

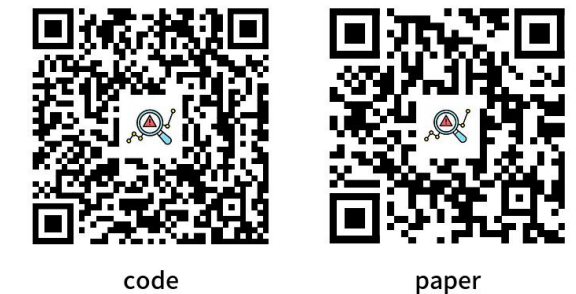


CATCH: Channel-Aware Multivariate Time Series Anomaly Detection via Frequency Patching

Xingjian Wu¹, Xiangfei Qiu¹, Zhengyu Li¹, Yihang Wang¹, Jilin Hu¹, Chenjuan Guo¹, Hui Xiong², Bin Yang^{1*}

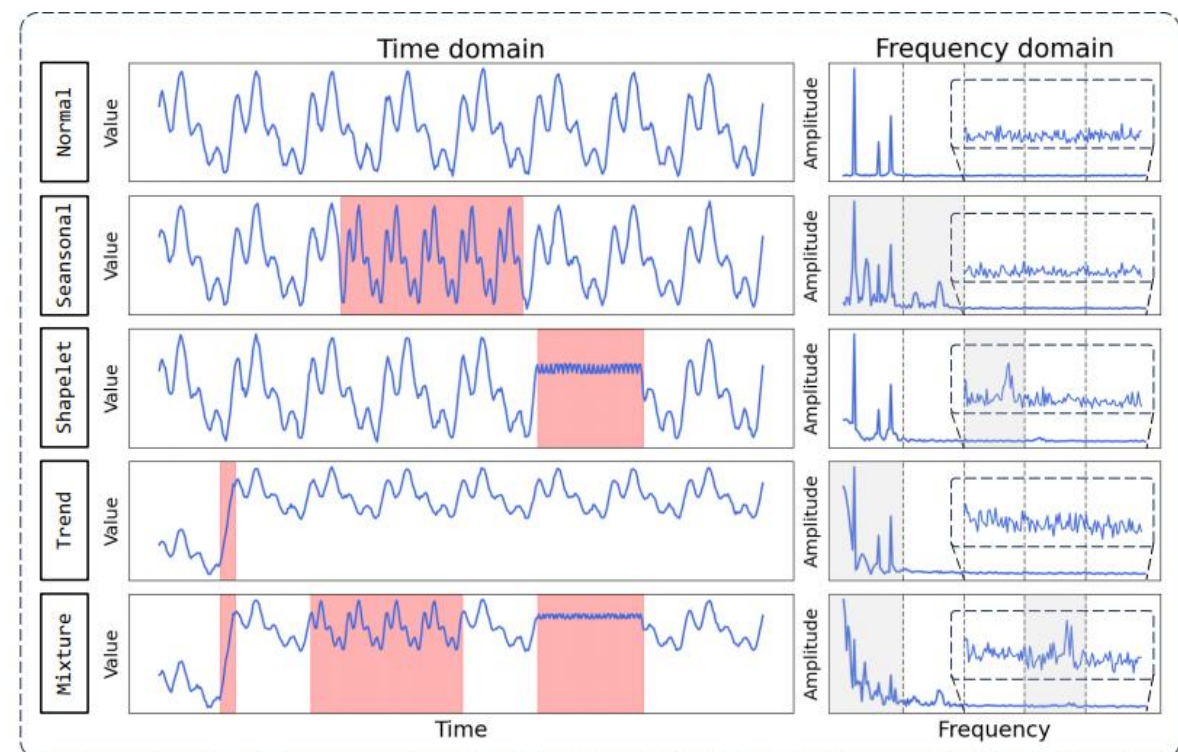
{xjwu,xfqiu,lizhengyu,yhwang}@stu.ecnu.edu.cn, {jlhu,cjguo,byang}@dase.ecnu.edu.cn, xionghui@ust.hk

¹East China Normal University ²The Hong Kong University of Science and Technology (Guangzhou)

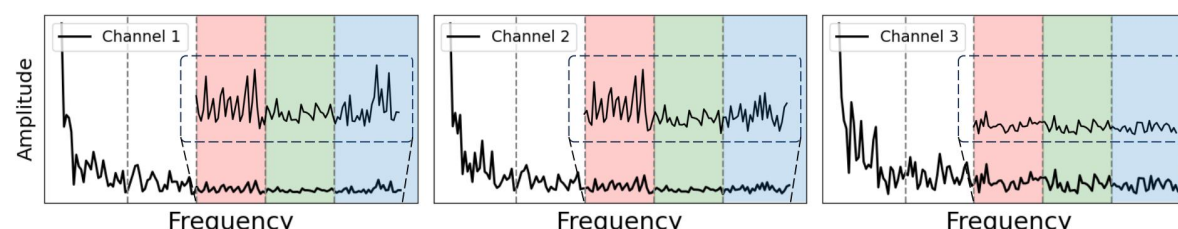


Introduction

Reconstruction-based Anomaly Detection methods achieve promising performance but still suffer from **heterogeneous subsequence anomalies** and **varying channel correlations** in Multivariate Time Series Anomaly Detection tasks.



(a) Different subsequence anomalies



(b) Varying Channel Correlations

Contribution

- We propose a general framework called **CATCH**, which enables simultaneous detection of heterogeneous point and subsequence anomalies via **frequency patch learning**. The framework enhances subsequence anomaly detection through frequency-domain patching and integrates **fine-grained adaptive channel correlations** across frequency bands.
- We design the Channel Fusion Module to fully utilize the fine-grained channel correlations. Driven by a bi-level multi-objective optimization algorithm, the Channel Fusion Module is able to iteratively discover appropriate channel correlations and facilitate the isolation of irrelevant channels and the clustering of relevant channels, which **provides both the capacity and robustness**.
- We conduct extensive experiments on 22 multivariate datasets. The results show that CATCH outperforms state-of-the-art baselines.

My WeChat:

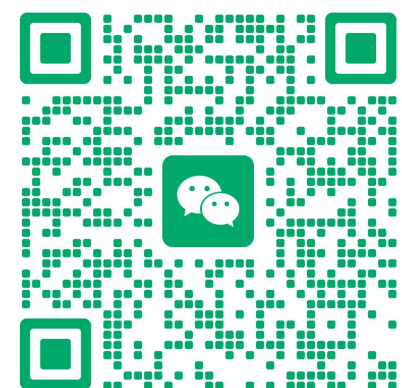
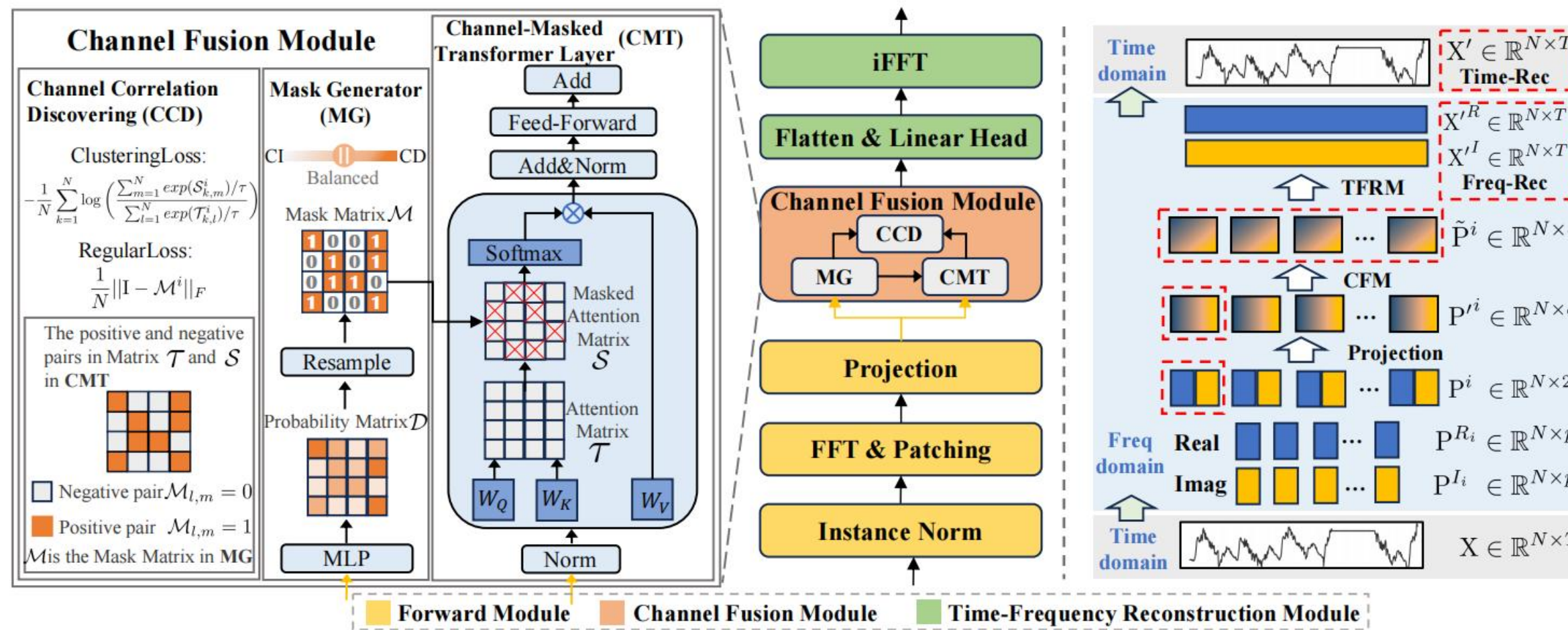


Table 1: Comparison of existing frequency-based TSAD methods

Property	Mmultivariate time serie anomaly detection	Time-frequency granularity alignment	Handle high-frequency information	Capture channel correlations
SR-CNN	✗	✗	✗	✗
PFT	✗	✗	✗	✗
TFAD	✓	✗	✗	✗
Dual-TF	✓	✓	✗	✗
CATCH	✓	✓	✓	✓

CATCH Framework



Forward Module

We utilize FFT to transform time series into orthogonal trigonometric signals in the frequency domain and apply the **frequency patching** operation to **create fine-grained frequency patches**.

Channel Fusion Module

Mask Generator (MG): By generating a binary mask matrix, it can perceive the association of each frequency band's channels, and **isolates the negative impacts of irrelevant channels**.

$$\mathcal{D}^i = \sigma(\text{Linear}(\mathbf{P}^i)), \mathcal{M}^i = \text{Resample}(\mathcal{D}^i)$$

Channel-Masked Transformer Layer (CMT): After the binary mask is generated by the Mask Generator, fine-grained channel correlations are further captured using Masked Transformer layers. **The gradient is kept through calculation in the attentional mechanism:**

$$\mathbf{Q}^i = \mathbf{P}^{*i} \cdot \mathbf{W}^Q, \mathbf{K}^i = \mathbf{P}^{*i} \cdot \mathbf{W}^K, \mathbf{V}^i = \mathbf{P}^{*i} \cdot \mathbf{W}^V,$$

$$\mathcal{T}^i = \mathbf{Q}^i \cdot (\mathbf{K}^i)^T, \mathcal{S}^i = \mathcal{T}^i \odot \mathcal{M}^i + (1 - \mathcal{M}^i) \odot (-\infty),$$

$$\text{MaskedScores}^i = \mathcal{S}^i / \sqrt{d}, \tilde{\mathbf{P}}^i = \text{Softmax}(\text{MaskedScores}^i) \cdot \mathbf{V}^i$$

Channel Correlation Discovering (CCD): ClusteringLoss is proposed by assuming that the channel correlation generated by the Mask Generator is locally optimal, where relevant and irrelevant channels are determined. This **encourages the attention mechanism to aggregate relevant channels**. RegularLoss is used to **limit the number of relevant channels**, preventing the Mask Generator from outputting a constant '1' matrix.

$$\text{ClusteringLoss} = -\frac{1}{N} \sum_{k=1}^N \log \left(\frac{\sum_{m=1}^N \exp(\mathcal{S}_{k,m}^i / \tau)}{\sum_{l=1}^N \exp(\mathcal{T}_{k,l}^i / \tau)} \right), \text{RegularLoss} = \frac{1}{N} \|\mathbf{I} - \mathcal{M}^i\|_F$$

Time-Frequency Reconstruction Module

After extracting the fine-grained channel correlations in each frequency band, the time series is reconstructed in both the frequency and time domain to detect heterogeneous anomalies.

$$\text{RecLoss}^{\text{time}} = \|\mathbf{X} - \mathbf{X}'\|_F^2, \text{RecLoss}^{\text{freq}} = \|\mathbf{X}^R - \mathbf{X}'^R\|_1 + \|\mathbf{X}^I - \mathbf{X}'^I\|_1$$

Experiments

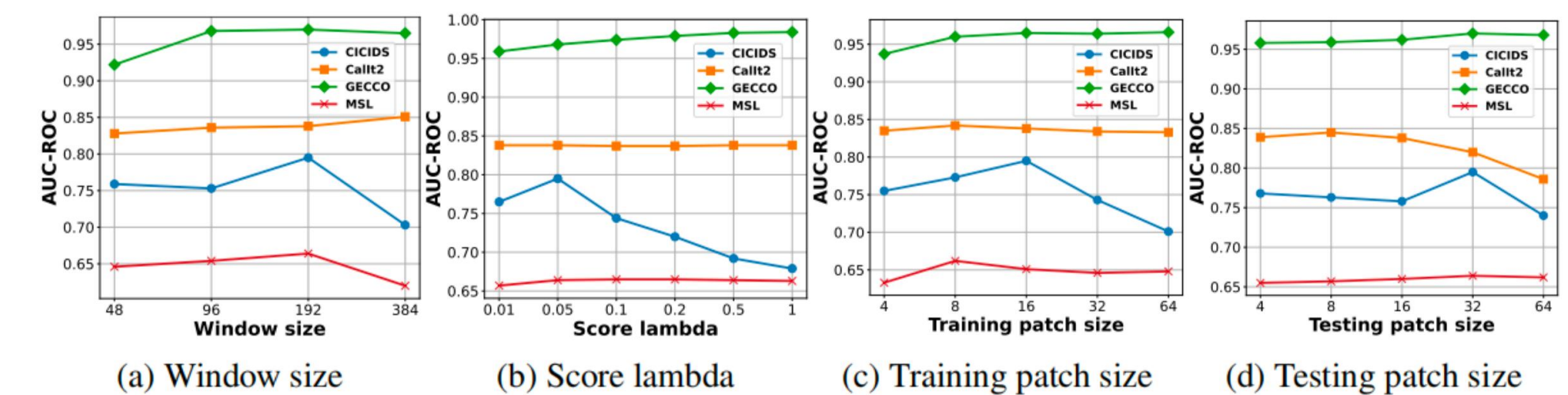
Main Results

Table 2: Average A-R (AUC-ROC) and aff-F (Affiliated-F1)

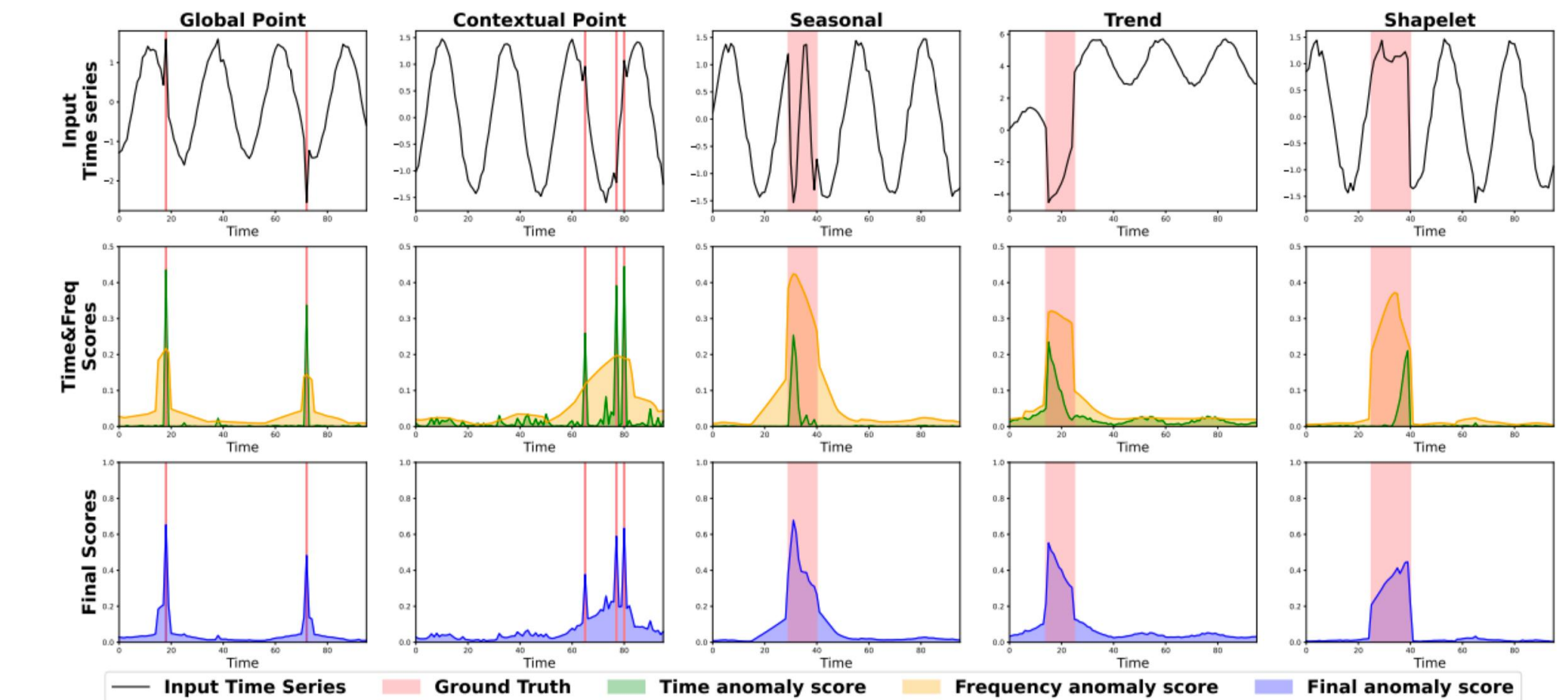
Dataset	Metric	CATCH	Modern	iTrans	DualTF	ATrans	DC	TsNet	Patch	DLin	NLin	AE	Ocsvm	IF	PCA	HBOS	TFAD
CICIDS	Aff-F	0.787	0.654	0.708	0.692	0.560	0.664	0.657	0.660	0.669	0.669	0.243	0.693	0.604	0.619	0.542	0.579
	A-R	0.795	0.697	0.692	0.603	0.528	0.638	0.732	0.716	0.751	0.691	0.629	0.537	0.787	0.601	0.760	0.504
	Aff-F	0.835	0.780	0.812	0.751	0.729	0.697	0.794	0.793	0.757	0.757	0.587	0.783	0.402	0.768	0.756	0.744
	A-R	0.838	0.676	0.791	0.574	0.533	0.527	0.771	0.808	0.752	0.695	0.767	0.804	0.775	0.790	0.756	0.504
Callit2	Aff-F	0.750	0.744	0.713	0.663	0.650	0.632	0.744	0.746	0.738	0.642	0.561	0.714	0.634	0.710	0.695	0.600
	A-R	0.958	0.957	0.934	0.703	0.552	0.504	0.957	0.957	0.954	0.948	0.909	0.953	0.860	0.871	0.951	0.500
	Aff-F	0.908	0.893	0.839	0.701	0.782	0.687	0.894	0.906	0.893	0.882	0.823	0.666	0.424	0.785	0.708	0.627
	A-R	0.970	0.952	0.795	0.714	0.516	0.555	0.954	0.949	0.947	0.936	0.769	0.804	0.619	0.711	0.557	0.499
GECCO	Aff-F	0.896	0.833	0.891	0.810	0.856	0.776	0.864	0.856	0.856	0.829	0.854	0.677	0.788	0.814	0.721	0.535
	A-R	0.974	0.676	0.690	0.937	0.947	0.659	0.913	0.685	0.696	0.755	0.931	0.733	0.549	0.815	0.897	0.497
	Aff-F	0.970	0.726	0.710	0.588	0.692	0.694	0.734	0.724	0.725	0.723	0.625	0.641	0.584	0.678	0.680	0.665
	A-R	0.664	0.633	0.611	0.576	0.508	0.507	0.613	0.637	0.624	0.592	0.562	0.524	0.524	0.552	0.574	0.500
MSL	Aff-F	0.994	0.769	0.684	0.768	0.853	0.862	0.794	0.776	0.828	0.819	0.689	0.667	0.648	0.680	0.675	0.689
	A-R	0.816	0.466	0.640	0.633	0.671	0.549	0.791	0.709	0.768	0.671	0.504	0.456	0.475	0.666	0.446	0.502
	Aff-F	0.859	0.825	0.854	0.725	0.710	0.682	0.842	0.831	0.831	0.843	0.707	0.531	0.620	0.702	0.658	0.628
	A-R	0.652	0.593	0.592	0.600	0.514	0.501	0.592	0.586	0.580	0.585	0.650	0.619	0.542	0.648	0.620	0.500
PSM	Aff-F	0.847	0.840	0.827	0.679	0.724	0.675	0.831	0.845	0.841	0.844	0.439	0.742	0.626	0.738	0.629	0.660
	A-R	0.811	0.722	0.745	0.631	0.508	0.502	0.727	0.736	0.728	0.738	0.774	0.602	0.664	0.679	0.626	0.500
	Aff-F	0.804	0.782	0.780	0.605	0.674	0.702	0.800	0.777	0.782	0.766	0.731	0.617	0.781	0.656	0.669	0.630
	A-R	0.824	0.692	0.759	0.579	0.506	0.520	0.805	0.760	0.739	0.690	0.704	0.688	0.618	0.656	0.603	0.502
Contextual	Aff-F	0.823	0.619	0.802	0.635	0.601	0.597	0.666	0.766	0.780	0.700	0.755	0.696	0.679	0.475	0.481	0.569
	A-R	0.910	0.562	0.905	0.598	0.546	0.525	0.908	0.854	0.700	0.530	0.896	0.711	0.821	0.538	0.464	0.504
	Aff-F	0.949	0.748	0.922	0.649	0.656	0.567	0.910	0.940	0.928	0.808	0.919	0.849	0.912	0.704	0.528	0.566
	A-R	0.997	0.873	0.976	0.595	0.564	0.514	0.989	0.992	0.979	0.675	0.996	0.996	0.938	0.758	0.608	0.500
Seasonal	Aff-F	0.997	0.681	0.992	0.776	0.788	0.989	0.992	0.993	0.951	0.927	0.805	0.938	0.637	0.673	0.686	0.500
	A-R	0.998	0.512	0.946	0.701	0.584	0.644	0.958	0.922	0.823	0.623	0.949	0.829	0.918	0.437	0.516	0.502
	Aff-F	0.985	0.675	0.961	0.692	0.699	0.737	0.941	0.933	0.961	0.759	0.871	0.771	0.887	0.683	0.640	0.684
	A-R	0.970	0.522	0.864	0.573	0.519	0.597	0.877	0.818	0.684	0.563	0.865	0.655	0.748	0.517	0.337	0.503
Shapelet	Aff-F	0.916	0.734	0.901	0.677	0.584	0.765	0.897	0.888	0.721	0.830	0.699	0.691	0.914	0.693	0.669	0.642
	A-R	0.892	0.612	0.847	0.524	0.300	0.569	0.858	0.835	0.671	0.642	0.482	0.471	0.878	0.484	0.468	0.502
	Aff-F	0.892	0.856	0.862	0.652	0.641	0.709	0.863	0.879	0.727	0.839	0.673	0.676	0.881	0.676	0.667	0.710
	A-R	0.931	0.763	0.854	0.570	0.522	0.516	0.861	0.863	0.767	0.749	0.493	0.475	0.911	0.517	0.531	0.501

- CATCH consistly **achieves state-of-the-art performance** on 10 real-world datasets and 12 synthetic datasets.
- Resources: <https://github.com/decisionintelligence/catch>

Parameter Sensitivity



Visualization



- CATCH shows outstanding performance in detecting both the point and subsequence anomalies.