Curved Spaces, Enhanced Diagnosis: Hyperbolic Neural Networks for Multi-Label ECG Classification

Pedro Dutenhefner[1], Diogo Tuler[1], Turi Rezende[1], Jose Fernandes[1], Luiz Facury [1], Luísa Porfírio[1], Yan Aquino[1], Arthur Buzelin[1], Pedro Bento[1], Gabriela Paixão[1], Gisele Pappa[1], Antônio Ribeiro[1], Wagner Meira Jr[1]

[1] Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

Abstract

In recent years, the use of deep learning models has established itself as a transformative approach in cardiology. In this context, disease categories often present hierarchical relationships, such as subtypes of arrhythmias. Moreover, Euclidean latent spaces struggle to encode hierarchical clinical structure, whereas hyperbolic geometry, with its negative curvature and exponential volume growth, naturally represents hierarchies with lower distortion, making it a promising basis for medical representation learning. In this work, we propose a model that integrates a Euclidean convolutional neural network backbone with a hyperbolic representation learning scheme. Guided by cardiology prior knowledge of class hierarchy, we introduce a hierarchy-aware loss that supervises superclasses, enforces parent–child consistency, and shapes embeddings geometry via radial/angular regularizers. Experimental results on multi-label ECG classification show consistent gains over a fully Euclidean baseline: macro F1 improves from 0.760 (ResNet) to 0.786 with a hyperbolic head (HypNet) and to 0.8509 with the proposed hierarchy-aware model (H-HypNet); macro recall rises from 0.662 to 0.7077 and to 0.787, respectively. These findings provide preliminary indications that hyperbolic representations, coupled with hierarchy-aware training, help model complex cardiovascular taxonomies in medical deep learning.

1. Introduction

Artificial intelligence has become a transformative tool in healthcare, particularly for cardiology. As cardiovascular diseases (CVDs) remain the leading cause of global mortality [1], deep learning models are increasingly applied to electrocardiograms (ECGs) for tasks such as disease classification [2, 3] and cardiological age prediction [4, 5], leveraging their ability to extract rich representations directly from raw signals.

Representation learning seeks to encode data into embeddings where semantic similarity is preserved: similar cases are mapped closer together, while dissimilar ones are pushed apart. This geometric property facilitates the use of simple classifiers and distance metrics. However, Euclidean space imposes inherent limitations on how well such relationships can be captured [6]. As an alternative, hyperbolic geometry—commonly modeled through the Poincaré ball—offers a more natural representation of hierarchical structures due to its negative curvature and exponential volume growth. This makes it especially suitable for domains where data are organized in taxonomies, such as cardiovascular disease hierarchies [7].

Building on these insights, we introduce a novel framework for ECG classification that integrates a Euclidean convolutional backbone with hyperbolic representation learning. In our method, embeddings are mapped onto the Poincaré ball, where classification is guided by hyperbolic decision boundaries. This allows the latent space to naturally capture the hierarchical structure of cardiac disorders. To further leverage domain knowledge, we design a hierarchy-aware loss function that jointly supervises leaf-level and superclass labels, enforces consistency between parent and child classes, and shapes the embedding geometry through simple radial and angular regularizers. This loss ensures that the learned representations remain aligned with the medical taxonomy of conduction blocks, rhythm disorders, and normal cases.

We evaluate our model on a large-scale, clinically validated ECG dataset. Results show that integrating hyperbolic embeddings improves performance compared to a Euclidean baseline (ResNet), and that adding the proposed hierarchy-aware loss yields further gains. In particular, our hierarchy-aware hyperbolic network (H-HypNet) achieves a macro F1-score of 0.8509 and macro recall of 0.787, surpassing the Euclidean and standard hyperbolic counterparts.

2. Related work

Several studies have proposed neural network models for cardiology-related tasks. For disease classification,

 ISSN: 2325-887X DOI: 10.22489/CinC.2025.444

Ribeiro. [2] developed a convolutional neural network for ECGs and reported predictive performance superior to that of cardiologists in identifying six heart conditions. Aiming to address the different temporal scales of ECG signals, Dutenhefner [8] and Buzelin [9] proposed hierarchical transformer architectures for disease classification.

Additionally, Tuler et al. [10] leveraged expert cardiologists' prior knowledge within a Mixture of Experts (MoE) model to improve hierarchical disease prediction, aligning model outputs with medically meaningful disease relationship. However, such approach may fragment the representation of inter-class dependencies, potentially limiting the global modeling of similarity structures.

In this regard, deep learning techniques in non-Euclidean spaces have gained attention in recent years. The pioneering work by Nickel and Kiela [6] demonstrated that embeddings in the Poincaré hyperbolic model can capture hierarchical data structures. In the field of computer vision, Khrulkov et al. [7] projected image embeddings onto the Poincaré ball, showing that ambiguous examples tend to be placed near the center while clearer examples are distributed toward the boundary—thus capturing latent hierarchies.

In the healthcare domain, Ayubcha et al. [11] proposed a hyperbolic neural network for classifying neuroimaging data and demonstrated how the network captured the semantic structure of diseases. However, to the best of our knowledge, no prior works have applied hyperbolic neural networks to ECG classification.

3. Preliminaries on Hyperbolic Geometry

The Poincaré ball is a model of hyperbolic space, defined as the Riemannian manifold $\mathbb{B}^{\mathbb{F}}$. Here, F denotes the dimension of the latent space. The ball consists of all points $x \in \mathbb{R}^{\mathbb{F}}$ lying inside a radius $1/\sqrt{c}$, with geometry defined by the metric tensor $g_x^c = \lambda_x^c I_F$ where $\lambda_x^c = \frac{2}{1-c\|x\|^2}$. Near the boundary, distances grow rapidly, as the metric stretches space when $\|x\| \to 1/\sqrt{c}$. This naturally accommodates hierarchies by placing general concepts near the center and specific ones toward the edge [7]. Mappings between the tangent space and the Poincaré ball are given by the exponential and logarithmic maps: the former projects Euclidean vectors onto hyperbolic space, and the latter brings points back. Both preserve direction while rescaling magnitudes, ensuring consistency between Euclidean and hyperbolic computations. For mathematical details, see [12].

4. H-HypNet

The proposed network combines the convolution backbone of [2] with a hyperbolic classification head. The ECG is encoded into a vector $h \in \mathbb{R}^{\mathbb{F}}$, which is then lifted to the Poincaré ball $\mathbb{B}^{\mathbb{F}}$ using the exponential map at the origin.

Classification is performed in the Poincaré ball $\mathbb{B}^{\mathbb{F}}$ using

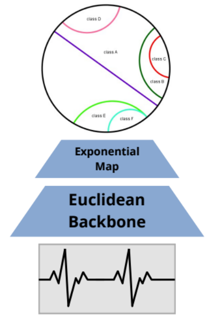

Figure 1. H-HypNet: an Euclidean CNN encodes the ECG, features are mapped to the Poincaré ball, and classification is done with hyperbolic gyroplanes.

one pair of learnable parameters (z_k, r_k) per class. Each class learns a *direction* z_k that points toward its region of the ball and a *bias* r_k that determines how far from the origin the boundary lies. The boundary itself is a *gyroplane*, the hyperbolic analogue of a Euclidean hyperplane: it bends with the curvature of the ball but still divides space into two half–spaces. Following the multinomial logistic regression approach of [13], logits are obtained as signed distances to these gyroplanes, so that points aligned with z_k and farther from the origin yield larger scores and more confident predictions. The overall architecture is illustrated in Fig. 1.

On top of standard leaf supervision, we add a lightweight hierarchical loss that (i) supervises superclasses, (ii) promotes parent $\geq$ child consistency, and (iii) shapes geometry by placing parents more centrally and separating top-level branches. The backbone is unchanged.

4.1. Hierarchically-Aware Hyperbolic Loss

The proposed loss leverages the label hierarchy to enhance generalization, calibration, and interpretability. It jointly supervises leaf and superclass labels, enforces logical parent–child consistency, and shapes the hyperbolic embedding so that radial depth encodes diagnostic specificity, while angular separation reflects diseases families.

We consider six leaf labels (1dAVb, RBBB, LBBB, SB, AF, ST) and three superlabels (*blockage*, *rhythm*, *normal*). Superlabels are derived directly from the leaves: *blockage* is positive if any of 1dAVb, RBBB, LBBB is present; *rhythm* if any of SB, AF, ST; and *normal* only when none of the six apply, following the hierarchy defined in [10]. This structure is encoded through edges $\mathcal{E}$ = *blockage*$\rightarrow$1dAVb, RBBB, LBBB, *rhythm*$\rightarrow$SB, AF, ST. Each class is represented by a prototype z_k in the Poincaré ball with direction u_k, while the sample radius r corresponds to the embedding's geodesic distance from the origin.

The total loss is a weighted sum $L = L_{\text{leaf}} + w_{\text{sup}} L_{\text{sup}} + w_{\text{con}} L_{\text{con}} + w_{\text{rad}} R_{\text{rad}} + w_{\text{ang}} R_{\text{ang}} + w_{\text{r}} R_{\text{radius}}$, where the $w_{\cdot} > 0$ are tunable weights.

(1) Leaf BCE. This term supervises fine-grained diagnoses and provides the primary discriminative signal. For leaf logits s_k and targets y_k:

$$L_{\text{leaf}} = \frac{1}{6} \sum_{k \in \mathcal{L}} \text{BCEWithLogits}(s_k,\, y_k).$$

(2) Superclass BCE. This term encourages the model to learn coarse groupings and provides stability when leaf-level evidence is ambiguous. Given logits s_s and targets y_s for the three superlabels:

$$L_{\text{sup}} = \frac{1}{3} \sum_{s \in \{\text{blockage,rhythm,normal}\}} \text{BCEWithLogits}(s_s,\, y_s).$$

(3) Consistency. This term enforces logical parent–child consistency by ensuring parent scores are never lower than those of their children:

$$L_{\text{con}} = \frac{1}{|\mathcal{E}|} \sum_{(p \to c) \in \mathcal{E}} \max\big(0,\, s_c - s_p\big).$$

(4) Radial Ordering. This term encodes hierarchical depth by encouraging parents to be more central than children in the Poincaré ball. Here z_k denotes the prototype of class k in the Poincaré ball:

$$R_{\text{rad}} = \frac{1}{|\mathcal{E}|} \sum_{(p \to c) \in \mathcal{E}} \max\big(0,\, \|z_p\| - \|z_c\|\big).$$

(5) Branch Separation. This term separates top-level families into distinct angular sectors by penalizing insufficient separation between their directions. Here $u_k = \log_0(z_k)/\|\log_0(z_k)\|$ is the unit direction of class k:

$$R_{\text{ang}} = \max\big(0,\, \cos(u_{\text{blockage}}, u_{\text{rhythm}}) - t\big),$$

where $t \geq 0$ controls the desired angular margin.
(6) Radius Prior. This term constrains "normal" cases to lie near the origin while preventing abnormal cases from being too central:

$$R_{\text{radius}} = \frac{a}{N} \sum_{i=1}^{N} \Big(y_i\, r_i \;+\; (1 - y_i)\, \max\{0,\, r_0 - r_i\}\Big),$$

where $y_i \in \{0, 1\}$ indicates "normal" for sample i, r_i is its radius, r_0 a margin, and $a > 0$ a prior weight.

5. Results

Experimental Setup: H-HypNet was developed using the CODE-15 dataset [2], containing 345,779 ECGs from 233,770 patients across six clinically relevant disorder types. Performance was evaluated on CODE-TEST, with 827 exams labeled by consensus of two or three cardiologists.

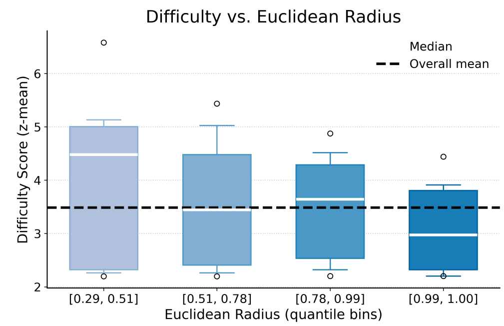

Figure 2. Relationship between cardiologist disagreement and the Euclidean radius of latent representations

Three models were compared: ResNet (Euclidean CNN), HypNet (CNN with hyperbolic head), and H-HypNet (CNN with hyperbolic head and hierarchical loss).
Results. With the same CNN backbone, moving from the Euclidean ResNet to the hyperbolic HypNet already improves overall discrimination: macro F1 and macro recall rise, indicating that the non-Euclidean geometry captures structure that aids multi-label ECG classification. These gains are distributed across most disorders (e.g., LBBB, SB, AF, ST), reflecting better separation of clinically coherent regions in representation space and fewer borderline errors near decision boundaries. Adding the proposed hierarchical loss (H-HypNet) yields the strongest and most consistent improvements: macro F1 increases from roughly 0.76 (ResNet) to about 0.85, and macro recall climbs from about 0.66 to nearly 0.79, while maintaining high precision. The recall gain is particularly important in medicine, as higher sensitivity reduces false negatives and the risk of missed diagnoses—critical for conditions such as atrial fibrillation or conduction blocks. The hierarchical terms reinforce coarse-to-fine supervision and logical consistency, shaping the hyperbolic embedding so that parent–child relations align with radial depth and branch directions, ultimately translating into more reliable detection without modifying the encoder.
Latent Space Analysis: The test set was annotated by multiple cardiologists, allowing us to derive a continuous *difficulty score* from inter–rater disagreement, with higher values indicating harder cases. Relating this score to the Euclidean radius of latent embeddings (Fig. 2), we find that smaller radii correspond to more difficult, less consensual exams, while difficulty decreases as radius grows, a trend confirmed by a negative Pearson correlation ($r = -0.21$). Intuitively, embeddings near the origin lie in indecisive regions close to class boundaries, whereas larger radii represent clearer, prototypical phenotypes with higher expert agreement. Geometrically, hyperbolic multinomial logistic regression defines class boundaries as *gyroplanes* in the Poincaré ball, where the conformal factor amplifies angular separation with radius, yielding greater margins and thus

Classes	Accuracy			Precision			Recall			F1 Score		
	ResNet	HypNet	H-HypNet	ResNet	HypNet	H-HypNet	ResNet	HypNet	H-HypNet	ResNet	HypNet	H-HypNet
1dAVb	0.981	0.9807	**0.9855**	**1.000**	0.9286	0.9444	0.425	0.4643	**0.6071**	0.591	0.619	**0.7391**
RBBB	0.987	**0.9952**	0.994	**0.971**	0.9412	0.9394	0.656	**0.9412**	0.9118	0.751	**0.9412**	0.9254
LBBB	0.990	0.9915	**0.994**	0.981	**1.000**	**1.000**	0.725	0.7667	**0.8333**	0.809	0.8679	**0.9091**
SB	0.990	0.9891	**0.9927**	**0.949**	0.7692	0.8125	0.714	0.625	**0.8125**	0.794	0.6897	**0.8125**
AF	0.991	0.9927	**0.9952**	0.959	0.8182	**1.000**	**0.708**	0.6923	0.6923	0.796	0.75	**0.8182**
ST	**0.992**	0.9879	0.9915	0.956	**0.9655**	0.9412	0.744	0.7568	**0.8649**	0.819	0.8485	**0.9014**
Macro	0.989	0.9895	**0.9921**	**0.969**	0.9038	0.9396	0.662	0.7077	**0.787**	0.760	0.786	**0.8509**

Table 1. Classification metrics per class comparing the baseline ResNet, the intermediate HypNet, and the optimized H-HypNet model.

more confident predictions.

6. Conclusion

We introduced H-HypNet, a hyperbolic head with a lightweight hierarchy-aware loss for multi-label ECG classification. On a clinically curated benchmark, our approach consistently outperformed a purely Euclidean baseline, notably improving macro F1 and recall, while a latent-space analysis linked larger radii to easier, more prototypical cases. These results suggest that hyperbolic geometry, coupled with hierarchical supervision, better captures cardiovascular taxonomies and yields more reliable predictions. Future work will extend to richer hierarchies, calibration and uncertainty modeling, and broader clinical validation.

Acknowledgments

This work was partially funded by CNPq, CAPES, FAPEMIG, CIIA-Saúde, and IAIA - INCT on AI. We thank the Telehealth Center of Minas Gerais for data access and fruitful discussions.

References

[1] Organization WH. Cardiovascular diseases, 2024. Online report, 2024. Accessed: 2024-08-16.

[2] Ribeiro AH, Ribeiro MH, Paixão GMM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MP, Andersson CR, Macfarlane PW, Meira Jr. W. Automatic diagnosis of the 12-lead ecg using a deep neural network. Nature Communications 2020;11(1):1760.

[3] Li X, Li C, Wei Y, Sun Y, Wei J, Li X, Qian B. Bat: Beat-aligned transformer for electrocardiogram classification. In 2021 IEEE International Conference on Data Mining (ICDM). 2021; 320–329.

[4] Robles Dutenhefner P, Lemos G, Rezende T, Fernandes JG, Tuler D, Pappa GL, Paixão GM, Ribeiro ALP, Meira Jr. W. Ecg-resnext: Age prediction in pediatric electrocardiograms and its correlations with comorbidities. In Proceedings of the XXI Encontro Nacional de Inteligência Artificial e Computacional (ENIAC). 2024; 49–60.

[5] Lima EM, Ribeiro AH, Paixão GMM, Ribeiro MH, Pinto-Filho MM, Gomes PR, Oliveira DM, Sabino EC, Duncan BB, Giatti L, et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. Nature Communications 2021;12(1):5117.

[6] Nickel M, Kiela D. Poincaré embeddings for learning hierarchical representations. In Advances in Neural Information Processing Systems (NeurIPS 30). 2017; .

[7] Khrulkov V, Mirvakhabova L, Ustinova E, Oseledets I, Lempitsky V. Hyperbolic image embeddings. In Proceedings of the IEEECVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020; 6418–6428.

[8] Dutenhefner PR, Rezende TAV, Pappa GL, de Matos Paixão GM, Ribeiro ALP, Jr WM. Um transformer hierárquico para classificação e diagnóstico de eletrocardiograma. In Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS). 2024; .

[9] Buzelin A, Dutenhefner PR, Rezende T, Porfirio LG, Bento P, Aquino Y, Fernandes J, Santana C, Miana G, Pappa GL, Ribeiro A, Jr WM. A cnn-based local-global self-attention via averaged window embeddings for hierarchical ecg analysis. In International Conference on Bioinformatics and Biomedical Engineering (iCBBE). 2024; .

[10] Tuler D, Dutenhefner PR, Fernandes JG, Rezende T, Lemos G, Pappa GL, Paixão G, Ribeiro A, Jr WM. Leveraging cardiologists prior-knowledge and a mixture of experts model for hierarchically predicting ecg disorders. In Computing in Cardiology (CinC). 2024; .

[11] Ayubcha C, Sajed S, Omara C, Singh SB, Lokesha YU, Liu A, Aziz-Sultan MA, Smith TR, Beam A. Geometric deep learning methods for improved generalizability in medical computer vision: Hyperbolic convolutional neural networks in multi-modality neuroimaging. medRxiv preprint, 2024.

[12] Ungar AA. A Gyrovector Space Approach to Hyperbolic Geometry, volume 4 of Synthesis Lectures on Mathematics and Statistics. San Rafael, CA, USA: Morgan & Claypool Publishers, 2009.

[13] van Spengler M, Berkhout E, Mettes P. Poincare resnet. Proceedings of the IEEECVF International Conference on Computer Vision ICCV October 2023;5419–5428.

Address for correspondence:

Pedro Robles Dutenhefner
DCC–ICEx–UFMG - Pampulha
Belo Horizonte, MG 31270-901
Brazil
pedroroblesduten@gmail.com