Machine Learning and Information Retrieval Techniques for Time Series Analysis

Bionda Rozin[1], and Daniel Carlos Guimarães Pedronette
Department of Statistics, Applied Mathematics and Computing,
State University of São Paulo (UNESP), Rio Claro, Brazil
Email: {bionda.rozin, daniel.pedronette}@unesp.br

***Abstract*—This work explores the intersection of Machine Learning and Information Retrieval for time series analysis, addressing key challenges in representation, classification, clustering, and retrieval. Four methodologies are proposed, covering univariate and multivariate time series across supervised, unsupervised, and semi-supervised scenarios. The approaches integrate contextual similarity learning, image-based representations, and domain-specific graph modeling, demonstrating competitive performance across multiple datasets.**

I. Introduction

Time series data play a crucial role across domains such as medicine, economics, and the sciences, with applications ranging from electrocardiograms (ECGs) [1], electroencephalograms (EEGs) [2], to sales trends [3], sensor readings [4], weather forecasts [5], cryptocurrency charts [6], and financial analysis [7].

Given their broad applicability, robust tools for time series analysis are essential. A wide range of computational methods, supervised, unsupervised, and semi-supervised learning, are employed depending on data availability and the specific objectives. Classification and clustering aim to group similar sequences, while forecasting predicts future values. These techniques have shown effectiveness in contexts such as early cancer detection [8], mobility pattern analysis [9], poultry welfare monitoring [10], and forecasting of stock market [7] and ECG trends [1].

The performance of machine learning in time series tasks depends on the feature extraction process, which defines data representation and influences similarity computations. Extracted features may capture aspects such as shape, temporal dependencies, and may employ strategies like bag-of-words or kernel-based models. In Information Retrieval (IR), both representation and similarity measures are fundamental, influencing the relevance ranking of retrieved results, and impacting data management, indexing, and retrieval. The success of IR tasks depends on how similarity is encoded and measured. Traditional metrics such as Euclidean distance perform only pairwise comparisons and overlook neighboring interactions, potentially limiting retrieval effectiveness [11].

Information Retrieval for time series has a wide range of applications, including identifying plant species from satellite imagery [12], tracking volcanic plume dispersion for aviation safety [13], and classifying eye states using EEG segments [2]. Designing IR methods that deliver high-quality results remains a complex challenge. A promising strategy is the use of unsupervised distance learning after retrieval. These algorithms consider global neighborhood structures to compute more refined distance measures, which improve ranking performance by incorporating contextual relationships [11].

Machine Learning and Information Retrieval are key tools in time series analysis, particularly when integrated. This work aims to investigate the applicability of combined strategies in diverse time series contexts and to evaluate how different representations and contextual similarity measures influence analytical performance.

The main contributions of this work are summarized below:

- Investigation of univariate time series feature extraction for retrieval tasks, considering different re-ranking approaches [14];
- Study on impact of approaching separately each dimension of a multivariate time series, and how contextual rank aggregation of dimensions impact in retrieval and classification tasks [15];
- Proposal of a generic and interpretable framework based on time segmentation criteria, clustering and evaluation with domain-specific variables [16];
- Semi-supervised classification of univariate time series considering image representations and label propagation [17].

Further sections will discuss in details each research question and its respective answers and contributions.

II. Related Work

We conducted an extensive literature review on machine learning and information retrieval techniques for univariate and multivariate time series, including image-based representations. The complete review is available in Chapter 3 of the dissertation. Due to space constraints, we summarize key literature gaps below:

- **Limited focus on information retrieval:** Most time series research emphasizes forecasting and classification [1], [5]–[8], [10], [18], [19], while content-based retrieval remains less explored [2], [4], [12], [13]. Our work addresses this with two content-based retrieval applications;
- **Limitations in deep learning adaptability:** Deep learning has become central in time series analysis [4], [18],

[1]This work is related to a M.Sc. dissertation.

[19], but many models are designed for specific tasks and require retraining or reconfiguration for each new dataset, which limits their reusability and increases computational cost. In this work, deep models are used solely for extracting features from image-based time series representations, while more traditional methods are employed for clustering, retrieval, and classification;

- **Limited transferable representations for feature extraction:** Traditional deep feature extractors are often coupled to the task, such as classification, retrieval or forecasting [2], [4], [6], [18], and training dataset, often requiring retraining and code-level adjustments to adapt to a new dataset or task. We address this by evaluating both traditional and image-based extractors based on transfer learning, comparing their performance across multiple tasks and data domains.

III. Main Contributions and Organization

This work explores the hypothesis that machine learning and information retrieval techniques can effectively support time series analysis, especially with scarcity or absence of labeled data. We employed **unsupervised**, **supervised**, and **semi-supervised** learning approaches for time series **clustering**, **similarity search**, and **classification**, while also evaluating the role of different **feature extractors** in these tasks.

Following, we discuss the research challenges and respective goals and contributions of this work:

- How can we **extract meaningful features** for a required task? Which features are more **suitable** for a machine learning or information retrieval task?
- There are diverse methods to **represent time series as images**. How the information contained in the images can be explored?
  - We compare diverse extractors, including **traditional and image-based approaches**, to identify which representations are most effective across learning tasks. **Multiple image encoding techniques were evaluated** across all scenarios (classification, retrieval, and clustering), extracting features using deep neural networks, pre-trained on huge datasets, such as ImageNet [20]. The effectiveness of transfer learning in this context is observed.
- How do we find **distance measurements** that effectively translate the **relationship between the dataset elements**?
- How can we explore **more global relationships** between time series to perform **information retrieval**?
  - Pairwise distance measurements often ignore contextual information available in the dataset. Four unsupervised distance learning methods were applied in univariate time series datasets, computing **more global distance measurements between elements** in a dataset, by incorporating **neighborhood relationships**, improving retrieval tasks.
- How to effectively explore **information encoded in the dimensions** of a multivariate time series?
  - To fully explore information present in the dimensions of multivariate time series, we process each dimension **individually** and use **rank aggregation methods** to obtain a final result for multivariate time series **retrieval**.
  - **The resulting similarity** measure defined by the rank aggregation step is considered for performing a **kNN classification** in a multivariate time series dataset.
- How do we properly **interpret** the **clustering results** in **unlabeled** time series datasets?
  - We introduce a **novel clustering-based framework based on time segmentation criteria**, encoding information as multivariate time series, and incorporating **domain-specific variables** to evaluate and enhance the **interpretability**. This framework was validated in ball possession analysis in football matches.
- How to properly **explore supervised** and **unsupervised information** in time series data?
  - We employ the label propagation algorithm for time series **semi-supervised classification**, exploring information available in both labeled information and structural unlabeled information present in time series datasets.

The next four Sections describe the proposed methodologies for time series analysis, Section VIII addresses the conclusions and future works, and Section IX presents the publications and partnerships. Due to space constraints, details on feature extractors, unsupervised distance learning and datasets can be found on Chapter 2 of the main work, as well as the complete description of results.

IV. Re-Ranking and Representations for Time Series Retrieval: A Comparative Study

This study presents a **comparative study of time series representations and re-ranking** methods for information retrieval, inspired by content-based image retrieval (CBIR) systems [21]. Currently under major revision for *IEEE Access* [14], this work evaluates ten feature extraction techniques—ranging from 1D signal-based descriptors to image-based approaches such as GAF, MTF, and RP with ResNet and ViT—and four unsupervised distance learning methods across six diverse datasets. Figure 1 illustrates the proposed pipeline, and Table I summarizes the methods, datasets, and best mAP results. Experimental results show that contextual re-ranking significantly enhances retrieval performance, with improvements in mean average precision (mAP) of up to 31.78%.

Our findings indicate that the **proper combination of feature representation and contextual re-ranking is crucial to effective time series retrieval**. While no single configuration outperforms all others, DWT-Detail and image-based representations, particularly GAF and RP with ResNet, consistently produce strong results. These results suggest that their selection is application-dependent.

TABLE I
THE FIRST COLUMN LISTS 1D AND IMAGE-BASED TIME SERIES FEATURE EXTRACTORS. THE SECOND SHOWS UNSUPERVISED DISTANCE LEARNING METHODS USED FOR RE-RANKING. THE THIRD PRESENTS THE DATASETS, AND THE LAST SHOWS THE BEST MAP ACHIEVED FOR EACH.

Feature	Re-Ranking	Datasets	Best mAP per Dataset
Beam Angle Statistics (BAS) Discrete Fourier Transform (DFT) Discrete Wavelet Transform (DWT) Random Convolutional Kernel Transform (ROCKET) Gramian Angular Fields (GAF) with CNN ResNet-152 Gramian Angular Fields (GAF) with Vision Transformers (ViT) Markov Transition Fields (MTF) with CNN ResNet-152 Markov Transition Fields (MTF) with Vision Transformers (ViT) Recurrence Plot (RP) with CNN ResNet-152 Recurrence Plot (RP) with Vision Transformers (ViT)	Log-based Hypergraph of Ranking References (LHRR) Breadth-First Search Tree of Ranking References (BFSTREE) Rank Diffusion Process with Assured Convergence (RDPAC) Rank Flow Embedding (RFE)	Beef CBF Electric Devices GunPoint Rock Yoga	50.99% (DWT-Detail + RFE) 95.43% (ROCKET + RDPAC) 41.96% (GAF + ResNet + RDPAC) 86.28% (RP + ResNet + RDPAC) 82.93% (DWT-Detail + LHRR) 31.86% (BAS + RFE)

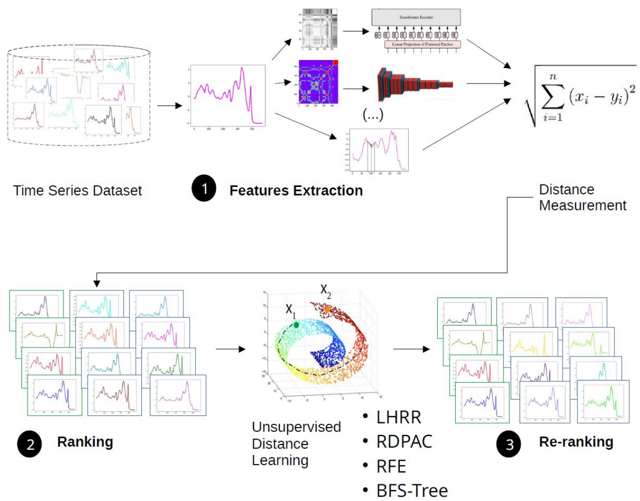

Fig. 1. Time series are first represented using different feature extractors. An initial ranking is generated based on Euclidean distance. Then, unsupervised distance learning methods are applied to refine the rankings by considering contextual relationships among the data.

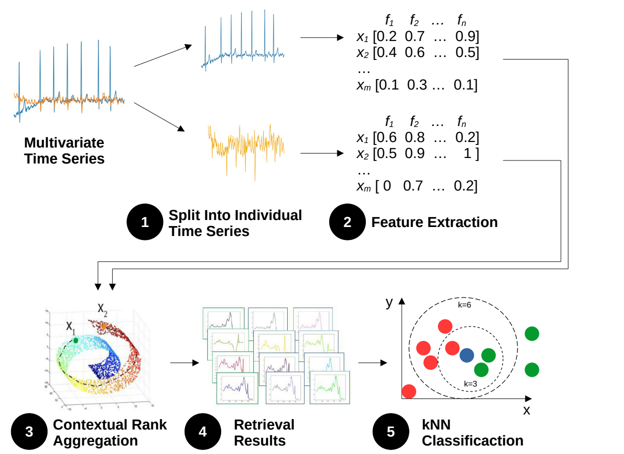

Fig. 2. Multivariate time series are decomposed into individual dimensions, each processed by a feature extractor. Contextual rank aggregation combines the results into a unified distance matrix, used for retrieval and kNN classification.

V. A Ranked-Based Framework based on Manifold Learning for Multivariate Time Series Retrieval and Classification

This work introduces a **novel framework for multivariate time series retrieval and classification using contextual rank aggregation**, currently under major review for *Pattern Recognition Letters (PRL)* [15]. Each dimension of the time series is processed individually using 1D and image-based extractors (notably RP with ResNet), and the results are fused via contextual rank aggregation to generate a unified similarity representation. By incorporating unsupervised distance learning as a pre-processing step, the framework effectively supports information retrieval and kNN classification, demonstrating versatility, and also fully exploring individual dimensions to find better similarity representations for multivariate time series datasets. Competitive performance compared to state-of-the-art methods is observed in the analyzed scenarios. Figure 2 depicts the pipeline, and Table II summarizes the methods and datasets.

The experimental evaluation indicates **promising results for retrieval tasks and competitive results for classification, outperforming, in many scenarios, recent baselines**, such as MF-Net and ROCKET.

VI. Framework for Data Analysis by Temporal Graph Encoding

In this work, we propose a **methodology for domain-specific temporal data analysis using graph-based representations and clustering**. This work is under review for *Multimedia Tools and Applications (MTAP)* [16]. The method models temporal data as graphs and extracts graph metrics, generating a multivariate time series. This series is split through a time segmentation criteria, and visual representations through Graph Visual Rhythms are obtained. Deep learning methods for feature extraction are applied, followed by dimensionality reduction and clustering. The pipeline (Figure 3) and algorithms (Table III) are detailed in Chapter 6 of the dissertation. The framework is validated in the context of football match analysis, using offensive ball possession sequences extracted from positional data of 10 Brazilian matches, analyzed to identify strategic patterns. The methodology proves effective for capturing temporal dependencies and provides interpretable insights into tactical play.

In Figure 4, we present a set of statistics derived from a specific clustering configuration that combines Average

TABLE II

THE FIRST COLUMN LISTS THE TIME SERIES FEATURE EXTRACTORS, FOLLOWED BY THE CONTEXTUAL RANK AGGREGATION METHODS IN THE SECOND COLUMN. THE THIRD COLUMN PRESENTS THE EVALUATED DATASETS, WHILE THE FOURTH AND FIFTH SHOW THE BEST MAP FOR RETRIEVAL AND THE HIGHEST CLASSIFICATION ACCURACY, RESPECTIVELY.

Feature	Rank Aggregation	Datasets	Best mAP per Dataset	Best Accuracy per Dataset
		ArticularyWordRecognition	97.01% (ROCKET + RDPAC)	99.60% (ROCKET)
		AtrialFibrillation	57.39% (BAS + RDPAC)	66.70% (BAS + LHRR + kNN)
		BasicMotions	99.09% (ROCKET + RFE)	100% (ROCKET + RFE + kNN, RP + ResNet + LHRR/RFE + kNN)
BAS	LHRR	Heartbeat	63.68% (ROCKET + RDPAC)	75.12% (ROCKET + RFE + kNN)
ROCKET	RDPAC	MotorImagery	51.89% (BAS + LHRR)	61.00% (ROCKET + RFE + kNN)
RP with CNN ResNet-152	RFE	NATOPS	74.11% (ROCKET + RDPAC + kNN)	92.70% (MF-Net)
		PenDigits	48.94% (ROCKET + RDPAC)	99.60% (ROCKET)
		StandWalkJump	55.95% (BAS + RDPAC)	53.33% (RP + ResNet + LHRR + kNN)

TABLE III

THE FIRST COLUMN LISTS THE GRAPH MEASUREMENTS, FOLLOWED BY THE FEATURE EXTRACTORS EMPLOYED IN THE GVR FIGURES. THE THIRD COLUMN PRESENTS THE DIMENSIONALITY REDUCTION TECHNIQUES, AND THE FOURTH DETAILS THE CLUSTERING METHODS APPLIED.

Graph Measurement	Feature Extractor	Dimensionality Reduction	Clustering
Average Path Lenght	CNN ResNet-152		
Betweeness Centrality	Vision Transformers	Uniform Manifold Approximation and Projection	Density-Based Spatial Clustering of Applications with Noise
Eccentricity	VGG-19	Principal Component Analysis	Affinity Propagation
Local Efficiency	EfficientNet V2	t-distributed Stochastic Neighbor Embedding	Hierarchical Density-Based Spatial Clustering of Applications with Noise
Vulnerability	DINO V2		

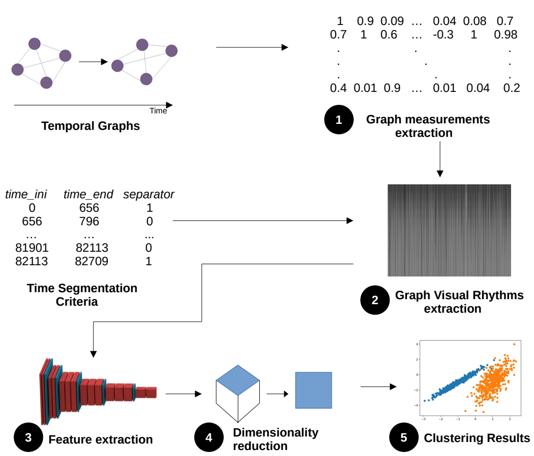

Fig. 3. Generic pipeline for data characterization based graph visual rhythm representations and transfer learning procedures. Domain-specific temporal graphs are represented via graph measurements, generating a multivariate time series. Based on a time segmentation criterion, each time series segment is used to generate a Graph Visual Rhythm image, followed by feature extraction, dimensionality reduction, and clustering.

Path Length, ResNet-152, UMAP, and Affinity Propagation. This configuration resulted in a **notable separation of ball possession sequences**, allowing the identification of clusters characterized by **distinct tactical behaviors**, and further enabling a **more refined and objective performance analysis**. The clusters highlighted in red and yellow are associated with ball possessions that exhibit a lower number of passes, shorter duration, reduced distance covered by the ball, and fewer occurrences of reaching the final third of the field. These characteristics suggest counterattacking situations or brief, ineffective plays. On the other hand, the cluster represented in purple displays possessions with noticeably different characteristics, including a higher number of passes, longer sequences, and a greater distance traveled by the ball. These features indicate a tactical approach where the team maintains a more spread positioning on the field, promoting increased variability and unpredictability during offensive actions. Such variability is also reflected in elevated values of Sample Entropy. Additionally, this cluster demonstrates a higher percentage of entries into the final third, reinforcing the interpretation of these sequences as more elaborate plays, characterized by greater variability of tactical behavior.

VII. Semi-Supervised Time Series Classification through Image Representations

We explored a **semi-supervised learning approach for time series classification**, addressing challenges posed by limited labeled data. This work was originally presented at the *International Conference on Computational Science and Its Applications – ICCSA 2023* [17]. The method converts time series into images, by employing *GAF*, *MTF* and *RP* as imaging methods, and extracts features using pre-trained neural networks (*ViT* and *CNN ResNet 152*). A kNN-based transductive model with label propagation is applied, achieving competitive results compared to fully supervised baselines across multiple datasets. Figure 5 illustrates the pipeline.

The methodology was evaluated on the *CBF*, *Electric Devices*, *ECG5000*, and *Yoga* datasets. Figures 6 and 7 showcase representative examples of promising and challenging outcomes.

The complete results, presented in [17], indicate that **image-based representations are effective for time series classification**, particularly outperforming raw-data approaches for the Electric Devices dataset. GAF and RP consistently delivered stronger results than MTF, while the choice of imaging method had more impact than the neural architecture itself. Employing raw data instead of imaging for datasets CBF, ECG5000,

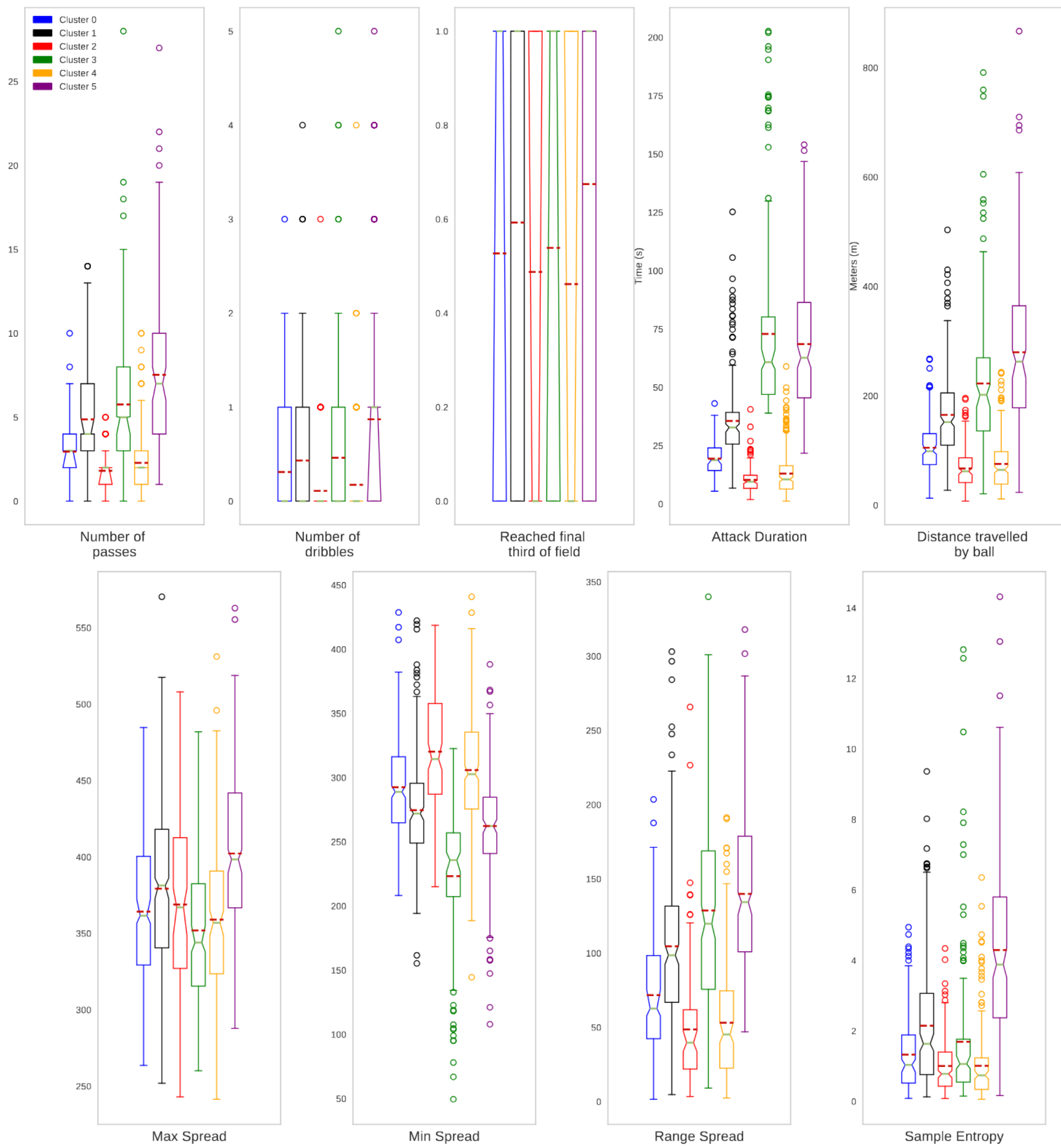

Fig. 4. Statistics for cluster configuration.

and Yoga proved to be the most effective approach in these scenarios.

VIII. Conclusions

This chapter presents the final considerations of the work, that explored four strategies for time series analysis — two focused on univariate and two on multivariate data. Rather than proposing a single model, the study investigated how different feature extraction impact performance across various tasks. It examined supervised, semi-supervised, and unsupervised scenarios for classification, retrieval, and clustering, including two methods based on contextual similarity.

Key contributions include:

- **Univariate Time Series Retrieval (Section IV):** Comparative study of feature extractors and re-ranking methods, achieving strong mAP, precision, and recall results [14];
- **Multivariate Time Series Retrieval and Classification (Section V):** Use rank aggregation to build distance matrices, leading to positive classification and retrieval results [15];
- **Unsupervised Framework for Multivariate Analysis (Section VI):** Clustering-based approach employing time segmentation, validated in football match analysis, revealing interpretable patterns [16];
- **Semi-Supervised Classification (Section VII):** Evaluation of imaging-based time series representations with label propagation, comparing with SOTA classification methods and providing valuable insights [17].

This research addressed fundamental questions related to

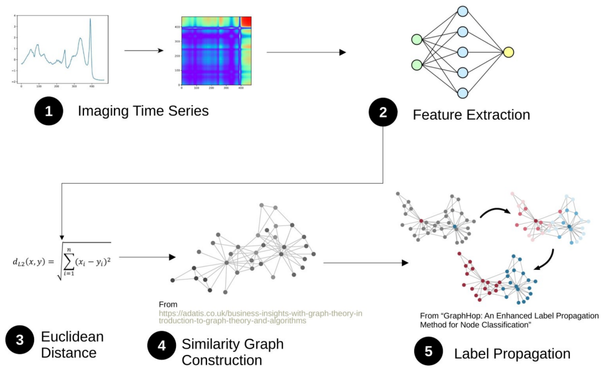

Fig. 5. Time series are transformed into images using GAF, MTF, and RP techniques. Features are then extracted with ResNet and ViT, followed by semi-supervised classification via label propagation.

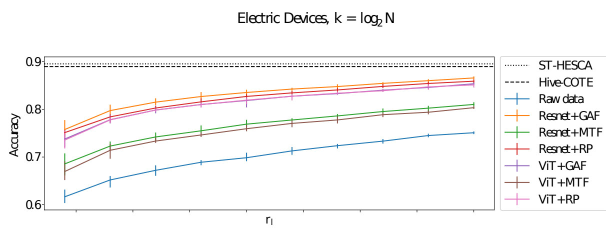

Fig. 6. Results for the ElectricDevices dataset as a function of r_l, where $0.025 \leq r_l \leq 0.2$. Supervised baselines were obtained from 8926 labeled points, corresponding to $r_l \approx 0.54$.

feature extraction, the use of image representations, the effectiveness of distance measures, and the interpretability of clustering results. While not aiming to deliver a universal solution, the work provides insights into key questions in time series learning. For example, it shows that representation choices (e.g., imaging techniques like GAF, RP, MTF) often affect results more than the choice of model. It also highlights that distance metrics and similarity measures are highly dataset and representation dependent, suggesting that no single best distance can be chosen without considering the domain. Overall, the study reinforces that representation is a critical design choice in time series tasks, that imaging approaches open promising directions for hybrid techniques combining traditional metrics and deep learning models, and that contextual similarity measures enhance retrieval performance.

A. *Future Works*

Future directions include: **Univariate Time Series Retrieval:** Perform parametric analysis and optimize the combination of descriptors and re-ranking methods; **Multivariate Time Series Retrieval and Classification:** Study the influence of parameter K, test new classifiers such as weighted kNN, and extend to other multichannel data; **Clustering with Temporal Graphs:** Apply the proposed pipeline to other domains, employ it to football match analysis, and use projection-based search systems for video retrieval; **Semi-Supervised Classification:** Investigate distance/similarity functions for label propagation and its properties; **General Enhancements:** Improve algorithm efficiency and explore texture-based features in time series image representations.

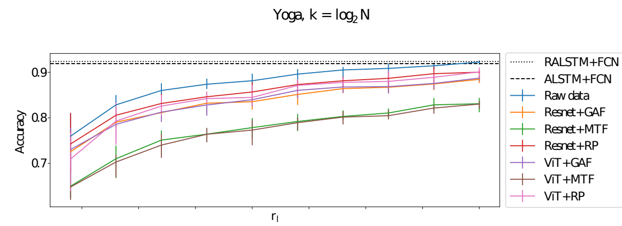

Fig. 7. Results for the Yoga dataset as a function of r_l, where $0.025 \leq r_l \leq 0.2$. Supervised baselines were obtained with initially 300 labeled points, corresponding to $r_l \approx 0.09$.

IX. Publications and International Fellowship

Chapters 4 to 7 of this dissertation led to four articles, one published in an international conference [17] and three submitted to high-impact journals in computer science:

1) **B. Rozin,** E. Bergamim, D. C. G. Pedronette, and F. A. Breve, "Semi-supervised time series classification through image representations," in International Conference on Computational Science and Its Applications – ICCSA 2023, pp. 48–65 [17] **(Qualis A3)**.
Status: Published. Cited by one.
2) **B. Rozin,** D. C. G. Pedronette, and R. S. Torres, "Re-Ranking and Representations for Time Series Retrieval: A Comparative Study" [14].
Status: In Major Revision for *IEEE Access* **(Qualis A1)**
3) **B. Rozin,** and D. C. G. Pedronette, "A Ranked-Based Framework based on Manifold Learning for Multivariate Time Series Retrieval and Classification" [15].
Status: In Minor Revision for *Pattern Recognition Letters* **(Qualis A2)**
4) **B. Rozin,** R. S. Torres, F. A. Moura, and D. C. G. Pedronette, "Ball Possession Analysis based on Temporal Network Properties" [16].
Status: Submitted to Journal *Multimedia Tools and Applications* **(Qualis A2)**

During the master's program, the student received a "*Bolsa Estágio de Pesquisa no Exterior (BEPE)*" fellowship and conducted a 3-month research internship (Oct–Dec 2023) at Wageningen University & Research (The Netherlands), under the supervision of Prof. Ricardo da Silva Torres. This internship supported the developments in Chapters 4 and 6.

Acknowledgments

The authors are grateful to São Paulo Research Foundation - FAPESP (grants #2018/15597-6, #2022/01359-1, and #2023/08087-0), Brazilian National Council for Scientific and Technological Development - CNPq (grants #313193/2023-1, and #422667/2021-8), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) and Petrobras (grant #2023/00095-3) for financial support. We also acknowledge Prof. Ricardo da Silva Torres for his guidance during the BEPE internship at Wageningen University & Research.

REFERENCES

[1] K. Ratna Prakarsha and G. Sharma, "Time series signal forecasting using artificial neural networks: An application on ecg signal," *Biomedical Signal Processing and Control*, vol. 76, p. 103705, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1746809422002270

[2] D. Zhu, D. Song, Y. Chen, C. Lumezanu, W. Cheng, B. Zong, J. Ni, T. Mizoguchi, T. Yang, and H. Chen, "Deep unsupervised binary coding networks for multivariate time series retrieval," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 02, pp. 1403–1411, Apr. 2020. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/5497

[3] J. Y. Campbell and N. G. Mankiw, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in *NBER Macroeconomics Annual 1989, V. 4*. National Bureau of Economic Research, Inc, November 1989, pp. 185–246. [Online]. Available: https://ideas.repec.org/h/nbr/nberch/10965.html

[4] D. Song, N. Xia, W. Cheng, H. Chen, and D. Tao, "Deep r -th root of rank supervised joint binary embedding for multivariate time series retrieval," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining*, ser. KDD '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 2229–2238. [Online]. Available: https://doi.org/10.1145/3219819.3220108

[5] P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu, "Temporal convolutional neural (tcn) network for an effective weather forecasting using time-series data from the local weather station," *Soft Computing*, vol. 24, no. 21, pp. 16 453–16 482, Nov 2020. [Online]. Available: https://doi.org/10.1007/s00500-020-04954-0

[6] M. Mudassir, S. Bennbaia, D. Ünal, and M. Hammoudeh, "Time-series forecasting of bitcoin prices using high-dimensional features: a machine learning approach," *Neural Computing and Applications*, 07 2020.

[7] Z. Zeng, T. Balch, and M. Veloso, "Deep video prediction for time series forecasting," in *Proceedings of the Second ACM International Conference on AI in Finance*, ser. ICAIF '21. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3490354.3494404

[8] H. Zhu, P. Zhao, Y.-P. Chan, H. Kang, and D. L. Lee, "Breast cancer early detection with time series classification," in *Proceedings of the 31st ACM International Conference on Information amp; Knowledge Management*, ser. CIKM '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 3735–3745. [Online]. Available: https://doi.org/10.1145/3511808.3557107

[9] Z. Zhang, D. Li, Z. Zhang, and N. Duffield, "A time-series clustering algorithm for analyzing the changes of mobility pattern caused by covid-19," in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Animal Movement Ecology and Human Mobility*, ser. HANIMOB '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 13–17. [Online]. Available: https://doi.org/10.1145/3486637.3489489

[10] A. Abdoli, A. C. Murillo, C.-C. M. Yeh, A. C. Gerry, and E. J. Keogh, "Time series classification to improve poultry welfare," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 635–642.

[11] L. P. Valem, D. C. G. Pedronette, and L. J. Latecki, "Rank flow embedding for unsupervised and semi-supervised manifold learning," *IEEE Transactions on Image Processing*, vol. 32, pp. 2811–2826, 2023. [Online]. Available: https://doi.org/10.1109%2Ftip.2023.3268868

[12] J. Almeida, D. C. G. Pedronette, B. C. Alberton, L. P. C. Morellato, and R. d. S. Torres, "Unsupervised distance learning for plant species identification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 12, pp. 5325–5338, 2016.

[13] S. Pailot-Bonnétat, A. J. L. Harris, S. Calvari, M. De Michele, and L. Gurioli, "Plume height time-series retrieval using shadow in single spatial resolution satellite images," *Remote Sensing*, vol. 12, no. 23, 2020. [Online]. Available: https://www.mdpi.com/2072-4292/12/23/3951

[14] B. Rozin, D. C. G. Pedronette, and R. S. Torres, "Re-ranking and representations for time series retrieval: A comparative study (in major revision)," *IEEE Access*, 2025.

[15] B. Rozin and D. C. G. Pedronette, "A ranked-based framework based on manifold learning for multivariate time series retrieval and classification (in major revision)," *Pattern Recognition Letters*, 2025.

[16] B. Rozin, R. S. Torres, F. A. Moura, and D. C. G. Pedronette, "Ball possession analysis based on temporal network properties (in revision)," *Multimedia Tools and Applications*, 2025.

[17] B. Rozin, E. Bergamim, D. C. G. Pedronette, and F. A. Breve, "Semi-supervised time series classification through image representations," in *International Conference on Computational Science and Its Applications – ICCSA 2023*, O. Gervasi, B. Murgante, D. Taniar, B. O. Apduhan, A. C. Braga, C. Garau, and A. Stratigea, Eds. Cham: Springer Nature Switzerland, 2023, pp. 48–65. [Online]. Available: https://doi.org/10.1007/978-3-031-36808-0_4

[18] C. Ji, M. Du, Y. Hu, S. Liu, L. Pan, and X. Zheng, "Time series classification based on temporal features," *Applied Soft Computing*, vol. 128, p. 109494, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494622005889

[19] D. Quoc Nguyen, M. Nguyet Phan, and I. Zelinka, "Periodic time series forecasting with bidirectional long short-term memory: Periodic time series forecasting with bidirectional lstm," in *2021 The 5th International Conference on Machine Learning and Soft Computing*, ser. ICMLSC'21. New York, NY, USA: Association for Computing Machinery, 2021, p. 60–64. [Online]. Available: https://doi.org/10.1145/3453800.3453812

[20] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[21] R. D. S. Torres and A. X. Falcão, "Content-based image retrieval: Theory and applications," *Revista de Informática Teórica e Aplicada*, vol. 13, pp. 161–185, 2006.