Computer Vision by Learning

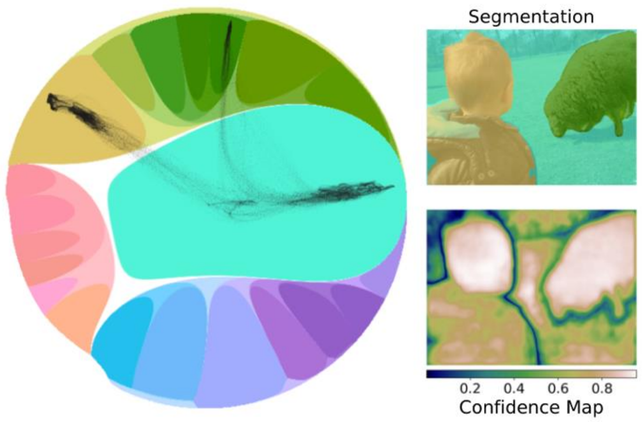
Hyperbolic Deep Learning

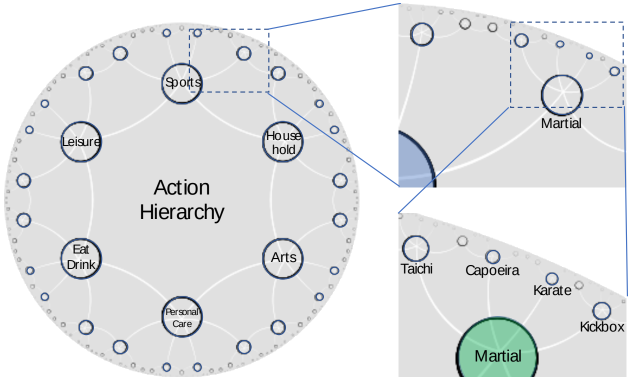
Pascal Mettes

University of Amsterdam

VIDEO & IMAGE SENSE LAB

Who am I





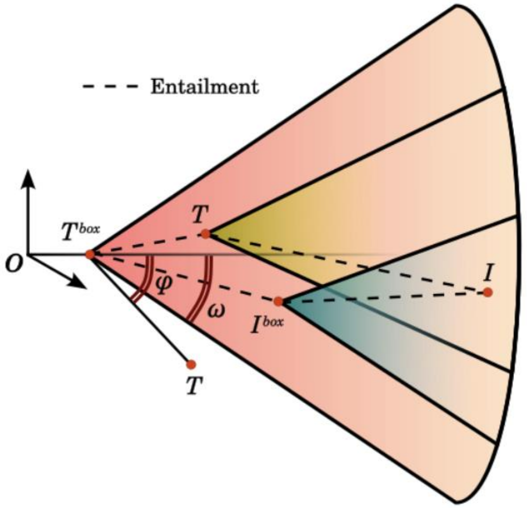
Pascal Mettes

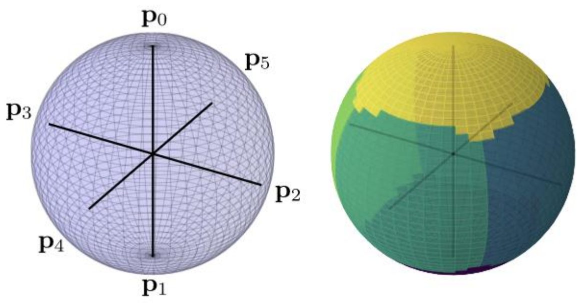
Assistant Professor - University of Amsterdam

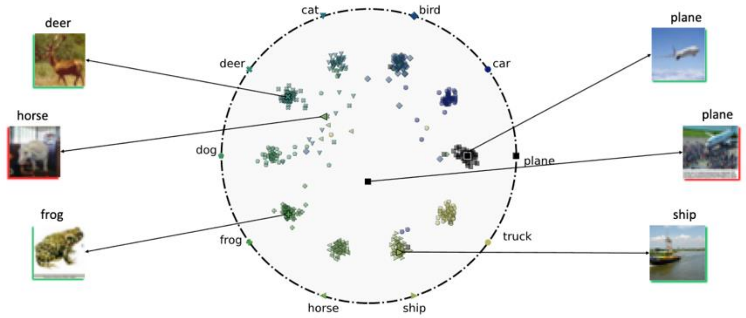
Computer vision

Hierarchical knowledge

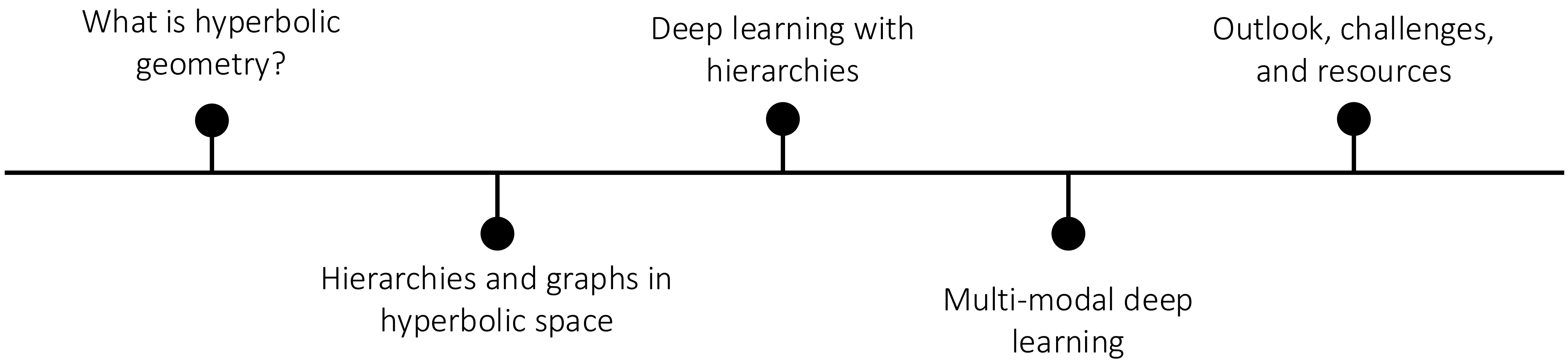
Hyperbolic geometry







Journey of the lecture



What is hyperbolic geometry?

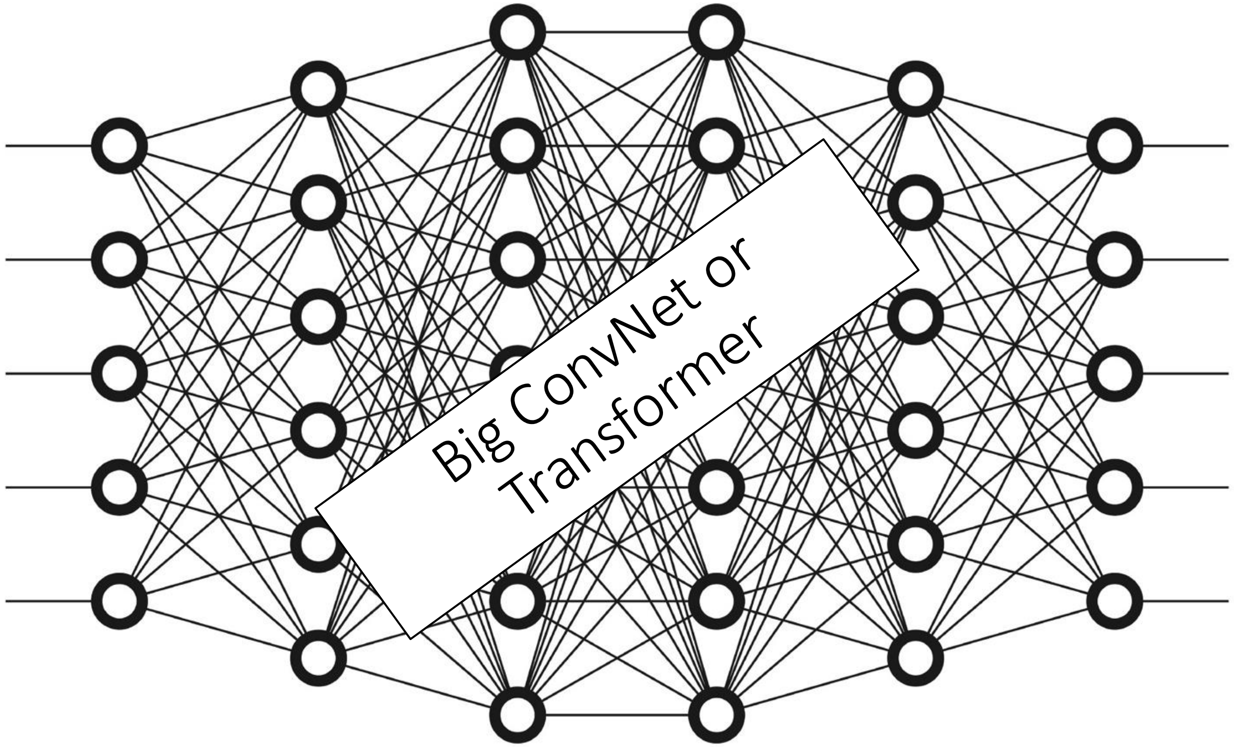
Hierarchies and graphs in hyperbolic space

Deep learning with hierarchies

Multi-modal deep learning

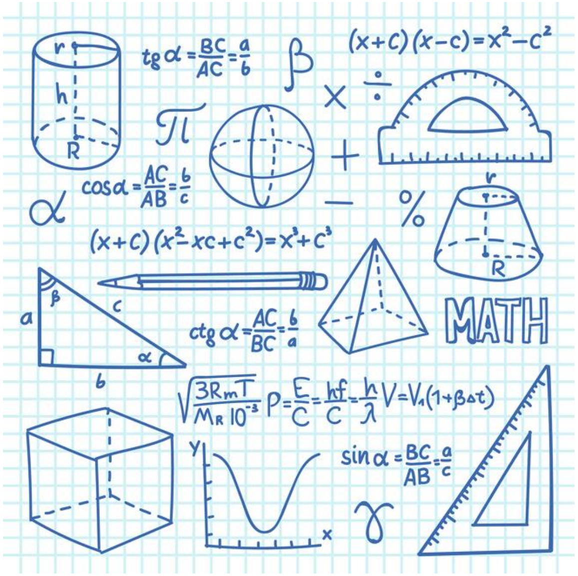
Outlook, challenges, and resources

Canonical deep learning



"person walking with
dog in the park"

Why are we so Euclidean?

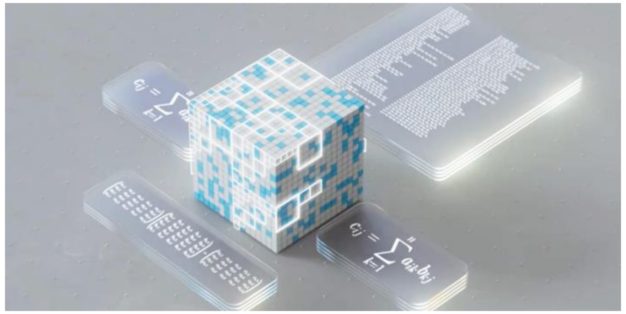


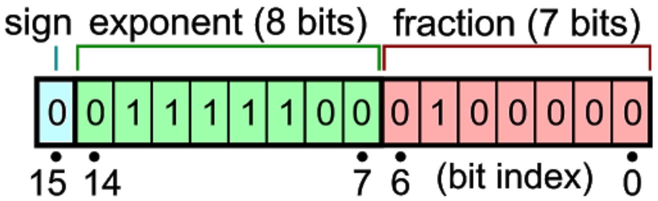
Our school curricula are Euclidean

PyTorch

TensorFlow

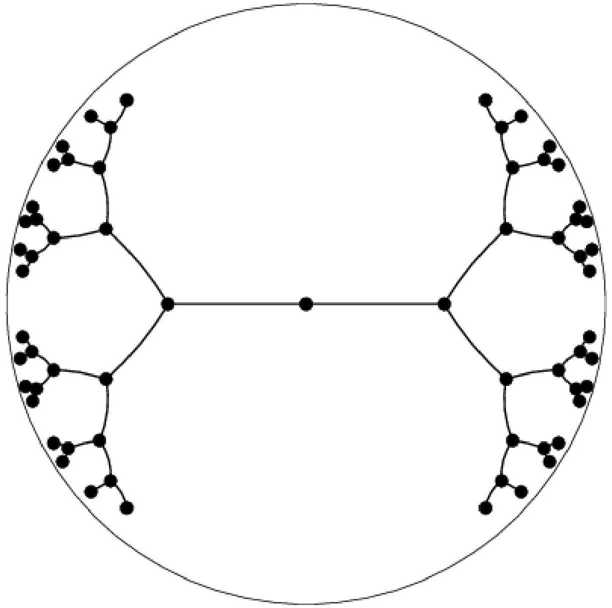
Our deep learning tools are Euclidean





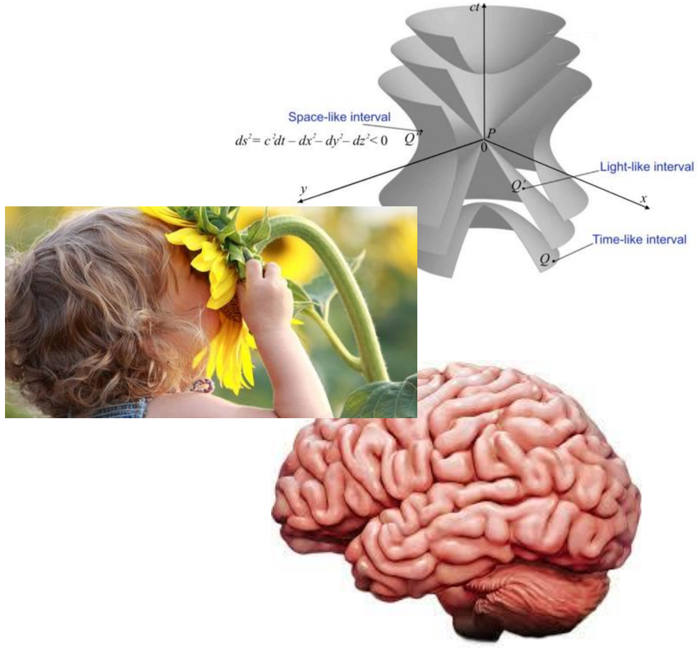
Our computers are built for Euclidean space

But is Euclidean always the answer?



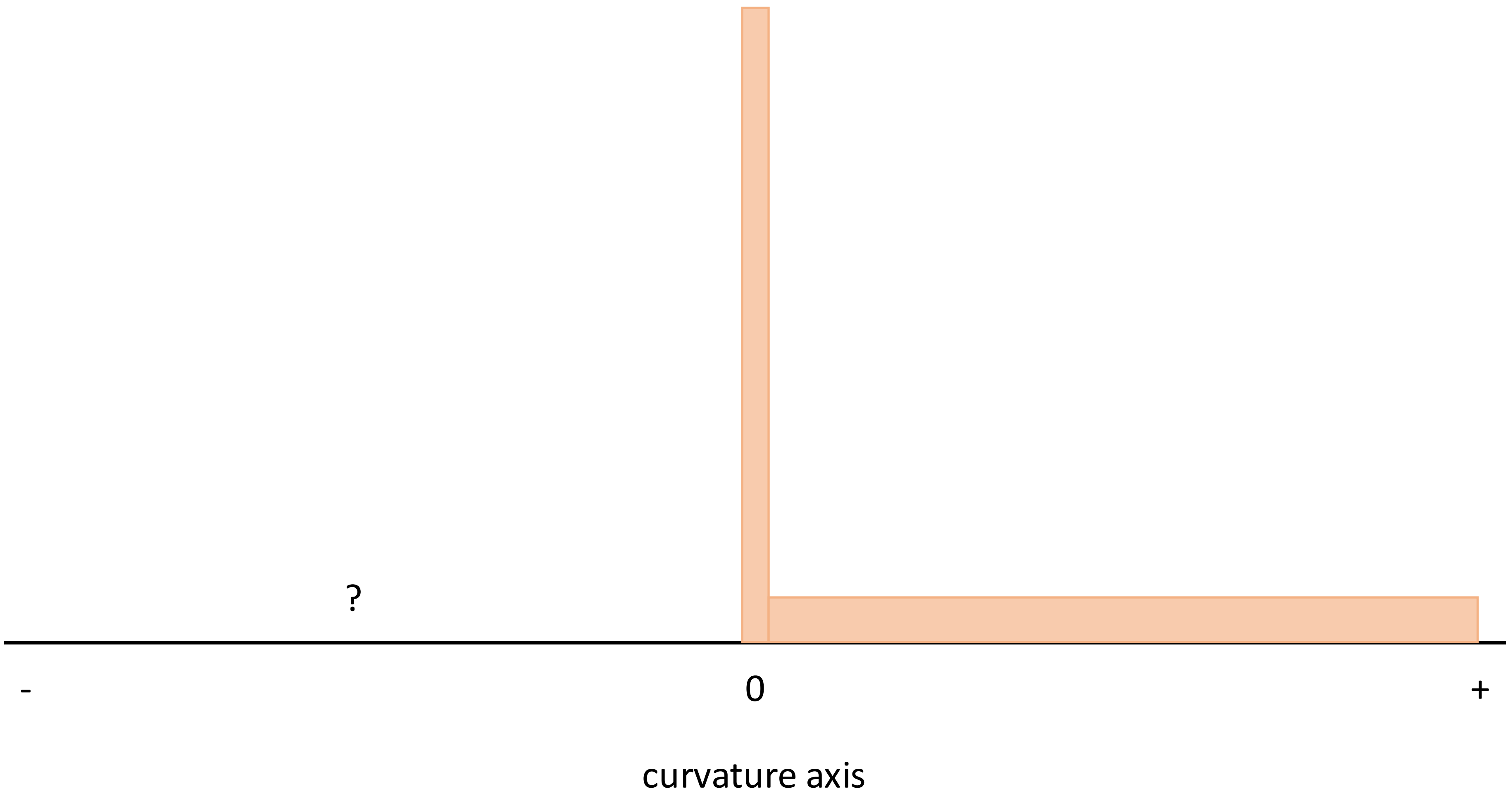
Euclidean is not hierarchical

Euclidean is not compact



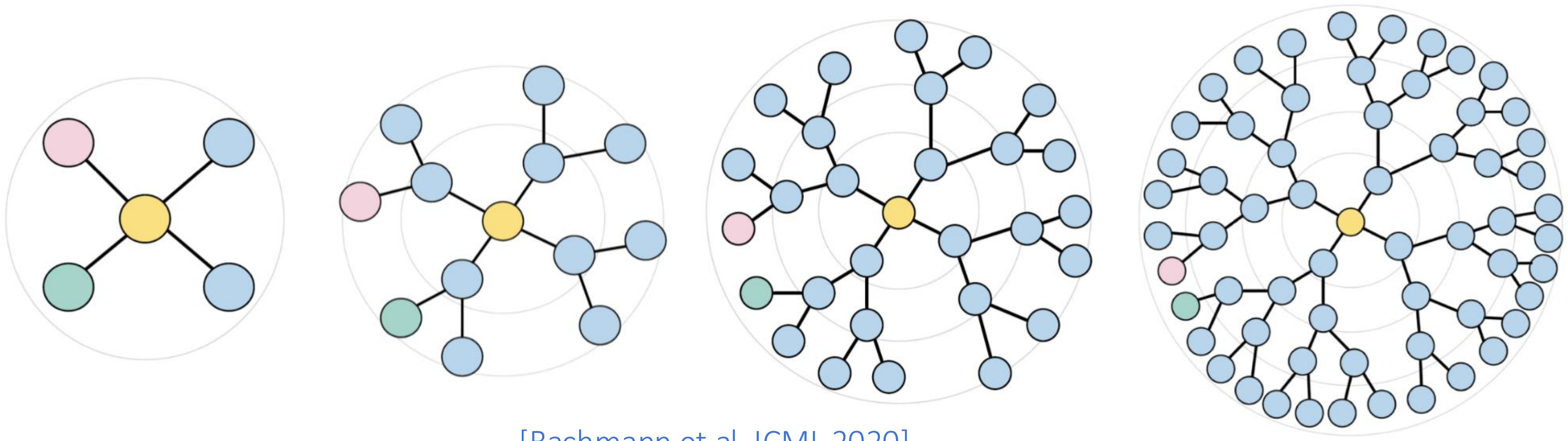
The world is not always Euclidean

The blindspots in deep learning research



The geometry of hierarchies

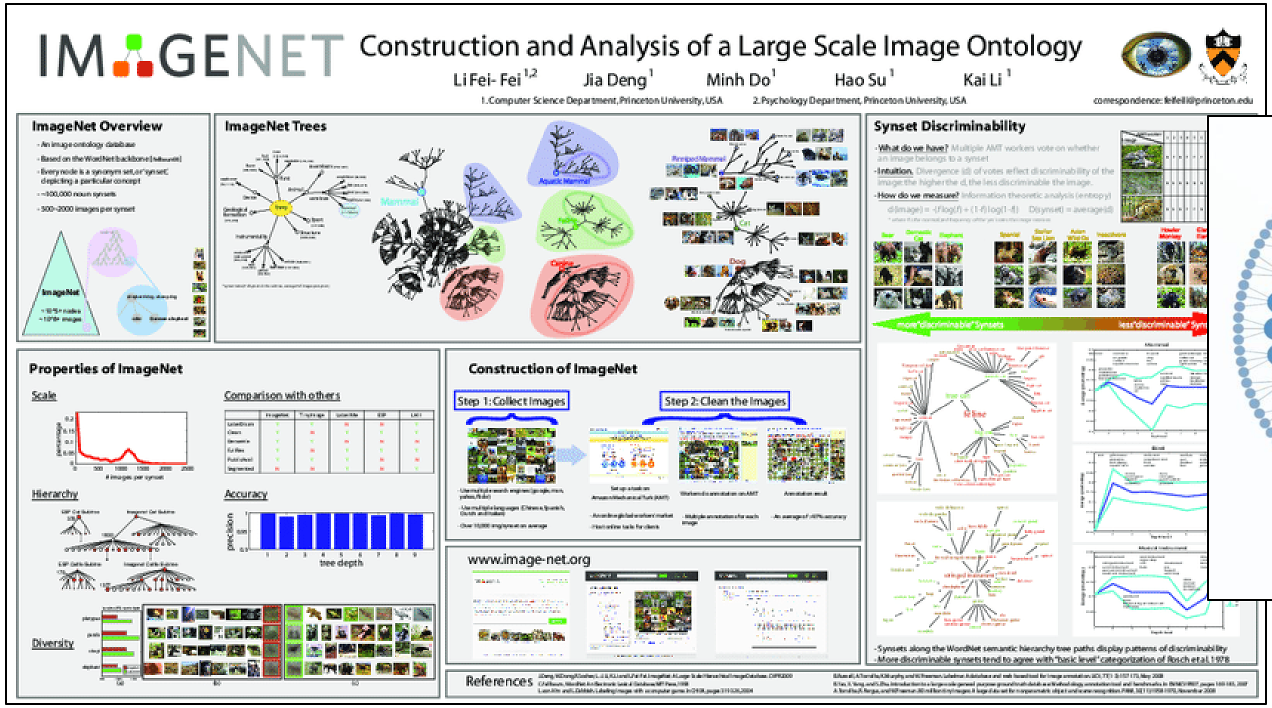
Euclidean space and hierarchies are a mismatch: linear vs. exponential growth.



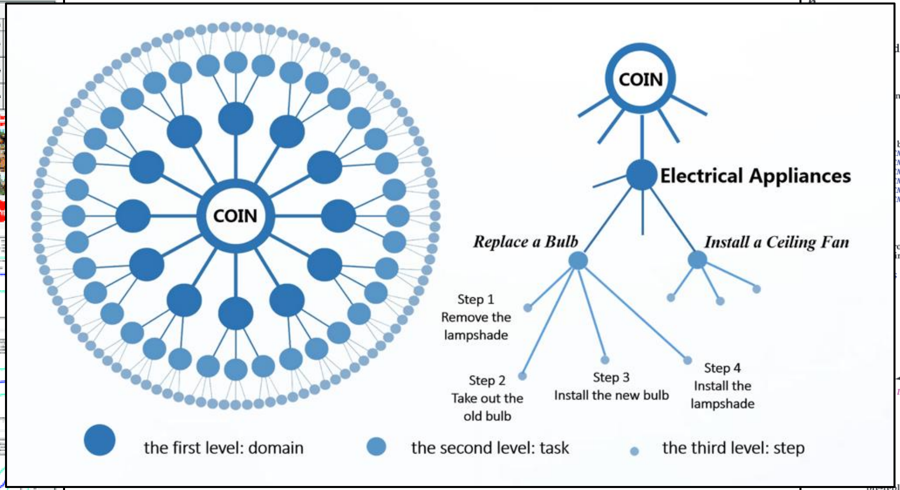
[Bachmann et al. ICML 2020]

What we need is a hierarchical geometry for representation learning!

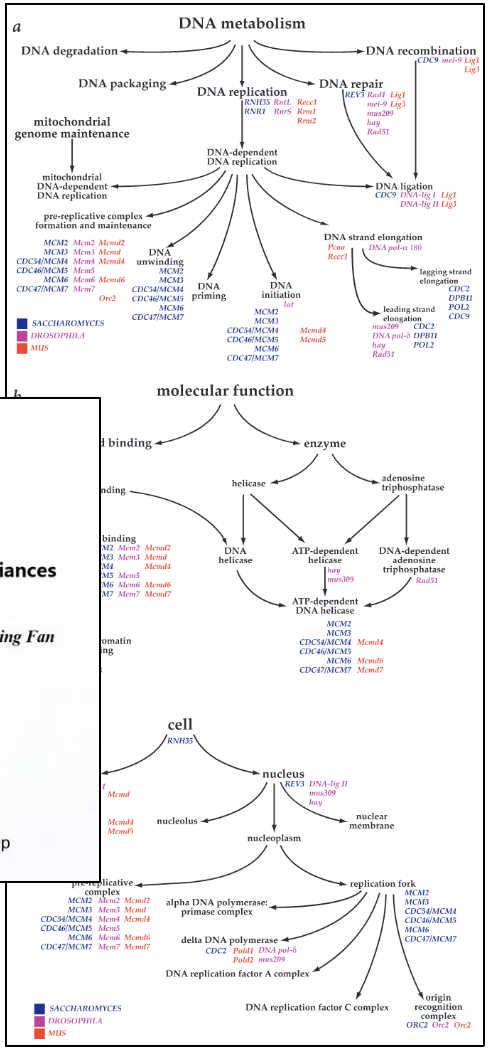
The importance of hierarchies



[Li et al. 2009]



[Tang et al. CVPR 2019]



[Ashburner et al. Nature 2000]

Hierarchies allow us to look beyond samples and their individual labels.

Origins of hyperbolic geometry

Euclid's 5 postulates:

1. *A straight-line segment can be drawn joining any two points.*

2. Any straight-line segment can be extended indefinitely in a straight line.

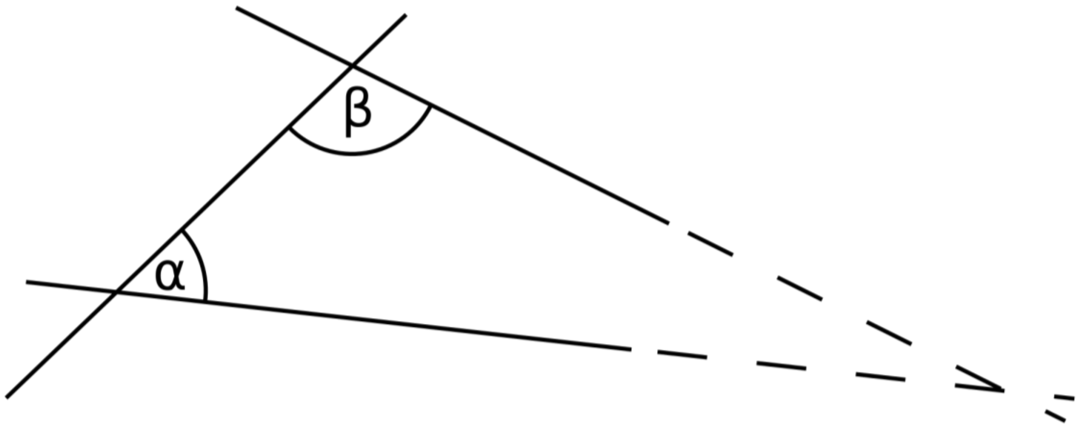
3. Given any straight lines segment, a circle can be drawn having the segment as radius and one endpoint as center.

4. All Right Angles are congruent.

5. If two lines are drawn which intersect a third in such a way that the sum of the inner angles on one side is less than two Right Angles, then the two lines inevitably must intersect each other on that side if extended far enough. This postulate is equivalent to what is known as the Parallel Postulate.

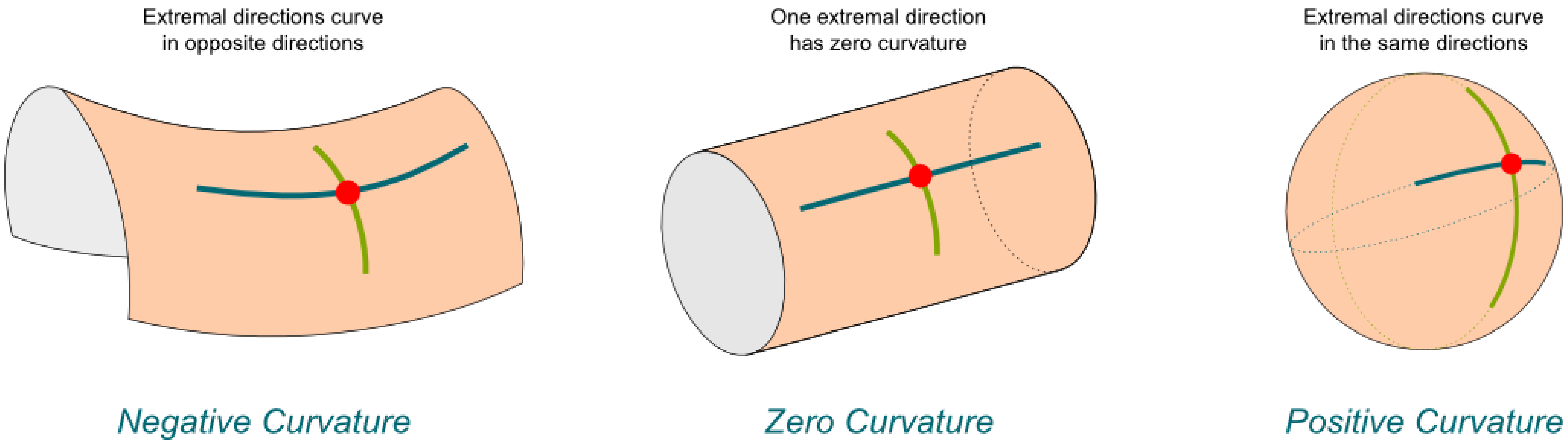
Origins of hyperbolic geometry

Euclid's 5th postulate:



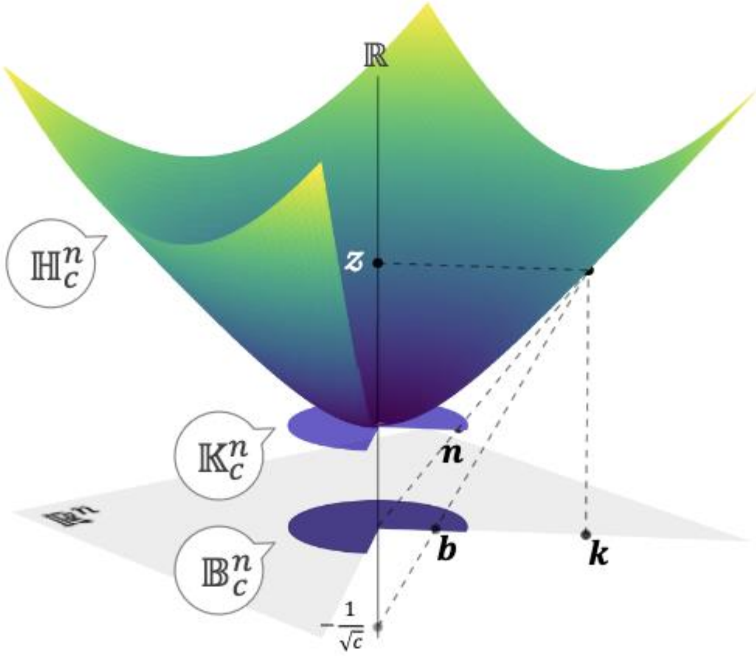
Did Euclid make a mistake by making it a postulate? Shouldn't it be a theorem?

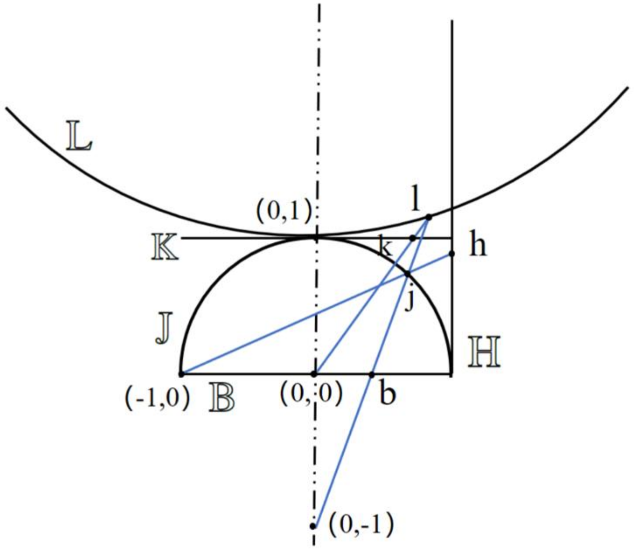
The rise of non-Euclidean geometry



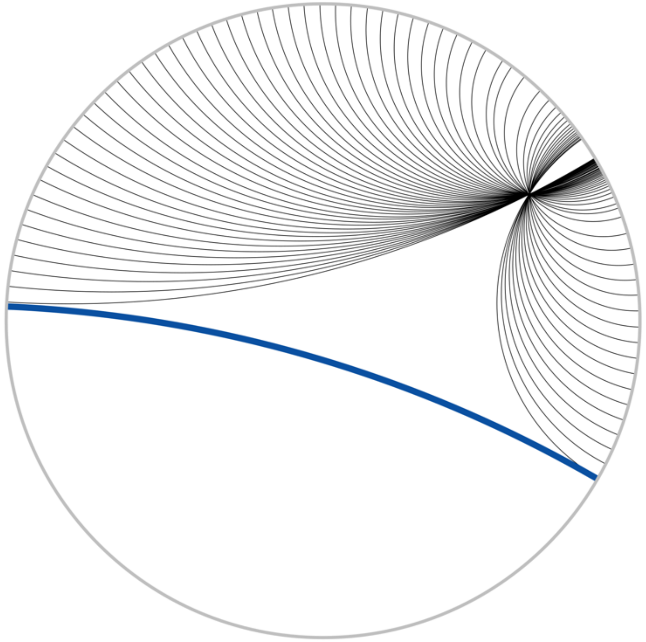
Models of hyperbolic geometry

To perform numerical operations, we need to operate in a model of hyperbolic geometry.

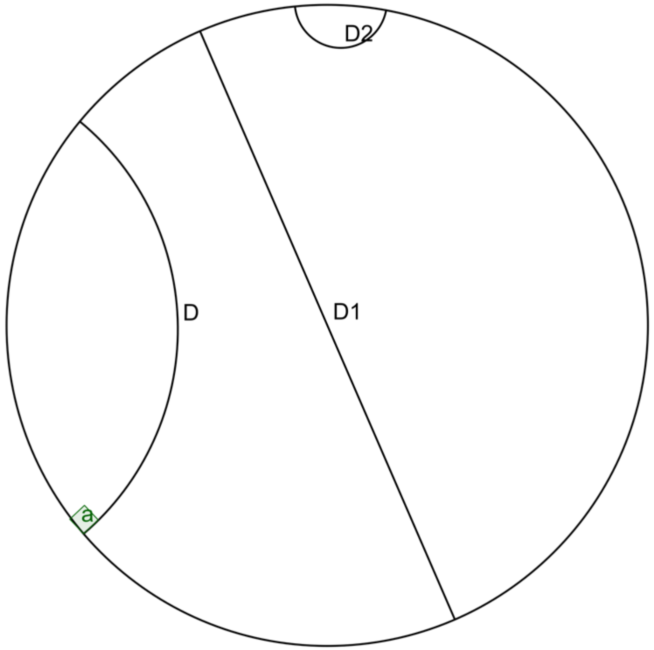




Multiple isometric models exist, with different pros and cons for numerical complexity, stability, and visualization prowess.



Poincaré ball model



Numerical operation in Poincaré model

Points inside unit ball

$$\mathbb{D}^n = \{x \in \mathbb{R}^n : \|x\| < 1\}$$

Tensor metric

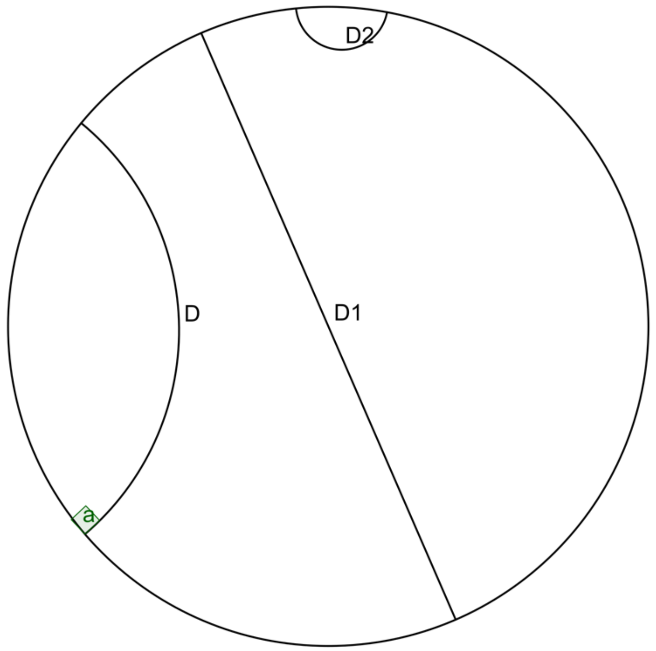
$$g_x^{\mathbb{D}} = \lambda_x^2 g^E, \quad \text{where } \lambda_x := \frac{2}{1 - \|x\|^2}$$

Distance between two points:

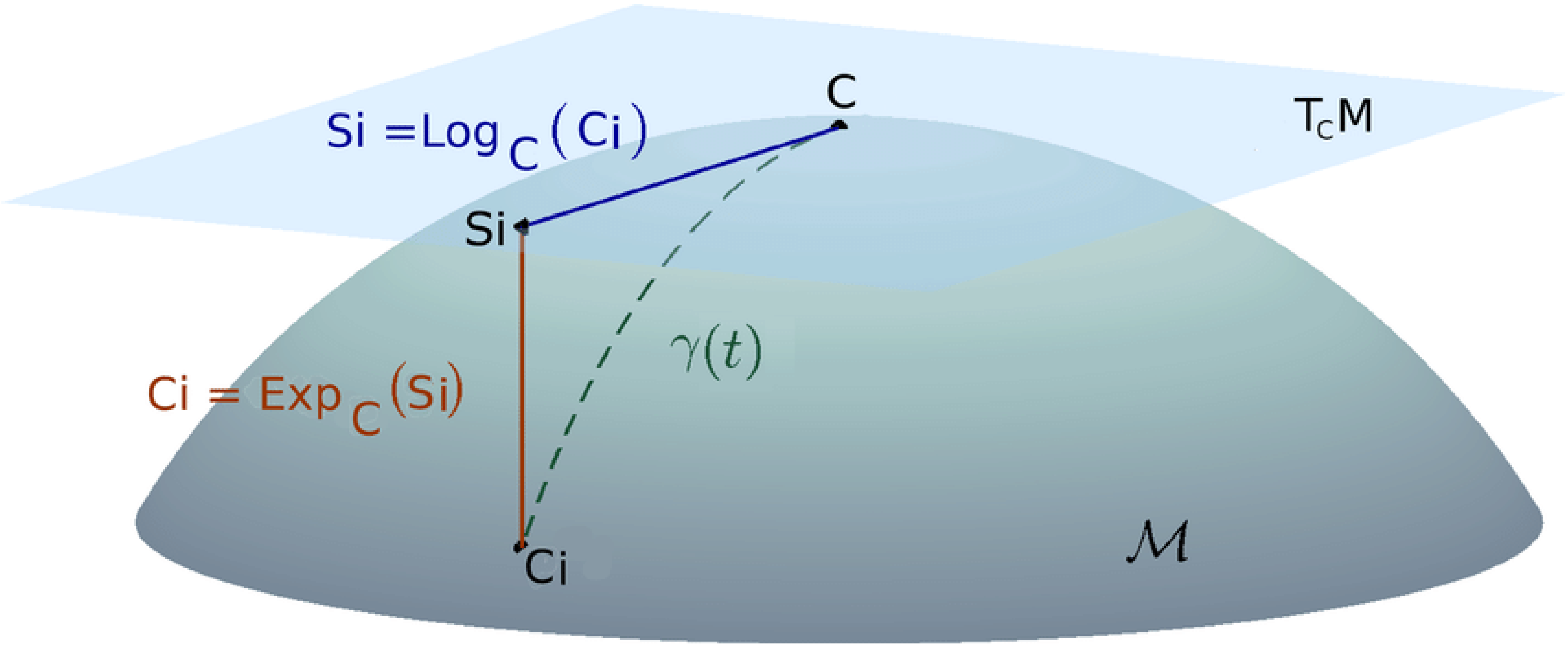
$$d_{\mathbb{D}}(x, y) = \cosh^{-1}\left(1 + 2\frac{\|x - y\|^2}{(1 - \|x\|^2)(1 - \|y\|^2)}\right)$$

Möbius addition:

$$x \oplus_c y := \frac{(1 + 2c\langle x, y\rangle + c\|y\|^2)x + (1 - c\|x\|^2)y}{1 + 2c\langle x, y\rangle + c^2\|x\|^2\|y\|^2}.$$



From tangent space to Poincaré ball (and back)



$$\exp_{\mathbf{0}}^{c}(v) = \tanh(\sqrt{c}\|v\|)\frac{v}{\sqrt{c}\|v\|}$$

$$\log_{\mathbf{0}}^{c}(y) = \tanh^{-1}(\sqrt{c}\|y\|)\frac{y}{\sqrt{c}\|y\|}$$

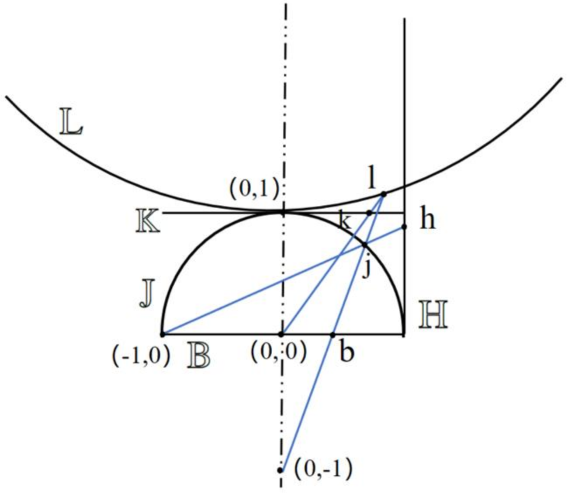
Models of hyperbolic geometry

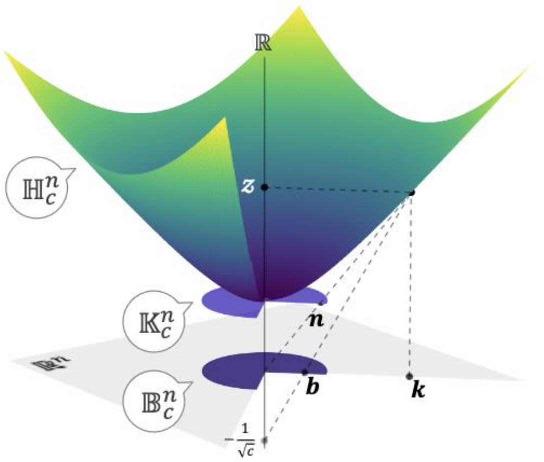
Different models shine in different aspects.

Poincaré most intuitive, Lorentz most stable, Klein fastest average operator

Many open questions about which models are best for deep learning.

Poincaré ball and Lorentz model default choices.





What is hyperbolic geometry?

Hierarchies and graphs in hyperbolic space

Deep learning with hierarchies

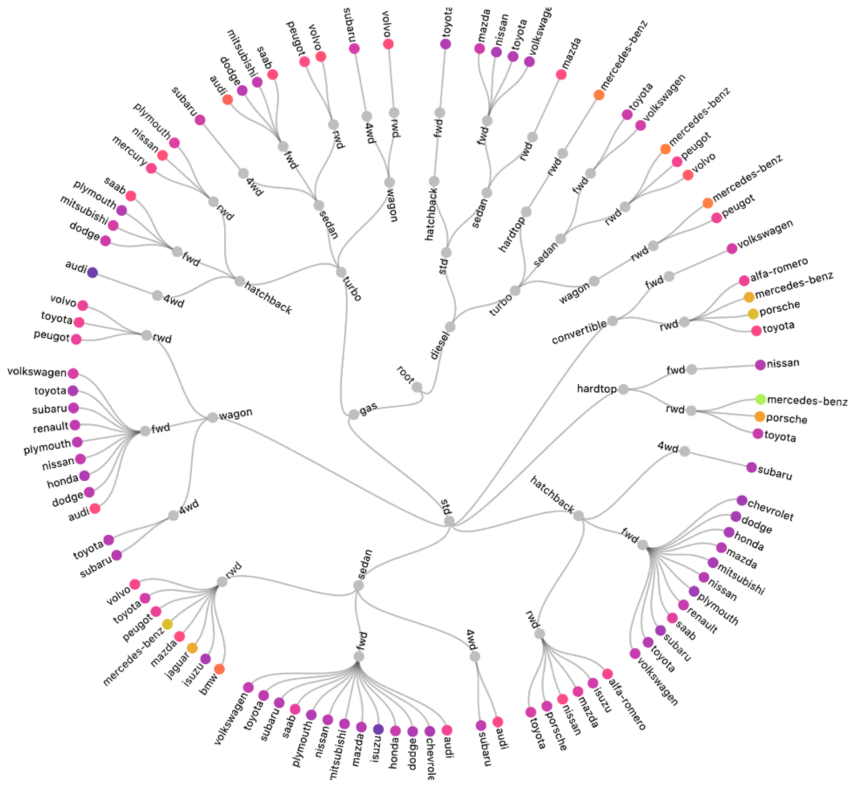
Multi-modal deep learning

Outlook, challenges, and resources

Poincaré Embeddings

[Nickel and Kiela. NeurIPS 2017]

Embed nodes as hyperbolic points and optimize with contrastive learning.



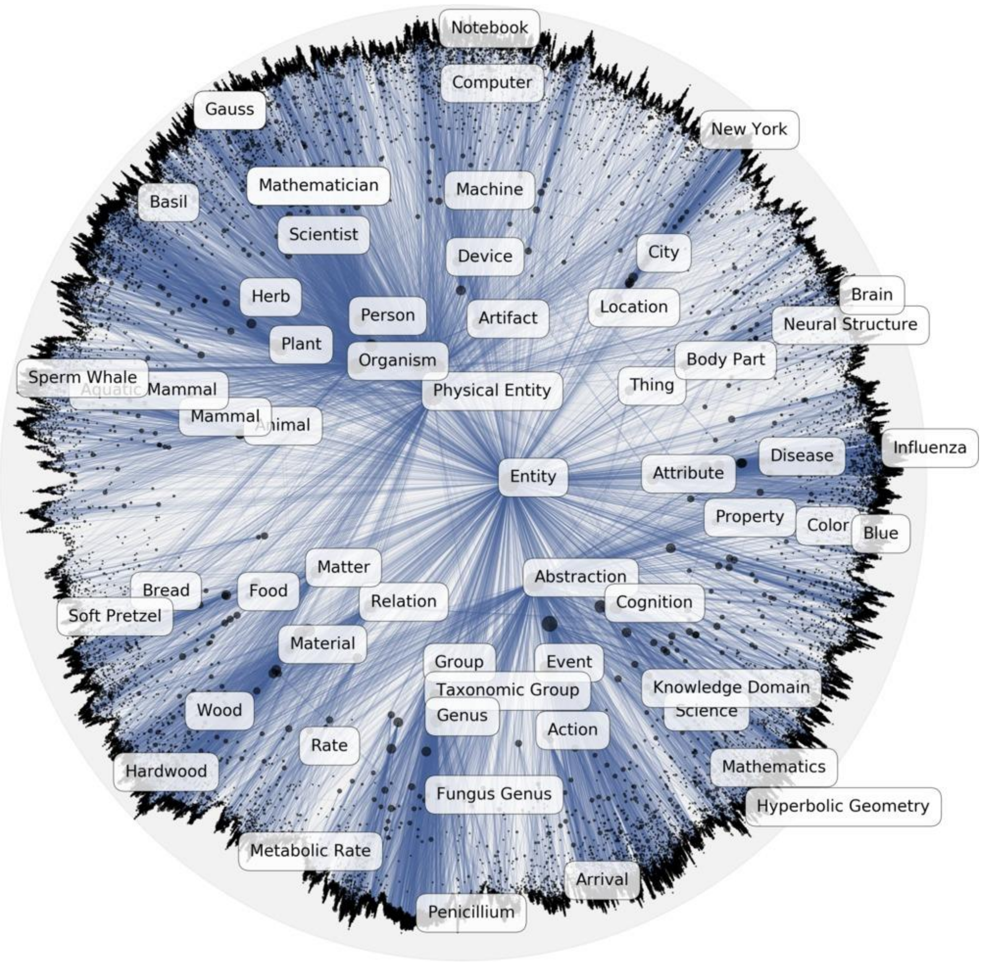
$$\mathcal{L}(\Theta) = \sum_{(u,v)\in\mathcal{D}} \log \frac{e^{-d(\boldsymbol{u},\boldsymbol{v})}}{\sum_{\boldsymbol{v}'\in\mathcal{N}(u)} e^{-d(\boldsymbol{u},\boldsymbol{v}')}},$$

Parent-child links

Hyperbolic distance

Non parent-child links

Poincaré Embeddings



			Dimensionality					
			5	10	20	50	100	200
WordNet Reconstruction	**Euclidean**	Rank	3542.3	2286.9	1685.9	1281.7	1187.3	1157.3
		MAP	0.024	0.059	0.087	0.140	0.162	0.168
	Translational	Rank	205.9	179.4	95.3	92.8	92.7	91.0
		MAP	0.517	0.503	0.563	0.566	0.562	0.565
	Poincaré	Rank	4.9	4.02	3.84	3.98	3.9	**3.83**
		MAP	0.823	0.851	0.855	0.86	0.857	**0.87**
WordNet Link Pred.	**Euclidean**	Rank	3311.1	2199.5	952.3	351.4	190.7	81.5
		MAP	0.024	0.059	0.176	0.286	0.428	0.490
	Translational	Rank	65.7	56.6	52.1	47.2	43.2	40.4
		MAP	0.545	0.554	0.554	0.56	0.562	0.559
	Poincaré	Rank	5.7	**4.3**	4.9	4.6	4.6	4.6
		MAP	0.825	0.852	0.861	**0.863**	0.856	0.855

Hyperbolic Entailment Cones

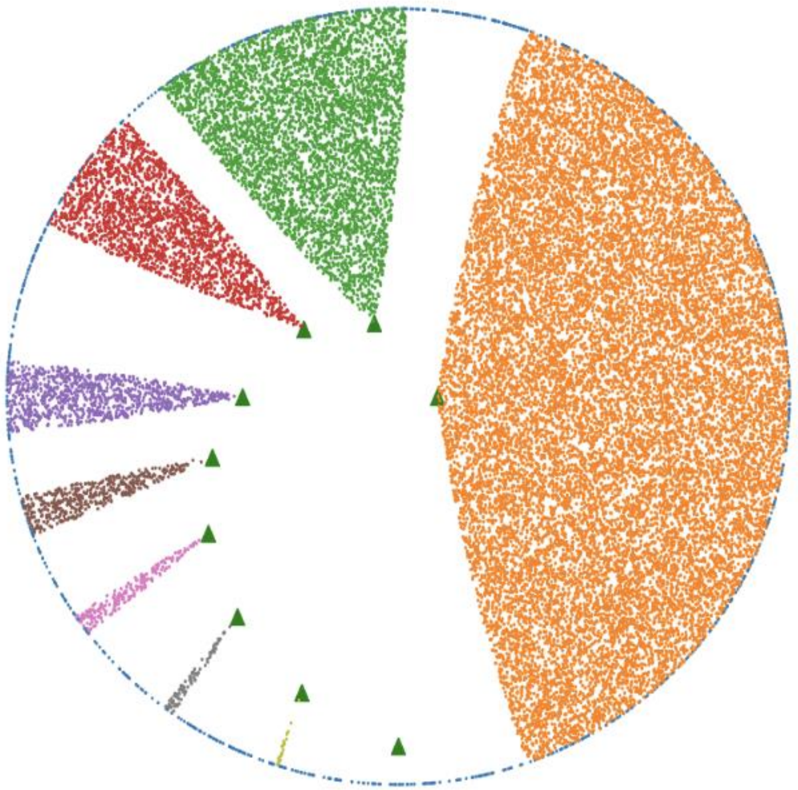
[Ganea et al. ICML 2018]

Pairwise contrastive learning has trouble enforcing hierarchical depth.

They propose to view points as cones of entailment.

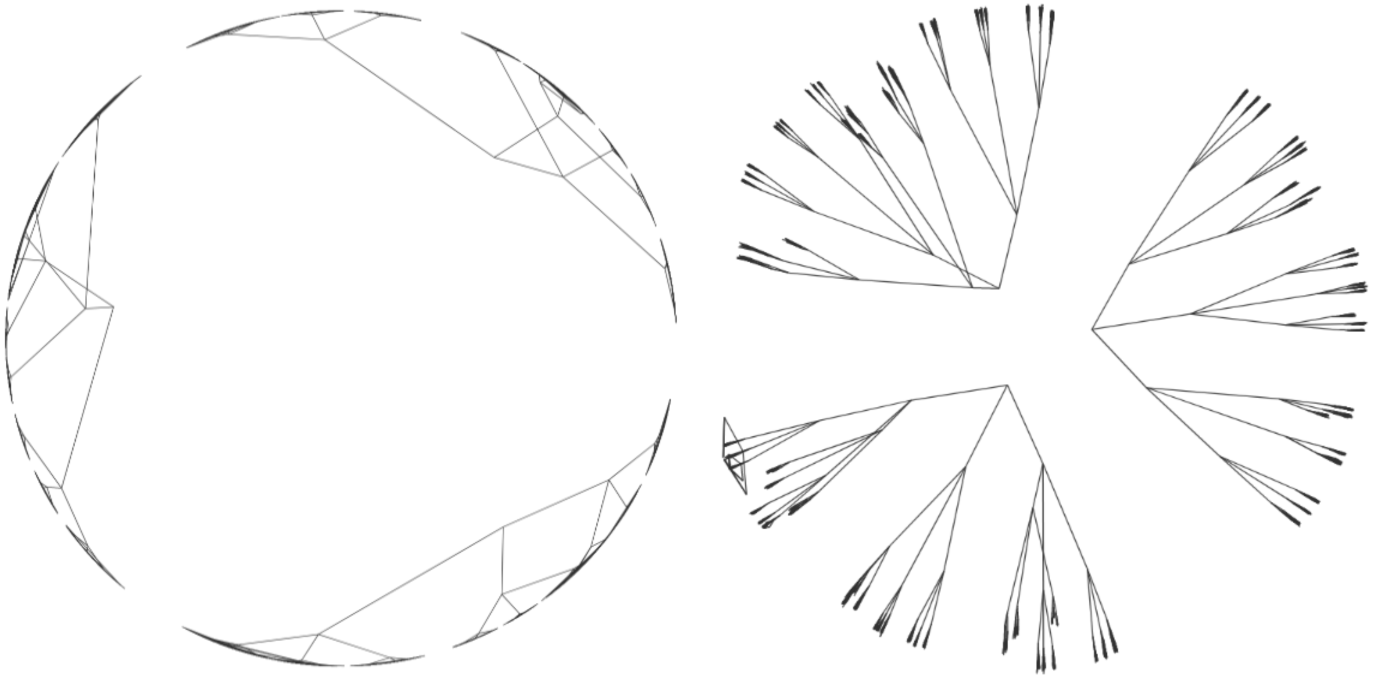
Required properties:

1. Axial symmetry
2. Rotation invariance
3. Aperture of cone is continuous function
4. Nested angular cones preserve partial order



Hyperbolic Entailment Cones

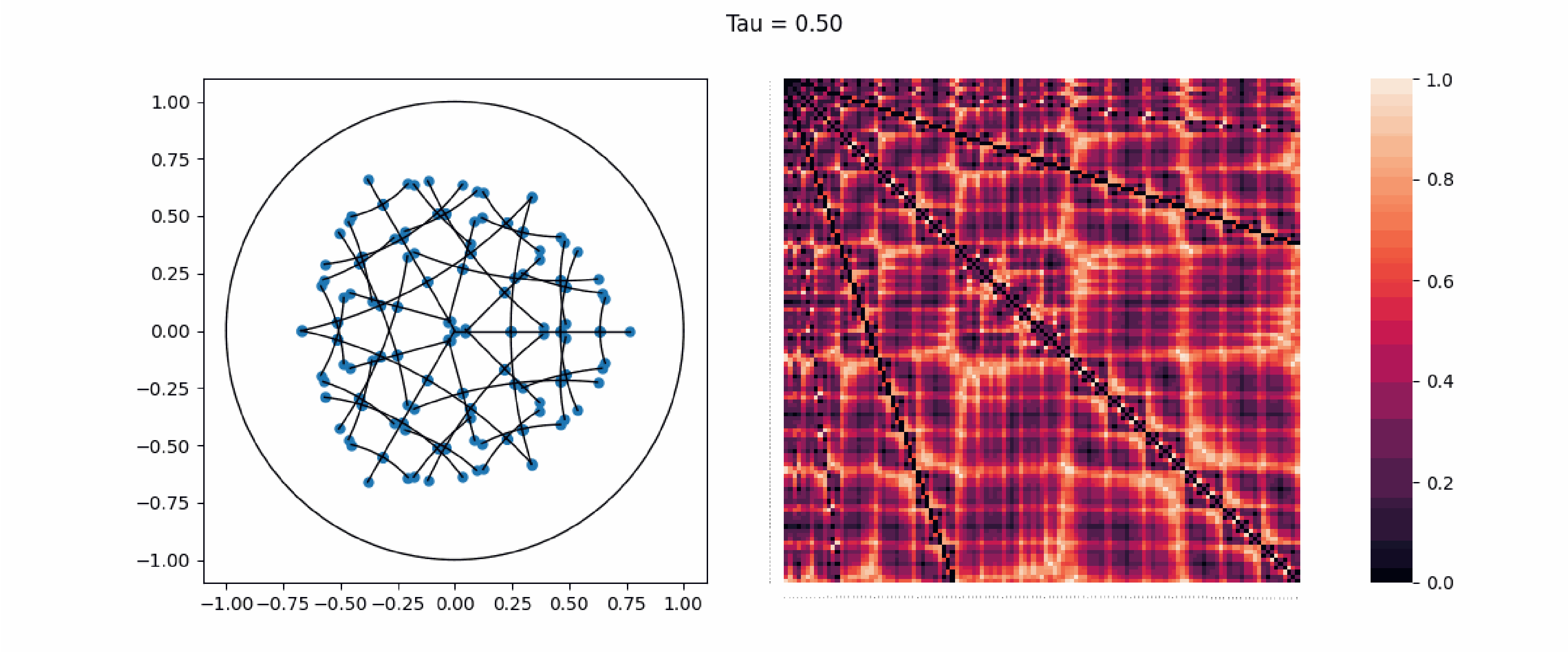
[Ganea et al. ICML 2018]



Poincaré Embeddings (left) vs Hyperbolic Entailment Cones (right)

Closed-form Embeddings

[Sarkar et al. 2012]



In 2D, it is possible to embed hierarchies through uniform spreading and projections.

No direct higher-dimensional generalization feasible, only approximate solutions.

What is hyperbolic geometry?

Hierarchies and graphs in hyperbolic space

Deep learning with hierarchies

Multi-modal deep learning

Outlook, challenges, and resources

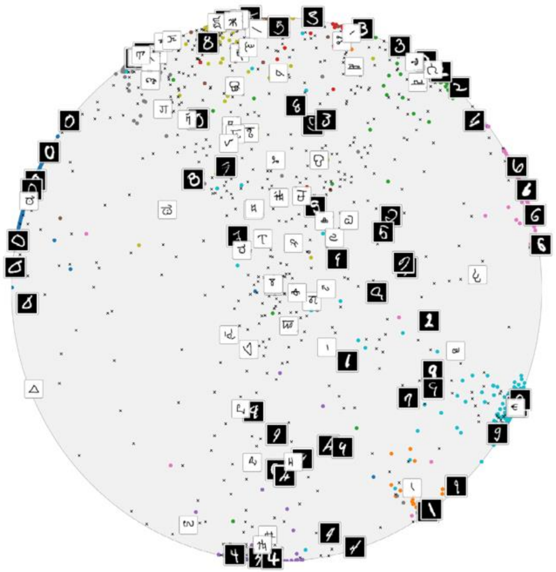
Hyperbolic embeddings for vision

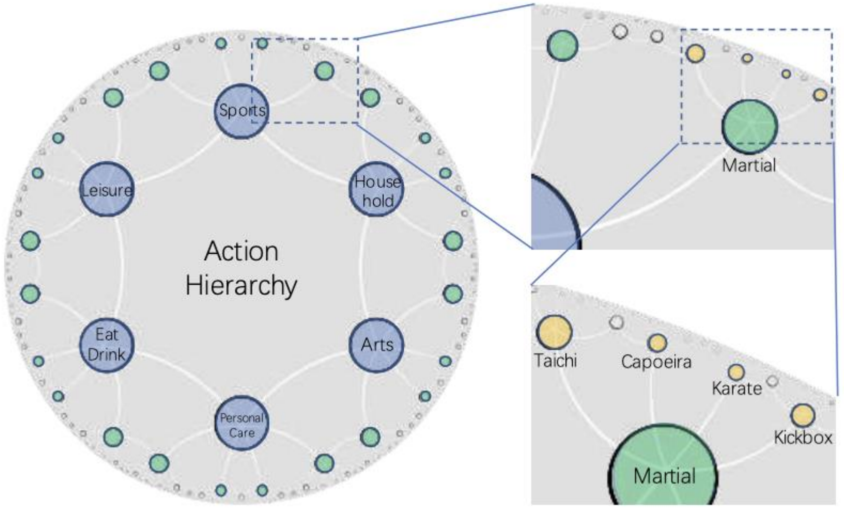
[Liu et al, Long et al., Khrulkov et al. CVPR 2020]

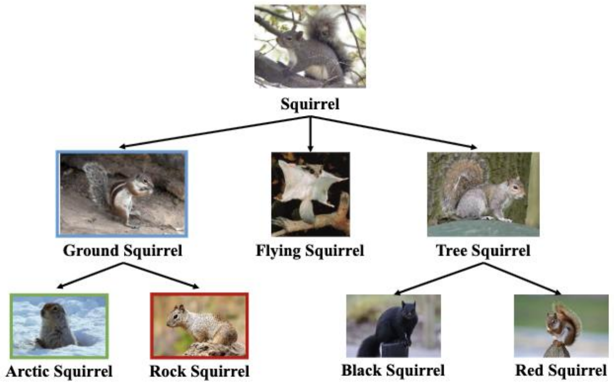
CVPR 2020

Three papers with the exact same question:

What happens when we place hyperbolic embeddings on top of deep networks?

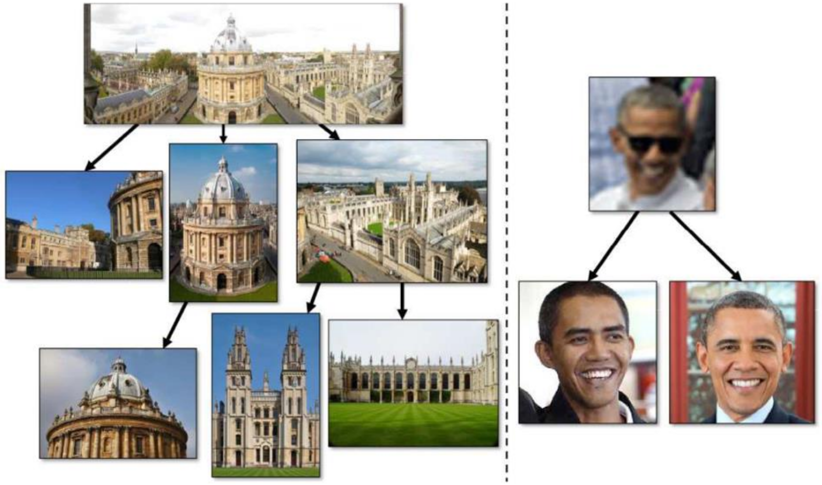


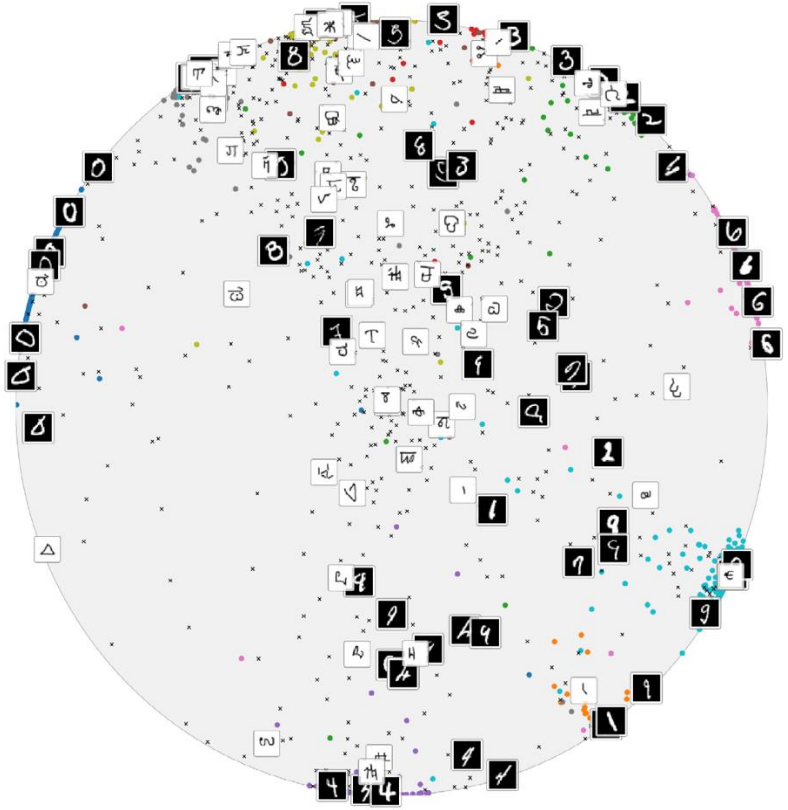




Hyperbolic Image Embeddings

[Khrulkov et al. CVPR 2020]



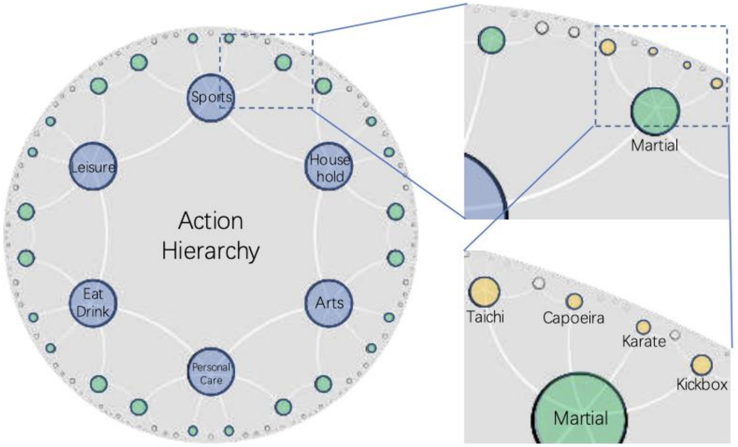


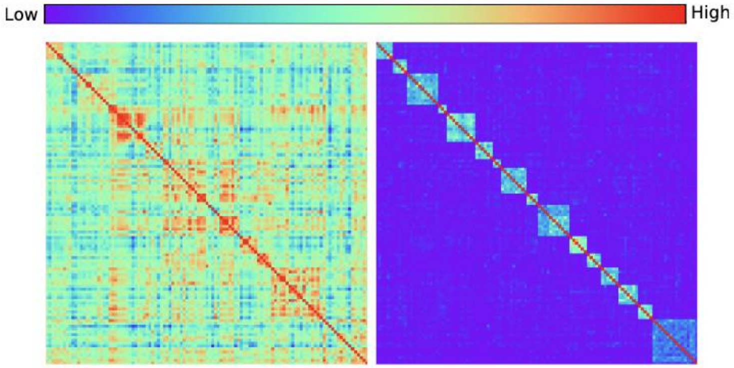
Encoder	Dataset			
	CIFAR10	CIFAR100	CUB	*Mini*ImageNet
Inception v3 [49]	0.25	0.23	0.23	0.21
ResNet34 [14]	0.26	0.25	0.25	0.21
VGG19 [42]	0.23	0.22	0.23	0.17

Images are naturally hierarchical, hyperbolic embeddings improve few-shot learning.

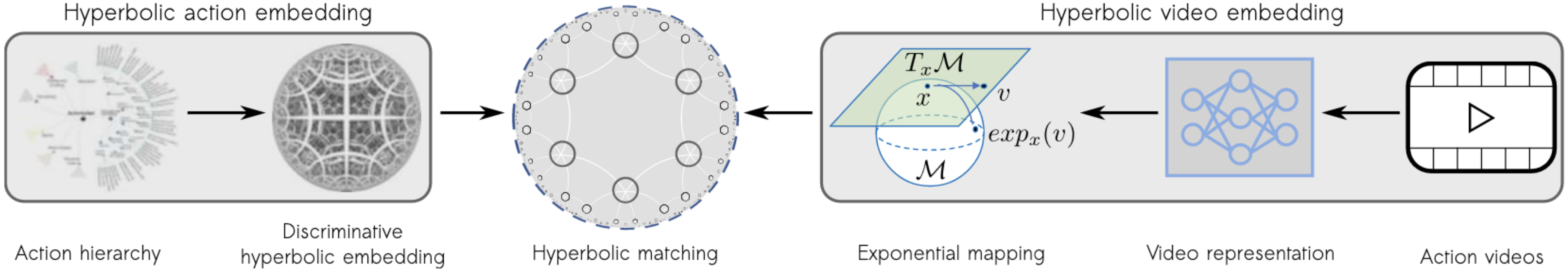
Hyperbolic actions

[Long et al.. CVPR 2020]





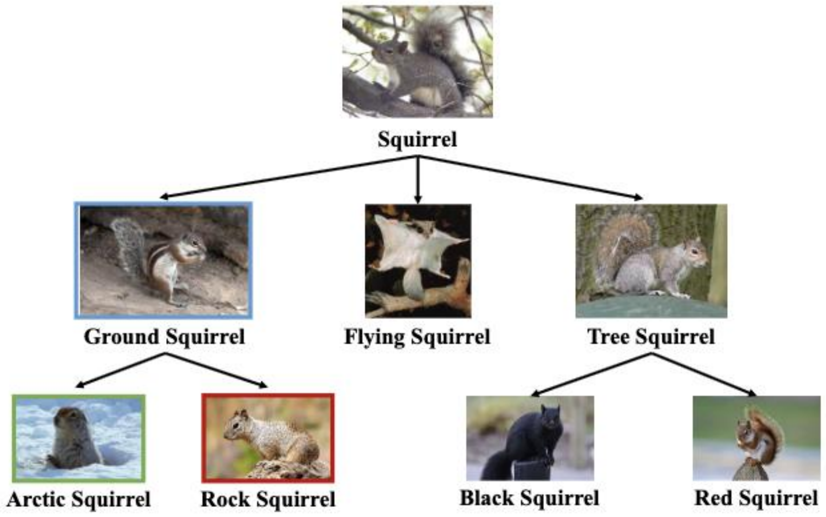
(a) One-hot. (b) Hyperbolic (*ours*).

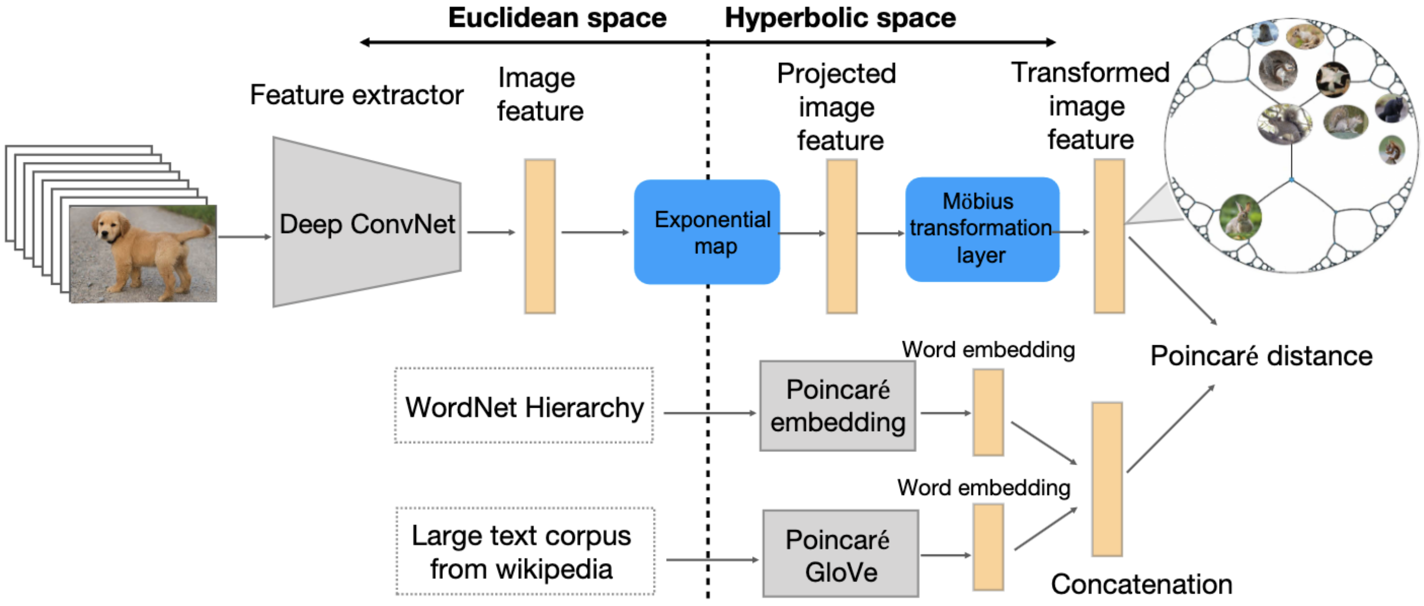


Videos are naturally hierarchical, hyperbolic embeddings improve action recognition.

Hyperbolic zero-shot learning

[Liu et al. CVPR 2020]

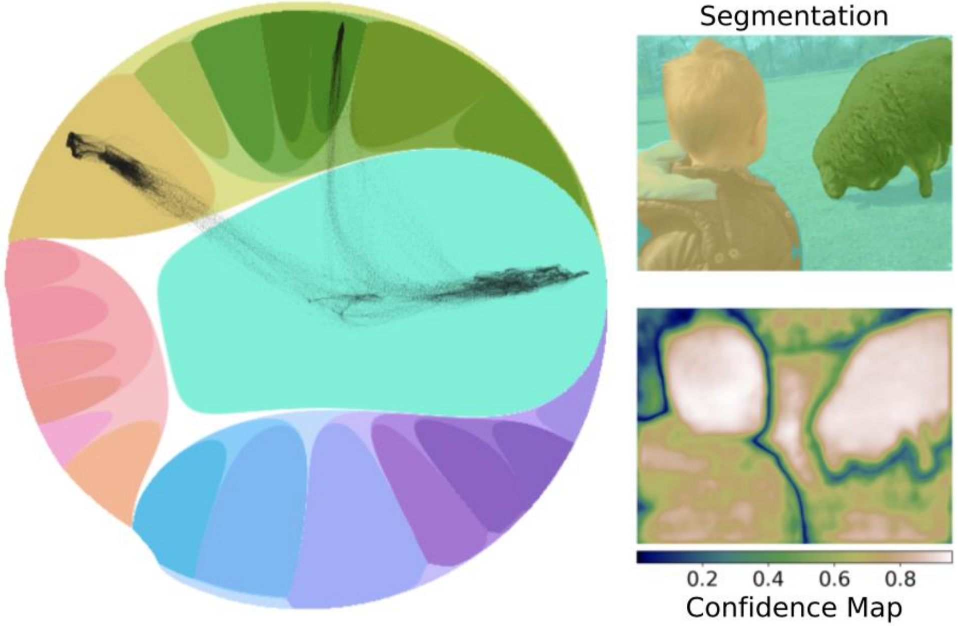




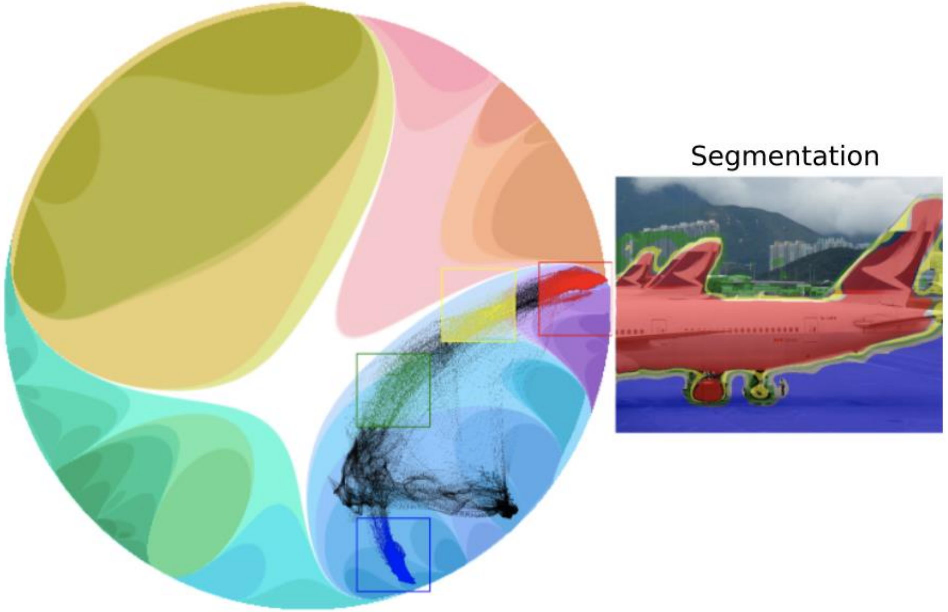
Semantics is naturally hierarchical, hyperbolic embeddings improve zero-shot recognition.

Hyperbolic Image Segmentation

[Ghadimi Atigh et al. CVPR 2022]



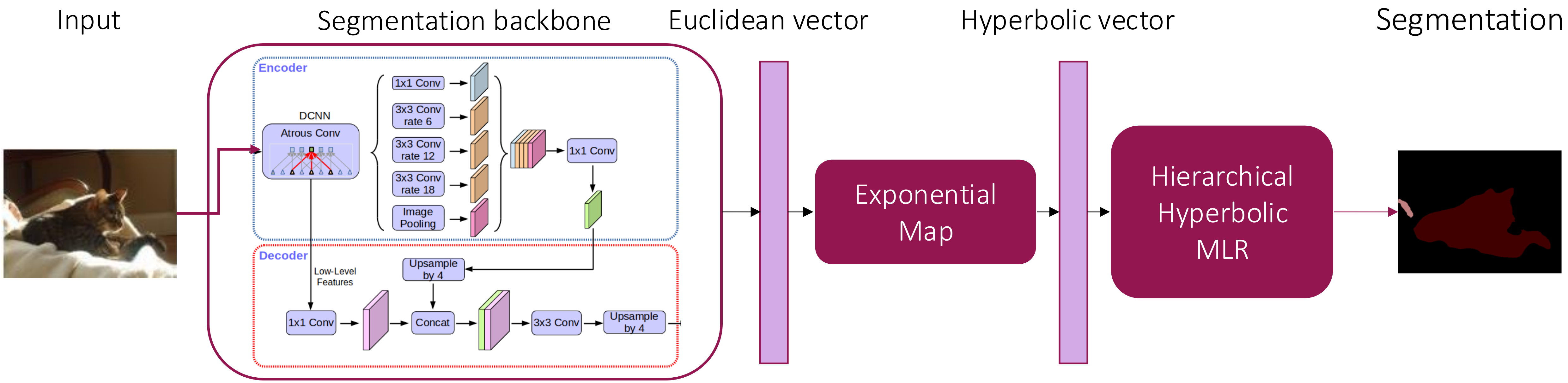
(a) Prediction uncertainty for free



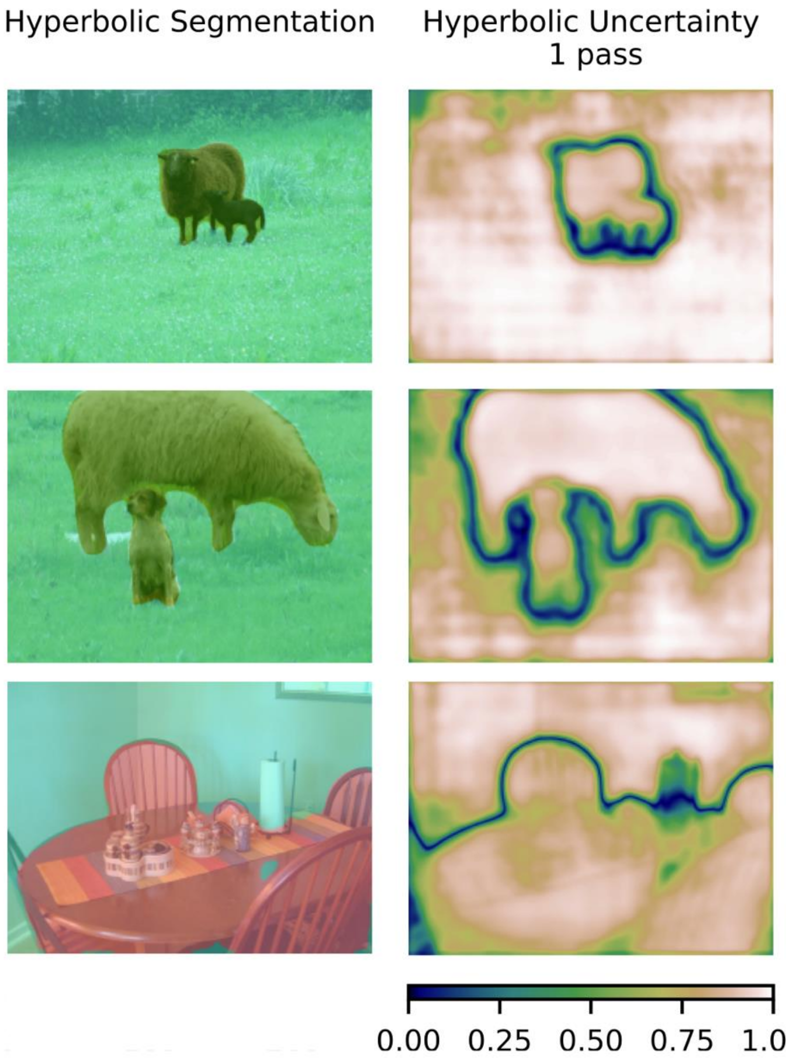
(b) Boundary information for free

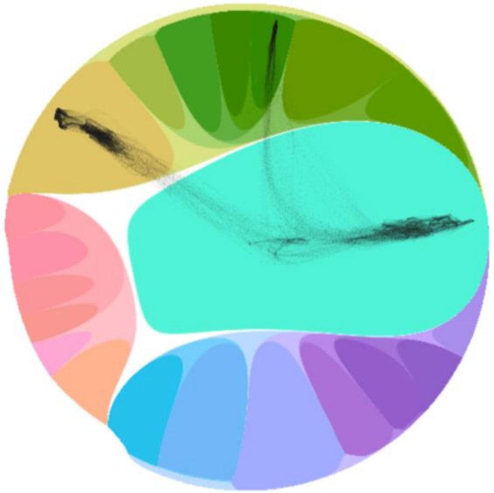
What happens when the final pixel classification is done in hyperbolic space?

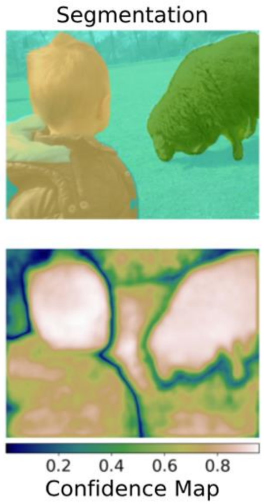
Hyperbolic Image Segmentation



Hyperbolic Image Segmentation



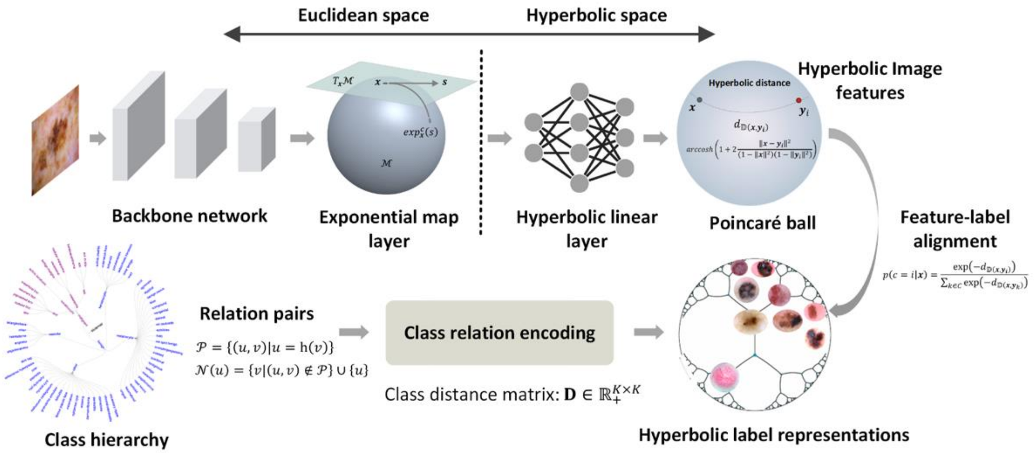




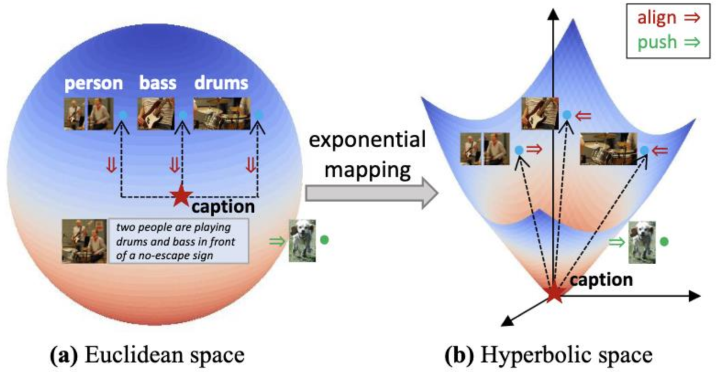
COCO-Stuff-10k				
Manifold	Hierarchical	Class Acc	Pixel Acc	mIOU
$\mathbb{R}$		0.44	0.33	0.23
$\mathbb{R}$	✓	3.29	48.65	18.53
$\mathbb{D}$	✓	**3.46**	**51.70**	**21.15**

Pascal VOC				
Manifold	Hierarchical	Class Acc	Pixel Acc	mIOU
$\mathbb{R}$		4.88	10.84	2.59
$\mathbb{R}$	✓	7.80	31.04	16.15
$\mathbb{D}$	✓	**12.15**	**47.92**	**34.87**

Too many hyperbolic vision papers to mention

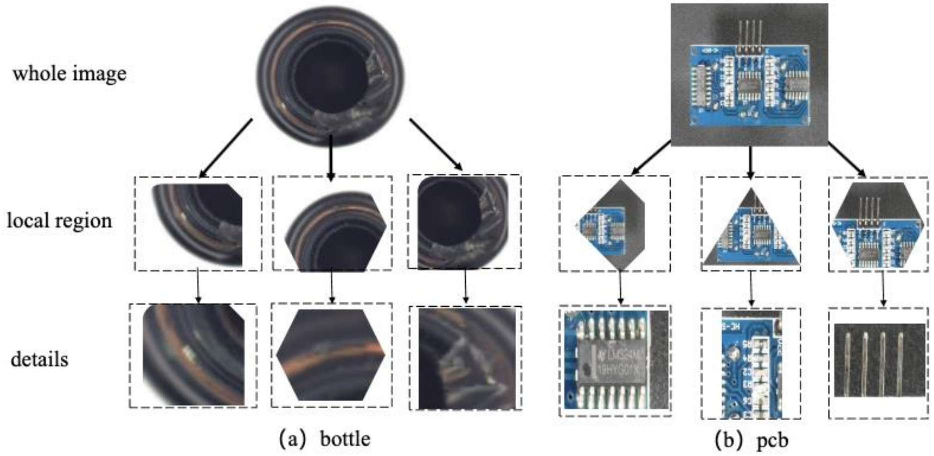


[Yu et al. MICCAI 2022]



[Kong et al. CVPR 2024]

[Ge et al. CVPR 2023]



[Li et al. CVPR 2024]

What is hyperbolic geometry?

Hierarchies and graphs in hyperbolic space

Deep learning with hierarchies

Multi-modal deep learning

Outlook, challenges, and resources

Hyperbolic vision-language models

[Desai et al. ICML 2023]

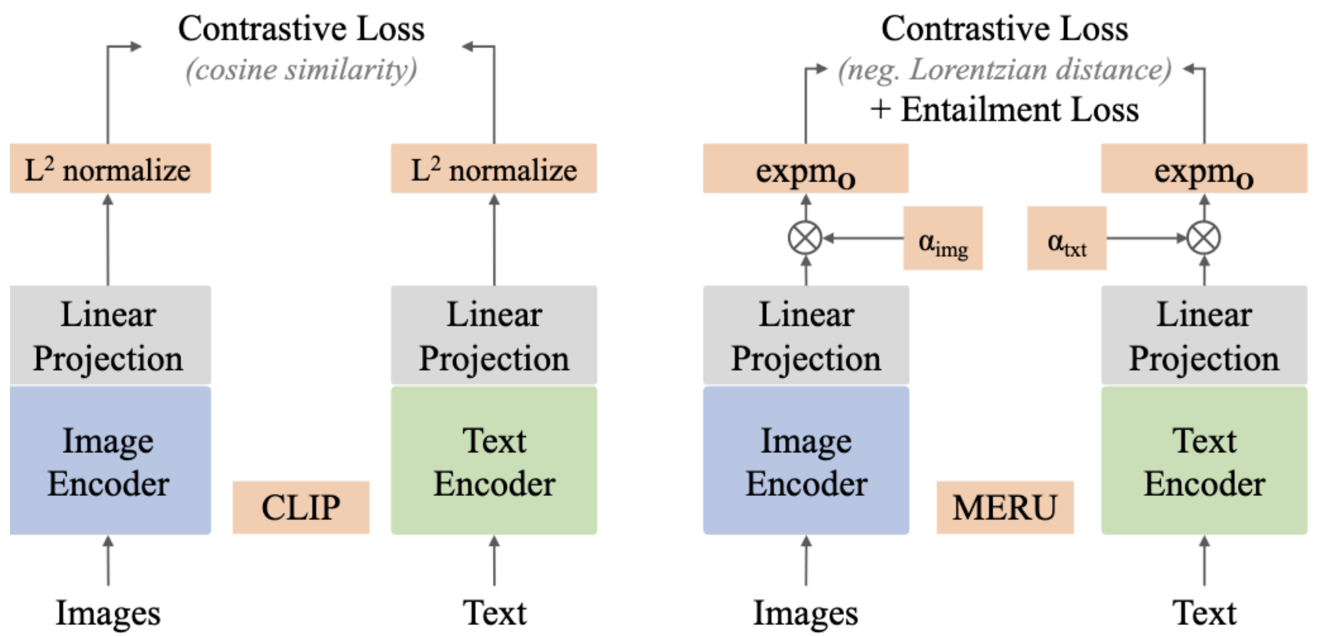
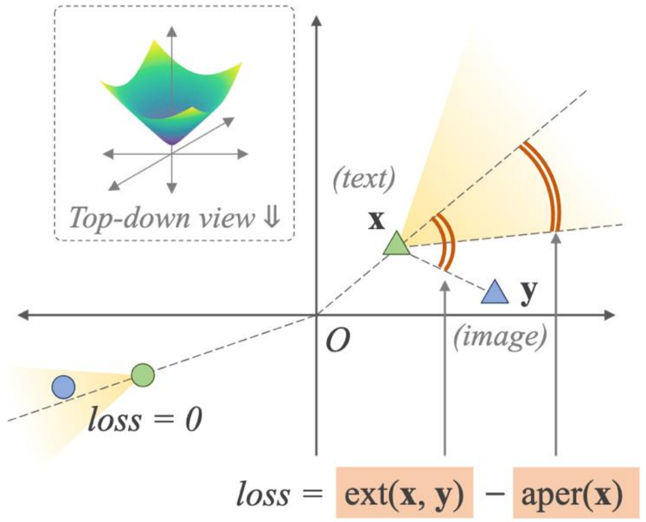


Image-text representation learning wants to collapse image and text embeddings.

Hyperbolic vision-language models

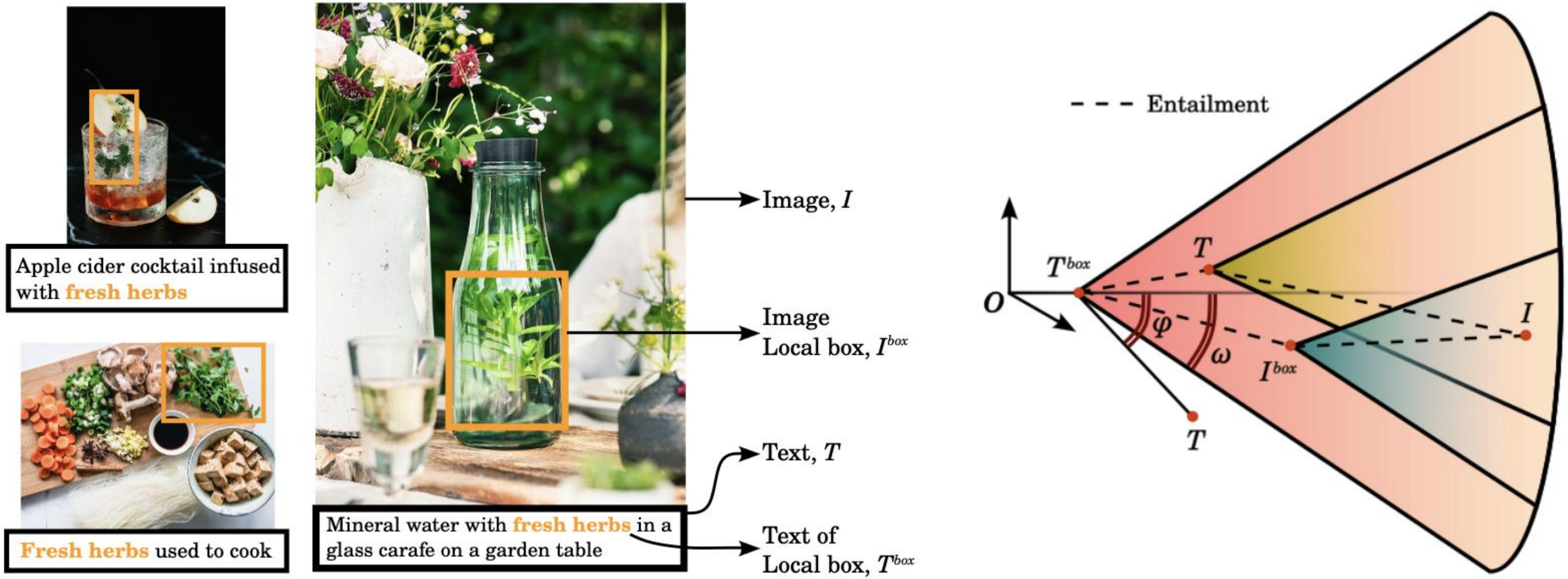
Intuitively, image and text embeddings are unequal!



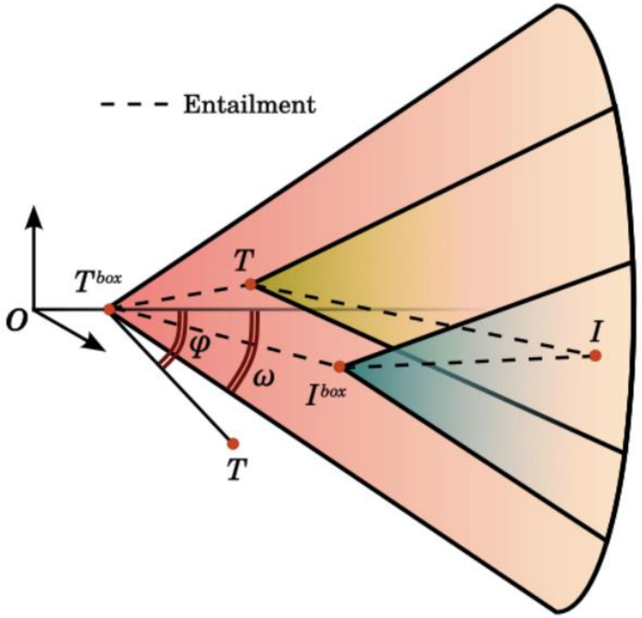
Hyperbolic entailments allow to model this imbalance
and learn the hierarchical nature of image-text representations.

Vision-language models and compositions

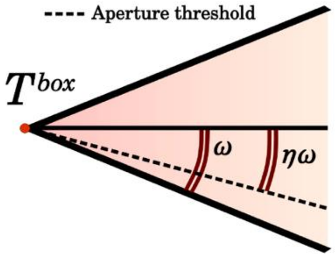
[under submission]



Vision-language models and compositions



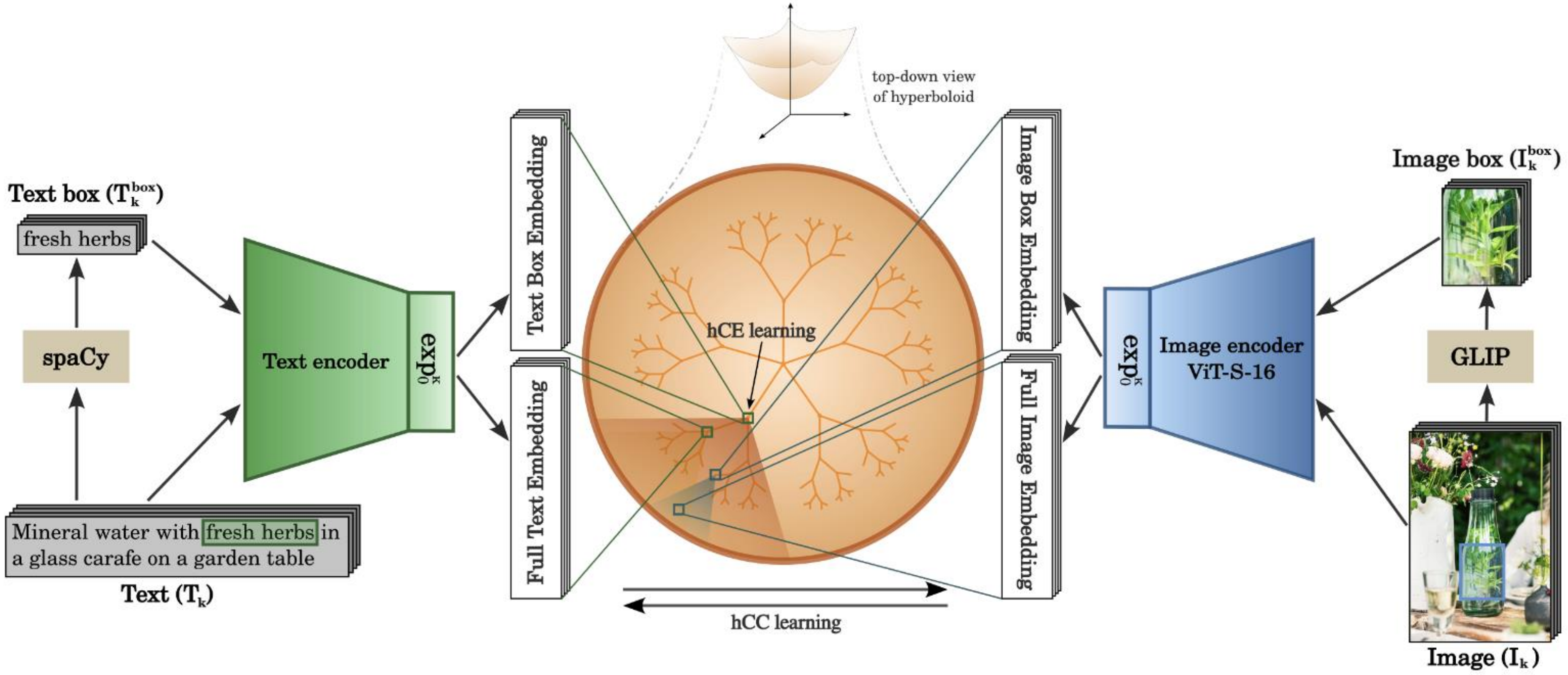
$$L^*_{entail}(p,q) = \max(0, \phi(\mathbf{p},\mathbf{q}) - \eta\omega(\mathbf{q})).$$

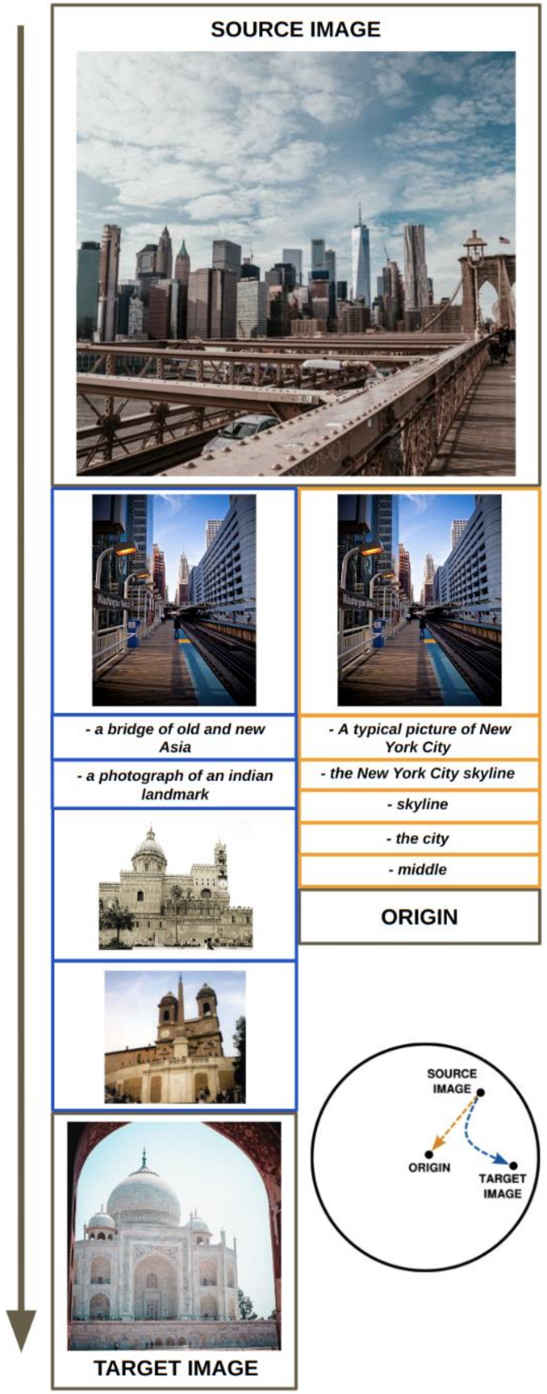


$$hCE(I,T,I^{box},T^{box}) = \underbrace{L^*_{entail}(I^{box},T^{box}) + L^*_{entail}(I,T)}_{\text{inter-modality entailment}}$$

$$+ \underbrace{L^*_{entail}(I,I^{box}) + L^*_{entail}(T,T^{box})}_{\text{intra-modality entailment}}.$$

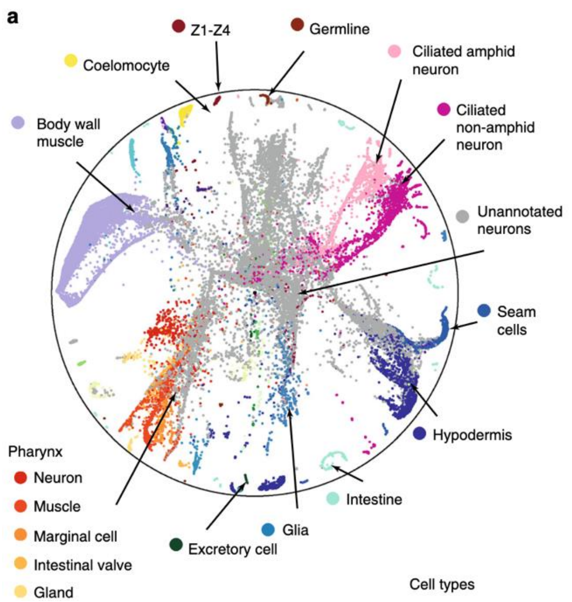
Multi-modal network





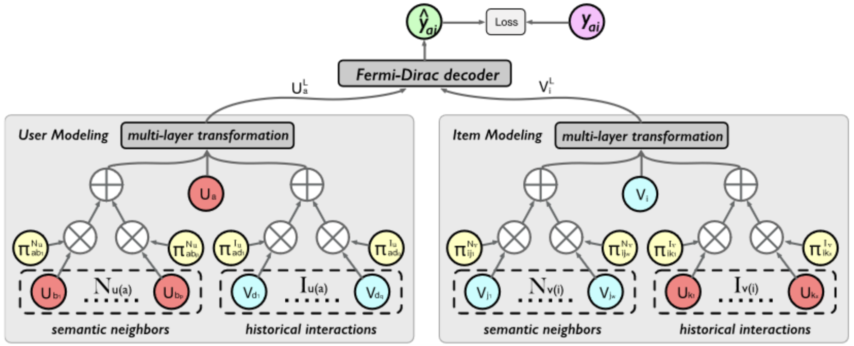
Hyperbolic embeddings for other data types

[Klimovskaia et al. Nature Comm. 2020]



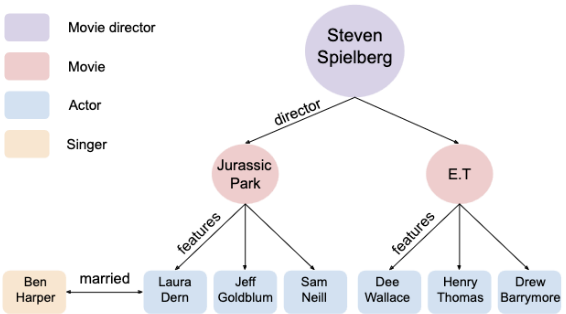
Hyperbolic embeddings of single-cell data.

[Li et al. TKDD 2023]



Hyperbolic recommender systems

[Chami et al. 2020]



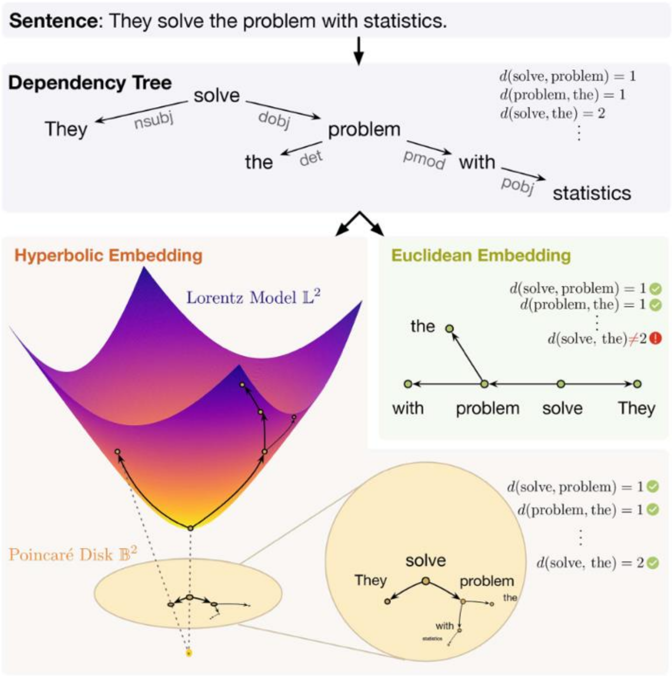
Hyperbolic knowledge graphs

And for text, music, 3D skeletons, phylogenetic placement, social networks, clustering...

Hyperbolic large language models?

[Chen et al. TASLP 2024]

Large language models are Euclidean, but should they be?



Very recent work is making the first step, but no big companies dare to make the step (yet?).

What is hyperbolic geometry?

Hierarchies and graphs in hyperbolic space

Deep learning with hierarchies

Multi-modal deep learning

Outlook, challenges, and resources

The big potential of hyperbolic learning

Hierarchical learning	model the hierarchies of semantics and data.
Robust learning	handle new distributions and adversarial samples.
Low-dimensional learning	hyperbolic space is dense, allowing for smaller networks.
Brain-like networks	brains are likely hyperbolic, big links with neuroscience

Zhang et al. Nature Communications 2022

The grand challenges of hyperbolic learning

Fully hyperbolic learning	which hyperbolic model is best? and how to optimize?
Computational challenges	numerical stability and speed of computation.
Open source community	where is hyperbolic PyTorch?
Learning at scale	we need an ImageNet/CLIP moment for hyperbolic learning.

Unexplored territories

Hyperbolic transformers

Hyperspectral data

Brain is hyperbolic?

Numerical stability

Hyperbolic ChatGPT

Optimization

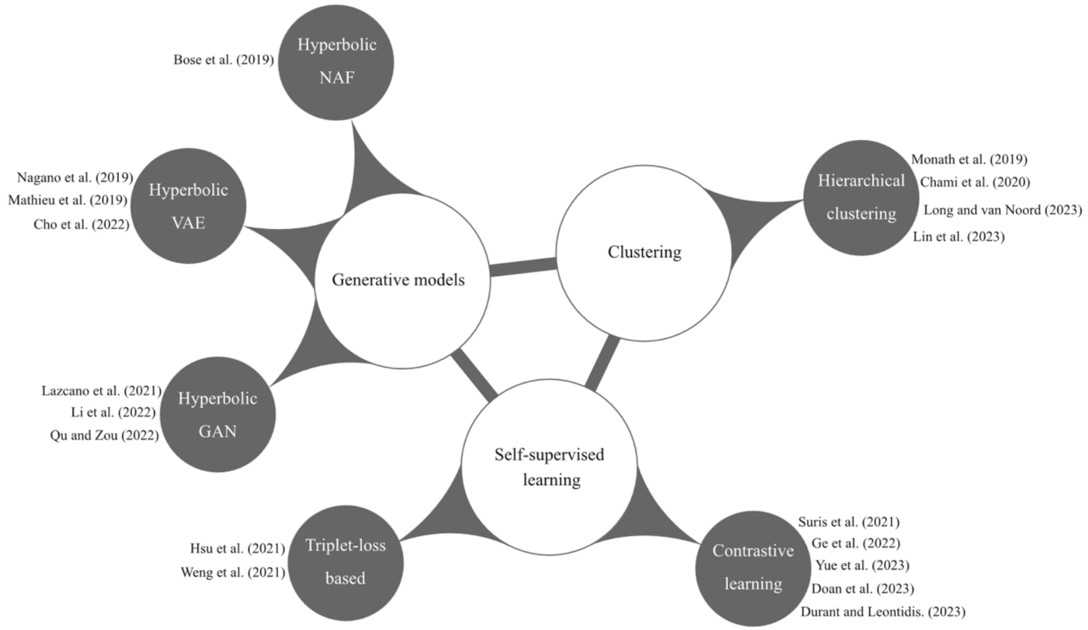
Hyperbolic diffusion models

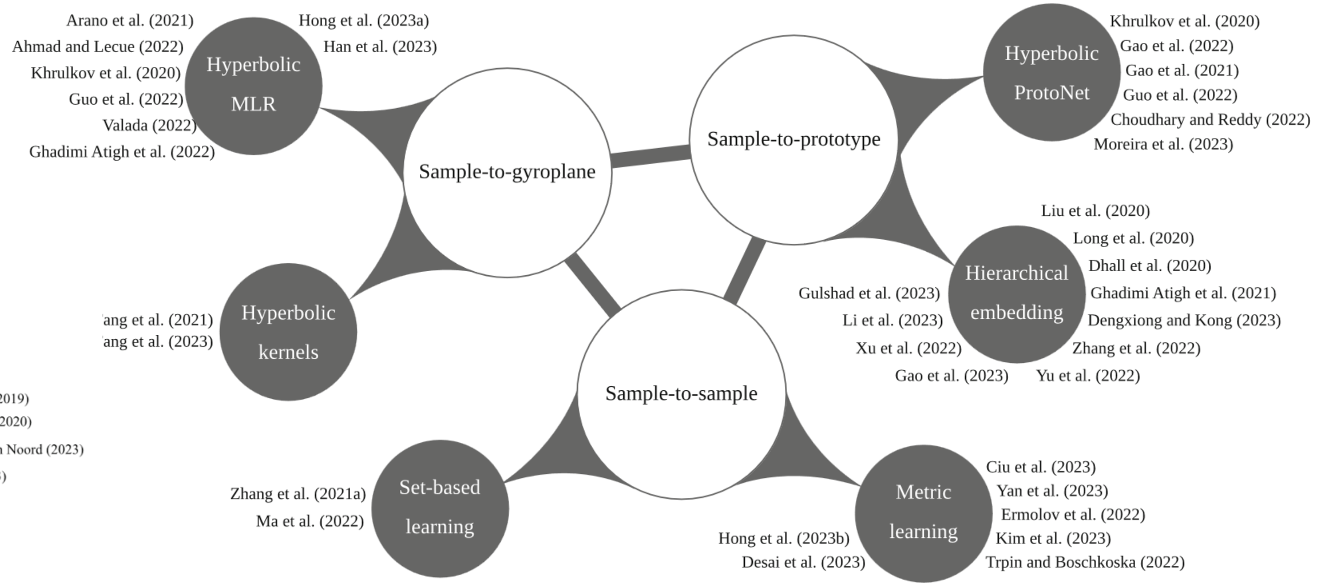
Which hyperbolic model is best?

Where to start?

Pascal Mettes, Mina Ghadimi, Martin Keller-Ressel, Jeffrey Gu, Serena Yeung. IJCV 2024

Hyperbolic Deep Learning in Computer Vision: A Survey





Where to start?

GitHub

Max van Spengler, Philipp Wirth, Pascal Mettes. ACM MM 2024

HypLL: The Hyperbolic Learning Library

```python
import hypll.nn as hnn
from hypll.manifolds.poincare_ball import (
    Curvature, PoincareBall
)

ball = PoincareBall(Curvature(1.0))
class HNet(nn.Module):
    def __init__(self):
        super().__init__()
        self.conv1 = hnn.HConvolution2d(3, 6, 5, ball)
        self.pool = hnn.HMaxPool2d(2, ball, 2)
        self.conv2 = hnn.HConvolution2d(6, 16, 5, ball)
        self.fc1 = hnn.HLinear(16 * 5 * 5, 120, ball)
        self.fc2 = hnn.HLinear(120, 84, ball)
        self.fc3 = hnn.HLinear(84, 10, ball)
        self.relu = hnn.HReLU(ball)

    def forward(self, x):
        x = self.pool(self.relu(self.conv1(x)))
        x = self.pool(self.relu(self.conv2(x)))
        x = x.flatten(1)
        x = self.relu(self.fc1(x))
        x = self.relu(self.fc2(x))
        x = self.fc3(x)
        return x
```

```python
import torch.nn as nn



class Net(nn.Module):
    def __init__(self):
        super().__init__()
        self.conv1 = nn.Conv2d(3, 6, 5)
        self.pool = nn.MaxPool2d(2, 2)
        self.conv2 = nn.Conv2d(6, 16, 5)
        self.fc1 = nn.Linear(16 * 5 * 5, 120)
        self.fc2 = nn.Linear(120, 84)
        self.fc3 = nn.Linear(84, 10)
        self.relu = nn.ReLU()

    def forward(self, x):
        x = self.pool(self.relu(self.conv1(x)))
        x = self.pool(self.relu(self.conv2(x)))
        x = x.flatten(1)
        x = self.relu(self.fc1(x))
        x = self.relu(self.fc2(x))
        x = self.fc3(x)
        return x
```

https://github.com/maxvanspengler/hyperbolic_learning_library

Where to start?

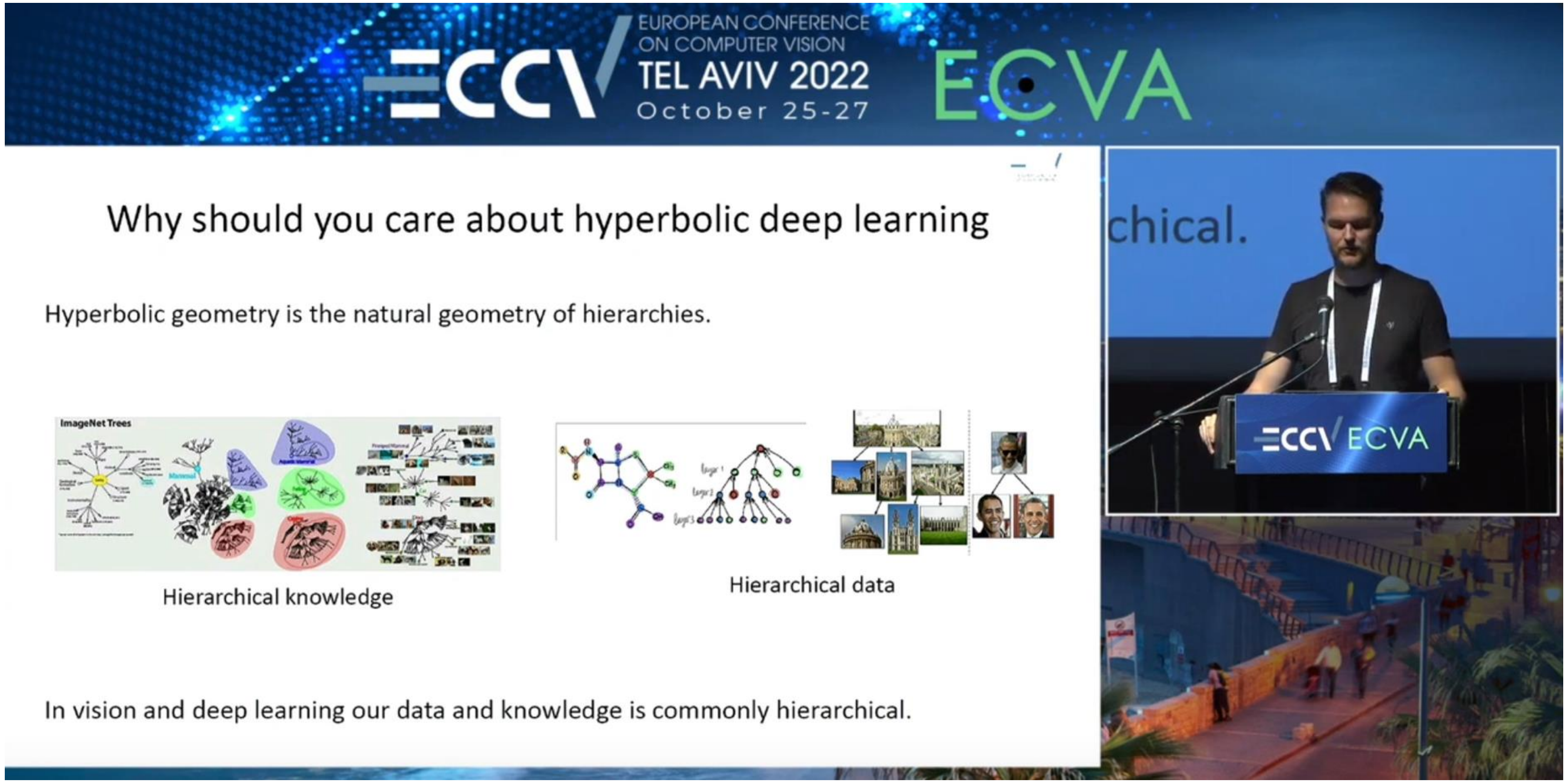
Max van Spengler, Philipp Wirth, Pascal Mettes. ACM MM 2024

HypLL: The Hyperbolic Learning Library

```python
from hypll.tensors import TangentTensor

for data in trainloader:
    inputs, labels = data

    tangents = TangentTensor(
        inputs, man_dim=1, manifold=ball
    )
    inputs_on_ball = ball.expmap(tangents)

    outputs = hnet(inputs_on_ball)
```

https://github.com/maxvanspengler/hyperbolic_learning_library

Where to start?



https://www.youtube.com/@hyperboliclearningforcv/playlists