LOG CONFERENCE | 2024

Hyperbolic Kernel Convolution: A Generic Framework

Eric Qu*, Lige Zhang#, Habib Debaya#, Yue Wu#, Dongmian Zou#

*: University of California, Berkeley #: Duke Kunshan University

Motivation

- Most generalizations of graph convolution in the hyperbolic space rely only on the topology of the underlying graph, neglecting the rich geometric structures in the hyperbolic space itself.

Hyperbolic Geometry

- Hyperbolic geometry is a special kind of Riemannian geometry with a constant negative curvature.
- For $x, y \in \mathbb{H}^n$, $d_{\mathcal{H}}(x, y)$ is the length of the geodesic between them, also called hyperbolic distance.
- For $x \in \mathbb{H}^n$, the tangent space at x is $T_x\mathbb{H}^n$.
- For $x, y \in \mathbb{H}^n$, and $v \in T_x\mathbb{H}^n$, $exp_x(v): T_x\mathbb{H}^n \to \mathbb{H}^n$ is the exponential map of v at x. The logarithmic map is denoted by $log_x: \mathbb{H}^n \to T_x\mathbb{H}^n$, where $log_x(exp_x(v)) = v$.
- For two points $x, y \in \mathbb{H}^n$, we use $PT_{x\to y}$ to denote the parallel transport map which "transports" a vector from $T_x\mathbb{H}^n$ to $T_y\mathbb{H}^n$ along the geodesic from x to y.

Hyperbolic Kernel Points

- To construct an HKConv layer, the first step is to fix K kernel points $\{\tilde{x}_k\}_{k=1}^K$ in the input space $\mathbb{H}^n$.
- The locations of kernel points are predetermined to ensure stable training.
- The kernel points can be regarded as located in near the hyperbolic origin o.
- Specifically, we minimize the following loss function with Riemannian gradient descent:

$$\mathcal{L}(\{\tilde{x}_k\}_{k=1}^K) = \sum_{k=1}^{K}\sum_{l\neq k}\frac{1}{d_{\mathcal{H}}(\tilde{x}_l, \tilde{x}_k)} + \sum_{k=1}^{K} d_{\mathcal{H}}(o, \tilde{x}_k).$$

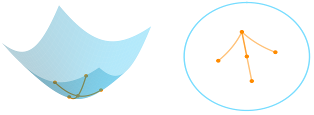

Illustration of Kernel Points

Hyperbolic Kernel Convolution

Step 1. Transformation of relative input features

- For $x, y, u \in \mathbb{H}^n$, we define:

$$T_{x\to y}(u) := exp_y\left(PT_{x\to y}(log_x(u))\right).$$

- We denote "translation by x" via:

$$u \ominus x := T_{x\to y}(u).$$

- For each $x_i \in \mathcal{N}(x)$, the translation $(x_i \ominus x)$ gives the relative input feature.
- The first step of HKConv is to extract features from translated input features via hyperbolic linear layers:

$$x_{ik} = HLinear_k(x_i \ominus x), \mathrm{k} = 1, \dots, \mathrm{K},$$

Step 2. Kernel Aggregation

- For each transformed feature x_{ik}, we take a "weighted average" with its associated kernel points by either:

$$x_i' = MEAN(\{x_{ik}\}_{k=1}^K, \{d_{\mathcal{H}}(x_i \ominus x, \tilde{x}_k)\}_{k=1}^K),$$

or

$$x_i' = MEAN(\{x_{ik}\}_{k=1}^K, \{d_{\mathcal{H}}(x_{ik}, \tilde{x}_k)\}_{k=1}^K).$$

Step 3 Neighbor Aggregation

- The last step of HKConv is a neighbor aggregation:

$$x' = MEAN(\{x_i'\}_{k=1}^K, \{w_i\}_{k=1}^K).$$

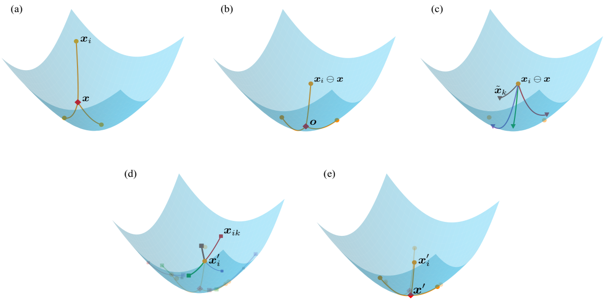

Illustration of HKConv

Properties

- Local Translation Invariance:

$$HKConv\left(T_{x\to y}(x); T_{x\to y}(\mathcal{N}(x))\right) = HKConv(x; \mathcal{N}(x)).$$

- Expressivity:

$$HKConv^{(1)}(x) = HKConv^{(0)}(x).$$

Experiments

	Cornell	Texas	Wisconsin	Chameleon	Squirrel	Actor
# of Nodes	183	183	251	2,277	5,201	7,600
# of Edges	280	295	466	31,421	198,493	26,752
# of Features	1,703	1,703	1,703	2,325	2,089	931
# of Classes	5	5	5	5	5	5
Hyperbolicity	$\delta = 1$	$\delta = 1$	$\delta = 1$	$\delta = 1.5$	$\delta = 1.5$	$\delta = 1.5$
MLP	81.89±6.40	80.81±4.75	85.29±3.31	46.21±2.99	28.77±1.56	36.53±0.70
GCN [27]	60.54±5.30	55.14±5.16	51.76±3.06	64.82±2.24	53.43±2.01	27.32±1.10
GAT [39]	61.89±5.05	52.16±6.63	49.41±4.09	60.26±2.50	40.72±1.55	27.44±0.89
GraphSAGE [40]	75.95±5.01	82.43±6.14	81.18±5.56	58.73±1.68	41.61±0.74	34.23±0.99
GCNII [41]	77.86±3.79	77.57±3.83	80.39±3.40	63.86±3.04	38.47±1.58	37.44±1.30
Geom-GCN [42]	60.54±3.67	66.76±2.72	64.51±3.66	60.00±2.81	38.15±0.92	31.59±1.15
WRGAT [43]	81.62±3.90	83.62±5.50	86.98±3.78	65.24±0.87	48.85±0.78	36.53±0.77
LINKX [44]	77.84±5.81	74.60±8.37	75.49±5.72	68.42±1.38	61.81±1.80	36.10±1.55
GloGNN [45]	83.51±4.26	84.32±4.15	87.06±3.53	69.78±2.42	57.54±1.39	37.35±1.30
GloGNN++ [45]	85.95±5.10	84.05±4.90	88.04±3.22	71.21±1.84	57.88±1.76	37.70±1.40
HGCN [6]	79.43±0.47	70.13±0.32	83.26±0.51	NaN	62.31±0.57	36.58±0.79
H2H-GCN [15]	75.52±0.82	71.47±0.63	88.71±0.82	78.71±0.96	66.85±0.72	42.73±0.86
HyboNet [8]	77.27±0.71	72.23±0.94	86.52±0.51	74.91±0.58	69.07±0.64	45.74±0.82
WLHN [46]	77.29±4.66	75.41±5.98	78.62±3.44	-	55.76±0.92	36.42±1.42
Ours (self-correlation (10))						
SKN	87.12±2.62	**93.94**±6.56	**92.59**±1.85	82.60±1.14	**73.05**±1.80	**66.42**±1.42
LKN	75.76±3.21	92.42±3.47	78.40±1.07	77.53±2.75	57.59±7.41	50.66±4.91
BKN	78.03±3.47	77.27±3.94	82.09±2.83	68.25±3.32	63.25±1.50	65.46±2.54
BKN-isoAgg	**93.94**±1.31	85.61±6.94	90.74±4.90	72.59±2.49	66.59±0.96	65.12±0.52
Ours (point-correlation (9))						
SKN	86.36±2.62	92.42±7.31	91.36±2.14	**83.33**±0.73	72.44±1.70	66.12±1.92
LKN	76.52±3.47	93.18±3.93	77.78±3.20	77.66±2.71	64.61±1.65	44.45±3.75
BKN	63.64±2.27	63.64±2.27	61.11±5.55	56.41±10.43	53.02±1.93	45.51±3.67
BKN-isoAgg	76.52±1.31	83.33±4.73	74.07±1.85	64.90±2.21	63.78±0.56	62.21±1.60

Node classification results(above), Embedding Visualization(below)