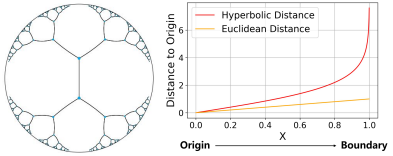
Clipped Hyperbolic Classifiers Are Super-Hyperbolic Classifiers

Yunhui Guo Xudong Wang Yubei Chen Stella X. Yu

Hyperbolic Space



Hyperbolic Distance

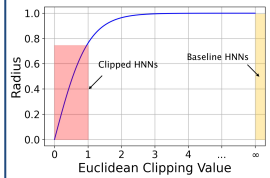
$$d_{\mathbb{B}^n}(\boldsymbol{u},\boldsymbol{v}) = \operatorname{arcosh}\left(1 + 2\frac{\|\boldsymbol{u}-\boldsymbol{v}\|^2}{(1-\|\boldsymbol{u}\|^2)(1-\|\boldsymbol{v}\|^2)}\right)$$

- **Non-Euclidean space with constant negative curvature**
- **Can embed tree-like data continuously with low distortion**

Contributions

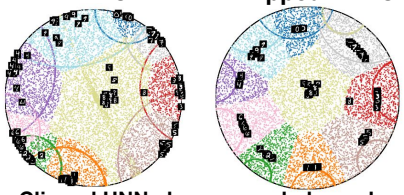
Feature Clipping: Clip the Euclidean embedding before the exponential map

$$\text{CLIP}(\mathbf{x}^E; r) = \min\{1, \frac{r}{||\mathbf{x}^E||}\} \cdot \mathbf{x}^E$$



- Escape the ill-conditioned region
- Preserving the hyperbolic property

Hyperbolic Feature Space



Clipped HNNs learn more balanced and discriminative feature space

OOD Detection

CIFAR10

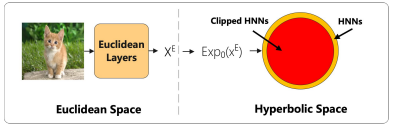
OOD Dataset	FPR95 ↓	AUROC ↑	AUPR ↑
iSUN	46.30±0.78 45.28±0.65	91.50±0.16 91.61±0.21	98.16±0.05 98.09±0.06
Place365	51.09±0.92 54.77±0.76	87.56±0.37 86.82±0.41	96.76±0.15 96.17±0.20
Texture	65.04±0.91 47.12±0.62	82.80±0.35 89.91±0.20	94.59±0.20 97.39±0.09
SVHN	71.66±0.84 49.89±1.03	86.58±0.21 91.34±0.22	97.06±0.06 98.13± 0.06
LSUN-Crop	22.22±0.78 23.87±0.73	96.05±0.10 95.65±0.22	99.16±0.03 98.98±0.07
LSUN-Resize	41.06±1.07 41.49±1.24	92.67±0.16 92.97±0.24	98.42±0.04 98.46 ±0.07
Mean	49.56 **43.74**	89.53 **91.38**	97.36 **97.87**

CIFAR100

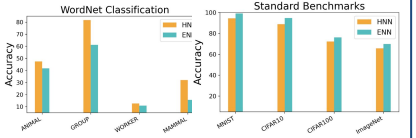
OOD Dataset	FPR95 ↓	AUROC ↑	AUPR ↑
iSUN	74.07±0.87 68.37±0.90	82.51±0.39 81.31±0.43	95.83±0.11 94.96±0.20
Place365	81.01±1.07 79.66±0.69	76.90±0.45 76.94±0.28	94.02±0.15 93.91±0.18
Texture	83.67±0.68 64.91±0.80	77.52±0.32 83.26±0.25	94.47±0.10 95.77±0.08
SVHN	84.56±0.78 53.11±1.04	84.32±0.22 89.53±0.26	96.69±0.07 97.71±0.07
LSUN-Crop	43.46±0.79 51.08±1.17	93.09±0.23 87.21±0.39	98.58±0.05 96.83±0.13
LSUN-Resize	71.50±0.73 63.86±1.10	82.12±0.40 82.36±0.42	95.69±0.13 95.16±0.13
Mean	73.05 **63.50**	82.74 **83.43**	**95.88** 95.72

HNNs show stronger OOD detection ability than ENNs

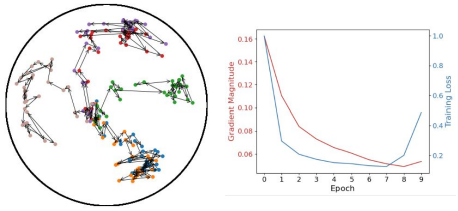
HNNs Underperform ENNs on Standard Benchmarks



HNNs consist of ENN feature extractors and hyperbolic classifiers



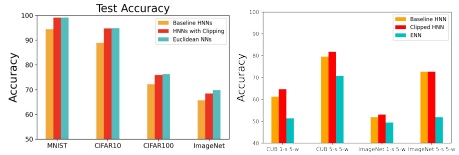
Training Dynamics of HNNs



$$\mathbf{x}_{t+1}^H = \mathbf{x}_t^H + C(E(\mathbf{w}_t^E)^T)\frac{(1-\|\mathbf{x}_t^H\|^2)^2}{4}\frac{\partial \ell}{\partial \mathbf{w}^E}$$

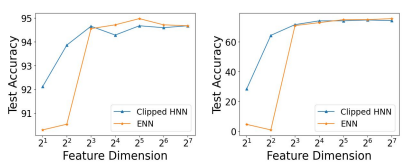
Gradients vanish when the embedding is close to the boundary

Standard Benchmarks and Few-shot Learning



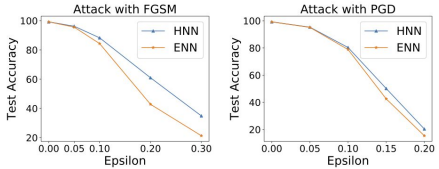
Clipped HNNs show better results compared with baseline HNNs

Impact of Dimensionality



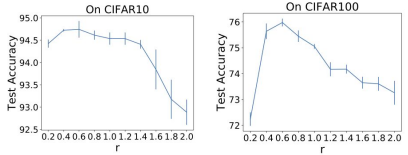
HNNs outperforms ENNs when the feature dimensionality is low

Adversarial Robustness



HNNs show stronger adversarial robustness

Impact of Clip Value



The clipping value should not be too large (causing vanishing gradient problem) or too small (no enough capacity).