±2.27	82.26±2.09	83.01±2.06	92.83±1.68	92.96±1.57
1	10	-0.90 ❄	0.40 ❄	83.35±2.22	84.75±2.58	81.09±2.41	82.00±2.42	92.22±2.03	92.41±1.94
1	10	-0.90 🔥	0.40 🔥	83.19±2.04	84.21±2.33	81.16±2.40	81.93±2.30	92.01±1.82	92.19±1.72
1	10	-0.90 ❄	0.50 ❄	82.78±2.16	83.87±2.50	80.63±2.41	81.45±2.40	91.52±2.15	91.72±2.05
1	10	-0.90 🔥	0.50 🔥	82.80±2.00	84.12±2.53	80.49±2.06	81.39±2.13	91.62±2.16	91.81±2.05
1	10	-0.90 ❄	0.60 ❄	83.05±2.85	84.44±3.08	80.73±3.14	81.63±3.19	91.27±3.08	91.47±2.97
1	10	-0.90 🔥	0.60 🔥	82.81±3.02	84.38±3.45	80.40±3.21	81.33±3.31	91.19±3.06	91.42±2.94
1	10	-0.90 ❄	0.70 ❄	81.83±2.43	83.46±2.89	79.29±2.59	80.21±2.69	90.11±3.29	90.38±3.17
1	10	-0.90 🔥	0.70 🔥	81.98±2.43	83.45±2.84	79.49±2.55	80.42±2.65	90.19±3.28	90.46±3.15
1	10	-0.90 ❄	0.80 ❄	82.12±1.91	83.59±2.32	79.61±1.95	80.57±2.05	90.15±2.89	90.52±2.76

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	10	-0.90 🔥	0.80 🔥	82.57±1.51	84.47±1.03	79.92±1.91	80.92±1.91	90.53±2.37	90.85±2.32
1	10	-0.90 ❄	0.90 ❄	81.33±2.10	82.93±2.05	78.65±2.43	79.58±2.46	89.00±2.80	89.45±2.63
1	10	-0.90 🔥	0.90 🔥	81.22±2.00	82.72±2.31	78.57±2.11	79.50±2.21	88.90±3.00	89.35±2.80
1	10	-0.90 ❄	1.00 ❄	80.70±2.22	82.20±2.75	78.02±2.29	78.93±2.39	88.22±3.63	88.72±3.36
1	10	-0.90 🔥	1.00 🔥	80.84±2.14	82.29±2.62	78.19±2.19	79.10±2.30	88.31±3.29	88.81±3.03
1	10	-0.80 ❄	-1.00 ❄	82.64±3.04	82.99±2.51	82.27±1.67	82.05±2.71	92.09±0.93	92.24±0.97
1	10	-0.80 🔥	-1.00 🔥	82.47±2.87	83.03±2.53	82.02±1.49	81.83±2.51	92.11±0.90	92.25±0.94
1	10	-0.80 ❄	-0.90 ❄	82.07±2.68	82.37±3.01	81.16±1.80	81.28±2.39	91.27±1.31	91.53±1.19
1	10	-0.80 🔥	-0.90 🔥	82.43±2.88	82.81±3.05	81.66±2.12	81.69±2.63	91.91±1.28	92.11±1.22
1	10	-0.80 ❄	-0.80 ❄	83.51±1.48	83.71±1.00	82.42±2.24	82.65±1.93	92.39±1.38	92.59±1.28
1	10	-0.80 🔥	-0.80 🔥	83.77±1.48	84.19±1.37	82.44±2.02	82.84±1.81	92.45±1.33	92.65±1.24
1	10	-0.80 ❄	-0.70 ❄	83.03±1.77	83.59±2.51	82.29±1.90	82.26±1.80	92.56±0.98	92.73±0.99
1	10	-0.80 🔥	-0.70 🔥	83.02±1.81	83.61±2.58	82.29±1.82	82.26±1.79	92.59±1.03	92.75±1.05
1	10	-0.80 ❄	-0.60 ❄	84.28±1.33	84.09±1.49	83.34±1.59	83.58±1.43	92.56±1.14	92.74±1.10
1	10	-0.80 🔥	-0.60 🔥	83.80±1.20	84.28±1.81	82.55±1.97	82.89±1.55	92.78±1.16	92.94±1.13
1	10	-0.80 ❄	-0.50 ❄	83.84±1.44	83.72±1.74	82.79±1.55	83.09±1.50	92.19±1.06	92.40±1.09
1	10	-0.80 🔥	-0.50 🔥	83.87±1.89	83.54±1.91	83.09±2.22	83.21±2.04	92.40±1.36	92.57±1.29
1	10	-0.80 ❄	-0.40 ❄	83.96±1.86	83.83±2.24	83.00±1.66	83.25±1.80	92.54±1.42	92.73±1.41
1	10	-0.80 🔥	-0.40 🔥	83.96±1.79	83.97±2.33	83.00±1.53	83.24±1.68	92.50±1.40	92.69±1.38
1	10	-0.80 ❄	-0.30 ❄	83.79±1.79	84.27±2.40	82.24±1.80	82.82±1.83	92.62±1.39	92.83±1.36
1	10	-0.80 🔥	-0.30 🔥	83.91±1.84	84.44±2.58	82.34±1.79	82.94±1.84	92.46±1.59	92.71±1.51
1	10	-0.80 ❄	-0.20 ❄	84.07±1.86	84.76±2.30	82.25±1.93	83.00±1.97	92.69±1.80	92.93±1.69
1	10	-0.80 🔥	-0.20 🔥	84.39±2.20	84.78±2.25	82.77±2.45	83.42±2.40	92.80±1.70	93.03±1.61
1	10	-0.80 ❄	-0.10 ❄	84.91±1.21	85.39±1.19	83.33±1.56	83.97±1.40	93.02±1.31	93.28±1.16
1	10	-0.80 🔥	-0.10 🔥	85.03±1.17	85.57±0.97	83.41±1.55	84.08±1.37	93.22±1.11	93.46±0.99
1	10	-0.80 ❄	0.10 ❄	85.05±1.42	86.16±1.20	83.04±1.73	83.93±1.66	93.39±1.18	93.60±1.09
1	10	-0.80 🔥	0.10 🔥	84.81±1.82	85.62±1.86	82.99±2.13	83.76±2.04	93.32±1.34	93.48±1.29
1	10	-0.80 ❄	0.20 ❄	84.86±2.12	85.58±2.18	83.14±2.48	83.85±2.37	93.29±1.67	93.43±1.60
1	10	-0.80 🔥	0.20 🔥	84.91±1.94	85.70±2.08	83.17±2.28	83.89±2.16	93.28±1.60	93.43±1.54
1	10	-0.80 ❄	0.30 ❄	83.83±2.15	84.81±2.52	81.92±2.43	82.67±2.38	92.32±1.75	92.47±1.68
1	10	-0.80 🔥	0.30 🔥	84.18±2.32	85.02±2.50	82.32±2.62	83.07±2.55	92.58±1.94	92.76±1.88
1	10	-0.80 ❄	0.40 ❄	83.80±2.71	84.93±2.96	81.71±2.94	82.56±2.96	92.10±2.30	92.30±2.22
1	10	-0.80 🔥	0.40 🔥	82.77±2.13	84.13±2.48	80.54±2.53	81.36±2.42	91.49±1.87	91.71±1.78
1	10	-0.80 ❄	0.50 ❄	82.92±2.35	84.52±2.66	80.48±2.54	81.44±2.61	91.33±2.48	91.54±2.42
1	10	-0.80 🔥	0.50 🔥	82.75±2.39	84.39±2.53	80.27±2.69	81.22±2.72	91.14±2.45	91.38±2.39
1	10	-0.80 ❄	0.60 ❄	82.82±2.70	84.30±2.92	80.42±2.94	81.35±3.00	90.82±2.86	91.08±2.77
1	10	-0.80 🔥	0.60 🔥	82.70±2.69	84.47±2.95	80.17±2.96	81.13±3.02	90.79±2.98	91.05±2.90
1	11	-1.00 ❄	-1.00 ❄	79.22±1.88	79.78±2.26	77.02±1.89	77.66±1.98	87.90±2.48	87.61±2.56
1	11	-1.00 🔥	-1.00 🔥	78.95±2.00	79.44±2.13	76.75±2.24	77.36±2.30	87.35±2.75	87.03±2.91
1	11	-1.00 ❄	-0.90 ❄	80.32±1.48	80.55±1.47	78.51±2.08	79.03±1.80	88.71±2.44	88.42±2.46
1	11	-1.00 🔥	-0.90 🔥	79.55±1.68	80.46±1.44	77.03±2.11	77.79±2.10	87.82±2.48	87.56±2.60
1	11	-1.00 ❄	-0.80 ❄	80.14±3.22	80.91±2.61	77.92±4.12	78.49±4.09	88.18±3.57	87.92±3.74
1	11	-1.00 🔥	-0.80 🔥	80.95±2.34	81.48±1.94	78.87±2.97	79.52±2.83	89.08±2.46	88.89±2.49
1	11	-1.00 ❄	-0.70 ❄	80.50±2.96	81.16±2.38	78.33±3.71	78.94±3.70	88.26±3.57	88.11±3.71
1	11	-1.00 🔥	-0.70 🔥	80.69±2.81	81.34±2.24	78.50±3.52	79.15±3.50	88.48±3.45	88.33±3.59
1	11	-1.00 ❄	-0.60 ❄	80.80±2.83	80.97±2.55	79.02±3.42	79.52±3.37	88.66±3.59	88.57±3.74
1	11	-1.00 🔥	-0.60 🔥	80.60±2.85	80.78±2.74	78.74±3.32	79.30±3.32	88.66±3.45	88.57±3.59
1	11	-1.00 ❄	-0.50 ❄	81.15±3.63	81.83±1.84	79.62±5.33	79.70±5.11	89.42±3.83	89.32±3.93
1	11	-1.00 🔥	-0.50 🔥	81.54±2.37	81.54±1.82	80.24±3.35	80.49±2.99	89.61±3.25	89.50±3.33
1	11	-1.00 ❄	-0.40 ❄	81.54±2.67	81.72±2.38	79.89±3.50	80.35±3.16	89.57±3.11	89.58±3.15
1	11	-1.00 🔥	-0.40 🔥	81.33±3.16	81.79±2.42	79.45±4.14	79.95±3.95	89.35±3.30	89.37±3.31
1	11	-1.00 ❄	-0.30 ❄	82.12±1.82	82.48±1.78	80.38±2.26	80.96±2.13	90.11±2.71	90.25±2.84
1	11	-1.00 🔥	-0.30 🔥	81.01±3.27	82.10±2.88	78.56±3.96	79.31±4.08	89.49±3.67	89.35±4.16

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	11	-1.00 ❄	-0.20 ❄	83.20±1.45	83.60±1.84	81.46±1.37	82.14±1.48	91.53±1.64	91.71±1.66
1	11	-1.00 🔥	-0.20 🔥	83.10±1.53	83.60±1.87	81.27±1.50	81.99±1.58	91.58±1.63	91.77±1.63
1	11	-1.00 ❄	-0.10 ❄	84.37±1.54	84.57±2.01	82.99±1.33	83.53±1.50	92.48±1.38	92.56±1.54
1	11	-1.00 🔥	-0.10 🔥	84.46±1.33	84.87±1.86	82.99±1.31	83.56±1.35	92.81±1.37	92.88±1.51
1	11	-1.00 ❄	0.10 ❄	83.73±1.69	84.85±2.98	82.21±1.05	82.73±1.36	93.46±1.44	93.53±1.44
1	11	-1.00 🔥	0.10 🔥	82.74±2.32	84.99±2.84	80.43±2.60	81.21±2.61	93.20±1.45	93.22±1.50
1	11	-1.00 ❄	0.20 ❄	83.62±2.29	84.44±3.21	82.16±1.83	82.66±2.13	92.80±1.99	92.81±2.07
1	11	-1.00 🔥	0.20 🔥	83.89±2.04	84.90±2.93	82.32±1.74	82.88±1.92	93.41±1.67	93.48±1.65
1	11	-1.00 ❄	0.30 ❄	83.82±2.04	84.80±2.93	82.24±1.38	82.82±1.79	93.31±1.81	93.37±1.80
1	11	-1.00 🔥	0.30 🔥	82.96±1.90	84.00±2.84	81.31±1.82	81.85±1.84	92.42±2.02	92.52±1.95
1	11	-1.00 ❄	0.40 ❄	83.59±2.03	84.92±3.01	81.71±1.68	82.43±1.93	93.09±2.10	93.15±2.08
1	11	-1.00 🔥	0.40 🔥	83.72±2.15	85.07±3.14	81.84±1.68	82.57±2.00	93.11±2.17	93.17±2.14
1	11	-1.00 ❄	0.50 ❄	81.58±2.82	83.88±2.85	79.04±3.47	79.77±3.58	92.37±1.95	92.40±1.99
1	11	-1.00 🔥	0.50 🔥	81.38±2.67	83.85±2.80	78.72±3.23	79.50±3.40	92.17±1.82	92.21±1.85
1	11	-1.00 ❄	0.60 ❄	81.27±2.91	83.59±3.45	78.66±3.11	79.46±3.33	91.70±2.11	91.73±2.11
1	11	-1.00 🔥	0.60 🔥	81.64±3.07	83.33±3.43	79.29±3.24	80.05±3.46	91.73±2.16	91.77±2.16
1	11	-1.00 ❄	0.70 ❄	82.84±1.59	85.24±1.04	80.22±2.58	81.13±2.37	92.97±0.53	93.03±0.47
1	11	-1.00 🔥	0.70 🔥	83.37±1.53	85.39±1.02	80.92±2.38	81.83±2.19	93.12±0.44	93.17±0.41
1	11	-1.00 ❄	0.80 ❄	83.21±1.26	85.12±1.01	80.78±2.07	81.69±1.82	92.67±0.56	92.72±0.50
1	11	-1.00 🔥	0.80 🔥	81.71±2.85	84.43±1.93	78.82±3.93	79.66±3.81	91.85±1.61	91.91±1.67
1	11	-1.00 ❄	0.90 ❄	81.12±1.64	83.78±1.27	78.20±2.68	79.06±2.50	91.78±0.87	91.84±0.80
1	11	-1.00 🔥	0.90 🔥	82.00±1.53	83.96±1.20	79.41±2.35	80.28±2.18	91.47±0.66	91.56±0.56
1	11	-1.00 ❄	1.00 ❄	81.36±1.50	83.66±1.26	78.52±2.24	79.43±2.21	91.25±0.61	91.32±0.55
1	11	-1.00 🔥	1.00 🔥	80.85±1.66	83.29±1.40	77.91±2.43	78.80±2.42	91.21±0.51	91.29±0.49
1	11	-0.90 ❄	-1.00 ❄	80.14±1.22	80.37±1.18	78.14±1.45	78.78±1.41	87.82±2.40	87.65±2.41
1	11	-0.90 🔥	-1.00 🔥	79.36±2.28	79.55±2.03	77.39±2.81	77.92±2.81	87.28±2.79	87.05±2.90
1	11	-0.90 ❄	-0.90 ❄	79.34±1.81	80.12±2.30	77.02±2.02	77.70±2.04	87.42±3.02	87.18±3.21
1	11	-0.90 🔥	-0.90 🔥	79.64±1.46	80.28±2.29	77.48±1.31	78.13±1.38	87.80±2.46	87.62±2.54
1	11	-0.90 ❄	-0.80 ❄	79.99±1.70	80.35±2.32	78.03±1.33	78.65±1.54	88.22±2.41	88.02±2.55
1	11	-0.90 🔥	-0.80 🔥	80.08±1.43	80.51±2.03	78.05±1.27	78.70±1.37	88.32±2.07	88.16±2.20
1	11	-0.90 ❄	-0.70 ❄	80.13±1.81	80.51±2.14	78.09±2.04	78.72±2.02	88.30±3.05	88.19±3.26
1	11	-0.90 🔥	-0.70 🔥	80.46±1.64	81.08±1.97	78.33±1.94	79.01±1.88	88.73±2.75	88.60±2.93
1	11	-0.90 ❄	-0.60 ❄	80.95±2.47	81.47±2.49	78.90±2.85	79.56±2.88	89.18±2.81	89.13±2.98
1	11	-0.90 🔥	-0.60 🔥	81.15±1.86	81.78±1.94	79.03±2.12	79.75±2.13	89.45±1.97	89.39±2.06
1	11	-0.90 ❄	-0.50 ❄	81.48±1.97	81.81±2.02	80.02±2.69	80.37±2.37	89.74±2.37	89.70±2.46
1	11	-0.90 🔥	-0.50 🔥	81.48±2.50	82.06±2.28	79.74±3.32	80.20±3.09	89.57±2.98	89.54±3.09
1	11	-0.90 ❄	-0.40 ❄	81.22±3.24	81.77±2.51	79.45±4.52	79.83±4.12	89.10±3.93	89.12±3.95
1	11	-0.90 🔥	-0.40 🔥	81.10±3.01	81.01±2.73	79.59±3.72	79.98±3.54	88.97±3.37	89.03±3.47
1	11	-0.90 ❄	-0.30 ❄	82.32±1.46	82.67±1.14	80.71±2.18	81.21±1.87	90.17±1.98	90.32±1.97
1	11	-0.90 🔥	-0.30 🔥	82.77±2.66	82.67±2.04	81.85±3.76	81.91±3.29	90.29±3.45	90.29±3.53
1	11	-0.90 ❄	-0.20 ❄	82.33±1.09	83.49±1.87	80.24±1.52	80.98±1.30	90.80±1.66	90.96±1.71
1	11	-0.90 🔥	-0.20 🔥	82.77±1.68	83.45±1.65	80.87±2.20	81.55±2.01	90.88±1.98	91.04±2.08
1	11	-0.90 ❄	-0.10 ❄	84.37±1.46	84.89±1.69	82.85±1.56	83.43±1.53	92.36±1.32	92.45±1.49
1	11	-0.90 🔥	-0.10 🔥	84.14±1.71	84.40±2.14	82.80±1.63	83.28±1.71	92.31±1.42	92.40±1.58
1	11	-0.90 ❄	0.10 ❄	85.19±1.31	86.08±1.29	83.39±1.68	84.16±1.53	93.98±0.84	94.08±0.91
1	11	-0.90 🔥	0.10 🔥	84.25±2.52	85.39±2.50	82.41±2.92	83.10±2.90	93.15±2.03	93.18±2.23
1	11	-0.90 ❄	0.20 ❄	83.24±1.93	84.15±2.38	81.61±2.19	82.15±2.14	92.52±1.61	92.61±1.56
1	11	-0.90 🔥	0.20 🔥	82.68±1.90	83.89±2.25	80.90±2.49	81.43±2.29	92.36±1.49	92.42±1.43
1	11	-0.90 ❄	0.30 ❄	81.52±1.86	83.80±2.72	79.08±2.15	79.83±2.12	92.12±1.77	92.19±1.72
1	11	-0.90 🔥	0.30 🔥	81.72±2.91	84.83±3.05	79.04±3.60	79.77±3.77	93.15±1.85	93.18±1.85
1	11	-0.90 ❄	0.40 ❄	82.68±2.38	84.76±3.01	80.42±2.63	81.19±2.65	92.73±1.87	92.79±1.89
1	11	-0.90 🔥	0.40 🔥	82.33±2.22	84.56±2.94	79.99±2.52	80.75±2.53	92.73±1.87	92.79±1.89
1	11	-0.90 ❄	0.50 ❄	80.22±2.26	83.74±3.01	77.06±2.85	77.86±3.00	91.89±1.90	91.96±1.91

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	11	-0.90 🔥	0.50 🔥	82.25±1.63	84.45±2.98	79.83±1.60	80.67±1.67	92.55±1.83	92.59±1.80
1	11	-0.90 ❄	0.60 ❄	81.94±3.27	85.51±1.67	78.75±4.37	79.66±4.48	93.08±1.12	93.08±1.12
1	11	-0.90 🔥	0.60 🔥	80.59±3.45	83.82±3.66	77.69±4.10	78.38±4.36	91.83±2.16	91.87±2.14
1	11	-0.90 ❄	0.70 ❄	81.48±3.27	84.64±1.36	78.39±4.45	79.21±4.53	92.24±1.32	92.25±1.22
1	11	-0.90 🔥	0.70 🔥	81.65±2.80	85.06±1.38	78.49±3.90	79.40±3.87	92.44±1.43	92.47±1.41
1	11	-0.90 ❄	0.80 ❄	82.43±2.13	84.70±1.00	79.76±3.07	80.65±2.99	92.18±0.69	92.20±0.61
1	11	-0.90 🔥	0.80 🔥	82.38±1.89	84.62±1.17	79.77±2.88	80.63±2.75	92.13±0.82	92.19±0.77
1	11	-0.90 ❄	0.90 ❄	80.80±1.61	83.74±1.34	77.75±2.67	78.60±2.53	91.39±1.01	91.46±0.95
1	11	-0.90 🔥	0.90 🔥	80.20±1.74	83.09±2.11	77.10±2.67	77.93±2.53	90.86±1.74	90.93±1.78
1	11	-0.90 ❄	1.00 ❄	80.77±1.59	83.06±1.41	77.92±2.45	78.76±2.39	90.62±0.65	90.70±0.66
1	11	-0.90 🔥	1.00 🔥	80.71±1.75	83.07±1.14	77.87±2.78	78.67±2.76	90.46±0.67	90.55±0.62
1	11	-0.80 ❄	-1.00 ❄	80.31±0.95	80.70±1.70	78.46±1.16	79.01±1.05	88.49±1.64	88.43±1.68
1	11	-0.80 🔥	-1.00 🔥	80.07±0.97	80.54±1.85	78.10±0.62	78.71±0.75	88.22±1.85	88.17±1.95
1	11	-0.80 ❄	-0.90 ❄	79.85±1.16	80.44±1.54	77.69±1.39	78.34±1.35	88.06±2.13	87.98±2.23
1	11	-0.80 🔥	-0.90 🔥	79.93±1.22	80.24±1.97	78.05±1.05	78.62±1.11	88.02±2.16	87.92±2.29
1	11	-0.80 ❄	-0.80 ❄	80.31±1.34	81.35±1.74	77.82±1.43	78.64±1.47	88.31±2.37	88.23±2.52
1	11	-0.80 🔥	-0.80 🔥	80.10±1.45	80.86±1.73	77.87±1.73	78.55±1.69	88.30±2.41	88.17±2.58
1	11	-0.80 ❄	-0.70 ❄	80.22±2.11	80.68±2.41	78.13±2.34	78.80±2.32	88.09±3.14	87.99±3.35
1	11	-0.80 🔥	-0.70 🔥	80.42±2.33	80.79±2.78	78.45±2.48	79.08±2.49	88.28±3.28	88.18±3.51
1	11	-0.80 ❄	-0.60 ❄	80.50±2.06	81.00±2.64	78.55±2.19	79.16±2.15	88.78±3.01	88.71±3.22
1	11	-0.80 🔥	-0.60 🔥	80.75±2.27	81.63±2.17	78.52±2.94	79.20±2.85	88.94±2.83	88.92±3.00
1	11	-0.80 ❄	-0.50 ❄	80.99±2.48	81.13±2.94	79.36±2.54	79.87±2.56	88.87±3.30	88.92±3.45
1	11	-0.80 🔥	-0.50 🔥	81.01±2.16	81.34±2.34	79.36±2.66	79.81±2.52	89.09±2.76	89.15±2.86
1	11	-0.80 ❄	-0.40 ❄	81.15±2.14	81.62±2.19	79.36±2.79	79.87±2.60	88.92±2.87	89.03±2.99
1	11	-0.80 🔥	-0.40 🔥	81.19±2.06	81.17±1.63	79.70±2.78	80.08±2.59	88.93±2.74	89.02±2.80
1	11	-0.80 ❄	-0.30 ❄	82.22±1.90	82.63±1.73	80.62±2.50	81.10±2.32	89.95±2.94	90.08±3.07
1	11	-0.80 🔥	-0.30 🔥	82.14±1.92	82.55±1.57	80.49±2.58	80.98±2.40	89.90±2.73	90.08±2.81
1	11	-0.80 ❄	-0.20 ❄	82.18±1.22	83.92±1.86	79.72±1.70	80.60±1.58	90.70±2.02	90.93±2.10
1	11	-0.80 🔥	-0.20 🔥	82.71±1.97	83.36±2.01	80.95±2.45	81.54±2.33	90.79±2.29	90.95±2.37
1	11	-0.80 ❄	-0.10 ❄	83.70±1.39	84.68±1.85	81.92±1.80	82.58±1.61	92.13±1.48	92.33±1.54
1	11	-0.80 🔥	-0.10 🔥	83.87±1.55	84.82±1.89	82.23±1.91	82.80±1.74	92.14±1.53	92.36±1.56
1	11	-0.80 ❄	0.10 ❄	84.17±1.31	85.52±2.56	82.50±1.43	83.09±1.31	93.34±1.36	93.43±1.40
1	11	-0.80 🔥	0.10 🔥	84.50±1.85	85.22±2.27	83.16±2.08	83.60±2.00	93.11±1.23	93.21±1.27
1	11	-0.80 ❄	0.20 ❄	84.18±2.35	85.30±2.13	82.11±2.81	82.96±2.71	93.22±1.82	93.26±1.85
1	11	-0.80 🔥	0.20 🔥	83.61±1.84	85.27±2.07	81.21±2.10	82.19±2.08	93.07±1.62	93.11±1.65
1	11	-0.80 ❄	0.30 ❄	81.97±3.23	83.74±2.62	79.74±4.14	80.34±4.16	92.14±1.82	92.17±1.83
1	11	-0.80 🔥	0.30 🔥	82.22±2.81	83.93±2.42	80.01±3.62	80.66±3.54	92.15±1.74	92.19±1.75
1	11	-0.80 ❄	0.40 ❄	82.52±2.69	85.46±1.58	79.52±3.50	80.54±3.50	93.40±1.17	93.43±1.26
1	11	-0.80 🔥	0.40 🔥	82.52±2.69	85.44±1.65	79.53±3.48	80.54±3.56	93.48±1.00	93.50±1.10
1	11	-0.80 ❄	0.50 ❄	82.04±2.79	84.87±1.69	79.13±3.81	80.02±3.81	92.83±0.97	92.87±1.01
1	11	-0.80 🔥	0.50 🔥	82.19±3.01	85.03±1.60	79.33±4.16	80.18±4.12	92.96±0.94	93.00±0.99
1	12	-1.00 ❄	-1.00 ❄	84.18±2.02	84.30±2.64	83.31±1.89	83.49±1.94	92.87±1.55	92.95±1.61
1	12	-1.00 🔥	-1.00 🔥	83.59±1.61	83.64±1.85	83.17±0.85	83.02±1.36	92.53±1.00	92.60±1.04
1	12	-1.00 ❄	-0.90 ❄	84.30±1.87	84.23±2.35	83.67±1.63	83.71±1.77	92.93±1.25	92.99±1.35
1	12	-1.00 🔥	-0.90 🔥	84.46±2.02	84.24±2.35	83.84±1.87	83.88±1.99	92.75±1.60	92.86±1.59
1	12	-1.00 ❄	-0.80 ❄	84.88±2.12	84.88±2.73	84.08±2.16	84.24±2.13	93.34±1.55	93.42±1.62
1	12	-1.00 🔥	-0.80 🔥	84.58±2.26	84.43±2.63	83.82±2.08	83.97±2.22	92.96±1.41	93.06±1.46
1	12	-1.00 ❄	-0.70 ❄	85.19±1.93	85.16±2.35	84.34±1.91	84.55±1.93	93.39±1.39	93.46±1.47
1	12	-1.00 🔥	-0.70 🔥	85.19±1.78	85.04±2.09	84.32±1.84	84.55±1.83	93.24±1.39	93.32±1.47
1	12	-1.00 ❄	-0.60 ❄	85.35±1.94	85.64±2.67	84.55±2.14	84.69±1.99	93.76±1.56	93.90±1.56
1	12	-1.00 🔥	-0.60 🔥	85.41±1.65	85.79±2.22	84.48±2.21	84.69±1.83	93.77±1.35	93.93±1.34
1	12	-1.00 ❄	-0.50 ❄	85.14±1.30	85.62±1.70	83.78±1.67	84.29±1.43	93.52±1.33	93.70±1.27
1	12	-1.00 🔥	-0.50 🔥	85.52±1.80	85.49±1.87	84.52±2.07	84.83±1.94	93.54±1.44	93.71±1.39

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	12	-1.00 ❄	-0.40 ❄	84.03±1.12	84.71±2.00	82.58±1.36	83.07±1.19	93.14±0.98	93.30±0.95
1	12	-1.00 🔥	-0.40 🔥	83.98±1.14	84.70±2.22	82.42±0.60	83.01±0.88	92.91±1.18	93.09±1.13
1	12	-1.00 ❄	-0.30 ❄	84.05±1.06	84.61±1.59	82.58±1.51	83.09±1.28	92.92±0.78	93.08±0.78
1	12	-1.00 🔥	-0.30 🔥	83.97±1.24	84.45±1.65	82.61±1.74	83.04±1.51	92.91±0.79	93.06±0.79
1	12	-1.00 ❄	-0.20 ❄	85.24±0.88	85.64±1.45	83.77±0.73	84.39±0.83	93.51±0.87	93.67±0.82
1	12	-1.00 🔥	-0.20 🔥	85.19±1.04	85.49±1.37	83.76±0.98	84.35±1.05	93.45±0.89	93.61±0.84
1	12	-1.00 ❄	-0.10 ❄	83.69±1.77	84.82±1.96	81.70±2.22	82.46±2.15	92.87±1.30	93.03±1.28
1	12	-1.00 🔥	-0.10 🔥	83.62±2.13	85.05±1.89	81.46±2.78	82.27±2.63	92.81±1.47	92.96±1.46
1	12	-1.00 ❄	0.10 ❄	84.25±1.76	84.32±2.16	82.92±1.65	83.42±1.79	92.53±1.45	92.65±1.46
1	12	-1.00 🔥	0.10 🔥	84.33±1.73	84.54±2.03	82.88±1.72	83.45±1.80	92.63±1.45	92.74±1.46
1	12	-1.00 ❄	0.20 ❄	82.74±1.78	83.24±1.66	80.88±2.12	81.57±2.03	91.24±1.56	91.40±1.55
1	12	-1.00 🔥	0.20 🔥	82.21±1.51	82.63±1.77	80.39±1.64	81.04±1.63	90.98±1.67	91.05±1.76
1	12	-1.00 ❄	0.30 ❄	80.52±2.72	80.56±3.10	79.14±2.51	79.49±2.72	89.65±2.26	89.66±2.42
1	12	-1.00 🔥	0.30 🔥	80.64±2.87	80.62±3.13	79.39±2.69	79.68±2.89	89.80±2.23	89.80±2.44
1	12	-1.00 ❄	0.40 ❄	79.37±2.80	79.60±2.74	78.00±2.38	78.26±2.69	88.59±2.08	88.71±2.21
1	12	-1.00 🔥	0.40 🔥	79.55±2.75	79.95±3.15	77.86±2.33	78.30±2.64	88.76±2.13	88.87±2.26
1	12	-1.00 ❄	0.50 ❄	77.61±2.79	78.16±3.04	75.85±2.19	76.17±2.56	86.92±1.93	87.05±2.14
1	12	-1.00 🔥	0.50 🔥	77.75±2.77	78.29±2.93	76.01±2.18	76.33±2.56	87.06±1.87	87.16±2.12
1	12	-1.00 ❄	0.60 ❄	76.59±2.74	76.89±2.96	74.58±2.54	74.97±2.77	85.51±1.86	85.68±2.08
1	12	-1.00 🔥	0.60 🔥	76.80±2.77	77.22±3.10	74.67±2.60	75.11±2.83	85.64±1.94	85.82±2.16
1	12	-1.00 ❄	0.70 ❄	74.58±1.66	75.08±1.90	72.82±1.73	72.87±1.76	83.55±0.69	83.69±0.95
1	12	-1.00 🔥	0.70 🔥	74.14±2.41	75.06±1.99	72.17±3.45	72.02±3.77	83.26±1.57	83.38±1.80
1	12	-1.00 ❄	0.80 ❄	72.86±0.93	73.07±1.82	70.88±1.55	70.94±1.47	81.88±0.69	81.97±0.94
1	12	-1.00 🔥	0.80 🔥	72.38±1.91	72.81±1.18	70.22±3.39	70.05±3.82	81.26±1.44	81.39±1.58
1	12	-1.00 ❄	0.90 ❄	72.69±1.54	72.52±1.88	70.35±1.86	70.63±1.91	80.94±1.12	81.00±1.17
1	12	-1.00 🔥	0.90 🔥	72.30±1.19	72.15±1.49	69.91±1.78	70.14±1.79	80.65±1.03	80.73±1.06
1	12	-1.00 ❄	1.00 ❄	72.82±1.82	72.76±1.85	70.16±2.39	70.48±2.47	80.34±1.35	80.52±1.41
1	12	-1.00 🔥	1.00 🔥	72.30±2.14	72.34±2.32	69.73±2.70	69.93±2.98	79.28±2.79	79.61±2.69
1	12	-0.90 ❄	-1.00 ❄	82.81±2.69	83.16±3.15	81.99±2.68	82.04±2.64	91.96±2.05	92.09±2.05
1	12	-0.90 🔥	-1.00 🔥	84.29±1.64	83.94±1.78	83.60±1.85	83.68±1.73	92.44±1.44	92.58±1.41
1	12	-0.90 ❄	-0.90 ❄	84.19±1.82	84.06±2.20	83.58±1.73	83.60±1.78	92.65±1.14	92.73±1.19
1	12	-0.90 🔥	-0.90 🔥	84.12±1.68	84.03±2.13	83.37±1.63	83.48±1.64	92.39±1.34	92.52±1.33
1	12	-0.90 ❄	-0.80 ❄	84.10±1.29	84.81±2.47	82.86±1.63	83.21±1.28	93.24±1.59	93.30±1.67
1	12	-0.90 🔥	-0.80 🔥	84.47±1.62	84.60±2.31	83.65±1.61	83.80±1.57	93.11±1.30	93.18±1.39
1	12	-0.90 ❄	-0.70 ❄	84.56±1.58	85.14±2.61	83.39±1.79	83.74±1.57	92.90±1.45	93.07±1.46
1	12	-0.90 🔥	-0.70 🔥	84.51±2.21	84.74±2.93	83.60±2.04	83.82±2.14	92.89±1.70	93.04±1.69
1	12	-0.90 ❄	-0.60 ❄	84.65±2.23	85.00±2.71	83.52±2.45	83.86±2.34	93.03±1.75	93.21±1.71
1	12	-0.90 🔥	-0.60 🔥	85.00±1.91	85.23±2.24	83.81±2.17	84.22±2.04	93.19±1.58	93.34±1.56
1	12	-0.90 ❄	-0.50 ❄	85.51±1.90	85.74±1.89	84.36±2.36	84.74±2.14	93.45±1.53	93.61±1.50
1	12	-0.90 🔥	-0.50 🔥	85.13±1.85	85.42±1.88	83.84±2.27	84.31±2.07	93.20±1.51	93.33±1.42
1	12	-0.90 ❄	-0.40 ❄	85.13±2.39	85.22±2.55	83.93±2.63	84.36±2.56	93.20±1.56	93.33±1.46
1	12	-0.90 🔥	-0.40 🔥	85.55±2.00	85.49±2.36	84.69±2.00	84.92±2.02	93.40±1.36	93.52±1.28
1	12	-0.90 ❄	-0.30 ❄	84.57±1.39	85.00±2.08	83.07±1.20	83.67±1.35	93.11±1.20	93.28±1.14
1	12	-0.90 🔥	-0.30 🔥	83.90±1.33	85.12±1.39	81.87±1.84	82.67±1.66	92.70±0.87	92.85±0.85
1	12	-0.90 ❄	-0.20 ❄	84.49±1.17	84.66±1.76	83.14±0.85	83.66±1.08	92.85±0.96	93.03±0.90
1	12	-0.90 🔥	-0.20 🔥	84.72±1.32	84.68±1.71	83.55±1.11	83.97±1.29	93.13±1.09	93.25±1.04
1	12	-0.90 ❄	-0.10 ❄	83.91±1.50	84.79±1.54	82.10±1.97	82.79±1.81	92.70±1.23	92.86±1.19
1	12	-0.90 🔥	-0.10 🔥	84.15±1.58	84.99±1.58	82.26±1.89	83.03±1.82	92.89±1.28	93.00±1.24
1	12	-0.90 ❄	0.10 ❄	84.04±1.85	84.19±1.99	82.61±1.99	83.14±1.99	92.08±1.52	92.21±1.55
1	12	-0.90 🔥	0.10 🔥	83.41±2.31	83.47±2.29	81.95±2.59	82.47±2.51	91.81±1.88	91.93±1.93
1	12	-0.90 ❄	0.20 ❄	81.77±2.66	81.92±2.74	80.48±2.58	80.81±2.71	90.51±2.00	90.50±2.17
1	12	-0.90 🔥	0.20 🔥	80.66±1.85	80.78±1.60	79.51±2.69	79.64±2.31	89.94±1.79	89.97±1.88
1	12	-0.90 ❄	0.30 ❄	80.22±2.29	80.42±2.57	78.60±2.23	79.04±2.34	89.09±2.03	89.06±2.27

Continued on next page

t	n	a	b	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
1	12	-0.90 🔥	0.30 🔥	80.24±2.39	80.57±2.60	78.68±2.22	79.07±2.39	89.24±2.15	89.19±2.41
1	12	-0.90 ❄	0.40 ❄	78.23±3.06	78.77±2.11	76.39±4.18	76.58±4.21	87.60±2.29	87.70±2.39
1	12	-0.90 🔥	0.40 🔥	78.21±3.00	78.94±2.08	76.13±4.20	76.41±4.14	87.42±2.65	87.53±2.74
1	12	-0.90 ❄	0.50 ❄	77.44±2.18	77.78±2.37	75.35±2.23	75.82±2.33	86.03±1.63	86.14±1.89
1	12	-0.90 🔥	0.50 🔥	76.73±2.99	77.57±2.51	74.30±3.78	74.64±4.06	85.69±2.08	85.84±2.31
1	12	-0.90 ❄	0.60 ❄	76.27±2.64	76.83±2.90	74.00±2.56	74.42±2.76	84.85±1.71	84.99±1.96
1	12	-0.90 🔥	0.60 🔥	75.74±2.75	76.63±2.59	73.23±3.24	73.55±3.53	84.34±1.82	84.46±2.11
1	12	-0.90 ❄	0.70 ❄	75.22±2.41	75.97±2.68	72.67±2.36	73.07±2.61	83.65±1.55	83.67±1.87
1	12	-0.90 🔥	0.70 🔥	74.03±1.66	74.65±2.27	71.90±1.85	72.03±1.81	83.04±0.82	82.99±1.05
1	12	-0.90 ❄	0.80 ❄	73.00±1.76	73.29±2.36	70.68±1.86	70.87±1.92	81.77±0.94	81.75±1.20
1	12	-0.90 🔥	0.80 🔥	73.11±1.55	73.26±2.04	70.86±1.66	71.07±1.69	81.76±0.93	81.69±1.23
1	12	-0.90 ❄	0.90 ❄	73.24±1.86	73.84±2.05	70.25±2.31	70.57±2.50	81.18±1.01	81.02±1.42
1	12	-0.90 🔥	0.90 🔥	72.12±1.04	72.84±2.07	69.16±1.45	69.33±1.54	80.58±0.57	80.50±0.91
1	12	-0.90 ❄	1.00 ❄	72.33±1.92	73.19±1.90	68.99±2.68	69.16±2.95	79.75±1.64	79.69±1.81
1	12	-0.90 🔥	1.00 🔥	71.73±2.72	72.00±3.35	68.88±2.88	69.10±3.16	79.28±2.16	79.35±2.25
1	12	-0.80 ❄	-1.00 ❄	84.43±1.79	84.12±2.00	83.83±1.94	83.85±1.83	92.42±1.46	92.51±1.44
1	12	-0.80 🔥	-1.00 🔥	84.28±1.52	83.87±1.75	83.75±1.53	83.73±1.52	92.46±1.43	92.51±1.39
1	12	-0.80 ❄	-0.90 ❄	83.47±1.22	83.33±1.42	83.08±0.90	82.91±1.05	92.20±0.98	92.28±1.00
1	12	-0.80 🔥	-0.90 🔥	83.87±1.72	83.65±2.13	83.45±1.51	83.34±1.64	92.49±1.40	92.54±1.40
1	12	-0.80 ❄	-0.80 ❄	84.07±1.47	83.83±1.68	83.54±1.65	83.49±1.51	92.49±1.19	92.61±1.26
1	12	-0.80 🔥	-0.80 🔥	84.08±1.57	84.21±2.13	83.25±1.69	83.38±1.56	92.63±1.37	92.71±1.43
1	12	-0.80 ❄	-0.70 ❄	84.35±1.59	84.69±2.35	83.18±1.55	83.54±1.54	92.81±1.58	92.87±1.68
1	12	-0.80 🔥	-0.70 🔥	84.22±2.03	84.39±2.67	83.07±2.07	83.43±2.04	92.61±1.85	92.76±1.81
1	12	-0.80 ❄	-0.60 ❄	84.91±2.08	85.43±2.01	83.42±2.56	83.98±2.35	92.77±1.68	92.97±1.59
1	12	-0.80 🔥	-0.60 🔥	84.65±1.82	84.98±2.11	83.30±2.03	83.80±1.96	92.74±1.63	92.90±1.58
1	12	-0.80 ❄	-0.50 ❄	84.33±2.12	85.09±1.94	82.86±2.92	83.33±2.52	92.82±1.78	92.99±1.71
1	12	-0.80 🔥	-0.50 🔥	84.57±1.91	85.18±1.95	83.12±2.52	83.62±2.23	93.02±1.69	93.19±1.63
1	12	-0.80 ❄	-0.40 ❄	84.30±1.52	85.33±2.20	82.42±1.70	83.20±1.64	92.90±1.53	93.01±1.44
1	12	-0.80 🔥	-0.40 🔥	84.05±1.41	84.94±2.21	82.42±1.82	83.01±1.57	92.92±1.57	93.05±1.46
1	12	-0.80 ❄	-0.30 ❄	84.23±1.66	85.00±1.84	82.54±1.99	83.19±1.89	92.87±1.22	93.03±1.15
1	12	-0.80 🔥	-0.30 🔥	84.30±1.38	85.64±1.53	82.33±2.16	83.10±1.72	93.14±1.40	93.27±1.32
1	12	-0.80 ❄	-0.20 ❄	84.46±1.62	85.23±1.53	82.63±1.94	83.39±1.87	92.85±1.15	92.98±1.07
1	12	-0.80 🔥	-0.20 🔥	84.26±1.58	85.56±1.83	82.18±1.88	83.05±1.82	92.96±1.30	93.11±1.25
1	12	-0.80 ❄	-0.10 ❄	83.69±0.93	84.36±1.98	81.99±0.32	82.65±0.67	92.33±1.16	92.44±1.28
1	12	-0.80 🔥	-0.10 🔥	83.42±1.18	84.76±1.61	81.21±1.44	82.10±1.39	92.74±0.96	92.85±0.96
1	12	-0.80 ❄	0.10 ❄	81.87±1.80	82.34±2.45	80.32±1.48	80.80±1.69	90.48±2.20	90.59±2.21
1	12	-0.80 🔥	0.10 🔥	82.39±1.66	82.48±1.35	80.87±2.25	81.35±1.99	90.82±1.77	90.87±1.86
1	12	-0.80 ❄	0.20 ❄	81.33±2.44	81.99±2.04	79.28±3.06	79.92±3.00	90.08±1.91	90.21±1.93
1	12	-0.80 🔥	0.20 🔥	81.08±2.25	81.50±2.14	79.16±2.70	79.76±2.64	89.78±1.94	89.90±1.97
1	12	-0.80 ❄	0.30 ❄	80.32±1.87	80.76±2.18	78.45±1.95	79.00±1.98	88.72±1.82	88.61±2.11
1	12	-0.80 🔥	0.30 🔥	79.55±1.02	79.72±1.23	78.22±1.77	78.42±1.32	88.34±1.65	88.22±1.85
1	12	-0.80 ❄	0.40 ❄	77.65±3.37	78.93±1.57	75.15±4.83	75.37±5.30	86.89±1.51	87.04±1.63
1	12	-0.80 🔥	0.40 🔥	78.07±2.81	79.33±1.61	75.53±4.00	75.93±4.22	87.00±1.53	87.14±1.61
1	12	-0.80 ❄	0.50 ❄	77.51±2.91	78.32±3.86	75.04±2.90	75.65±3.12	86.13±3.08	86.21±3.24
1	12	-0.80 🔥	0.50 🔥	77.33±2.74	77.67±2.78	75.09±3.15	75.59±3.17	85.76±2.27	85.88±2.52

M.2 Generalizability of Hyperparameters

We transfer the key hyperparameters obtained from HMTCN—such as the fusion coefficients (a, b), and segment length (n)—to other baseline models. This experiment evaluates whether these hyperparameters retain effectiveness when applied to different architectures, indicating their robustness and broader applicability.

Table 25: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
APAVA (2-Classes) Reproduced	Autoformer	73.18±7.33	73.87±6.72	73.01±6.10	72.40±7.03	81.64±7.24	81.10±7.75
	Autoformer + CIF	75.96±2.68	76.33±3.47	75.11±1.10	74.97±1.90	83.42±1.99	83.15±2.63
	Crossformer	72.76±2.04	79.64±2.45	67.41±2.62	66.88±3.61	71.81±4.06	71.64±3.74
	Crossformer + CIF	82.32±2.60	85.35±1.83	79.21±3.17	80.29±3.30	90.39±1.58	90.02±1.47
	FEDformer	75.16±1.67	74.98±0.69	73.34±2.97	73.50±2.90	83.89±1.54	83.27±1.62
	FEDformer + CIF	77.20±2.17	76.97±2.07	77.02±2.60	76.55±2.40	86.70±1.73	86.53±1.76
	Informer	72.20±2.78	73.92±4.80	68.48±2.51	68.74±2.70	70.14±3.43	70.84±3.80
	Informer + CIF	79.78±2.07	82.29±3.03	76.55±2.05	77.53±2.23	78.56±1.33	78.58±1.14
	iTransformer	74.55±1.66	74.77±2.10	71.76±1.72	72.30±1.79	85.59±1.55	84.39±1.57
	iTransformer + CIF	74.95±0.87	74.40±0.75	73.79±1.78	73.81±1.43	84.30±0.88	82.49±0.96
	MTST	69.24±1.24	75.87±2.80	63.28±1.81	61.62±2.75	66.09±3.27	68.08±2.93
	MTST + CIF	76.20±2.39	81.46±1.11	71.67±3.06	72.06±3.73	77.65±3.37	77.98±2.73
	Nonformer	71.81±4.20	71.31±4.40	70.15±3.38	70.38±3.74	71.54±2.73	72.79±2.50
	Nonformer + CIF	71.89±3.81	71.80±4.58	69.44±3.56	69.74±3.84	70.55±2.96	70.78±4.08
	Reformer	78.42±2.85	80.89±4.52	75.20±2.28	76.09±2.54	75.48±2.79	77.52±2.64
	Reformer + CIF	81.51±0.57	**84.90±0.86**	78.18±0.60	79.32±0.64	79.10±2.50	80.77±2.21
	Transformer	75.53±4.28	76.90±5.05	72.14±4.87	72.64±5.44	72.30±6.04	73.04±7.15
	Transformer + CIF	77.96±2.82	79.34±3.86	75.07±2.52	75.87±2.73	74.75±1.62	74.76±2.33
	Medformer	77.85±2.42	80.31±3.21	74.38±2.49	75.21±2.67	80.85±3.80	81.62±3.24
	Medformer + CIF	81.06±1.58	82.97±2.23	78.26±1.52	79.23±1.65	85.74±1.85	86.32±1.48
	MedGNN	77.40±5.77	82.77±4.46	73.24±7.06	73.29±9.01	81.31±2.94	82.80±2.91
	MedGNN + CIF	81.02±1.51	84.21±2.62	77.76±1.52	78.83±1.64	86.27±2.79	87.45±2.40
	ADformer	79.59±2.78	81.75±3.61	76.51±2.95	77.42±3.13	84.70±6.23	85.41±6.00
	ADformer + CIF	82.39±0.53	85.63±1.47	79.29±0.99	80.42±0.90	89.12±1.45	89.69±1.24
	BIOT	70.10±0.29	73.17±1.51	65.10±0.52	64.60±0.81	68.58±1.21	69.15±1.59
	BIOT + CIF	73.12±0.60	80.95±0.48	67.59±0.74	67.13±0.96	75.35±1.09	75.80±1.47
	Conformer	82.21±4.85	86.69±2.48	78.71±6.06	79.59±6.61	90.70±2.53	91.09±2.44
	Conformer + CIF	84.74±3.10	86.77±3.23	82.48±3.59	83.40±3.53	91.69±2.48	92.00±2.36
	EEGInception	77.96±6.15	83.49±5.29	74.41±7.36	74.29±9.21	86.30±1.96	86.95±1.89
	EEGInception + CIF	84.72±2.37	88.02±1.09	81.84±3.10	83.04±2.97	88.47±3.90	89.19±3.80
	EEGNet	72.10±4.33	78.40±4.58	66.77±5.38	65.83±7.06	83.28±5.97	83.44±6.00
	EEGNet + CIF	81.33±4.50	84.71±4.34	78.23±5.57	78.98±5.92	90.83±5.40	91.15±5.15
	Inception1d	81.08±2.85	82.63±4.69	78.78±2.28	79.58±2.65	85.65±5.17	86.53±4.86
	Inception1d + CIF	84.04±1.90	87.80±2.11	80.90±2.08	82.20±2.19	88.79±3.99	89.16±4.08
	ResNet1d	80.11±3.36	80.69±4.28	78.44±2.87	78.92±3.15	83.07±6.04	83.99±6.00
	ResNet1d + CIF	81.84±1.63	83.60±2.24	79.28±1.89	80.20±1.90	86.50±4.14	86.94±4.16
	XResNet1d	66.25±5.95	66.58±5.94	63.76±6.04	63.01±6.58	70.66±6.51	70.55±6.85
	XResNet1d + CIF	74.84±3.93	75.83±5.85	72.70±3.11	73.02±3.53	81.21±5.31	81.60±5.23
	HMTCN	82.49±1.40	82.38±1.79	81.20±1.32	81.60±1.39	91.10±1.63	91.30±1.71
	HMTCN + CIF	**86.30±1.05**	86.16±1.09	**85.47±1.12**	**85.71±1.09**	**94.26±0.54**	**94.42±0.49**

Table 26: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
TDBrain (2-Classes) Reproduced	Autoformer	90.38±3.03	91.16±2.42	90.38±3.03	90.31±3.09	95.83±2.14	95.43±2.31
	Autoformer + CIF	**93.42±2.49**	**93.71±2.27**	**93.42±2.49**	**93.40±2.51**	97.46±1.08	97.21±1.15
	Crossformer	82.15±2.60	82.81±2.11	82.15±2.60	82.04±2.70	91.20±2.23	91.47±2.16
	Crossformer + CIF	89.40±1.26	89.83±1.18	89.40±1.26	89.37±1.27	96.76±0.62	96.80±0.62
	FEDformer	77.60±1.23	78.25±1.52	77.60±1.23	77.48±1.19	86.31±1.23	86.48±1.36
	FEDformer + CIF	78.87±1.94	79.19±2.00	78.88±1.94	78.82±1.94	88.12±1.87	88.39±1.92
	Informer	88.42±2.99	89.01±2.45	88.42±2.99	88.36±3.05	96.54±0.90	96.66±0.85
	Informer + CIF	89.38±1.66	89.71±1.37	89.38±1.66	89.35±1.68	96.92±0.66	97.03±0.64
	iTransformer	74.69±1.02	74.76±1.04	74.69±1.02	74.67±1.02	83.35±1.24	83.65±1.41
	iTransformer + CIF	72.79±1.12	72.89±1.16	72.79±1.12	72.76±1.11	81.08±1.03	81.31±0.93
	MTST	77.67±3.58	78.97±4.37	77.67±3.58	77.45±3.55	86.47±4.84	84.99±6.43
	MTST + CIF	87.21±1.99	87.30±2.06	87.21±1.99	87.20±1.99	94.14±1.83	93.24±2.63
	Nonformer	88.10±2.39	88.76±1.74	88.10±2.39	88.04±2.47	96.56±0.91	96.36±1.21
	Nonformer + CIF	88.00±2.06	88.91±1.68	88.00±2.06	87.92±2.12	97.10±1.10	97.17±1.09
	PatchTST	77.98±2.64	79.30±3.73	77.98±2.64	77.76±2.65	86.67±4.03	84.93±5.47
	PatchTST + CIF	79.58±0.86	80.22±0.82	79.58±0.86	79.47±0.88	87.20±1.57	85.81±1.17
	Reformer	88.50±2.30	89.01±1.80	88.50±2.30	88.45±2.35	96.10±0.63	96.19±0.55
	Reformer + CIF	89.08±1.19	89.47±0.70	89.08±1.19	89.05±1.23	96.88±0.35	97.00±0.35
	Transformer	85.13±1.86	86.39±1.56	85.13±1.86	84.99±1.93	95.61±1.05	95.63±0.91
	Transformer + CIF	89.96±1.57	90.53±1.26	89.96±1.57	89.92±1.60	97.73±0.45	97.77±0.46
	Medformer	88.77±1.24	88.91±1.11	88.77±1.24	88.76±1.25	96.38±0.34	96.44±0.30
	Medformer + CIF	90.88±0.87	90.94±0.81	90.88±0.87	90.87±0.87	97.39±0.30	97.46±0.28
	ADformer	84.19±1.33	85.83±0.44	84.19±1.33	83.99±1.51	91.72±0.27	91.61±0.39
	ADformer + CIF	86.90±1.08	87.61±0.95	86.90±1.09	86.83±1.12	92.93±1.10	93.06±1.04
	BIOT	81.15±5.30	81.52±5.68	81.15±5.30	81.11±5.27	90.03±4.16	90.29±3.98
	BIOT + CIF	82.81±2.14	83.41±1.74	82.81±2.14	82.72±2.21	91.91±1.87	92.28±1.76
	Conformer	90.38±1.66	90.48±1.59	90.38±1.66	90.37±1.67	96.87±0.95	96.79±1.08
	Conformer + CIF	93.69±1.46	94.01±1.43	93.69±1.46	93.68±1.46	99.04±0.65	99.06±0.61
	EEGInception	87.42±4.28	88.55±3.37	87.42±4.28	87.29±4.41	96.52±1.83	96.63±1.77
	EEGInception + CIF	93.81±2.26	93.93±2.19	93.81±2.26	93.81±2.26	98.85±0.50	98.88±0.48
	EEGNet	78.46±4.62	79.25±4.96	78.46±4.62	78.32±4.63	89.86±4.37	89.83±4.92
	EEGNet + CIF	86.29±4.51	87.29±3.91	86.29±4.51	86.17±4.64	95.75±1.75	95.93±1.69
	Inception1d	88.71±3.80	89.57±3.28	88.71±3.80	88.63±3.86	96.70±1.86	96.81±1.80
	Inception1d + CIF	91.60±3.17	91.86±2.98	91.60±3.17	91.59±3.19	97.79±1.30	97.89±1.23
	ResNet1d	85.40±2.30	87.42±1.52	85.40±2.30	85.18±2.41	96.11±1.27	96.24±1.21
	ResNet1d + CIF	88.73±1.06	88.96±1.08	88.73±1.06	88.71±1.06	95.91±0.91	95.99±0.86
	XResNet1d	87.88±3.45	88.21±3.16	87.88±3.45	87.84±3.49	95.38±2.52	95.55±2.44
	XResNet1d + CIF	91.38±2.10	91.55±2.06	91.38±2.10	91.36±2.10	97.63±1.14	97.72±1.09
	HMTCN	84.90±2.60	86.02±2.00	84.90±2.60	84.76±2.74	93.94±1.92	94.20±1.85
	HMTCN + CIF	93.13±1.41	93.33±1.37	93.13±1.41	93.12±1.42	**98.62±0.66**	**98.68±0.63**

Table 27: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
ADFTD (3-Classes) Reproduced	Autoformer	46.90±2.89	45.59±2.37	44.91±2.23	44.34±2.52	63.49±2.44	45.63±2.29
	Autoformer + CIF	45.92±2.78	44.81±2.44	44.66±2.25	44.26±2.38	62.69±2.21	44.93±2.68
	Crossformer	50.18±1.97	45.97±1.84	46.30±1.73	45.90±1.84	66.68±1.67	48.65±1.89
	Crossformer + CIF	54.58±1.22	47.85±0.67	48.96±0.89	48.22±0.72	69.10±1.05	52.23±0.98
	FEDformer	45.75±0.78	45.71±1.29	44.27±1.28	43.51±1.00	62.64±1.64	45.88±1.35
	FEDformer + CIF	48.63±1.99	46.97±1.58	46.87±1.63	46.69±1.54	65.56±2.14	48.30±2.30
	Informer	48.42±1.99	46.94±1.60	46.41±0.99	45.76±0.43	65.99±1.14	47.49±1.07
	Informer + CIF	50.12±0.66	47.23±0.47	46.77±0.39	46.62±0.38	65.13±0.52	46.84±0.54
	iTransformer	52.85±1.36	46.97±1.05	47.31±1.03	46.84±0.78	67.46±0.96	49.90±0.89
	iTransformer + CIF	50.76±0.50	47.11±0.67	47.29±0.59	47.10±0.66	67.00±0.55	49.60±0.69
	MTST	45.77±1.70	44.39±1.73	43.70±1.82	43.36±1.98	61.38±1.57	44.01±1.60
	MTST + CIF	46.36±0.93	45.04±0.65	45.47±0.88	44.87±0.81	63.57±1.24	47.35±1.36
	Nonformer	50.81±1.06	48.71±1.40	48.55±1.47	48.36±1.38	66.95±1.54	48.08±1.82
	Nonformer + CIF	51.81±2.11	49.93±0.81	49.66±0.36	49.10±0.57	68.59±0.60	50.44±0.94
	Reformer	51.28±2.60	49.68±2.75	49.64±2.02	48.45±2.06	69.20±2.53	51.74±3.24
	Reformer + CIF	52.50±0.91	52.10±2.91	49.01±2.74	47.53±4.23	68.25±1.79	50.44±2.59
	Transformer	50.53±0.94	49.31±0.87	48.57±1.23	48.42±1.28	67.98±0.90	49.07±1.35
	Transformer + CIF	52.84±2.50	51.27±2.33	51.53±2.37	51.10±2.21	70.25±2.42	52.35±2.94
	Medformer	53.70±1.18	51.51±1.32	50.49±1.48	50.35±1.53	70.48±1.17	50.91±1.13
	Medformer + CIF	55.88±0.82	51.91±1.90	50.80±1.63	50.29±1.92	70.45±1.33	53.73±1.70
	MedGNN	50.22±3.21	48.65±3.72	47.50±4.57	47.33±4.40	67.18±4.39	48.84±4.11
	MedGNN + CIF	54.89±1.23	51.57±1.57	51.46±1.68	50.85±2.05	71.98±1.59	53.89±1.73
	ADformer	52.46±3.18	51.11±3.15	50.80±2.87	50.52±3.10	72.27±2.44	54.36±3.10
	ADformer + CIF	54.59±2.91	52.45±2.60	51.43±2.71	51.36±2.61	72.53±2.32	54.97±2.99
	BIOT	42.49±0.63	42.17±0.63	41.44±0.83	41.66±0.72	58.61±0.80	42.76±1.65
	BIOT + CIF	45.02±1.40	43.53±1.57	41.81±1.21	41.97±1.33	59.47±1.35	42.55±1.58
	Conformer	49.37±2.34	48.64±1.33	48.88±1.61	48.03±1.92	68.00±2.16	49.38±1.41
	Conformer + CIF	52.97±1.97	49.12±3.03	49.28±2.84	48.81±3.05	69.28±2.52	51.40±2.53
	EEGInception	50.75±1.49	47.55±1.30	47.36±1.27	46.31±1.75	67.52±1.01	48.55±1.16
	EEGInception + CIF	55.70±3.24	52.74±4.49	51.31±3.54	50.10±4.09	73.38±3.98	57.52±3.92
	EEGNet	51.71±1.84	50.37±1.43	50.04±1.27	49.66±1.44	68.26±1.61	50.66±1.89
	EEGNet + CIF	51.46±3.77	50.14±3.57	48.58±2.11	48.36±2.05	67.94±2.86	50.73±3.84
	Inception1d	49.85±1.73	48.40±1.31	48.33±1.19	47.22±1.05	68.06±1.16	49.62±0.96
	Inception1d + CIF	55.16±1.81	52.60±2.35	52.33±2.10	52.02±2.14	72.61±2.07	56.18±1.95
	ResNet1d	48.75±2.95	47.12±2.83	47.23±3.12	46.88±3.02	65.99±3.07	47.88±2.85
	ResNet1d + CIF	53.59±1.28	49.02±1.52	49.34±1.33	48.25±1.10	69.40±1.43	52.42±1.67
	XResNet1d	49.89±1.57	46.69±1.75	46.93±1.48	46.45±1.57	65.93±1.84	47.86±1.94
	XResNet1d + CIF	52.36±0.96	48.37±1.21	48.69±1.09	47.77±1.38	68.08±1.43	50.66±1.67
	HMTCN	52.05±2.22	50.45±3.00	50.40±2.55	49.48±2.70	69.43±2.84	50.99±3.15
	HMTCN + CIF	**58.56±0.93**	**55.65±0.81**	**55.86±0.79**	**55.42±0.82**	**76.07±0.59**	**59.75±0.67**

During the process of transferring the optimized CIF hyperparameter configurations from the HMTCN architecture to other model structures, we observed significant improvements in performance across most models and evaluation metrics. This result strongly demonstrates that the CIF module not only exhibits high generalizability but also possesses well-transferable hyperparameters that can be effectively adapted to various architectures. These hyperparameters enhance feature extraction and representation capabilities across different model designs. This finding further highlights the potential of the CIF architecture in diverse tasks and provides valuable insights for future model design and parameter sharing.

Table 28: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
PTB (2-Classes)	**Autoformer**	71.99±2.74	69.60±3.85	61.50±4.23	61.43±5.07	74.29±1.89	70.26±2.00
	Autoformer + CIF	77.71±0.63	77.15±1.29	70.32±1.62	71.77±1.60	81.20±4.15	78.13±4.80
	Crossformer	78.06±3.44	81.53±3.13	68.62±5.63	69.76±6.53	88.31±2.07	85.81±2.43
	Crossformer + CIF	84.55±4.80	85.96±3.63	78.64±7.49	80.25±7.89	92.17±2.99	91.16±3.70
	FEDformer	74.54±2.27	77.99±4.10	63.14±3.29	63.28±4.36	84.63±4.27	80.91±5.55
	FEDformer + CIF	77.93±1.86	80.65±3.50	68.59±2.20	70.23±2.63	86.00±3.51	83.42±4.76
	Informer	79.59±0.65	83.33±0.77	70.58±0.95	72.58±1.08	92.77±0.48	90.89±0.57
	Informer + CIF	83.63±2.32	85.31±1.80	77.28±3.67	79.37±3.45	93.87±0.48	92.29±0.63
	iTransformer	83.43±1.19	88.06±1.47	75.64±1.55	78.29±1.70	91.38±1.41	91.08±1.30
	iTransformer + CIF	81.62±2.96	87.48±2.03	72.74±4.41	74.99±4.91	90.25±3.29	89.80±3.28
	MTST	75.53±2.45	78.72±1.87	64.78±4.06	65.30±4.81	87.76±4.09	83.60±3.92
	MTST + CIF	81.80±1.60	84.19±2.05	74.28±2.05	76.52±2.25	91.30±2.48	88.41±3.30
	Nonformer	78.93±1.46	82.48±1.53	69.68±2.15	71.50±2.56	90.54±0.59	87.78±1.46
	Nonformer + CIF	71.89±3.81	71.80±4.58	69.44±3.56	69.74±3.84	70.55±2.96	70.78±4.08
	PatchTST	75.28±2.44	77.05±2.44	64.86±4.05	65.41±5.28	88.11±2.59	82.65±2.87
	PatchTST + CIF	80.80±1.90	82.46±1.75	73.23±2.81	75.24±2.95	92.45±0.91	88.23±1.61
	Reformer	78.11±1.65	82.70±0.80	68.17±2.68	69.68±3.34	90.77±1.56	88.14±1.20
	Reformer + CIF	81.95±2.09	84.63±1.07	74.44±3.53	76.55±3.64	93.36±0.87	91.53±1.05
	Transformer	76.43±1.98	81.25±1.15	65.64±3.29	66.44±4.39	90.21±1.24	87.28±1.49
	Transformer + CIF	78.80±2.44	82.45±2.36	69.42±3.62	71.10±4.33	92.50±1.17	89.07±1.39
	Medformer	80.99±0.75	83.01±0.72	73.35±1.16	75.47±1.21	93.10±1.18	90.69±1.04
	Medformer + CIF	78.74±0.64	81.11±0.84	75.40±0.66	76.31±0.71	83.20±0.91	83.66±0.92
	ADformer	77.61±1.94	84.81±1.12	66.80±3.20	67.90±4.14	89.64±0.90	88.93±0.81
	ADformer + CIF	82.25±1.25	85.60±0.82	74.51±1.97	76.87±2.05	91.19±1.62	89.76±1.84
	BIOT	77.32±2.04	81.37±2.35	67.13±2.99	68.42±3.72	91.74±0.98	87.57±1.45
	BIOT + CIF	81.91±1.13	84.28±0.94	74.47±1.74	76.70±1.81	91.91±0.66	89.28±1.35
	Conformer	77.27±1.79	81.33±1.14	67.22±3.30	68.39±3.93	89.13±2.11	86.44±2.40
	Conformer + CIF	79.49±1.36	81.58±1.59	71.12±1.90	73.05±2.10	93.06±1.29	90.08±1.92
	EEGInception	81.75±1.74	84.07±2.71	74.28±2.04	76.51±2.28	91.60±1.33	89.98±1.85
	EEGInception + CIF	81.09±2.38	83.29±2.47	73.37±3.50	75.43±3.78	89.11±1.64	87.30±1.18
	EEGNet	78.68±0.81	81.46±1.06	69.66±1.24	71.47±1.41	92.70±0.42	88.98±0.61
	EEGNet + CIF	80.19±2.29	82.47±2.22	72.02±3.33	73.99±3.60	94.27±0.66	90.08±2.00
	Inception1d	79.91±2.37	84.92±1.79	70.76±4.03	72.56±4.59	89.25±3.46	88.15±2.62
	Inception1d + CIF	85.12±2.78	87.34±2.49	79.03±4.19	81.25±4.17	92.25±1.65	91.80±1.68
	ResNet1d	76.02±1.37	79.36±1.42	65.50±2.30	66.41±2.93	87.96±1.66	85.09±1.82
	ResNet1d + CIF	82.19±1.67	84.60±3.33	75.00±2.34	77.20±2.43	87.94±2.45	85.52±4.06
	XResNet1d	80.46±2.26	85.58±1.79	71.34±3.42	73.42±3.81	90.19±2.13	89.32±2.49
	XResNet1d + CIF	82.96±3.12	85.85±2.18	75.72±4.83	77.86±5.18	92.72±1.19	91.19±1.15
	HMTCN	81.87±1.87	86.50±1.24	73.49±2.90	75.84±3.20	94.20±0.29	93.04±0.45
	HMTCN + CIF	**88.29**±1.45	**90.66**±1.48	**83.21**±2.02	**85.59**±1.96	**94.28**±0.93	**93.78**±1.11

M.3 Fine-tuning Transferred Parameters

For models where the transferred hyperparameters from HMTCN do not yield optimal results, we perform additional fine-tuning. This step investigates the adaptability of CIF-related parameters and explores how they can be optimized for other model structures.

Table 29: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
TDBrain (2-Classes) Reproduced	iTransformer	74.69±1.02	74.76±1.04	74.69±1.02	74.67±1.02	83.35±1.24	83.65±1.41
	iTransformer + CIF(HMTCN)	72.79±1.12	72.89±1.16	72.79±1.12	72.76±1.11	81.08±1.03	81.31±0.93
	iTransformer + CIF(New)	76.10±0.76	76.16±0.75	76.10±0.76	76.09±0.76	84.79±1.15	85.13±1.11
	Nonformer	88.10±2.39	88.76±1.74	88.10±2.39	88.04±2.47	96.56±0.91	96.36±1.21
	Nonformer + CIF(HMTCN)	88.00±2.06	88.91±1.68	88.00±2.06	87.92±2.12	97.10±1.10	97.17±1.09
	Nonformer + CIF (New)	88.67±1.30	89.34±0.93	88.67±1.30	88.61±1.33	96.89±0.31	96.91±0.37

Table 30: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
ADFTD (3-Classes) Reproduced	Autoformer	46.90±2.89	45.59±2.37	44.91±2.23	44.34±2.52	63.49±2.44	45.63±2.29
	Autoformer + CIF((HMTCN))	45.92±2.78	44.81±2.44	44.66±2.25	44.26±2.38	62.69±2.21	44.93±2.68
	Autoformer + CIF(New)	48.49±1.31	46.42±1.07	45.39±1.80	44.94±1.97	63.75±1.46	46.10±1.84
	iTransformer	52.85±1.36	46.97±1.05	47.31±1.03	46.84±0.78	67.46±0.96	49.90±0.89
	iTransformer + CIF(HMTCN)	50.76±0.50	47.11±0.67	47.29±0.59	47.10±0.66	67.00±0.55	49.60±0.69
	iTransformer + CIF(New)	53.93±1.52	48.44±0.89	49.13±1.17	48.56±0.84	68.70±0.94	50.88±1.20
	Reformer	51.28±2.60	49.68±2.75	49.64±2.02	48.45±2.06	69.20±2.53	51.74±3.24
	Reformer + CIF(HMTCN)	52.50±0.91	52.10±2.91	49.01±2.74	47.53±4.23	68.25±1.79	50.44±2.59
	Reformer + CIF(New)	53.11±1.71	52.14±1.42	52.42±1.90	51.73±2.01	70.47±1.64	51.88±1.92
	Medformer	53.70±1.18	51.51±1.32	50.49±1.48	50.35±1.53	70.48±1.17	50.91±1.13
	Medformer + CIF(HMTCN)	55.88±0.82	51.91±1.90	50.80±1.63	50.29±1.92	70.45±1.33	53.73±1.70
	Medformer + CIF(New)	55.21±1.42	52.84±2.36	52.60±2.26	52.43±2.65	72.11±2.42	54.10±3.09

Table 31: **Results of Subject-Independent Setup.** Green cells indicate performance improvement with CIF.

Datasets	Models	Accuracy	Precision	Recall	F1 score	AUROC	AUPRC
PTB (2-Classes) Reproduced	iTransformer	83.43±1.19	88.06±1.47	75.64±1.55	78.29±1.70	91.38±1.41	91.08±1.30
	iTransformer + CIF(HMTCN)	81.62±2.96	87.48±2.03	72.74±4.41	74.99±4.91	90.25±3.29	89.80±3.28
	iTransformer + CIF(New)	83.56±1.15	86.00±1.03	76.78±1.94	79.10±1.87	90.40±0.75	89.80±0.93
	Nonformer	78.93±1.46	82.48±1.53	69.68±2.15	71.50±2.56	90.54±0.59	87.78±1.46
	Nonformer + CIF(HMTCN)	71.89±3.81	71.80±4.58	69.44±3.56	69.74±3.84	70.55±2.96	70.78±4.08
	Nonformer + CIF(New)	80.46±2.20	84.22±2.44	71.79±3.10	73.89±3.50	92.74±0.97	90.30±1.26
	Medformer	80.99±0.75	83.01±0.72	73.35±1.16	75.47±1.21	93.10±1.18	90.69±1.04
	Medformer + CIF(HMTCN)	78.74±0.64	81.11±0.84	75.40±0.66	76.31±0.71	83.20±0.91	83.66±0.92
	Medformer + CIF(New)	83.98±0.81	85.21±0.62	78.00±1.30	80.11±1.22	93.32±0.97	90.97±1.69

For models that did not initially benefit from the direct transfer of CIF hyperparameters optimized for HMTCN, we conducted further fine-tuning of the key CIF parameters. The results demonstrate that, after targeted adjustment, these models also achieved notable performance improvements. This process not only reinforces the adaptability of the CIF architecture across diverse models but also highlights the high tunability and flexibility of its hyperparameters. By appropriately modifying the number of selected channels, fusion direction, and scaling factors, as shown in Tables 25, 26, 27, 28 and Figures 11, 12, 13, 14, CIF can be effectively tailored to different network structures, thereby maximizing its strengths in feature modeling and discriminative capability.

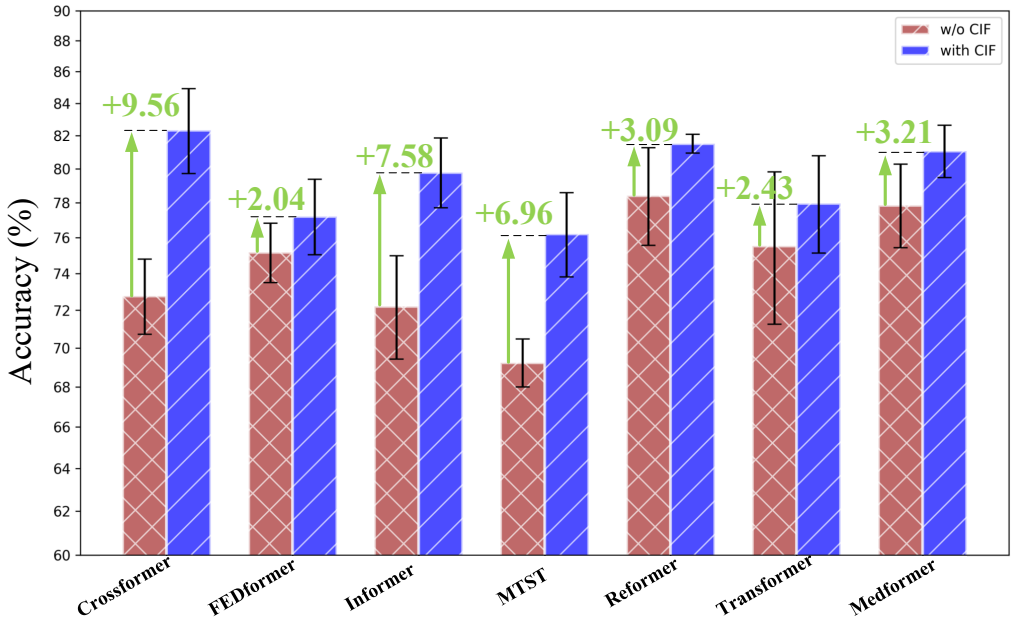


Figure 11: Performance improvement on the APAVA dataset using the CIF method.

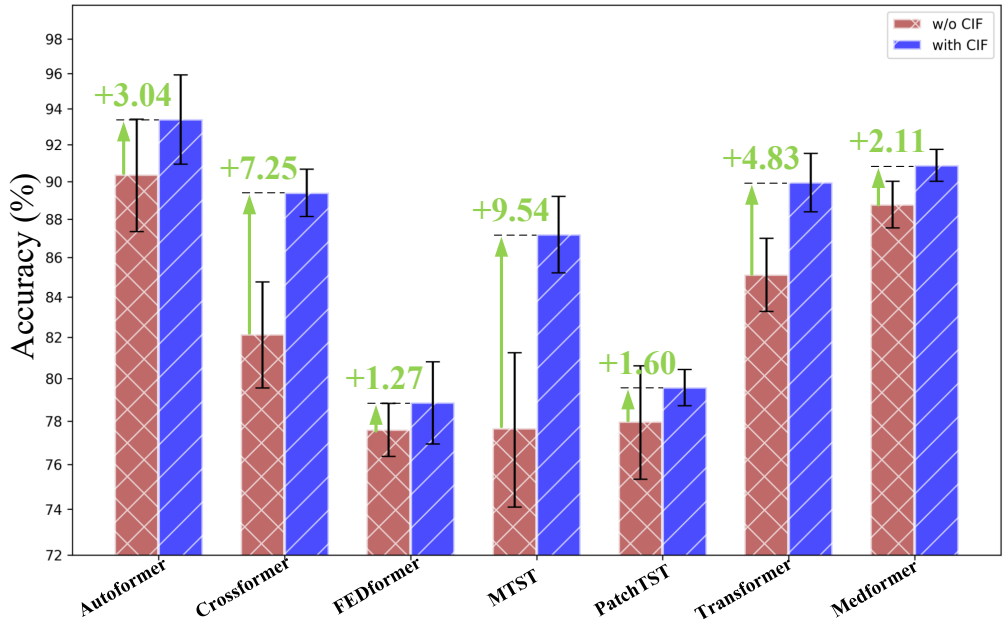


Figure 12: Performance improvement on the TDBRAIN dataset using the CIF method.

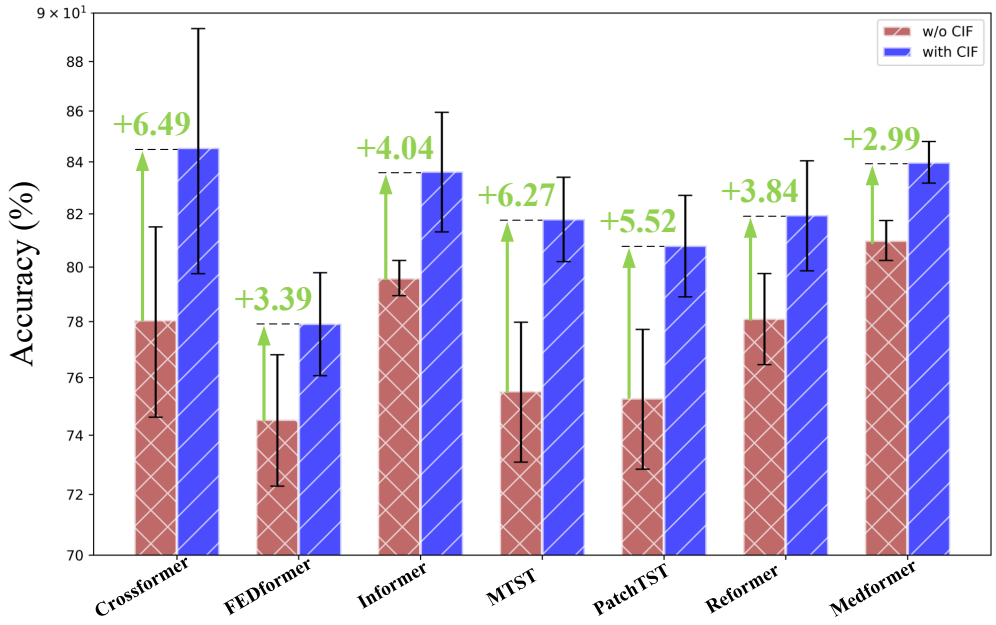


Figure 13: Performance improvement on the PTB dataset using the CIF method.

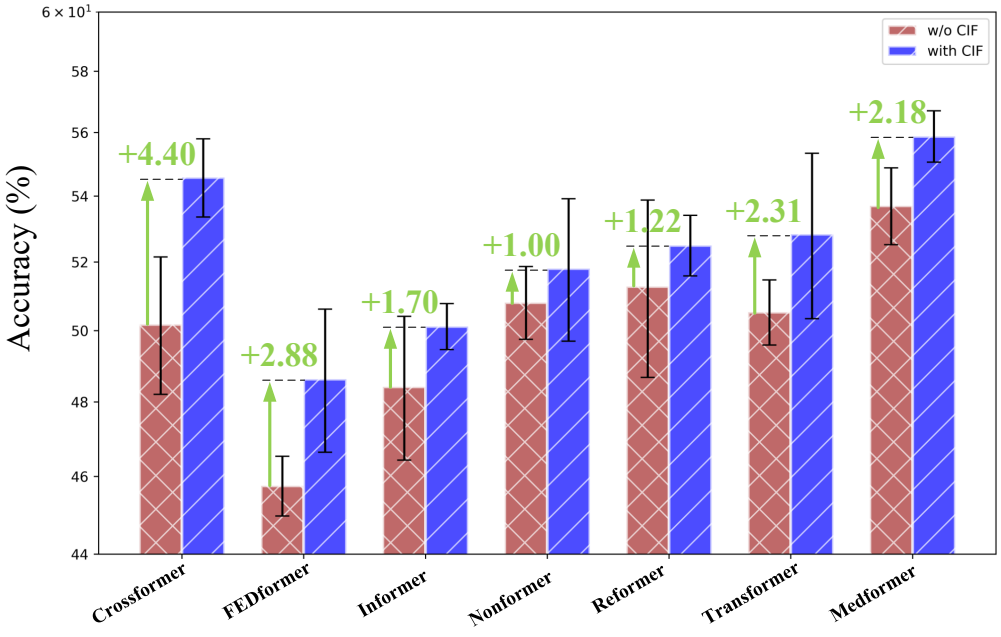


Figure 14: Performance improvement on the ADFTD dataset using the CIF method.

N LIMITATIONS AND FUTURE WORK

Limitations:

Biomedical time series exhibit complex modal characteristics, which lead to significant efficiency bottlenecks when manually adjusting the prior parameters (t, n, a, b) with clear medical interpretations in the CIF model based on empirical experience. This limitation highlights the urgent need for developing novel automated hyperparameter optimization frameworks.

Future Work:

We plan to explore a more universal and generalizable time-series analysis approach, incorporating domain knowledge, structural modeling, and automated hyperparameter optimization. This integration should foster both deeper theoretical insights and stronger practical applicability, providing robust solutions for real-world medical problems.

Furthermore, incorporating domain-specific prior knowledge into medical time-series analysis can more precisely reveal and model relationships between channels. By integrating medical expertise, clinical experience, and existing pathological data, the interpretability and predictive performance of models can be enhanced, thereby supporting clinical decision-making and interventions. On this basis, frequency-domain analysis Hu et al. (2025); Nason & Sachs (1999); Yi et al. (2025) offers an additional perspective: by applying Fourier transform or wavelet decomposition to the signals, physiological features at different frequency components can be identified, revealing patterns that are difficult to capture in the time domain. This is particularly valuable for noise reduction, extraction of periodic signals, and detection of pathological events, and can also provide richer feature representations for model inputs. Future research could further explore how to combine time-domain and frequency-domain information, integrating domain priors to improve the accuracy and robustness of intelligent medical analytics.

Finally, we must acknowledge that the development trends in the field of artificial intelligence highlight the importance of architectural innovation. Future research should focus on designing novel architectures that align more closely with the CIF method, combining the strengths of existing models. For example, the local feature extraction capabilities of CNNs LeCun et al. (1989), the temporal stability of TCNs Bai et al. (2018) for long sequences, the long-term dependency modeling of RNNs Rumelhart et al. (1986) and LSTMs Hochreiter & Schmidhuber (1997), the global modeling efficiency of Transformers Vaswani et al. (2017), the resource-efficient computation of Mamba Gu & Dao (2023), and the hybrid recurrence-attention structure of RWKV Peng et al. (2023). By adapting and integrating these methods, we aim to build a powerful model that is not only deeply compatible with the CIF framework but also capable of efficiently handling complex medical time-series data.

O DETAILED ANALYSIS OF CIF

O.1 SENSITIVITY TO THE FUSION RATIO n

We systematically evaluate the sensitivity of CIF to the fusion ratio n across different datasets. As reported in Tables 22, 23, and 24, performance varies with n, confirming that the fusion ratio is an important hyperparameter. For example, on ADFTD, the accuracy reaches 58.56% at $n = 10$, while suboptimal choices such as $n = 7$ and $n = 11$ reduce the accuracy to below 55%. Across datasets, the optimal values are $n = 9$ for APAVA ($C = 16$), $n = 8$ for PTB ($C = 15$), and $n = 10$ for ADFTD ($C = 19$). Although these values are generally close to $C/2$, the exact optimum is affected by task difficulty, signal quality, and inter-channel correlation structure.

O.2 STABILITY AND STRUCTURAL CONSTRAINTS OF THE LEARNED COEFFICIENTS

Although the fusion coefficients a and b are learned end-to-end, CIF remains fundamentally different from a fully connected linear layer. A linear layer learns an unconstrained $C \times C$ mixing matrix with $O(C^2)$ free parameters, whereas CIF only learns coefficients associated with predefined physiological interactions. Therefore, optimization can adjust the interaction strength but cannot arbitrarily change channel connectivity or channel semantics.

To further examine the stability of the learned coefficients, we evaluate CIF across multiple random seeds. On APAVA, the learned coefficients converge to a narrow range, with representative values around $a \approx -0.81$ and $b \approx -0.607$. This stable convergence suggests that the learned coefficients do not collapse into an unconstrained channel-mixing operator, but instead consistently modulate the predefined physiological interactions.

O.3 Preservation of Asymmetric Channel Information

We further analyze whether CIF suppresses asymmetric physiological patterns. CIF performs channel-preserving pairwise fusion rather than global averaging, and therefore does not impose symmetry across channels. The channel-wise effects are heterogeneous rather than uniform, suggesting that CIF selectively adapts to channel-specific signal characteristics. These observations indicate that CIF can preserve asymmetric physiological information while exploiting meaningful inter-channel relationships.

O.4 Controlled Comparison with Spatial Preprocessing Methods

To isolate the contribution of CIF from the backbone architecture, we fix HMTCN as the backbone and replace only the spatial preprocessing module. We compare CIF with eight representative spatial preprocessing methods, including Bipolar, CAR, CSP, ICA, Laplacian, Linear, SSP, and SSS. As reported in Table 1, CIF achieves the best performance on most dataset–metric combinations, providing an apple-to-apple comparison with alternative spatial preprocessing strategies.

P Broader Impacts

In this study, we take the first step toward a paradigm shift in medical time-series analysis, transitioning from a focus on model architecture improvements to emphasizing the intrinsic structure of the data. Our experimental results, which outperform several state-of-the-art methods across multiple datasets, validate the effectiveness of this approach. We emphasize that uncovering the essential characteristics in medical time-series data is a valuable research direction, particularly given the critical need for model interpretability in medical applications. However, practitioners should be cognizant of the method's limitations.