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

Sucheta Ghosh SUCHETA.GHOSH@UNI-HEIDELBERG.DE
IWR, Heidelberg University, Germany

Zahra Monfared ZAHRA.MONFARED@UNI-HEIDELBERG.DE
IWR and Department of Mathematics & Computer Science, Heidelberg University, Germany

Abstract

Learning robust representations from physiological time series remains challenging under patient variability, class imbalance, and signal perturbation. We investigate Koopman operator theory and EDMD as structured dynamical representations for studying representation geometry, robustness, and stability in high-dimensional physiological time series. Using patient-wise PTB-XL classification as a controlled testbed, we compare Koopman-only, deep neural, and hybrid spectral–deep representations. Our analysis reveals a stability–performance tradeoff. Deep neural models achieve the strongest predictive accuracy, whereas hybrid spectral–deep representations slightly improve Macro-F1, indicating that Koopman spectral descriptors provide complementary information for class-balanced learning. Although less competitive in predictive performance, Koopman representations capture temporal structure through compact spectral descriptors. Robustness experiments show that recomputed EDMD-based representations become unstable under perturbation, producing eigenspectrum drift and degradation of hybrid representations. These results highlight limitations of fixed spectral feature extraction and suggest that Koopman-inspired representations are most useful as complementary tools for studying temporal organization, robustness, interpretability, and representation geometry rather than as direct replacements for deep neural representations.

1. Introduction

Learning robust representations from high-dimensional physiological time series remains challenging due to measurement noise, class imbalance, and cross-entity variability in clinical signals [4, 11]. Deep neural models achieve strong performance for time-series analysis [15, 17, 20], but often learn representations that are difficult to interpret and sensitive to perturbation [1, 5]. An alternative perspective treats temporal signals as realizations of an underlying dynamical system. Koopman operator theory models nonlinear dynamics through linear evolution in a lifted function space [7, 10], while Dynamic Mode Decomposition (DMD) and Extended Dynamic Mode Decomposition (EDMD) extract compact spectral descriptors from trajectories [13, 18]. These representations provide interpretable views of temporal organization that are difficult to obtain from standard deep embeddings. We investigate Koopman-based spectral representations for patient-wise PTB-XL classification, comparing Koopman-only, deep neural, and hybrid spectral–deep representations in terms of predictive performance, robustness, interpretability, and representation stability. Our results reveal a stability–performance tradeoff. Deep neural models achieve the strongest predictive accuracy, whereas hybrid spectral–deep representations slightly improve Macro-F1, indicating that Koopman spectral descriptors provide complementary information for class-balanced classification. Although Koopman-only representations remain less competitive in predictive performance, they capture

© S. Ghosh & Z. Monfared.

temporal structure through compact spectral descriptors. We further show that recomputed EDMD-based representations become unstable under perturbation, producing eigenspectrum drift, fragmented representation geometry, and degradation of hybrid representations. From a high-dimensional learning perspective, these results suggest that Koopman representations act as structured low-rank projections linking representation geometry, robustness, and predictive performance. Overall, Koopman-inspired spectral representations appear most useful as complementary analytical tools rather than direct replacements for deep neural representations.

1.1. Related Work

Deep neural architectures achieve strong performance for time-series representation learning but often produce implicit representations that remain sensitive to noise and distribution shifts in high-dimensional settings. Recent self-supervised and contrastive approaches, including SimCLR and supervised contrastive learning, improve representation quality under limited supervision [3, 6], but depend heavily on augmentation design and do not explicitly model temporal dynamics. 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 [7, 10], while DMD and EDMD extract spectral descriptors directly from trajectories [12, 19]. These methods have been applied across fluid dynamics, control, and machine learning [2], and have recently been integrated with neural representation learning through learned lifting functions and latent linear dynamics [9, 14]. In contrast, we study Koopman-based spectral representations as structured projections for noisy high-dimensional physiological time series, emphasizing robustness, interpretability, representation geometry, 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

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

We estimate the Koopman operator using Extended Dynamic Mode Decomposition (EDMD) [8]. 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. We use a reduced-order EDMD approximation with identity observables and extract compact spectral descriptors from independently estimated trajectory-level operators. From a high-dimensional learning perspective, these descriptors induce a low-rank representation that facilitates analysis of representation geometry and stability.

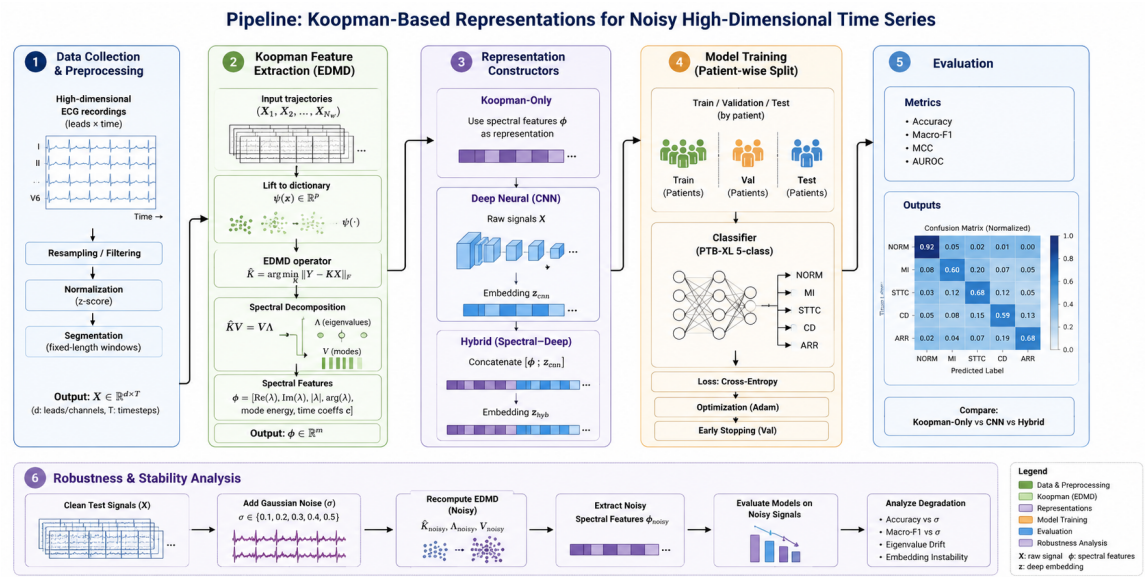

Figure 1: Overview of the Koopman-based representation learning framework. EDMD spectral features are combined with deep neural representations and evaluated under patient-wise perturbations.

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. The Koopman representation promotes **noise attenuation** and **low-rank organization** through dominant spectral components, yielding compact 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

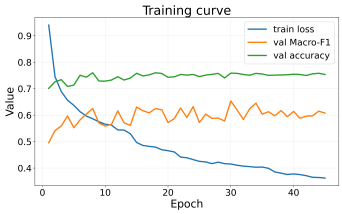

Figure 2: Training dynamics of the hybrid CNN–Koopman model on PTB-XL. Training loss decreases steadily, while validation Macro-F1 stabilizes after approximately 30 epochs.

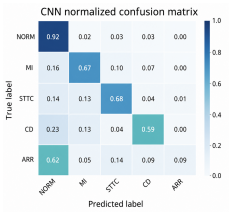

Figure 3: Normalized confusion matrix of the CNN baseline on PTB-XL. The model performs well on the majority NORM class but struggles with the minority ARR class.

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 [16], 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. Details are 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.

Training Dynamics Training dynamics are shown in Figure 2. Validation Macro-F1 stabilizes after approximately 30 epochs, suggesting stable convergence under patient-wise evaluation.

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.

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

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

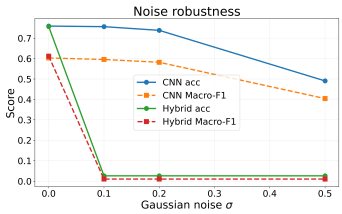

Figure 4: Noise robustness under additive Gaussian perturbation. CNN degrades gradually, and the hybrid model becomes unstable when Koopman features are recomputed from noisy signals.

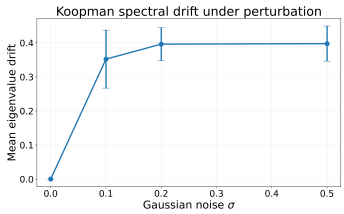

Figure 5: Koopman spectral drift under perturbation. Increasing noise causes substantial eigenspectrum deformation, indicating instability of EDMD-derived spectral features.

4.1 Predictive Performance and Error Analysis Table 1 compares Koopman-only, deep neural, and hybrid spectral–deep representations. The CNN baseline achieves the highest 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 class imbalance and weaker minority-class recognition. Although Koopman-only representations are less competitive than deep models, they capture discriminative structure using compact spectral descriptors. The hybrid model slightly improves Macro-F1, suggesting that Koopman features provide complementary temporal information. 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. Ablation results are provided in Appendix F.

4.2 Robustness Under Noise Perturbation. We evaluate robustness under increasing additive Gaussian perturbations. 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. It suggests that EDMD-derived spectral features are highly sensitive to perturbation and depend strongly on stable operator estimation. Figure 4 illustrates this. As 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

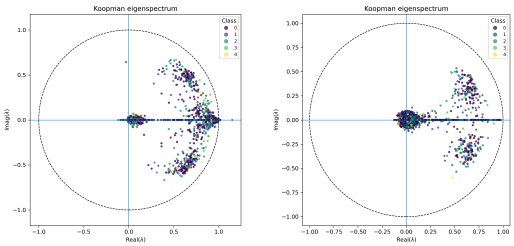

Figure 6: Koopman eigenspectrum under clean and noisy conditions. Noise substantially distorts the estimated spectral structure.

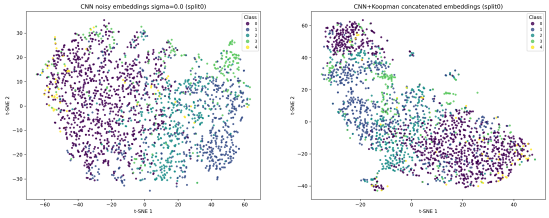

Figure 7: t-SNE visualization of CNN (left) and hybrid CNN–Koopman (right) embeddings. The hybrid representation exhibits stronger global organization induced by spectral descriptors.

under clean and noisy conditions. Strong perturbation distorts the dominant spectral modes and increases dispersion, explaining the instability observed in the hybrid representation.

4.3 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 organization induced by Koopman spectral descriptors. Under strong perturbation, the hybrid geometry becomes fragmented, reflecting instability from the spectral component.

4.4 Limitations. Koopman-only representations remain weaker than deep neural models for predictive accuracy, suggesting that fixed spectral descriptors may not fully capture nonlinear discriminative structure. Perturbation experiments further show that EDMD-derived spectral features can become unstable under noise. Nevertheless, Koopman representations remain 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, rather to analyze representation stability, spectral behavior across controlled representation regimes.*

5. Conclusion

We investigated Koopman-based spectral representations for high-dimensional physiological time-series learning using EDMD, comparing Koopman-only, deep neural, and hybrid spectral–deep representations under patient-wise PTB-XL classification. Our results reveal a stability–performance tradeoff: deep neural models achieved the strongest predictive accuracy, whereas hybrid representations slightly improved Macro-F1, indicating that Koopman spectral descriptors provide complementary information for class-balanced classification. Robustness experiments showed that recomputed EDMD-based representations become unstable under perturbation, producing eigenspectrum drift and degradation of hybrid representations. From a high-dimensional learning perspective, Koopman representations can be viewed as structured low-rank projections linking temporal organization, representation geometry, robustness, and predictive performance. Increased representation structure improves interpretability but may also increase sensitivity to instability in spectral estimation. Overall, Koopman representations appear most useful as complementary analytical tools rather than direct replacements for deep neural representations. Future work will investigate adaptive and regularized Koopman estimation, self-supervised spectral learning, and scaling behavior with respect to trajectory length, representation dimension, dictionary complexity, and operator rank, connecting Koopman-based representations to broader questions in high-dimensional learning dynamics.

Acknowledgements

Z.M. is grateful to the Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR, Federal Ministry for Research, Technology and Space) for funding through project OIDLITDSM, No. 01IS24061.

References

[1] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8): 1798–1828, 2013.

[2] Steven L. Brunton, Bingni W. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Koopman invariant subspaces and finite linear representations of nonlinear dynamical systems for control. *PLoS ONE*, 11(2):e0150171, 2016.

[3] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020.

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

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep learning. *Nature*, 521(7553): 436–444, 2015.

[6] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. In *NeurIPS*, 2020.

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

[8] Daniel Lehmberg, Felix Dietrich, Gerta Köster, and Hans-Joachim Bungartz. 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.

[9] Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. In *Nature Communications*, 2018.

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

[11] Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, and Samy Bengio. Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems*, 2019.

[12] Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.

[13] Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, 656:5–28, 2010.

[14] Naoya Takeishi, Yoshinobu Kawahara, and Takehisa Yairi. Learning koopman invariant subspaces for dynamic mode decomposition. In *NeurIPS*, 2017.

[15] Ashish Vaswani et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

[16] Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dirk Kreiseler, Felix Isaac Lunze, Wojciech Samek, and Tobias Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset. *Scientific Data*, 7(1):154, 2020.

[17] Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. *International Joint Conference on Neural Networks*, 2017.

[18] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data-driven approximation of the koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.

[19] Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data-driven approximation of the koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015.

[20] George Zerveas, Srishti Jayaraman, Dhaval Patel, Anuradha Bhamidipaty, and Carsten Eickhoff. A transformer-based framework for multivariate time series representation learning. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2021.

Appendix A. Extended Dynamic Mode Decomposition

To estimate the Koopman operator, we use Extended Dynamic Mode Decomposition (EDMD) [8]. 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.

Appendix 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:

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

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.

Appendix C. Additional Tables

Appendix D. Additional Experimental Details

D.1. Dataset Description

We evaluate the proposed framework on the PTB-XL dataset [16], 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.

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	**75.90**	60.29	**74.53**	**0.647**	**0.921**
Hybrid	75.67	**61.17**	73.30	0.643	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	75.90	75.07	73.35	46.01
Macro-F1	60.29	59.13	57.63	37.86

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.

Appendix 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

Appendix F. Additional Ablation Results

Appendix 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.

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	0.7590	0.6029	0.647	0.921
CNN + eigenvalues	0.7492	0.6191	0.6109	0.9138
CNN + singular values	0.7556	0.621	0.6121	0.9135
CNN + coefficient statistics	**0.7599**	**0.631**	**0.6532**	0.914
CNN + all Koopman features	0.7567	0.6117	0.643	0.911

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.

Appendix 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.

More broadly, these observations connect to questions in high-dimensional learning dynamics. The Koopman representation acts as an explicit dynamical prior that constrains the geometry of the representation space. Our results suggest that increasing representation structure improves interpretability and temporal organization, but may reduce robustness when spectral estimation becomes unstable. This highlights a fundamental trade-off between representation flexibility and representation stability.

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.

Appendix I. Figure: Koopman Noise Geometry

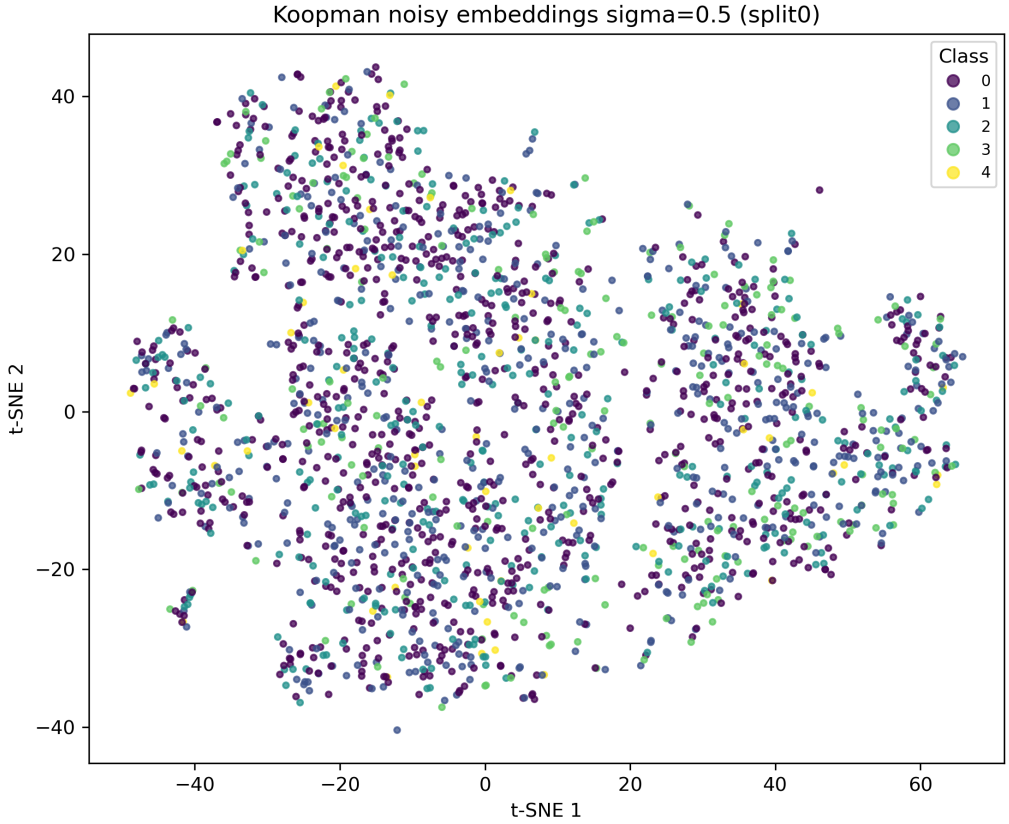

Figure 8: Koopman embedding geometry under strong perturbation ($\sigma = 0.5$). The representation loses coherent organization, illustrating the instability of spectral operator estimation under noisy conditions.

Appendix J. Figure:Hybrid noise geometry

Appendix 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.

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

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

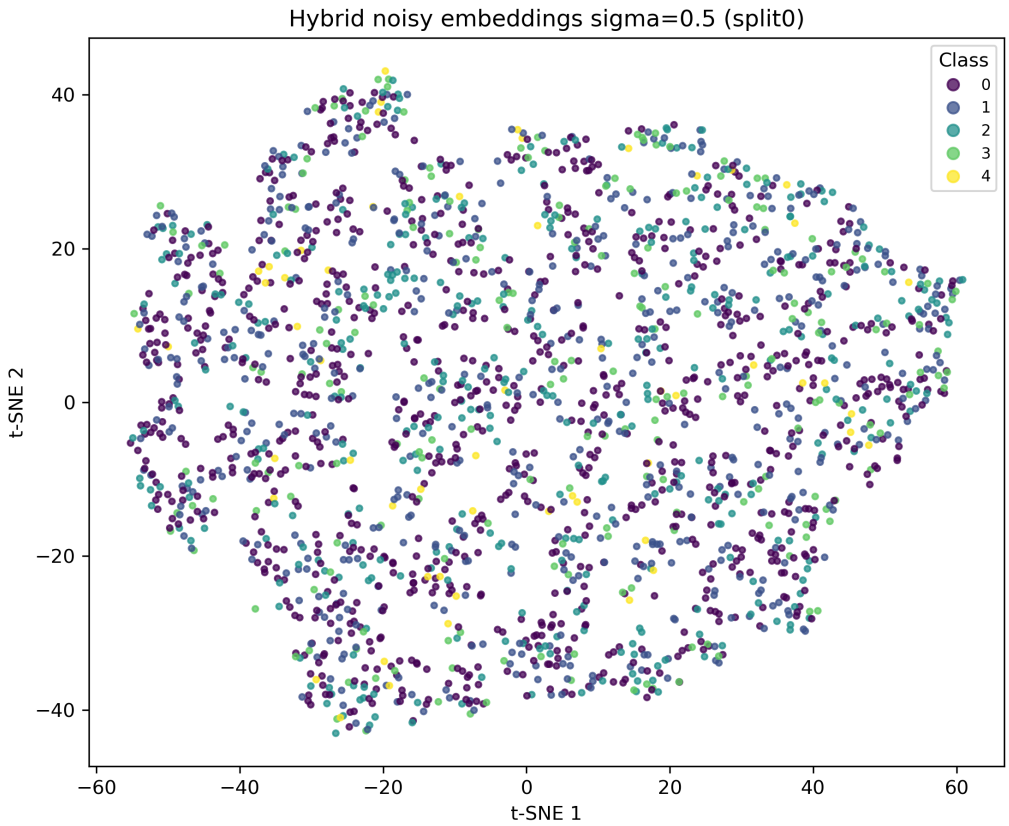

Figure 9: Hybrid CNN–Koopman embeddings under strong perturbation ($\sigma = 0.5$). The latent geometry becomes fragmented due to instability in the Koopman spectral descriptors.

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

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

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.