Representation Stability in High-Dimensional Noisy Time Series via Koopman-Based Features

Anonymous Authors[1]

Abstract

Learning robust representations from high-dimensional noisy physiological time series remains challenging under patient variability, class imbalance, and signal perturbation. In this work, we investigate Koopman operator theory and Extended Dynamic Mode Decomposition (EDMD) as structured spectral representations for ECG time-series learning under patient-wise PTB-XL classification. Rather than proposing a new state-of-the-art classifier, we perform a systematic analysis of Koopman-only, deep neural, and hybrid spectral–deep representations, focusing on predictive performance, robustness, interpretability, and representation stability. Deep neural models achieve the strongest predictive accuracy, while hybrid spectral–deep representations slightly improve Macro-F1, suggesting that Koopman spectral descriptors provide complementary information for class-balanced learning. Although Koopman-only representations remain weaker in predictive performance, they capture meaningful temporal structure through compact and interpretable spectral descriptors. Robustness experiments further reveal that recomputed EDMD-based spectral representations become unstable under noisy perturbation, producing substantial eigenspectrum drift and degradation of hybrid representations. These findings highlight an important limitation of fixed spectral feature extraction in high-dimensional physiological settings. Overall, our results suggest that Koopman-inspired spectral representations are most useful as complementary analytical tools for studying temporal organization, robustness, and interpretability rather than as direct replacements for deep neural representations in physiological time-series learning.

[1]. Correspondence to: <>.

ICML.

1. Introduction

Learning robust representations from high-dimensional noisy physiological time series remains challenging due to measurement noise, class imbalance, and cross-entity variability commonly observed in clinical signals (Raghu et al., 2019; Goldberger et al., 2000). Deep learning approaches, including convolutional and transformer-based architectures, achieve strong predictive performance for time-series analysis (Vaswani et al., 2017; Zerveas et al., 2021; Wang et al., 2017), but their learned representations are often difficult to interpret and sensitive to perturbations and distribution shifts (Goodfellow et al., 2015; Bengio et al., 2013). An alternative perspective models temporal signals as realizations of an underlying dynamical system. Koopman operator theory represents nonlinear dynamics through linear evolution in a lifted function space (Koopman, 1931; Mezić, 2005). Data-driven methods such as Dynamic Mode Decomposition (DMD) and Extended Dynamic Mode Decomposition (EDMD) extract compact spectral descriptors including eigenvalues, modes, and temporal coefficients (Schmid, 2010b; Williams et al., 2015a), providing interpretable views of temporal organization that are difficult to obtain from standard deep embeddings alone. In this work, we investigate Koopman-based spectral representations for physiological time-series learning under patient-wise PTB-XL classification. Rather than proposing a new state-of-the-art classifier, we perform a systematic analysis of Koopman-only, deep neural, and hybrid spectral–deep representations, focusing on predictive performance, robustness under additive perturbation, interpretability, and representation stability in noisy physiological settings. Our experiments reveal a clear stability–performance tradeoff. Deep neural models achieve the strongest predictive accuracy, whereas hybrid spectral–deep representations slightly improve Macro-F1, suggesting that Koopman spectral descriptors provide complementary information for class-balanced classification. Although Koopman-only representations remain weaker in predictive performance, they still capture meaningful temporal structure through compact and interpretable spectral descriptors. We further show that recomputed EDMD-based spectral representations become unstable under noisy perturbation, producing eigenspectrum drift, fragmented representation geometry, and degradation of hybrid representa-

tions. Overall, our findings suggest that Koopman-inspired spectral representations are most useful as complementary analytical tools for studying temporal organization, robustness, and interpretability rather than as direct replacements for deep neural representations in physiological time-series learning.

1.1. Related Work

Deep neural architectures, including convolutional, recurrent, and transformer-based models, achieve strong performance for time-series representation learning, but often learn implicit representations that remain sensitive to noise, limited samples, and cross-entity distribution shifts in high-dimensional settings. Recent self-supervised and contrastive learning approaches, including SimCLR and supervised contrastive learning, further improve representation quality under limited supervision (Chen et al., 2020; Khosla et al., 2020). However, these methods depend heavily on augmentation design, which can distort physiological temporal structure and does not explicitly model temporal dynamics. Classical signal-processing methods such as wavelet-based time–frequency analysis provide interpretable multi-scale representations for physiological signals (Mallat, 1999), but are not designed to capture underlying temporal system evolution. An alternative perspective treats time-series data as observations of a dynamical system. Koopman operator theory models nonlinear dynamics through linear evolution in a lifted function space (Koopman, 1931; Mezić, 2005). Data-driven methods including Dynamic Mode Decomposition (DMD) and Extended Dynamic Mode Decomposition (EDMD) extract spectral descriptors directly from trajectories (Schmid, 2010a; Williams et al., 2015b), with applications across fluid dynamics, control, and machine learning (Brunton et al., 2016). More recent studies integrate Koopman theory with neural representation learning via learned lifting functions or latent linear dynamics (Lusch et al., 2018; Takeishi et al., 2017). Unlike prior work, we study Koopman-based spectral representations as structured projections for noisy high-dimensional physiological time series, emphasizing robustness, interpretability, and patient-wise generalization rather than predictive performance alone.

2. Method

2.1. High-Dimensional Dynamical Representation

We model multivariate physiological time series as high-dimensional observations of an underlying dynamical system. Let $x_t \in \mathbb{R}^d$ denote the system state at time t. Rather than relying purely on discriminative feature learning, we adopt a Koopman-based dynamical systems perspective in which nonlinear temporal evolution is approximated by a linear operator acting in a lifted function space. Given a nonlinear lifting function

$$\phi : \mathbb{R}^d \to \mathbb{R}^m,$$

we estimate a linear operator K satisfying

$$\phi(x_{t+1}) \approx K\phi(x_t).$$

To estimate the Koopman operator, we use Extended Dynamic Mode Decomposition (EDMD) (Lehmberg et al., 2020). See Appendix A for details.

From the estimated operator, we extract compact spectral descriptors including eigenvalues, eigenvalue magnitudes/phases, singular values, and Koopman coefficient (that is, scalar or vector values that appear in the spectral decomposition of the Koopman operator) statistics. These quantities form interpretable spectral embeddings summarizing dominant temporal dynamics. In practice, we use a reduced-order EDMD approximation with identity observables to isolate the behavior of explicit spectral descriptors without introducing additional nonlinear lifting complexity. The Koopman operator is estimated independently for each trajectory using truncated singular value decomposition, and compact spectral descriptors are extracted for downstream representation analysis.

2.2. Representation as Structured Projection

Koopman spectral embeddings can be interpreted as structured projections of high-dimensional trajectories into a temporally organized representation space. Unlike arbitrary latent embeddings learned purely through discriminative optimization, the Koopman representation is constrained by spectral consistency and temporal evolution. This induces two desirable properties: **(1) Noise attenuation:** temporally incoherent perturbations contribute weakly to dominant spectral modes, suppressing high-frequency noise. **(2) Low-rank organization:** the effective representation is governed by a relatively small number of dominant spectral components, yielding compact and structured embeddings. However, fixed spectral approximations may become unstable under aggressive perturbation, particularly when noise significantly alters the estimated eigenspectrum.

2.3. Representation Regimes

We evaluate three representation paradigms: **1. Koopman-only representation:** spectral descriptors are used directly for classification using lightweight classifiers. **2. Deep neural representation:** a convolutional neural network is trained directly on raw multivariate time-series inputs. **3. Hybrid spectral–deep representation:** Koopman spectral features are concatenated with deep neural embeddings to analyze complementary effects between structured dynamical information and learned representations. This setup

enables a controlled comparison between explicit dynamical inductive bias and purely data-driven representation learning. Figure 1 presents the overall workflow of the proposed framework. Multichannel ECG signals are first preprocessed through normalization and segmentation. Koopman spectral representations are then extracted using EDMD, producing compact descriptors based on eigenvalues, modes, and spectral statistics. In parallel, deep neural representations are learned directly from raw trajectories using a CNN encoder. We evaluate three representation regimes: Koopman-only, deep neural, and hybrid spectral–deep representations obtained through feature concatenation. Models are trained and evaluated under patient-wise splits, followed by robustness analysis under additive Gaussian perturbations to study representation stability and degradation. *Complete implementation details, including EDMD observables, spectral truncation strategy, Koopman feature dimensions, and hybrid representation construction, are provided in Appendix K for reproducibility*.

3. Experimental Setup

Dataset and Task. We evaluate the proposed framework on the PTB-XL dataset (Wagner et al., 2020), a large-scale benchmark for electrocardiography (ECG) analysis. Each sample is represented as a multivariate time series $X \in \mathbb{R}^{C\times T}$, where $C = 12$ denotes ECG leads and T denotes the temporal dimension. Following standard PTB-XL superclass aggregation, we perform five-class ECG classification using the categories: Normal rhythm (NORM), Myocardial infarction (MI), ST/T changes (STTC), Conduction disturbance (CD), and Arrhythmia (ARR). Evaluation is performed under patient-wise splits to avoid subject leakage and ensure cross-entity generalization. Additional dataset details are provided in Appendix D.

Representation Regimes We compare three representation paradigms: **(1) Koopman-only representation:** spectral descriptors extracted from EDMD are used with lightweight classifiers. **(2) Deep neural representation:** a lightweight 1D convolutional neural network is trained directly on raw ECG trajectories. **(3) Hybrid spectral–deep representation:** Koopman spectral features are concatenated with learned deep embeddings.

Evaluation Protocol All signals are standardized using train-set z-score normalization. We evaluate models using classification accuracy, macro-F1, weighted-F1, Matthews Correlation Coefficient (MCC), and AUROC. Macro-F1 is emphasized due to class imbalance. To analyze robustness, we evaluate model behavior under additive Gaussian perturbations:

$$x_{\text{noisy}} = x + \mathcal{N}(0, \sigma^2),$$

where $\sigma \in \{0.0, 0.1, 0.2, 0.5\}$. Additional preprocessing details, Koopman feature construction, implementation settings, and training hyperparameters are provided in Appendix D and Appendix E.

Table 1. Predictive performance on PTB-XL under patient-wise evaluation. Macro-F1 is emphasized because of class imbalance.

Method	Accuracy (%)	Macro-F1 (%)	Weighted-F1 (%)	MCC
Koopman-LR	39.39	30.82	40.94	0.184
Koopman-RF	52.27	27.18	43.48	0.248
CNN	**75.9**	60.29	**74.53**	**0.6473**
Hybrid	75.67	**61.17**	73.30	0.6434

Training Dynamics Figure 2 shows the optimization behavior of the hybrid CNN–Koopman model during training. The training loss decreases steadily, while validation accuracy and Macro-F1 stabilize after approximately 25–30 epochs. Early stopping therefore prevents overfitting while preserving stable generalization performance. The validation Macro-F1 exhibits moderate fluctuations despite relatively stable accuracy, reflecting the class imbalance and patient-wise variability of PTB-XL. This further motivates the use of Macro-F1 as a primary evaluation metric.

4. Results and Discussion

We evaluate Koopman-based spectral representations for patient-wise ECG classification on PTB-XL, focusing on predictive performance, robustness, and representation stability under perturbation.

4.1 Predictive Performance. Table 1 compares Koopman-only, deep neural, and hybrid spectral–deep representations. The CNN baseline achieves the highest overall accuracy, whereas the hybrid model achieves the best Macro-F1, indicating improved class-balanced behavior. Additional metrics are provided in Appendix 3. The gap between accuracy and Macro-F1 highlights strong class imbalance and weaker minority-class recognition. While Koopman-only representations are less competitive than deep models, they still capture meaningful discriminative structure using compact spectral descriptors. The hybrid model slightly improves Macro-F1, suggesting that Koopman features provide complementary temporal information rather than replacing deep embeddings.

4.2 Class-wise Error Analysis. Figure 3 shows the normalized confusion matrix of the CNN baseline. The majority NORM class is recognized well (recall 0.92), whereas ARR remains difficult (recall 0.09) and is frequently confused with NORM. This explains the discrepancy between high accuracy and lower Macro-F1 and motivates emphasizing Macro-F1 under patient-wise PTB-XL evaluation.

Ablation of spectral feature groups. Coefficient-based Koopman statistics provided the strongest complementary improvement in Macro-F1 and MCC, whereas full spectral concatenation slightly reduced performance. Complete

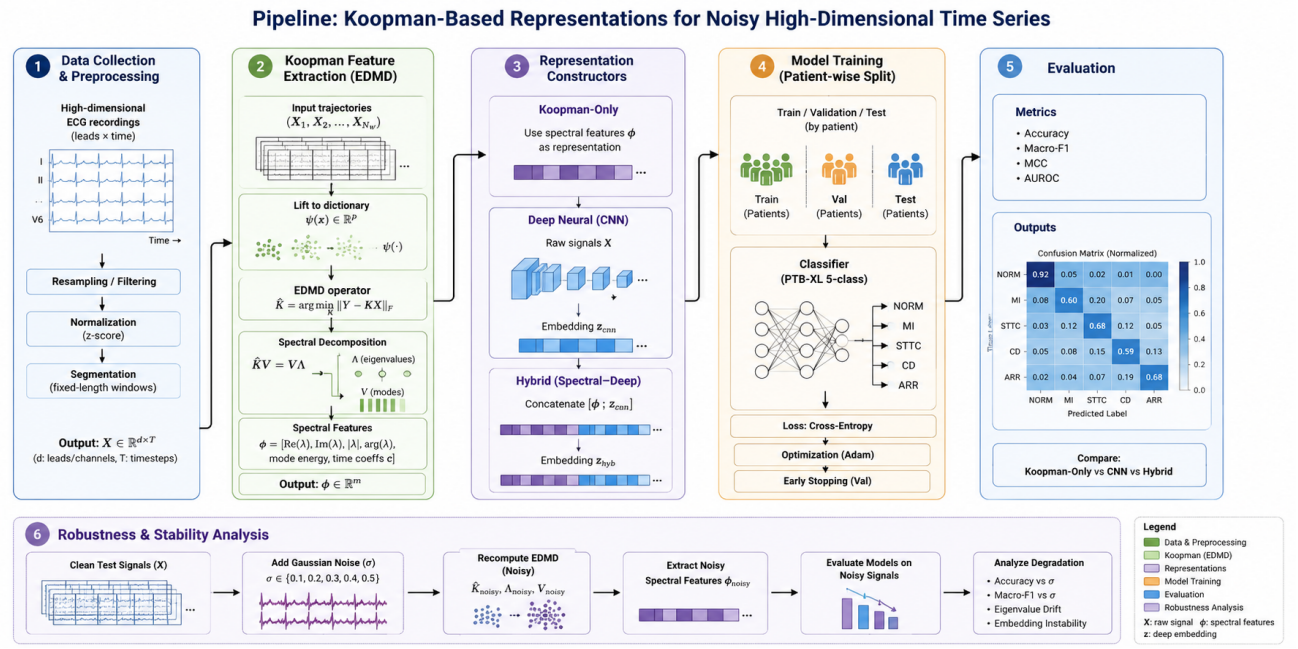

Figure 1. Overview of the proposed Koopman-based representation learning framework for noisy high-dimensional physiological time series. The pipeline combines EDMD-based spectral feature extraction with deep neural representations and evaluates robustness under patient-wise perturbation analysis.

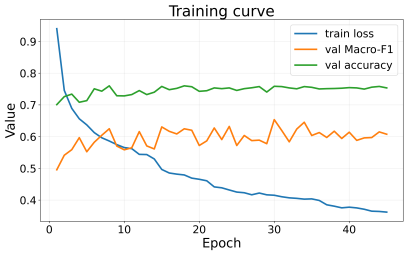

Figure 2. Training dynamics of the hybrid CNN–Koopman model on PTB-XL. Training loss decreases steadily, while validation Macro-F1 stabilizes after approximately 30 epochs.

Table 2. Noise robustness under additive Gaussian perturbation.

Metric	$\sigma = 0.0$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.5$
CNN Accuracy	75.90	75.07	73.35	46.01
CNN Macro-F1	60.29	59.13	57.63	37.86

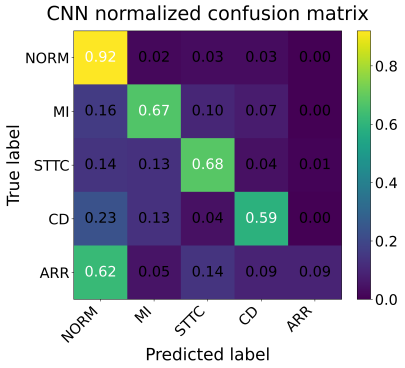

Figure 3. Normalized confusion matrix of the CNN baseline on PTB-XL. The model performs well on the majority NORM class but struggles with the minority ARR class.

ablation results are provided in Appendix F.

4.3 Robustness Under Noise Perturbation. To evaluate robustness, we apply additive Gaussian perturbations of increasing magnitude. Table 2 summarizes the degradation behavior. The CNN baseline remains relatively stable under mild perturbation but degrades substantially at higher noise levels. In contrast, the hybrid representation becomes unstable when Koopman features are recomputed from noisy signals, often collapsing toward degenerate predictions. This suggests that EDMD-derived spectral features are highly sensitive to perturbation and depend strongly on stable operator estimation. Figure 4 illustrates this behavior. While the CNN degrades gradually, the hybrid model collapses under noisy spectral recomputation. Figure 5 further explains the instability. Even moderate perturbation substantially shifts the estimated Koopman eigenspectrum, indicating that spectral descriptors are sensitive to corruption in high-dimensional ECG trajectories. Figure 6 visualizes the eigenspectrum under clean and noisy conditions. Strong perturbation distorts the dominant spectral modes and increases dispersion, explaining the instability observed in the hybrid representation.

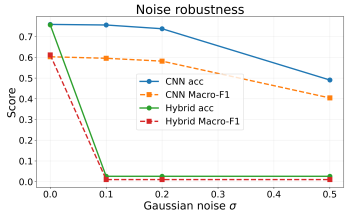

Figure 4. Noise robustness under additive Gaussian perturbation. The CNN degrades gradually, while the hybrid model becomes unstable when Koopman features are recomputed from noisy signals.

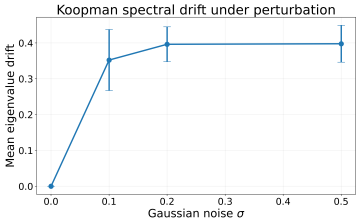

Figure 5. Koopman spectral drift under perturbation. Increasing noise causes substantial eigenspectrum deformation, indicating instability of EDMD-derived spectral features.

4.4 Representation Geometry. Figure 7 visualizes CNN and hybrid embeddings using t-SNE. The CNN embeddings exhibit partial class separation, whereas the hybrid representation shows stronger large-scale organization induced by Koopman spectral descriptors. However, under strong perturbation the hybrid geometry becomes fragmented, reflecting instability inherited from the spectral component.

4.5 Limitations. Koopman-only representations remain weaker than deep neural models for predictive accuracy, suggesting that fixed spectral descriptors alone may not capture highly nonlinear discriminative structure. Moreover, perturbation experiments show that EDMD-derived spectral features can become unstable under noisy conditions. Overall, Koopman representations appear useful as complementary tools for analyzing temporal structure, robustness, and interpretability in high-dimensional physiological time-series learning. *Our goal is not to establish state-of-the-art ECG classification performance, but rather to analyze representation stability and spectral behavior across controlled representation regimes.*

5. Conclusion

In this work, we investigated Koopman-based spectral representations for high-dimensional physiological time-series learning using Extended Dynamic Mode Decomposition (EDMD). We compared Koopman-only, deep neural, and hybrid spectral–deep representations under patient-wise PTB-XL classification, focusing on predictive performance, robustness, interpretability, and representation stability. Our

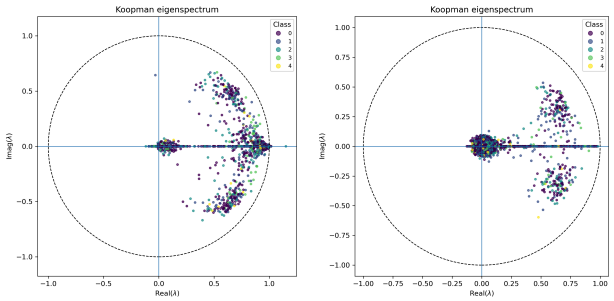

Figure 6. Koopman eigenspectrum under clean and noisy conditions. Noise substantially distorts the estimated spectral structure.

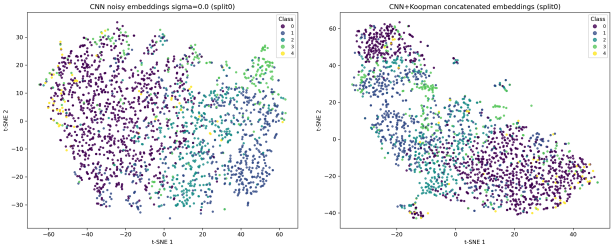

Figure 7. t-SNE visualization of CNN (left) and hybrid CNN–Koopman (right) embeddings. The hybrid representation exhibits stronger global organization induced by spectral descriptors.

results reveal a stability–performance tradeoff. Deep neural models achieved the strongest predictive accuracy, while hybrid spectral–deep representations slightly improved Macro-F1, suggesting that Koopman spectral descriptors provide complementary information for class-balanced classification. Although Koopman-only representations were weaker in predictive performance, they captured meaningful temporal structure through compact and interpretable spectral descriptors. Robustness experiments further exposed limitations of fixed EDMD-based spectral extraction. Under noisy perturbations, recomputed Koopman representations became unstable, producing eigenspectrum drift and degradation of hybrid representations. These findings indicate that Koopman spectral features are highly sensitive to perturbation quality and operator estimation in noisy physiological trajectories. Overall, our analysis suggests that Koopman representations are most useful as complementary analytical tools for studying temporal organization, robustness, and interpretability rather than as direct replacements for deep neural representations. Future work will investigate adaptive lifting functions, more robust Koopman estimation, and integration with self-supervised and contrastive learning for stable spectral representations under cross-subject variability and noise.

References

Bengio, Y., Courville, A., and Vincent, P. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (8):1798–1828, 2013.

Brunton, S. L., Brunton, B. W., Proctor, J. L., and Kutz, J. N. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE*, 11(2):e0150171, 2016.

Chen, T., Kornblith, S., Norouzi, M., and Hinton, G. A simple framework for contrastive learning of visual representations. In *ICML*, 2020.

Goldberger, A. L. et al. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220, 2000.

Goodfellow, I., Bengio, Y., and Courville, A. Deep learning. *Nature*, 521(7553):436–444, 2015.

Khosla, P., Teterwak, P., Wang, C., Sarna, A., Tian, Y., Isola, P., Maschinot, A., Liu, C., and Krishnan, D. Supervised contrastive learning. In *NeurIPS*, 2020.

Koopman, B. O. Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931.

Lehmberg, D., Dietrich, F., Köster, G., and Bungartz, H.-J. datafold: data-driven models for point clouds and time series on manifolds. *Journal of Open Source Software*, 5(51):2283, 2020. doi: 10.21105/joss.02283. URL https://doi.org/10.21105/joss.02283.

Lusch, B., Kutz, J. N., and Brunton, S. L. Deep learning for universal linear embeddings of nonlinear dynamics. In *Nature Communications*, 2018.

Mallat, S. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

Mezić, I. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41 (1):309–325, 2005.

Raghu, M., Zhang, C., Kleinberg, J., and Bengio, S. Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, 2019.

Schmid, P. J. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656: 5–28, 2010a.

Schmid, P. J. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656: 5–28, 2010b.

Takeishi, N., Kawahara, Y., and Yairi, T. Learning koopman invariant subspaces for dynamic mode decomposition. In *NeurIPS*, 2017.

Vaswani, A. et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

Wagner, P., Strodthoff, N., Bousseljot, R.-D., Kreiseler, D., Lunze, F. I., Samek, W., and Schaeffter, T. Ptb-xl, a large publicly available electrocardiography dataset. *Scientific Data*, 7(1):154, 2020.

Wang, Z., Yan, W., and Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. *International Joint Conference on Neural Networks*, 2017.

Williams, M. O., Kevrekidis, I. G., and Rowley, C. W. A data-driven approximation of the koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015a.

Williams, M. O., Kevrekidis, I. G., and Rowley, C. W. A data-driven approximation of the koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015b.

Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., and Eickhoff, C. A transformer-based framework for multivariate time series representation learning. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2021.

A. Extended Dynamic Mode Decomposition

To estimate the Koopman operator, we use Extended Dynamic Mode Decomposition (EDMD) (Lehmberg et al., 2020). Given temporally shifted trajectory matrices

$$X_0 = [x_1, x_2, \ldots, x_{T-1}], \qquad X_1 = [x_2, x_3, \ldots, x_T],$$

we define a dictionary of observables $\psi(x)$ and construct the lifted data matrices

$$\Psi_0 = [\psi(x_1), \psi(x_2), \ldots, \psi(x_{T-1})],$$

$$\Psi_1 = [\psi(x_2), \psi(x_3), \ldots, \psi(x_T)].$$

The EDMD approximation of the Koopman operator is then computed as

$$K_{\mathrm{EDMD}} = G^{\dagger} A,$$

where

$$G = \frac{1}{T-1} \Psi_0 \Psi_0^{\top}, \qquad A = \frac{1}{T-1} \Psi_0 \Psi_1^{\top},$$

and $G^{\dagger}$ denotes the Moore–Penrose pseudoinverse.

B. Additional Experimental Details

B.1. Implementation Details

All experiments were implemented in Python using PyTorch for deep neural models and NumPy-based implementations for Koopman spectral feature extraction. Deep neural baselines were trained using the Adam optimizer with cross-entropy loss. Unless otherwise stated, experiments used a learning rate of 10^{-3}, batch size of 64, and 50 training epochs.

For Koopman representations, spectral features were extracted independently for each trajectory using a reduced-order EDMD approximation. The resulting feature vector included:

- real eigenvalue components,
- imaginary eigenvalue components,
- eigenvalue magnitudes,
- eigenvalue phases,
- singular values,
- Koopman coefficient statistics.

The Koopman-only setting used lightweight downstream classifiers including Logistic Regression and Random Forests. In the hybrid setting, spectral features were concatenated with deep latent embeddings produced by the convolutional encoder.

B.2. Noise Perturbation Protocol

To evaluate robustness under perturbation, additive Gaussian noise was applied to the test signals:

$$x_{\mathrm{noisy}} = x + \mathcal{N}(0, \sigma^2),$$

where

$$\sigma \in \{0.0, 0.1, 0.2, 0.5\}.$$

Noise perturbations were applied only during evaluation to study representation degradation behavior under progressively stronger corruption.

B.3. Discussion of Class Imbalance

The PTB-XL superclass classification setup exhibits substantial class imbalance. Consequently, overall classification accuracy may overestimate representation quality, particularly for minority classes. For this reason, macro-F1 was treated as an important complementary evaluation metric throughout the experiments.

The observed discrepancy between accuracy and macro-F1 indicates that models primarily optimized dominant diagnostic categories while remaining less effective for underrepresented classes. Interestingly, the hybrid spectral–deep representation achieved slightly improved macro-F1 despite slightly lower overall accuracy, suggesting that Koopman spectral descriptors may provide complementary information for class-balanced representation learning.

B.4. Hybrid Representation Instability Under Noise

During aggressive perturbation experiments, recomputed Koopman spectral features in the hybrid setting frequently collapsed toward degenerate predictions. This behavior likely arises from instability in eigenspectrum estimation under heavily corrupted trajectories.

These observations highlight an important limitation of fixed EDMD-based feature extraction in noisy physiological settings. While spectral descriptors provide interpretable structure, their robustness depends strongly on stable operator estimation and sufficiently coherent temporal trajectories.

Future work may address this issue through:

- learned lifting functions,
- adaptive spectral dictionaries,
- regularized Koopman estimation,
- contrastive or self-supervised spectral learning.

C. Additional Tables

Table 3. Detailed predictive performance under patient-wise PTB-XL evaluation.

Method	Accuracy	Macro-F1	Weighted-F1	MCC	AUROC
Koopman-LR	39.39	30.82	40.94	0.184	0.684
Koopman-RF	52.27	27.18	43.48	0.248	0.726
CNN	**75.49**	60.29	**74.53**	**0.643**	**0.921**
Hybrid	74.51	**59.02**	73.30	0.628	0.911

Table 4. Noise robustness evaluation under additive Gaussian perturbation.

Metric	$\sigma = 0.0$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.5$
Accuracy	75.49	75.07	73.35	46.01
Macro-F1	58.85	59.13	57.63	37.86

D. Additional Experimental Details

D.1. Dataset Description

We evaluate the proposed framework on the PTB-XL dataset (Wagner et al., 2020), a publicly available large-scale ECG benchmark containing 12-lead clinical recordings from thousands of patients. The dataset exhibits substantial inter-subject variability, measurement noise, and clinically realistic class imbalance.

Each sample is represented as a multivariate time series

$$X \in \mathbb{R}^{C \times T},$$

where $C = 12$ denotes the number of ECG leads and T denotes the temporal dimension.

To reduce label sparsity, diagnostic labels are grouped into five superclass categories:

- Normal rhythm (NORM),
- Myocardial infarction (MI),
- ST/T changes (STTC),
- Conduction disturbance (CD),
- Arrhythmia (ARR).

Evaluation is performed under patient-wise splits to avoid subject leakage between training and test partitions.

D.2. Preprocessing

Signals are standardized using z-score normalization computed exclusively on the training partition:

$$x' = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}} + \epsilon}.$$

To analyze robustness under perturbation, noisy variants are generated using additive Gaussian noise:

$$x_{\text{noisy}} = x + \mathcal{N}(0, \sigma^2),$$

where

$$\sigma \in \{0.0, 0.1, 0.2, 0.5\}.$$

D.3. Koopman Feature Construction

Koopman representations are estimated independently for each trajectory using EDMD with truncated singular value decomposition. Given temporally shifted matrices X_0 and X_1, a reduced Koopman operator is estimated and used to derive compact spectral descriptors.

Extracted features include:

- real and imaginary eigenvalue components,
- eigenvalue magnitudes and phases,
- singular values,
- Koopman coefficient statistics.

These descriptors form compact spectral embeddings summarizing temporal dynamics.

E. Training and Implementation Details

Deep neural models are implemented in PyTorch and trained using Adam optimization with cross-entropy loss. Koopman representations are computed using NumPy-based EDMD approximations. All experiments use identical patient-wise splits and preprocessing settings for fair comparison.

Table 5. Training and implementation settings.

Parameter	Value
Optimizer	Adam
Learning rate	10^{-3}
Batch size	64
Epochs	50
Loss function	Cross-entropy
Normalization	Z-score
Noise levels	$\sigma \in \{0.0, 0.1, 0.2, 0.5\}$
Deep model	1D CNN
Koopman method	EDMD
Evaluation split	Patient-wise
Metrics	Accuracy, Macro-F1, MCC, AUROC

Table 6. Ablation of Koopman feature groups in the hybrid representation. Coefficient statistics provide the strongest improvement in Macro-F1 and MCC, whereas concatenating all spectral descriptors slightly reduces performance.

Representation	**Accuracy**	**Macro-F1**	**MCC**	**AUROC**
CNN only	0.7261	0.6288	0.6172	0.9150
CNN + eigenvalues	0.7192	0.6242	0.6091	0.9138
CNN + singular values	0.7206	0.6266	0.6101	0.9135
CNN + coefficient statistics	**0.7266**	**0.6315**	**0.6182**	0.9139
CNN + all Koopman features	0.7183	0.6242	0.6085	0.9116

F. Additional Ablation Results

G. Interpretability and Spectral Structure

An important advantage of Koopman-based representations lies in their interpretability. Unlike standard deep embeddings, the spectral decomposition provides direct access to temporal organization through eigenvalue distributions and mode structure. Dominant eigenvalues correspond to characteristic temporal scales, while Koopman modes encode coherent spatial-temporal patterns across ECG channels.

The observed class-dependent variation in spectral descriptors suggests that Koopman features capture intrinsic temporal organization rather than purely superficial waveform characteristics. This provides a meaningful analytical perspective on representation geometry that is difficult to obtain from standard deep neural embeddings alone.

From a high-dimensional learning perspective, the Koopman representation can also be interpreted as a low-rank temporal projection. Although the lifted feature space may itself be high-dimensional, the effective representation is governed by a relatively small number of dominant spectral components. This implicit low-rank organization likely explains the compactness and structured behavior of the resulting embeddings.

H. Stability–Performance Tradeoff

A central observation throughout the experiments is the tradeoff between predictive flexibility and representation structure. Deep neural representations achieve the strongest predictive accuracy due to their ability to model highly nonlinear discriminative patterns directly from raw signals. However, these representations remain largely implicit and difficult to interpret.

In contrast, Koopman-based representations impose explicit temporal structure through spectral decomposition. Although this constraint reduces predictive flexibility, it introduces interpretable organization through eigenvalues, modes, and temporal coefficients. The resulting representations behave as structured projections emphasizing coherent temporal dynamics rather than arbitrary discriminative correlations.

The hybrid representation partially bridges these two paradigms. By combining learned deep embeddings with explicit spectral descriptors, the hybrid model slightly improves macro-F1 performance, suggesting that dynamical structure provides complementary information for minority-class discrimination. At the same time, the instability observed under noisy spectral recomputation highlights the limitations of fixed EDMD-style approximations in highly perturbed settings.

I. Figure: Koopman Noise Geometry

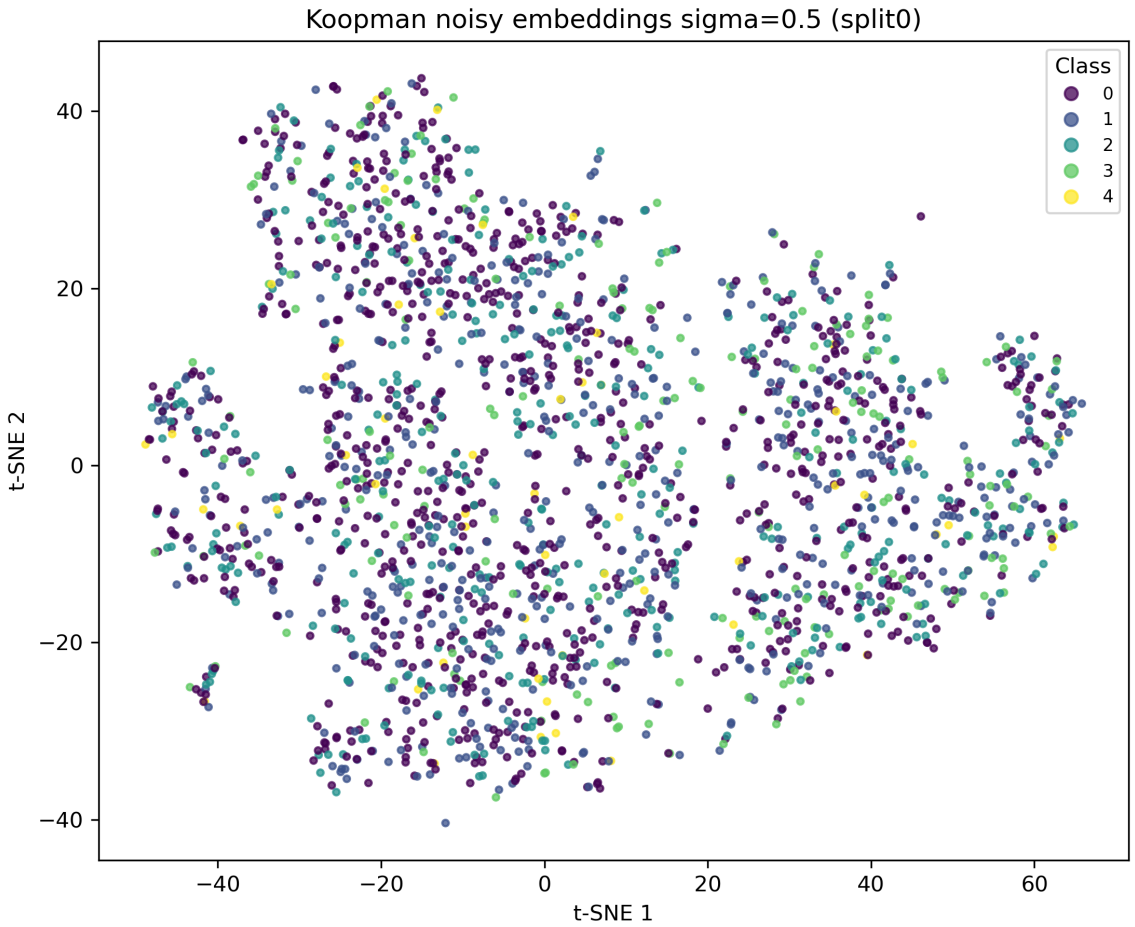

Figure 8. Koopman embedding geometry under strong perturbation ($\sigma = 0.5$). The representation loses coherent organization, illustrating the instability of spectral operator estimation under noisy conditions.

J. Figure:Hybrid noise geometry

K. Detailed EDMD and Hybrid Representation Setup

K.1. EDMD Observable Construction

For each multichannel ECG trajectory

$$X \in \mathbb{R}^{C \times T},$$

temporally shifted matrices

$$X_0 = [x_1, x_2, \ldots, x_{T-1}], \quad X_1 = [x_2, x_3, \ldots, x_T]$$

were constructed.

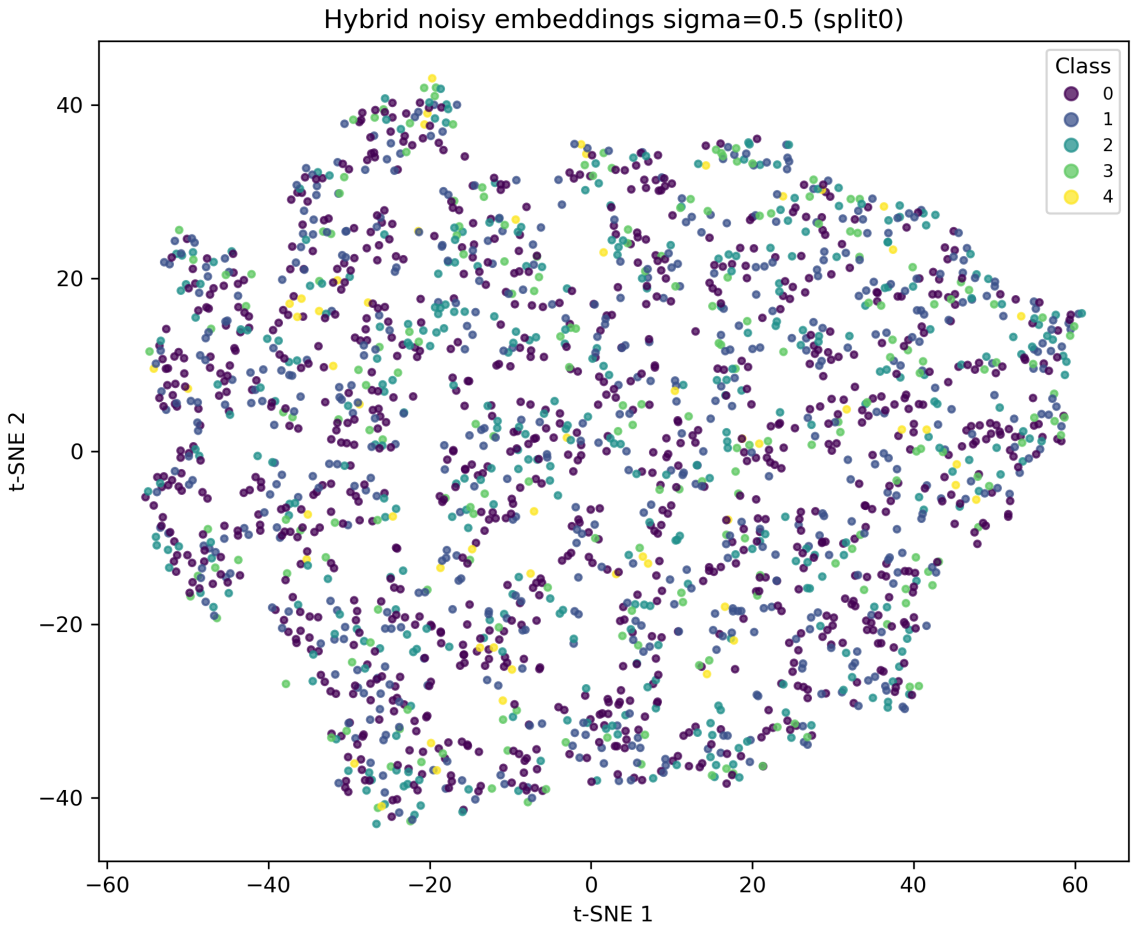

Figure 9. Hybrid CNN–Koopman embeddings under strong perturbation ($\sigma = 0.5$). The latent geometry becomes fragmented due to instability in the Koopman spectral descriptors.

We used a linear observable dictionary corresponding to the identity lifting function

$$\phi(x) = x,$$

which produces a standard reduced-order EDMD approximation. This choice was adopted to isolate the behavior of spectral representations without introducing additional nonlinear lifting complexity.

K.2. Reduced Koopman Operator Estimation

The Koopman operator was estimated using truncated singular value decomposition:

$$X_0 \approx U_r \Sigma_r V_r^\top,$$

followed by

$$\tilde{K} = U_r^\top X_1 V_r \Sigma_r^{-1}.$$

The truncation rank was selected as

$$r = \min(64, \operatorname{rank}(X_0)),$$

to stabilize spectral estimation while preserving dominant temporal structure.

K.3. Spectral Feature Extraction

For each trajectory, the following spectral descriptors were extracted from the estimated Koopman operator:

- real eigenvalue components,
- imaginary eigenvalue components,
- eigenvalue magnitudes,
- eigenvalue phases,
- singular values,
- Koopman coefficient statistics.

Coefficient statistics included:

$$\{\mu_c, \sigma_c, \max|c|, \mathrm{median}(|c|)\},$$

where c denotes Koopman spectral coefficients.

The resulting Koopman feature vector contained approximately $4r + r + 4$ scalar descriptors per trajectory.

K.4. CNN Representation

The deep neural baseline used a lightweight 1D convolutional neural network operating directly on raw ECG trajectories. The architecture consisted of:

- three 1D convolutional blocks,
- ReLU nonlinearities,
- batch normalization,
- temporal global average pooling,
- a fully connected classification layer.

The final latent embedding dimension before classification was 128.

K.5. Hybrid Spectral–Deep Representation

In the hybrid setting, Koopman spectral descriptors were concatenated with the CNN latent embedding:

$$z_{\mathrm{hybrid}} = [z_{\mathrm{CNN}}; z_{\mathrm{Koopman}}],$$

followed by a fully connected classifier trained using cross-entropy loss.

This setup enabled controlled analysis of complementary effects between explicit dynamical structure and learned deep representations.

K.6. Implementation Details

All deep learning experiments were implemented in PyTorch. Koopman spectral estimation and feature extraction were implemented using NumPy and SciPy routines.

Training used:

- Adam optimizer,
- learning rate 10^{-3},

- batch size 64,
- 50 epochs,
- early stopping based on validation Macro-F1.