

## **Geometric Deep Learning Methods for Improved Generalizability in Medical Computer Vision:**

### **Hyperbolic Convolutional Neural Networks in Multi-Modality Neuroimaging**

Cyrus Ayubcha Ph.D. <sup>1,2,3</sup>, Sulaiman Sajed B.S. <sup>3,4</sup>, Chady Omara M.D. <sup>3,5</sup>, Anna B. Veldman M.D. <sup>3,5</sup>, Shashi Bhushan Singh M.D. <sup>6</sup>, Yashas Ullas Lokesha M.D. <sup>7</sup>, Alex Liu B.S. <sup>8</sup>, Mohammad Ali Aziz-Sultan M.D. <sup>3</sup>, Timothy R. Smith M.D. Ph.D. <sup>3</sup>, Andrew Beam Ph.D. <sup>2</sup>

1. Harvard Medical School, Boston, MA, United States
2. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States
3. Department of Neurosurgery, Cushing Neurosurgical Outcomes Center, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
4. Department of Medical Science, Boston University, Boston, USA
5. Department of Neurosurgery, Leiden University Medical Center, Leiden, the Netherlands
6. Stanford University School of Medicine, Stanford, CA, 94305
7. Department of Radiology, Sri Devaraj Urs Medical College, Karnataka, India
8. Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA

Corresponding Author:

Andrew Beam

Harvard T.H. Chan School of Public Health,

677 Huntington Ave, Boston, MA 02115

[Andrew\\_Beam@hms.harvard.edu](mailto:Andrew_Beam@hms.harvard.edu)

Word Count: 3123

## **Abstract**

*Objective:* This study investigates the potential advantages of hyperbolic convolutional neural networks (HCNNs) over traditional convolutional neural networks (CNNs) in neuroimaging tasks.

*Materials and Methods:* We conducted a comparative analysis of HCNNs and CNNs across various medical imaging modalities and diseases, with a focus on a compiled multi-modality neuroimaging dataset. The models were assessed for performance parity, robustness to adversarial attacks, semantic organization of embedding spaces, and generalizability. Zero-shot evaluations were also performed with ischemic stroke non-contrast CT images.

*Results:* HCNNs matched CNN performance on less complex settings and demonstrated superior semantic organization, and robustness to adversarial attacks. While HCNNs equaled CNNs in out-of-sample datasets identifying Alzheimer's disease, in zero-shot evaluations, HCNNs outperformed CNNs and radiologists.

*Discussion:* HCNNs deliver enhanced robustness and organization in the neuroimaging data. This likely underlies why while HCNNs perform similarly to CNNs with respect to in-sample tasks, they confer improved generalizability. Nevertheless, HCNNs encounter efficiency and performance challenges with larger, complex datasets. These limitations underline the need for further optimization of HCNN architectures.

*Conclusion:* HCNNs present promising improvements in generalizability and resilience for medical imaging applications, particularly in neuroimaging. Despite challenges with larger datasets, HCNNs enhance performance under adversarial conditions and offer better semantic organization, suggesting valuable potential in generalizable deep learning models in medical imaging and neuroimaging diagnostics.

**Keywords:** Hyperbolic Neural Networks, Euclidean, Convolutional Neural Networks, Lorentz, Neuroimaging, Medical Imaging, Generalizability, Adversarial Robustness, Hierarchical Data Structures

## Background and Significance:

Advances in computational power have facilitated the expansion of predictive deep learning models in many fields. While deep learning models are idealized as universal approximators, not all tasks are not equally appropriate for all neural network architectures; such that, misalignment may result in subpar empirical task performance.<sup>1</sup> One notable instance of this may involve data structures with inherent hierarchical structure.<sup>2</sup> Tree-like or hierarchical data structures have been shown to be represented with superior fidelity with less distortion in hyperbolic space as compared to Euclidian space.<sup>3</sup> Specifically, the hyperbolic manifold allows for exponential scaling from the radial axis which mirrors the distance relationships found in hierarchical structures and accordingly, prevents distortion and information loss.<sup>4</sup>

There has been notable effort to utilize the superior data representation observed in hyperbolic spaces with deep learning algorithms. Over the last few years, many have attempted to operationalize constructs of hyperbolic space in a computationally efficient manner with the goal of reconstructing the fundamental functionality of neural network operations consistent with hyperbolic space.<sup>2</sup> Thus, Hyperbolic Neural Networks (HNNs) were developed as an alternative to standard Euclidean neural networks.<sup>5</sup> While most papers explore densely connected HNNs, there have been efforts to use alternative neural network architectures. For instance, Khrulkov et al developed a hybrid HNN that was applied to image datasets by utilizing a traditional Euclidean convolution structure and connected layer prior to mapping the resulting embeddings into hyperbolic space and conducting a multi-class logistic regression.<sup>6</sup> Khrulkov et al showed superior performance to analogous Euclidean models in certain datasets (Caltech-UCSD Birds, DukeMTMC-reID dataset) as well as in few-shots classification.<sup>6</sup> They also showed that HNNs provided more intuitive organization of the classes in the embedding space likely explaining their improved out-of-sample and out-of-distribution performance.<sup>6</sup>

Given issues with numerical stability during training, most literature has moved to make HNNs more numerically stable.<sup>7,8,9</sup> Guo et al proposed a clipping mechanism that would bound the embedding space so coerce numerically stabilize during training.<sup>9</sup> These clipped HNNs were found to outperform standard HNNs on various benchmarks, including CIFAR10, CIFAR100, and ImageNet, and demonstrated better adversarial robustness and out-of-

distribution detection.<sup>9</sup> Unlike unclipped HNNs, clipped HNNs achieved performance on par with ENNs in data settings without natural tree structures. The use of the Lorentz model of hyperbolic spaces, which has different mathematical constructions for the computational operations of the neural network as compared to the Poincaré ball model, has also been proven to reduce numerical instability.<sup>8,10</sup> Finally, there have also been efforts to translate common Euclidean convolutional neural network operations into hyperbolic space where the fully hyperbolic convolutional neural networks (HCNN) could be contained to hyperbolic space in an end-to-end fashion.<sup>11</sup> This prevents the need for mapping between Euclidean and hyperbolic spaces so to limit numerical instability.

Nevertheless, the literature remains uncertain in the fully exploring the value of HNNs. The original seminal paper in hyperbolic imaging embeddings suggests that most imaging datasets have some degree of implicit hierarchical structure.<sup>6</sup> Other studies found poor performance in settings without any natural hierarchical structures as compared to Euclidean counterparts until Guo et al suggested that clipped HNNs achieved similar performance in certain settings.<sup>9</sup> In totality, there is a clear need to further robust evaluation of HCNNs in various domains to weight the possible benefits and limitation of these evolving models.

One field of application includes medical imaging where the successful application of computer vision algorithms has been keenly appreciated.<sup>12</sup> To our knowledge, only one study has been performed utilizing HNNs for classification of medical imaging data. Utilizing Khrulkov et. al's hybrid paradigm, Yu et al introduces a hyperbolic prototype network capable of jointly learning image embeddings and class prototypes in a shared hyperbolic space, guided by an error construction mechanism derived from a prior known class hierarchy.<sup>13</sup> Their approach preserved the semantic class relationships of dermatoscope images in the hyperbolic embedding space and found superior performance in classification as compared to analogous Euclidean approaches, though with lower space curvature hyperparameters.<sup>13</sup>

Given the success of present day convolutional neural networks in imaging recognition tasks, there has been a greater interest in developing models with both wide spanning capabilities across a variety of tasks or classes; and those durable to a variety of clinical settings where models may encounter rare patient morphologies or imperfect images.<sup>14-17</sup> We explore the value of HCNNs by evaluating the performance of clipped HCNNs relative to their

Euclidean counterparts in neuroimaging settings as well as other multi-modality medical imaging databases.

Agnostic to any prior data hierarchy, we conduct an evaluation of in sample performance across three medical imaging datasets of varying complexity. We evaluate the neuroimaging models not only in their respective classification performance but also in their organization of embedding spaces, durability to adversarial attacks, and out-of-sample as well as zero-shot generalizability.

## Materials and Methods:

### *Data Sources:*

We develop a core neuro-imaging dataset to evaluate the HCNN and CNN models: Multi-Modality Neuroimaging (MMN) Dataset composed of 72,634 images of 42 total classes. The images were acquired from previously open-source databases which included various neurological diseases including ischemic stroke<sup>18</sup>, hemorrhagic stroke<sup>19</sup>, metastasis<sup>20</sup>, tumor<sup>21</sup>, schizophrenia (COBRE, MCICShare)<sup>22</sup>, and Alzheimer's disease (AD)<sup>23</sup> across computed tomography (CT) and Magnetic Resonance Imaging (MRI) modalities. Images not from a peer-reviewed source, were independently confirmed to be correctly classified according to a qualified radiologist.

Certain classes in the neuroimaging dataset were manually composed due to either overlapping classes or, occasionally, to better capture disease signals. Specifically, the “normal” categories are composed from multiple appropriate non-diseased classes in the datasets above. For instance, patients in the schizophrenia databases defined as “normal” patients was moved into the respective normal classes once we excluded the possibility of them joining another diseased class in the database. Schizophrenia positive scans were restricted to four of the median axial scans for each patients to better capture morphological abnormalities commonly conferred by patients with schizophrenia.

31

We also constructed two additional datasets of other medical imaging types to conduct further analyses in larger and smaller class settings: Miniature Multi-Disease Dataset (MMD), and Multi-Disease Dataset (MD). The MD is composed of 89,496 images of 78 total classes. These images were acquired from previously published open-source databases which include Chest X-Rays<sup>24</sup>, Fundoscopy<sup>25</sup>, Gastrointestinal Scopes<sup>26</sup>, Musculoskeletal X-Rays<sup>27</sup>, Neuroimaging and Dermatoscopy.<sup>28–30</sup> Finally, the MMD was restricted to a smaller subset of the MMD with a total of 19,880 unique images across 16 total classes. The classes and their respective imaging data can be observed in

### **Table 1.**

Out-of-sample data were acquired from the T-1 MRI sequences of the OASIS-1 cross-sectional cohort imaging study<sup>32</sup>. We derived one mid-brain axial slice from each patient in the study and defined an AD positive case as an

individual with a Mini-Mental State Examination (MMSE) score of below 25 and negative otherwise. In total, this dataset entailed a total of 436 subjects where 40 subjects were AD positive and 396 were AD negative. The zero-shot dataset was acquired from the Mass General Brigham (MGB) Research Patient Data Registry (RPDR) with MGB Institutional Review Board approval. We retrospectively selected individuals from January 2004 to December 2022 who received Non-Contrast Computed Tomography (NCCT) brain scan upon presenting to the emergency department. We further restricted ourselves to those diagnosed with acute ischemic strokes on Magnetic Resonance Imaging (MRI) brain scans within 7 days of their original presentation to the emergency department. We include the full axial scans of the 151 patients for a total of 23,371 images. We also document the radiologist reports and their ability to correctly able to diagnose the ischemic stroke on NCCT. We label the axial images as positive if they included regions implicated by the positive MRI reads for that patient where all other images are deemed negative for that patient.

#### *Processing and Models:*

All training data underwent uniform pre-processing transformations prior to their use in the model including normalization, rotation, and horizontal flip. Furthermore, to homogenize the diverse presentation of brain imaging data, we conducted skull-stripping when appropriate. For all the models in this study, we utilize a Res-Net-18 (RN18) backbone where we define the baseline Euclidean CNNs as the RN18. The HCNN models are constructed as hybrid models with an identical convolution structure as the CNN. However, the HCNN model acquires the subsequent embeddings and translates them into the Lorentz space with an exponent mapping procedure where the remaining dense layers of the RN18 backbone reside. Other features of our hybrid HCNN construct include trainable curvature, feature-clipping<sup>9</sup>, and Euclidean reparameterizations<sup>10</sup> which are documented by the Bdeir et al<sup>33</sup> code base we utilize for our study.

#### *Evaluation:*

To examine the performance of each model in the respective dataset, we reproduce cross-entropy loss, top-1 accuracy, and top-5 accuracy. Utilizing the average embedding position of each class we construct a low dimensional representation using T-SNE algorithms with the Euclidean and Lorentz models, respectively. We also

utilize hierarchical clustering procedures with the average class embedding positions for each model which allow us to derive a dendrogram of the inter-class relationships learned by the respective models.

To examine the interpretability of the embedding space, we compare the geodesic distance matrices of average class embeddings from the respective Euclidean and Euclidean-Lorentz models to a ground truth hierarchical distance matrix. We derive the ground truth distance matrix based on the sets of categorical variables as descriptors of the imaging category and disease type (refer to Appendix). We first normalize all distance matrices and compute the absolute pairwise difference between the model distance matrix and the ground truth distance matrix. We also derive the Spearman's rank correlation coefficient between the respective model and ground truth distances matrices based on the ranked distance of the pairwise classes.

To evaluate out-of-sample accuracy, we utilize a single median axial slice in each OASIS-1 subject and consider a positive diagnosis as one that ascribes an AD-related class to the scan and a negative diagnosis as a normal T1 MRI predicted class. Next, we evaluate zero-shot accuracy by defining a true positive diagnosis read as having an axial NCCT image predicted as an ischemic or hemorrhagic stroke class by the model where the ground truth was positive. At a patient level, we identify accuracy as a patient having at least one true positive. We define a true negative as a normal CT class. Then we identify overall image-based accuracy as the number of true positives and true negatives over the overall number of images. Finally, we conduct a Projected Gradient Descent (PGD) adversarial attack to assess the comparative durability of each model to more extreme distortions in data. Note that all 95% confidence intervals are derived from a respective bootstrapping procedure ( $n=1000$ ).



Table 1: Dataset Characteristics

Miniature	Modality	Case	Image	Image
Derm: Actinic Keratosis	AD Moderate MRI T1	867	896	Derm: Actinic Keratosis
Derm: Basal Cell Carcinoma	AD Severe MRI	3323	64	Pulm: Bacterial Pneumonia
Derm: Benign Keratosis	AD Mild MRI T1	2624	2240	Derm: Basal Cell Carcinoma
Derm: Dermatofibroma	Hemorrhagic Stroke Epidural CT Bone	239	167	Derm: Benign Keratosis
Derm: Melanoma	Hemorrhagic Stroke Intraparenchymal CT Bone	4522	52	Derm: Dermatofibroma
Derm: Melanocytic Nevi	Hemorrhagic Stroke Intraventricular CT Bone	12875	13	Neuro: Epidural Hemorrhagic Stroke CT Bone
Derm: Squamous Cell Carcinoma	Hemorrhagic Stroke Subarachnoid CT Bone	628	9	Neuro: Intraparenchymal Hemorrhagic Stroke CT Bone
Derm: Vascular Lesion	Hemorrhagic Stroke Subdural CT Bone	253	52	Neuro: Subdural Hemorrhagic Stroke CT Bone
Gastro: Dyed Lifted Polyps	Hemorrhagic Stroke Epidural CT Brain	1000	167	Neuro: Epidural Hemorrhagic Stroke CT Brain
Gastro: Dyed Resection Margins	Hemorrhagic Stroke Intraparenchymal CT Brain	1000	52	Neuro: Intraparenchymal Hemorrhagic Stroke CT Brain
Gastro: Esophagitis	Hemorrhagic Stroke Intraventricular CT Brain	1000	13	Neuro: Subdural Hemorrhagic Stroke CT Brain
Ophtho: Normal Fundus	Hemorrhagic Stroke Subarachnoid CT Brain	2873	9	Neuro: Ischemic Stroke DWI
Gastro: Normal Cecum	Hemorrhagic Stroke Subdural CT Brain	1000	52	Neuro: Ischemic Stroke Flair
Gastro: Normal Pylorus	Ischemic Stroke MRI DWI	1000	1012	Derm: Melanoma
Gastro: Normal Z-Line	Ischemic Stroke MRI Flair	1000	1002	Neuro: Metastasis Flair
Gastro: Polyps	Metastasis MRI Flair	1000	4248	Neuro: Metastasis T1
	Metastasis MRI TIC		4248	Neuro: Metastasis TIC+
	Metastasis MRI T1		4248	Pulm: Normal CXR
	Normal CT Bone		1495	Derm: Melanocytic Nevi
	Normal CT Brain		1494	Neuro: Normal CT Bone
	Normal MRI DWI		1406	Neuro: Normal CT Brain
	Normal MRI Flair		14949	Neuro: Normal DWI
	Normal MRI T1		17925	Neuro: Normal_Flair
	Normal MRI TIC+		13941	Neuro: Normal T1
	Normal MRI T2		18	Neuro: Normal TIC+
	Schizophrenia MRI DWI		471	Neuro: Normal T2
	Schizophrenia MRI T1		1314	Derm: Squamous Cell Carcinoma
	Glioma MRI TIC+		152	Neuro: Schizophrenia DWI
	Meningioma MRI TIC+		233	Neuro: Schizophrenia T1
	Neurocitoma MRI TIC+		76	Neuro: Glioma T1
	Other Lesions MRI TIC+		9	Neuro: Meningioma T1
	Schwannoma MRI TIC+		36	Neuro: Neurocitoma T1
	Glioma MRI T1		65	Neuro: Schwannoma T1
	Meningioma MRI T1		141	Neuro: Glioma TIC+
	Neurocitoma MRI T1		39	Neuro: Meningioma TIC+
	Other Lesions MRI T1		27	Neuro: Neurocitoma TIC+
	Schwannoma MRI T1		31	Neuro: Schwannoma TIC+
	Glioma MRI T2		9	Neuro: Glioma T2
	Meningioma MRI T2		67	Neuro: Meningioma T2
	Neurocitoma MRI T2		145	Neuro: Schwannoma T2

	Other Lesions MRI T2	14	Derm: Vascular Lesion	253
	Schwannoma MRI T2	33	Pulm: Viral Pneumonia	1493
			Ophtho: Branch Retinal Vein Occlusion	19
			Ophtho: Cataract	287
			Ophtho: Diabetic Retinopathy	53
			Ophtho: Drusen	148
			Ophtho: Dry Age-Related Macular Degeneration	202
			Gastro: Dyed Lifted Polyps	1000
			Gastro: Dyed Resection Margins	1000
			Ophtho: Epiretinal Membrane	140
			Gastro: Esophagitis	1000
			Ophtho: Glaucoma	213
			MSK: Hand Normal	877
			MSK: Hand Fractured	379
			MSK: Hand Shoulder Normal	180
			MSK: Hand Shoulder Fractured	53
			MSK: Hip Normal	169
			MSK: Hip Fractured	13
			Ophtho: Hypertensive Retinopathy	123
			Ophtho: Macular Epiretinal Membrane	140
			Ophtho: Maculopathy	23
			Ophtho: Mild Nonproliferative Retinopathy	464
			Ophtho: Moderate Non-Proliferative Retinopathy	798
			Ophtho: Myelinated Nerve Fibers	68
			Ophtho: Normal Fundus	2873
			Gastro: Normal Cecum	1000
			Gastro: Normal Pylorus	1000
			Gastro: Normal Z-Line	1000
			Ophtho: Pathological Myopia	231
			Gastro: Polyps	1000
			Ophtho: Refractive Media Opacity	54
			Ophtho: Severe Nonproliferative Retinopathy	144
			Gastro: Ulcerative Colitis	1000
			Ophtho: Vitreous Degeneration	58
			Ophtho: Wet Age-Related Macular Degeneration	41

## Results:

The performance of the Euclidean model was generally matched by the Euclidean-Lorentz model in the MMN and the MMD datasets with minimal differences in Top-1 accuracy and identical Top-5 accuracy (**Table 2**). More broadly, however, the Euclidean model generally began to outperform the Euclidean-Lorentz model as the number of images and class size of the dataset increased; this is most prominent in the MD dataset (**Figure 1**). Interpreting the low-dimensional T-SNE representation of the average embedding class as well as the respective dendrogram, the Euclidean-Lorentz model appears to have a more reasonable distribution of the embedding space based on prior understanding of the inter-relationships between the imaging classes in the MMN (**Figures 2 and 3**).

When compared to the known ground truth distance matrix, the distinction between the two models becomes more apparent. The mean absolute difference between the respective Euclidean and Euclidean-Lorentz models compared to the ground truth distance matrix for the MMN dataset was highest in the Euclidean model ( $0.290 \pm 0.005$ ). In contrast, the mean absolute difference in the Euclidean-Lorentz model was significantly lower ( $0.158 \pm 0.003$ ), with a two-sample t-test p-value of  $< 0.0001$ . Spearman's rank correlation findings show that the Euclidean model exhibited a weak correlation with the ground truth ranking (correlation coefficient = 0.021,  $p = 0.3783$ ), while the Euclidean-Lorentz model showed a stronger correlation (correlation coefficient = 0.328,  $p < 0.0001$ ). These results indicate that the Euclidean-Lorentz model not only has a lower mean absolute difference compared to the ground truth but also demonstrates a stronger correlation to the ground-truth class ranks.

Compared to the radiologist's performance, which identified 82 of the 151 patients (0.53), the Euclidean model performed worse, identifying only 62 stroke patients (0.41), while the Euclidean-Lorentz model outperformed by identifying 94 (0.62). Across all images, the Euclidean-Lorentz model achieved a higher overall accuracy (0.50) than the Euclidean model (0.45) (**Figure 4**).

In the out-of-sample dataset, both models were able to correctly identify the modality of the axial images with a 100% identification rate. The Euclidean-Lorentz model and the Euclidean model achieved a Top-1 accuracy of 0.54 (95% CI: 0.44, 0.64) and 0.55 (95% CI: 0.45, 0.65), respectively, suggesting statistically indistinguishable performance in this out-of-sample dataset. Within the NCCT ischemic stroke dataset, the negative cases were

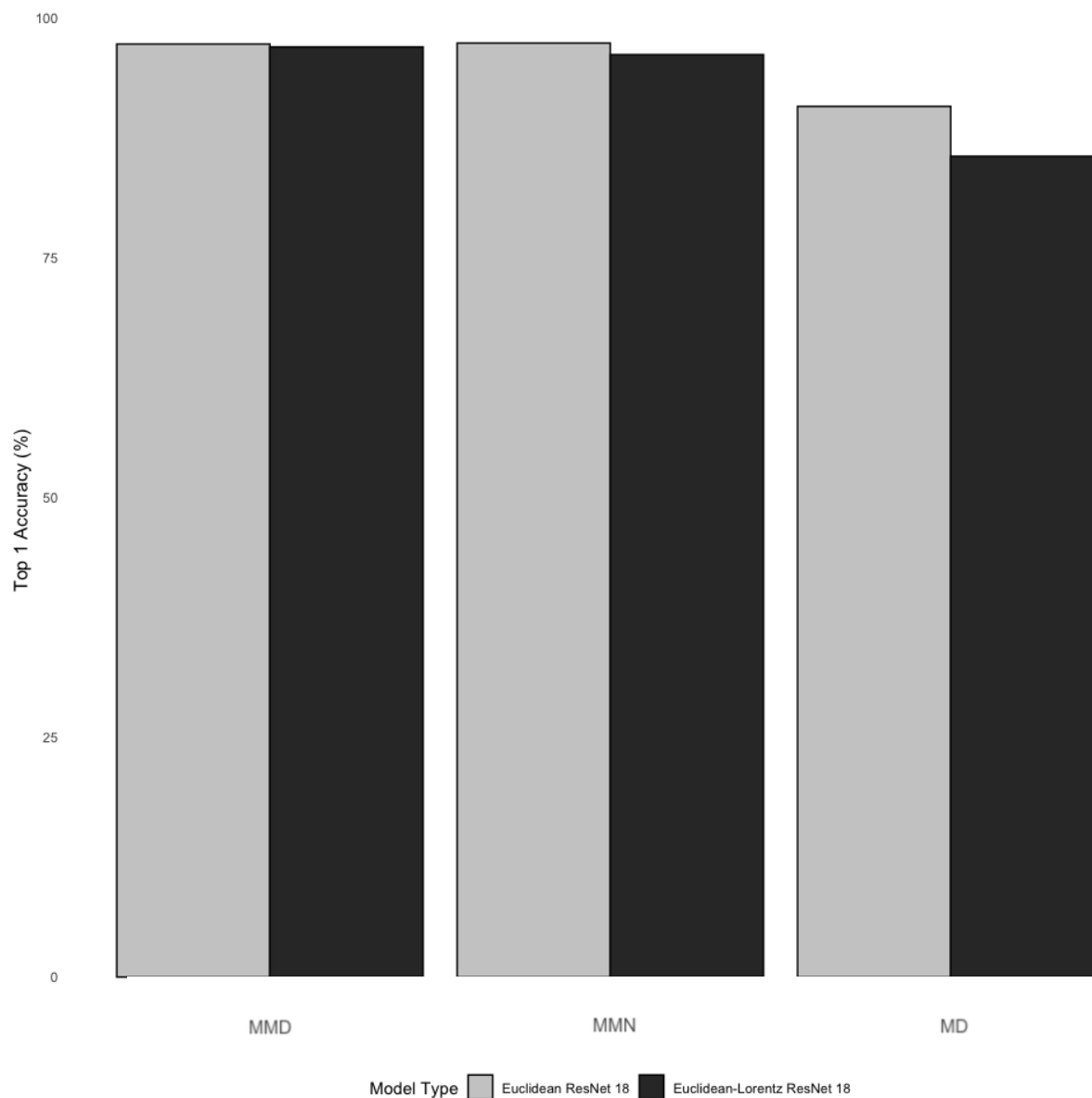
technically a class already observed by the models, so we used this as an additional out-of-sample experiment where the models were tasked with correctly identifying negative NCCT slices as normal CT images. The Euclidean model achieved an accuracy among the negative axial NCCT images of 0.81 (95% CI: 0.81 - 0.82), which was statistically comparable to the Euclidean-Lorentz model, which reached an accuracy of 0.82 (95% CI: 0.82 - 0.83).

Interestingly, the PGD adversarial attack analysis suggests that the Euclidean-Lorentz model often outperforms its Euclidean counterpart in the larger MMN and MD datasets with respect to Top-1 and Top-5 accuracy metrics (**Table 3**). The performance becomes more similar across the two models in the smaller MMD dataset.

**Table 2: Model Characteristics and Performance**

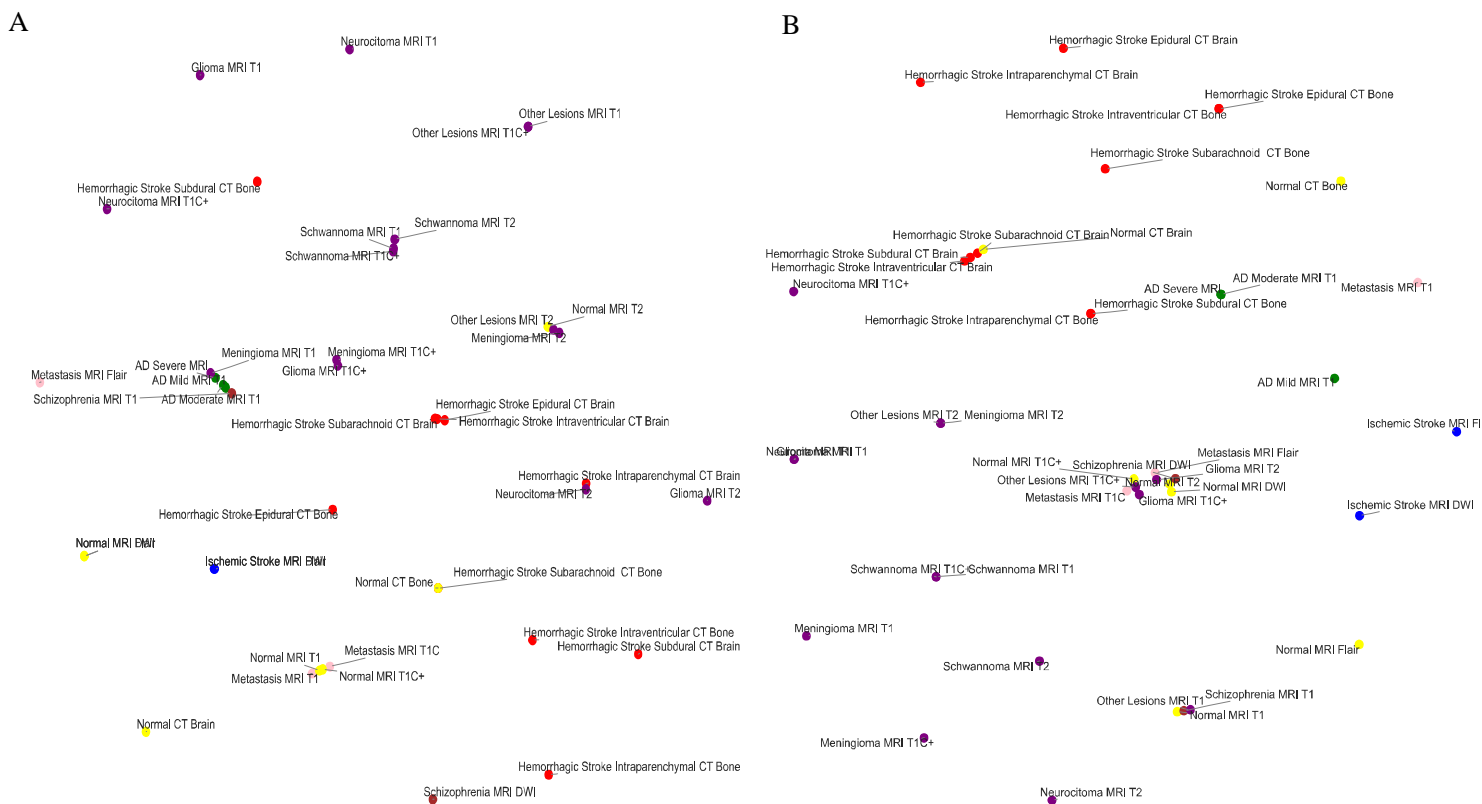
<b>Dataset</b>	<b>Class Size</b>	<b>Sample Size</b>	<b>Model Type</b>	<b>Trainable Parameters</b>	<b>Cross Entropy Loss</b>	<b>Top 1 Accuracy</b>	<b>Top 1 Accuracy: 95% Confidence Intervals</b>	<b>Top 5 Accuracy</b>
MMN	42	72,634	Euclidean ResNet 18	11,189,226	<b>0.05</b>	<b>97.4</b>	<b>97.3 - 97.5</b>	<b>99.9</b>
MMN	42	72,634	Euclidean-Lorentz ResNet 18	11,189,227	0.12	96.2	96.1 - 96.4	<b>99.9</b>
MD	75	89,496	Euclidean ResNet 18	11,207,307	<b>0.26</b>	<b>90.8</b>	<b>90.6 - 91.0</b>	<b>99.6</b>
MD	75	89,496	Euclidean-Lorentz ResNet 18	11,207,308	0.47	85.6	85.4 - 85.9	99.0
MMD	16	19,880	Euclidean ResNet 18	11,177,040	<b>0.11</b>	<b>97.3</b>	<b>97.1 - 97.4</b>	<b>99.9</b>
MMD	16	19,880	Euclidean-Lorentz ResNet 18	11,177,041	<b>0.11</b>	97.0	<b>96.8 - 97.1</b>	<b>99.9</b>

Figure 1: Relative Model Performance Across Datasets



The bar plot above shows the top-1 accuracy metrics with 95% confidence intervals for the Euclidean ResNet 18 and the Euclidean-Lorentz ResNet 18 across the three datasets increasing in size from left to right (i.e., Miniature Multi-Disease (MDD) Dataset, Multi-Modality Neuroimaging (MMN) Dataset, and Multi-Disease (MD) Dataset).

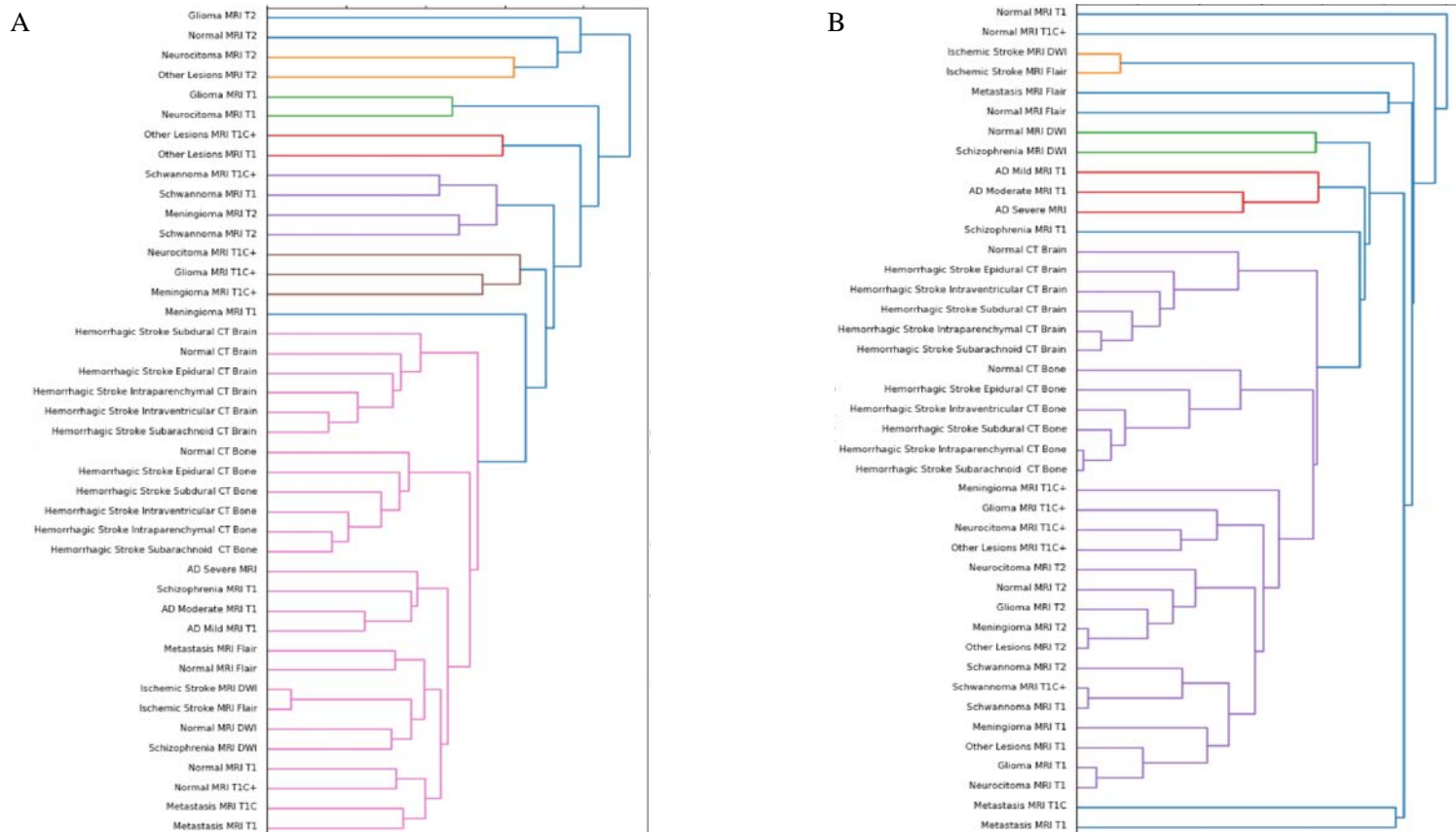
**Figure 2: Euclidean and Hyperbolic Models T-SNE in the Neuroimaging Dataset**



The figure shows the low dimensional representation T-SNE of the average class embedding space from the Euclidean ResNet 18 (A) and the Euclidean-Lorentz ResNet 18 (B) in the Multi-Modality Neuroimaging (MMN) Dataset



Figure 3: Euclidean and Hyperbolic Models Dendrograms in the Neuroimaging Dataset



The figure illustrates the hierarchical clustering dendrogram of the average class embedding space of the Euclidean ResNet 18 (A) and the Euclidean-Lorentz ResNet 18 (B) in the Multi-Modality Neuroimaging (MMN) Dataset

Table 3: Projected Gradient Descent Adversarial Attack

<b>Dataset</b>	<b>Model Type</b>	<b>Epsilon</b>	<b>Top 1 Accuracy</b>	<b>Top 5 Accuracy</b>
MMN	Euclidean ResNet 18	0.003	6.50	53.12
MMN	Euclidean-Lorentz ResNet 18	0.003	<b>45.72</b>	<b>91.15</b>
MMN	Euclidean ResNet 18	0.006	0.08	6.55
MMN	Euclidean-Lorentz ResNet 18	0.006	<b>4.87</b>	<b>41.91</b>
MMN	Euclidean ResNet 18	0.012	0.00	2.81
MMN	Euclidean-Lorentz ResNet 18	0.012	<b>0.52</b>	<b>18.37</b>
MD	Euclidean ResNet 18	0.003	3.78	56.38
MD	Euclidean-Lorentz ResNet 18	0.003	<b>12.47</b>	<b>69.74</b>
MD	Euclidean ResNet 18	0.006	0.02	14.55
MD	Euclidean-Lorentz ResNet 18	0.006	<b>2.64</b>	<b>38.03</b>
MD	Euclidean ResNet 18	0.012	0.01	3.50
MD	Euclidean-Lorentz ResNet 18	0.012	<b>0.17</b>	<b>23.16</b>
MMD	Euclidean ResNet 18	0.003	13.48	<b>74.63</b>
MMD	Euclidean-Lorentz ResNet 18	0.003	<b>13.93</b>	48.04
MMD	Euclidean ResNet 18	0.006	1.29	<b>33.55</b>
MMD	Euclidean-Lorentz ResNet 18	0.006	<b>2.22</b>	28.92
MMD	Euclidean ResNet 18	0.012	0.01	16.05
MMD	Euclidean-Lorentz ResNet 18	0.012	<b>0.27</b>	<b>23.95</b>

Figure 4: Zero-Shot Identification of Stroke Patients



The diagram above shows the how many of the zero-shot stroke patients were able to be identified across in the Euclidean and Euclidean-Lorentz models as well as by human radiologists with their emergent non-contrast brain CT imaging. We also note that 26 patients were not identified in any of the three approaches.

## **Discussion:**

Our empirical analysis study elucidates several important insights into the comparative performance of clipped Euclidean-Lorentz HCNNs and Euclidean CNNs in neuroimaging tasks as well as other medical imaging settings. The results suggest parity in performance between the two neural network approaches in smaller, less complex datasets. We further note distinct semantic organization within the respective embedding spaces, where the HCNN aligned better with ground truth relations between the neuroimaging classes. In assessing generalizability, the HCNN achieved similar out-of-sample performance in identifying AD and normal NCCT images but greatly improved zero-shot performance in identifying ischemic stroke on NCCT.

The cross-entropy loss and top-1 accuracy metrics followed a similar trend across the three medical imaging datasets. Notably, these metrics were identical or similar in datasets where the CNN achieved higher performance (>95% accuracy). However, as the complexity and size of the datasets grew larger, there was a precipitous drop in HCNN performance compared to the CNN. Interestingly, despite the difference in loss in the MMN dataset, the top-1 accuracy between the two models was more similar, unlike in the MD dataset. In both settings of performance parity and disparity, the top-5 accuracy metric across the two models was nearly identical in all three datasets, perhaps due to the improved generalizability of HCNNs, which we explore further.

Achieving parity replicates the findings from Guo et al., which demonstrate that clipped HCNNs achieve similar performance in data settings without strong hierarchy.<sup>9</sup> Nevertheless, we illustrate that the performance of the HCNN suffers compared to the CNN when applied to larger datasets with seemingly more difficult tasks. Given the similarity in model size across the three datasets, our findings may suggest that the HCNNs, as currently constructed, are less efficient with their trainable parameters, contrary to prior literature.<sup>5</sup>

One of the known features of HCNNs is the improved preservation of hierarchical data structures, as reflected by the organized embedding space.<sup>2,6</sup> Low-dimensional T-SNE representations of the embedding space suggest a stratification of classes in the MMN dataset, often by modality first and then disease type, in both models. Similarities in class grouping may be starker in the Euclidean-Lorentz model, as observed in the hierarchical clustering dendrogram from the embedding space.

Nevertheless, the noted limitations of low-dimensional representations may offer a distorted view of the true geodesic distances between the average class embeddings.<sup>34</sup> To explore whether the two models developed meaningfully distinct organizations of embedding space in a more robust fashion, we derived a respective geodesic distance matrix for the average class embedding in both MMN models. We then compared the pairwise distance matrix from each model against a constructed ground truth difference matrix between the classes. Using the pairwise distance differences as well as the Spearman's rank correlation coefficient, we show that the HCNN, and not the CNN, better aligned with our known semantic understanding of the class relationships.

While we observe superior learning and conservation of known class structures in neuroimaging data, we further explored the tangible value of this distinguishing feature. One of the more important aspects of any diagnostic medical imaging algorithm is its ability to function with out-of-sample and out-of-distribution imaging data.<sup>35</sup> We specifically find that the MMN HCNN performed similarly to the CNN in the OASIS I and stroke-negative NCCT datasets, despite poorer HCNN performance in top-1 accuracy and loss in the MMN dataset.

We also find that in zero-shot performance, the HCNN unequivocally outperforms not only the CNN but also the trained radiologists by a significant margin. Finally, the HCNN has shown consistently increased durability to adversarial attacks, which may be relevant when confronted with imaging artifacts, image quality disparities between scanners, or image corruption that may or may not be perceptible.<sup>36,37</sup> As suggested by prior studies in non-medical imaging settings<sup>6,8,9</sup>, we show that the neuroimaging HCNN has improved generalizability with respect to out-of-sample, zero-shot, and adversarial attack performance.

Our study should be interpreted with certain limitations. While recent studies have attempted to move convolutional functions into hyperbolic space,<sup>11,33</sup> we observed significant numerical instability with these methods. Moreover, the computational efficiency of HCNNs was dramatically lower, with convergence requiring three to four times more epochs with similar hyperparameters on larger datasets, limiting our ability to use datasets larger than those in this study. We tested task complexity and size concurrently across the three datasets, but future work should explore scalability and complexity further. Additionally, we are restricted to speculating on the performance of clipped

hybrid HCNNs in the context of our included medical imaging classes and can only speak to generalizability in terms of the out-of-sample and zero-shot datasets used.

## **Conclusion:**

In this study, we show that agnostic HCNN models demonstrated superior ability to learn and retain the native hierarchical structure in the neuroimaging dataset compared to Euclidean CNNs. Importantly, the HCNN achieved disproportionately superior performance in adversarial attack experiments and zero-shot settings, outperforming both radiologists and CNNs. The neuroimaging HCNN also achieved parity in in-sample performance with out-of-sample data but showed depreciating performance in more complex medical imaging task settings with larger datasets. These findings suggest that improvements in the efficiency and scalability of HCNNs are needed to achieve parity with CNNs. However, HCNNs provided notable value in their generalizability in medical imaging settings with multi-modal and multi-disease neuroimaging data.

## References

1. Lu Y, Lu J. A Universal Approximation Theorem of Deep Neural Networks for Expressing Probability Distributions. In: *Advances in Neural Information Processing Systems*. Vol 33. Curran Associates, Inc.; 2020:3094-3105. Accessed November 8, 2023. [https://proceedings.neurips.cc/paper\\_files/paper/2020/hash/2000f6325dfc4fc3201fc45ed01c7a5d-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2020/hash/2000f6325dfc4fc3201fc45ed01c7a5d-Abstract.html)
2. Ganea O, Becigneul G, Hofmann T. Hyperbolic Neural Networks. In: *Advances in Neural Information Processing Systems*. Vol 31. Curran Associates, Inc.; 2018. Accessed November 8, 2023. [https://proceedings.neurips.cc/paper\\_files/paper/2018/hash/dbab2adc8f9d078009ee3fa810bea142-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2018/hash/dbab2adc8f9d078009ee3fa810bea142-Abstract.html)
3. Sala F, Sa CD, Gu A, Re C. Representation Tradeoffs for Hyperbolic Embeddings. In: *Proceedings of the 35th International Conference on Machine Learning*. PMLR; 2018:4460-4469. Accessed November 8, 2023. <https://proceedings.mlr.press/v80/sala18a.html>
4. Sarkar R. Low Distortion Delaunay Embedding of Trees in Hyperbolic Plane. In: van Kreveld M, Speckmann B, eds. *Graph Drawing*. Lecture Notes in Computer Science. Springer; 2012:355-366. doi:10.1007/978-3-642-25878-7\_34
5. Peng W, Varanka T, Mostafa A, Shi H, Zhao G. Hyperbolic Deep Neural Networks: A Survey. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(12):10023-10044. doi:10.1109/TPAMI.2021.3136921
6. Khrukov V, Mirvakhabova L, Ustinova E, Oseledets I, Lempitsky V. Hyperbolic Image Embeddings. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2020:6417-6427. doi:10.1109/CVPR42600.2020.00645
7. Shimizu R, Mukuta Y, Harada T. Hyperbolic Neural Networks++. Published online March 17, 2021. Accessed November 8, 2023. <http://arxiv.org/abs/2006.08210>
8. Chen W, Han X, Lin Y, et al. Fully Hyperbolic Neural Networks. Published online March 15, 2022. Accessed November 8, 2023. <http://arxiv.org/abs/2105.14686>
9. Guo Y, Wang X, Chen Y, Yu SX. Clipped Hyperbolic Classifiers Are Super-Hyperbolic Classifiers. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE; 2022:1-10. doi:10.1109/CVPR52688.2022.00010
10. Mishne G, Wan Z, Wang Y, Yang S. The Numerical Stability of Hyperbolic Representation Learning. Published online June 27, 2023. doi:10.48550/arXiv.2211.00181
11. van Spengler M, Berkhout E, Mettes P. Poincaré ResNet. Published online December 19, 2023. doi:10.48550/arXiv.2303.14027
12. Rana M, Bhushan M. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimed Tools Appl*. 2023;82(17):26731-26769. doi:10.1007/s11042-022-14305-w
13. Yu Z, Nguyen T, Gal Y, et al. Skin Lesion Recognition with Class-Hierarchy Regularized Hyperbolic Embeddings. Published online September 13, 2022. Accessed November 8, 2023. <http://arxiv.org/abs/2209.05842>
14. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*. 2017;42:60-88. doi:10.1016/j.media.2017.07.005



15. Shi P, Qiu J, Abaxi SMD, Wei H, Lo FPW, Yuan W. Generalist Vision Foundation Models for Medical Imaging: A Case Study of Segment Anything Model on Zero-Shot Medical Segmentation. *Diagnostics*. 2023;13(11):1947. doi:10.3390/diagnostics13111947
16. Li X, Zhang L, Wu Z, et al. Artificial General Intelligence for Medical Imaging. Published online July 2, 2023. doi:10.48550/arXiv.2306.05480
17. Hirano H, Minagi A, Takemoto K. Universal adversarial attacks on deep neural networks for medical image classification. *BMC Med Imaging*. 2021;21(1):9. doi:10.1186/s12880-020-00530-y
18. Tasci B, Tasci I. Deep feature extraction based brain image classification model using preprocessed images: PDRNet. *Biomedical Signal Processing and Control*. 2022;78:103948. doi:10.1016/j.bspc.2022.103948
19. Hssayeni M. Computed Tomography Images for Intracranial Hemorrhage Detection and Segmentation. Published online 2019. doi:10.13026/W8Q8-KY94
20. Grøvik E, Yi D, Iv M, Tong E, Rubin D, Zaharchuk G. Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI. *Journal of Magnetic Resonance Imaging*. 2020;51(1):175-182. doi:10.1002/jmri.26766
21. Brain Tumor MRI Images 17 Classes. Accessed July 24, 2024. <https://www.kaggle.com/datasets/fernando2rad/brain-tumor-mri-images-17-classes>
22. Ambite JL, Tallis M, Alpert K, et al. SchizConnect: Virtual Data Integration in Neuroimaging. In: Ashish N, Ambite JL, eds. *Data Integration in the Life Sciences*. Springer International Publishing; 2015:37-51. doi:10.1007/978-3-319-21843-4\_4
23. Hernandez RM, Sison DK, Nolasco NC, Melo J, Castillo R. Application of Machine Learning on MRI Scans for Alzheimer's Disease Early Detection. In: *Proceedings of the 8th International Conference on Sustainable Information Engineering and Technology*. SIET '23. Association for Computing Machinery; 2023:143-149. doi:10.1145/3626641.3627609
24. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*. 2018;172(5):1122-1131.e9. doi:10.1016/j.cell.2018.02.010
25. Li N, Li T, Hu C, Wang K, Kang H. A Benchmark of Ocular Disease Intelligent Recognition: One Shot for Multi-disease Detection. Published online February 16, 2021. doi:10.48550/arXiv.2102.07978
26. Pogorelov K, Randel KR, Griwodz C, et al. KVASIR: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In: *Proceedings of the 8th ACM on Multimedia Systems Conference*. MMSys'17. Association for Computing Machinery; 2017:164-169. doi:10.1145/3083187.3083212
27. Abedeen I, Rahman MA, Protyasha FZ, Ahmed T, Chowdhury TM, Shatabda S. FracAtlas: A Dataset for Fracture Classification, Localization and Segmentation of Musculoskeletal Radiographs. *Sci Data*. 2023;10(1):521. doi:10.1038/s41597-023-02432-4
28. Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data*. 2018;5(1):180161. doi:10.1038/sdata.2018.161
29. Codella NCF, Gutman D, Celebi ME, et al. Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC). Published online January 8, 2018. doi:10.48550/arXiv.1710.05006
30. Combalia M, Codella NCF, Rotemberg V, et al. BCN20000: Dermoscopic Lesions in the Wild. Published online August 30, 2019. doi:10.48550/arXiv.1908.02288

31. Shenton ME, Dickey CC, Frumin M, McCarley RW. A review of MRI findings in schizophrenia. *Schizophrenia Research*. 2001;49(1):1-52. doi:10.1016/S0920-9964(01)00163-3
32. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience*. 2007;19(9):1498-1507. doi:10.1162/jocn.2007.19.9.1498
33. Bdeir A, Schwethelm K, Landwehr N. Fully Hyperbolic Convolutional Neural Networks for Computer Vision. Published online February 7, 2024. doi:10.48550/arXiv.2303.15919
34. Sorzano COS, Vargas J, Montano AP. A survey of dimensionality reduction techniques. Published online March 12, 2014. doi:10.48550/arXiv.1403.2877
35. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. *Science*. 2019;363(6433):1287-1289. doi:10.1126/science.aaw4399
36. Apostolidis KD, Papakostas GA. A Survey on Adversarial Deep Learning Robustness in Medical Image Analysis. *Electronics*. 2021;10(17):2132. doi:10.3390/electronics10172132
37. Shi X, Peng Y, Chen Q, et al. Robust convolutional neural networks against adversarial attacks on medical images. *Pattern Recognition*. 2022;132:108923. doi:10.1016/j.patcog.2022.108923

## Appendix

### I. Neuroimaging Class Table Characteristics Tables

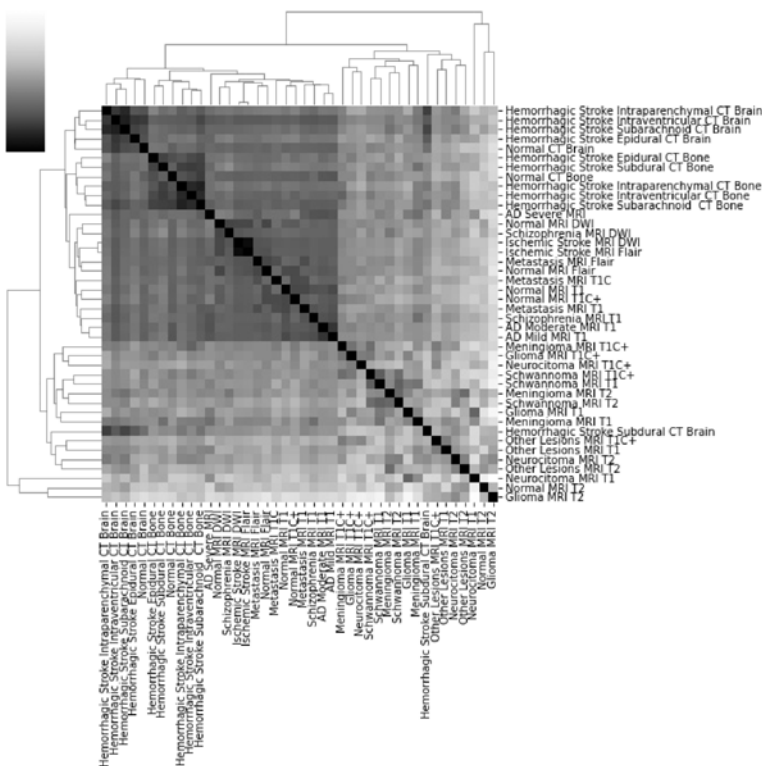
Multi-Modality Neuroimaging: Classes	Modality General	Modality Specific	Disease State	Disease Specific
AD Moderate MRI T1	mri	t1	diffuse	ad
AD Severe MRI	mri	t1	diffuse	ad
AD Mild MRI T1	mri	t1	diffuse	ad
Hemorrhagic Stroke Epidural CT Bone	ct	ctbo	lesion	hemm
Hemorrhagic Stroke Intraparenchymal CT Bone	ct	ctbo	lesion	hemm
Hemorrhagic Stroke Intraventricular CT Bone	ct	ctbo	lesion	hemm
Hemorrhagic Stroke Subarachnoid CT Bone	ct	ctbo	lesion	hemm
Hemorrhagic Stroke Subdural CT Bone	ct	ctbo	lesion	hemm
Hemorrhagic Stroke Epidural CT Brain	ct	ctbr	lesion	hemm
Hemorrhagic Stroke Intraparenchymal CT Brain	ct	ctbr	lesion	hemm
Hemorrhagic Stroke Intraventricular CT Brain	ct	ctbr	lesion	hemm
Hemorrhagic Stroke Subarachnoid CT Brain	ct	ctbr	lesion	hemm
Hemorrhagic Stroke Subdural CT Brain	ct	ctbr	lesion	hemm
Ischemic Stroke MRI DWI	mri	dwi	lesion	isch
Ischemic Stroke MRI Flair	mri	flair	lesion	isch
Metastasis MRI Flair	mri	flair	lesion	met
Metastasis MRI T1C	mri	t1c	lesion	met
Metastasis MRI T1	mri	t1	lesion	met
Normal CT Bone	ct	ctbo	normal	normal
Normal CT Brain	ct	ctbr	normal	normal
Normal MRI DWI	mri	dwi	normal	normal
Normal MRI Flair	mri	flair	normal	normal
Normal MRI T1	mri	t1	normal	normal
Normal MRI T1C+	mri	t1c	normal	normal
Normal MRI T2	mri	t2	normal	normal
Schizophrenia MRI DWI	mri	dwi	diffuse	schiz
Schizophrenia MRI T1	mri	t1	diffuse	schiz
Glioma MRI T1C+	mri	t1c	lesion	tumor
Meningioma MRI T1C+	mri	t1c	lesion	tumor
Neurocitoma MRI T1C+	mri	t1c	lesion	tumor
Other Lesions MRI T1C+	mri	t1c	lesion	lesion
Schwannoma MRI T1C+	mri	t1c	lesion	tumor
Glioma MRI T1	mri	t1	lesion	tumor
Meningioma MRI T1	mri	t1	lesion	tumor
Neurocitoma MRI T1	mri	t1	lesion	tumor

Other Lesions MRI T1	mri	t1	lesion	tumor
Schwannoma MRI T1	mri	t1	lesion	tumor
Glioma MRI T2	mri	t2	lesion	tumor
Meningioma MRI T2	mri	t2	lesion	tumor
Neurocitoma MRI T2	mri	t2	lesion	tumor
Other Lesions MRI T2	mri	t2	lesion	lesion
Schwannoma MRI T2	mri	t2	lesion	tumor

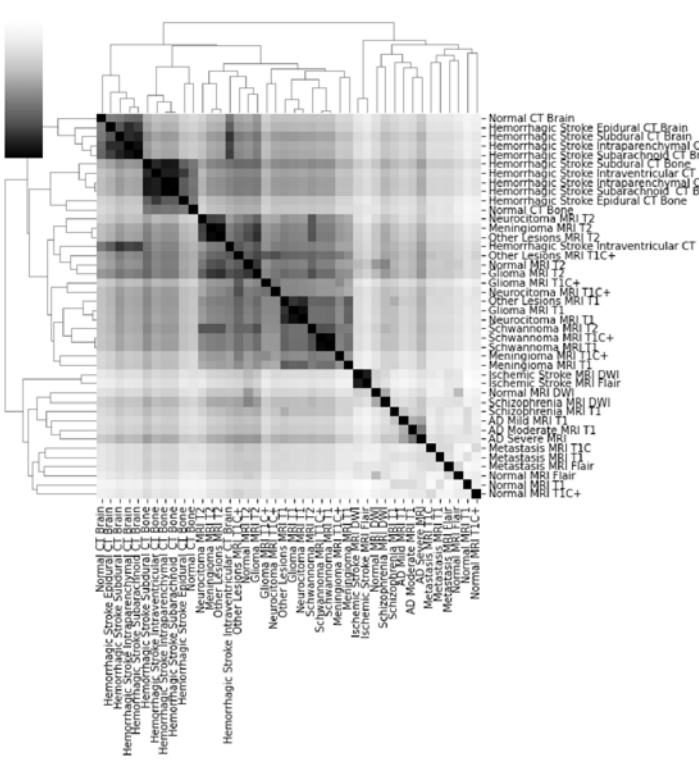
† T1 MRI and T1 Contrast Enhanced MRI were designated a 0.5-unit difference. T2 MRI, DWI MRI, and Flair MRI were designated a 0.5-unit difference. All other difference between a class value was designated as a one-unit difference.

## II. Average Class Embedding Space Distance Heat Map

A



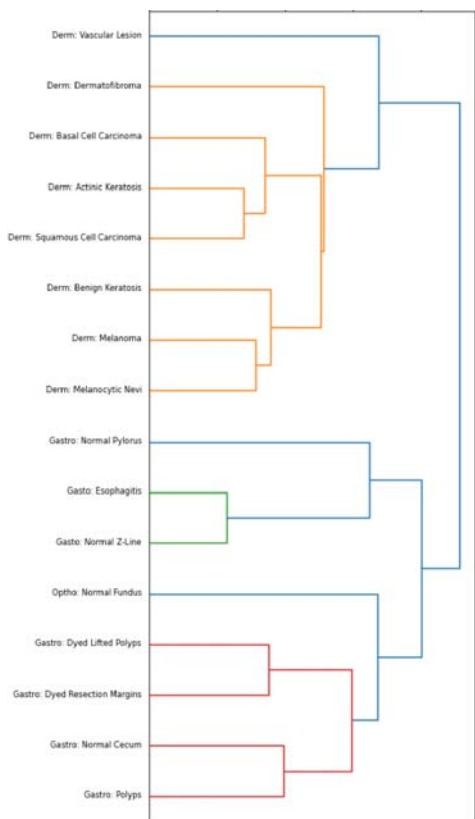
B



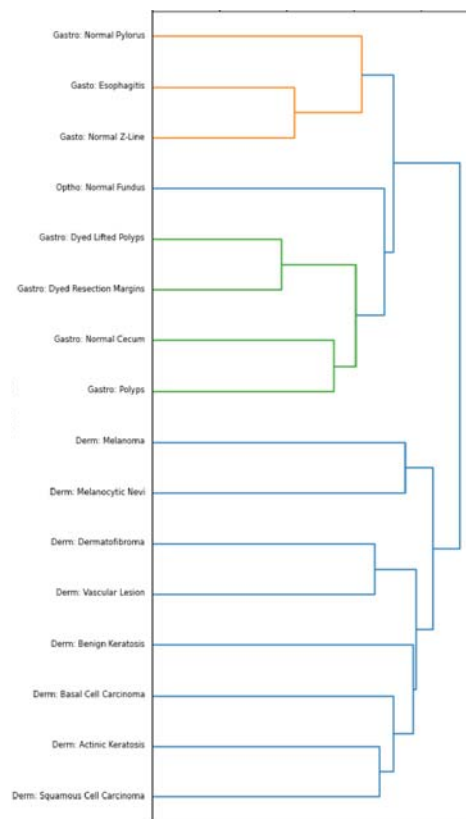
The figure illustrates the hierarchical clustering dendrogram of the average class embedding space of the Euclidean ResNet 18 (A) and the Euclidean-Lorentz ResNet 18 (B) in the Multi-Modality Neuroimaging (MMN) Dataset. The dendrogram is accompanied by a heatmap where the darker voxels represent a closer distance between the respective classes.

### III. Hierarchical Dendrograms in the Other Disease Datasets

A

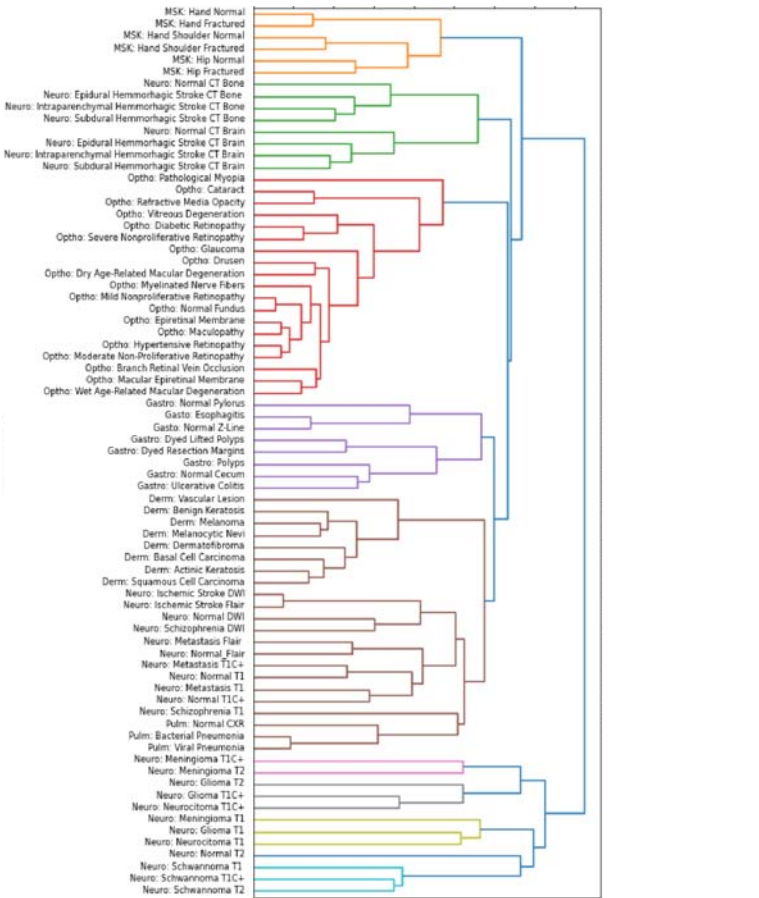


B

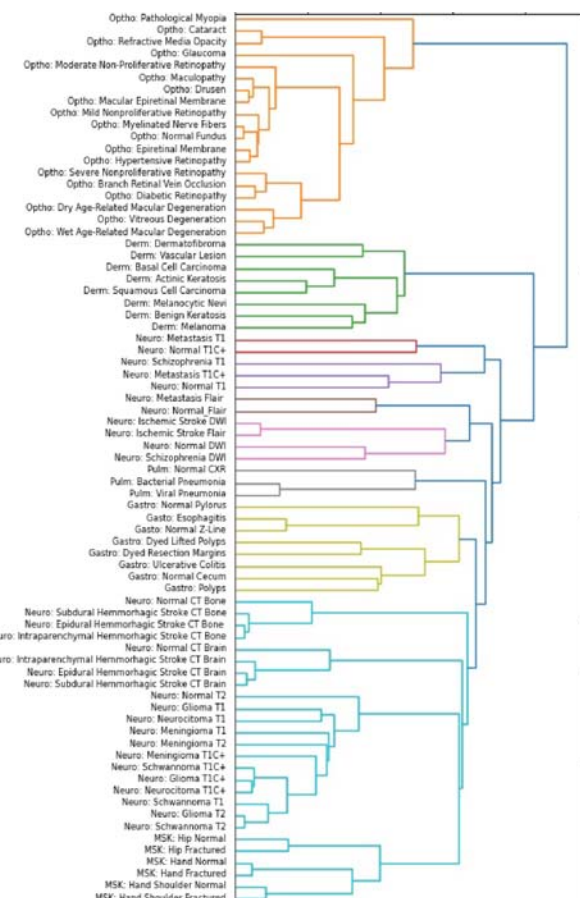


The figure illustrates the hierarchical clustering dendrogram of the average class embedding space of the Euclidean ResNet 18 (A) and the Euclidean-Lorentz ResNet 18 (B) in the Miniature Multi-Disease (MMD) Dataset.

A



B



The figure illustrates the hierarchical clustering dendrogram of the average class embedding space of the Euclidean ResNet 18 (A) and the Euclidean-Lorentz ResNet 18 (B) in the Multi-Disease (MD) Dataset.