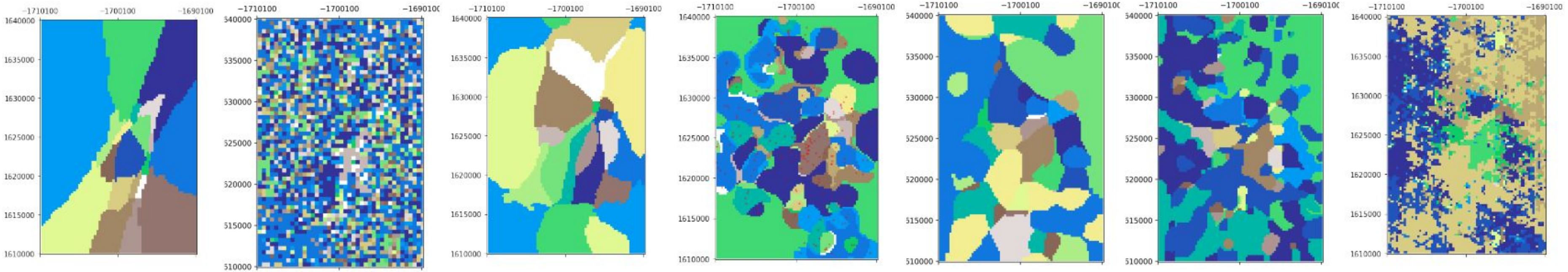
Location Encoding and Spatially-Explicit Machine Learning

Gengchen Mai

Ph.D. Candidate, Space Time Knowledge Organization Lab, UC Santa Barbara
http://www.geog.ucsb.edu/~gengchen_mai/



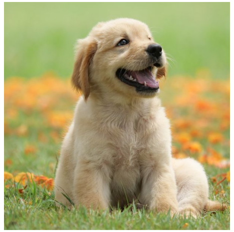
(a) *direct* (b) *tile* (c) *wrap* (d) *rbf* (σ=1k) (e) λ_{min}=1k (f) λ_{min}=500 (g) λ_{min}=50

Embedding clustering of different location encoding models:
(a)-(d) baselines (e)-(f) **Space2Vec**

ICLR 2020 paper: https://arxiv.org/abs/2003.00824 GitHub Repo: https://github.com/gengchenmai/space2vec

Why Spatial is Special?

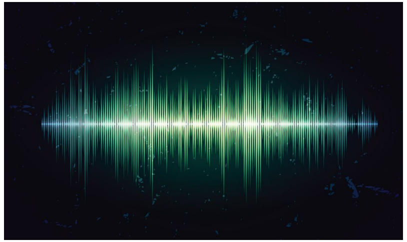
Spatial data have more **irregular structures**



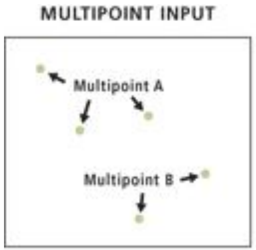
Image

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under **gravity**. The main forms of precipitation include drizzle, rain, sleet, snow, **graupel** and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called "showers".

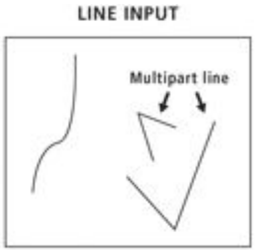
Text



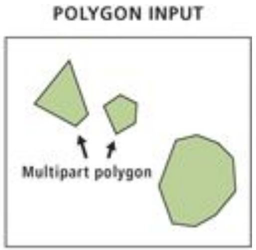
Audio



Point



Line



Polygon

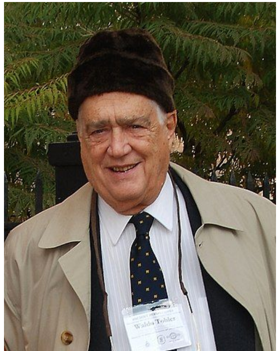
Why Spatial is Special?

Spatial Autocorrelation:

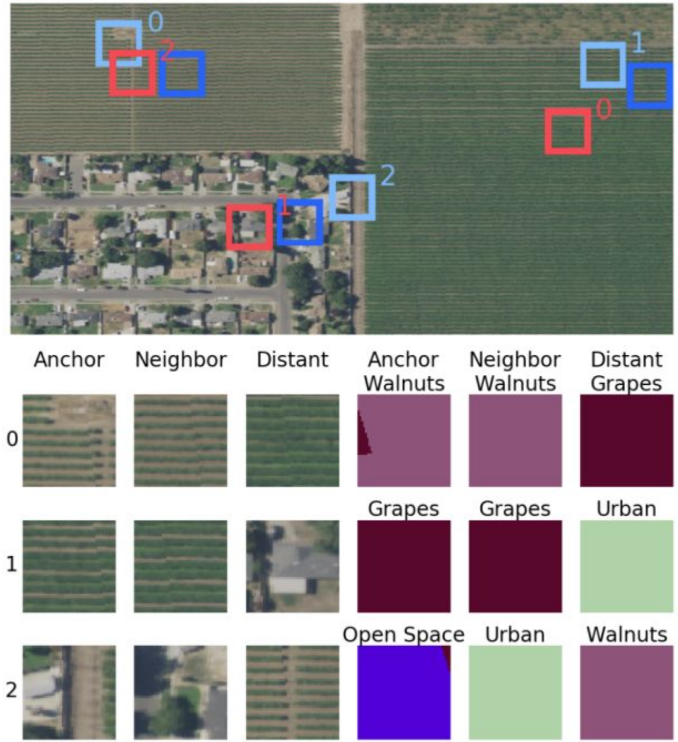
Tobler's First Law of Geography:

"Everything is related to everything else, but near things are more related than distant things."

-- Waldo R. Tobler (Tobler 1970)



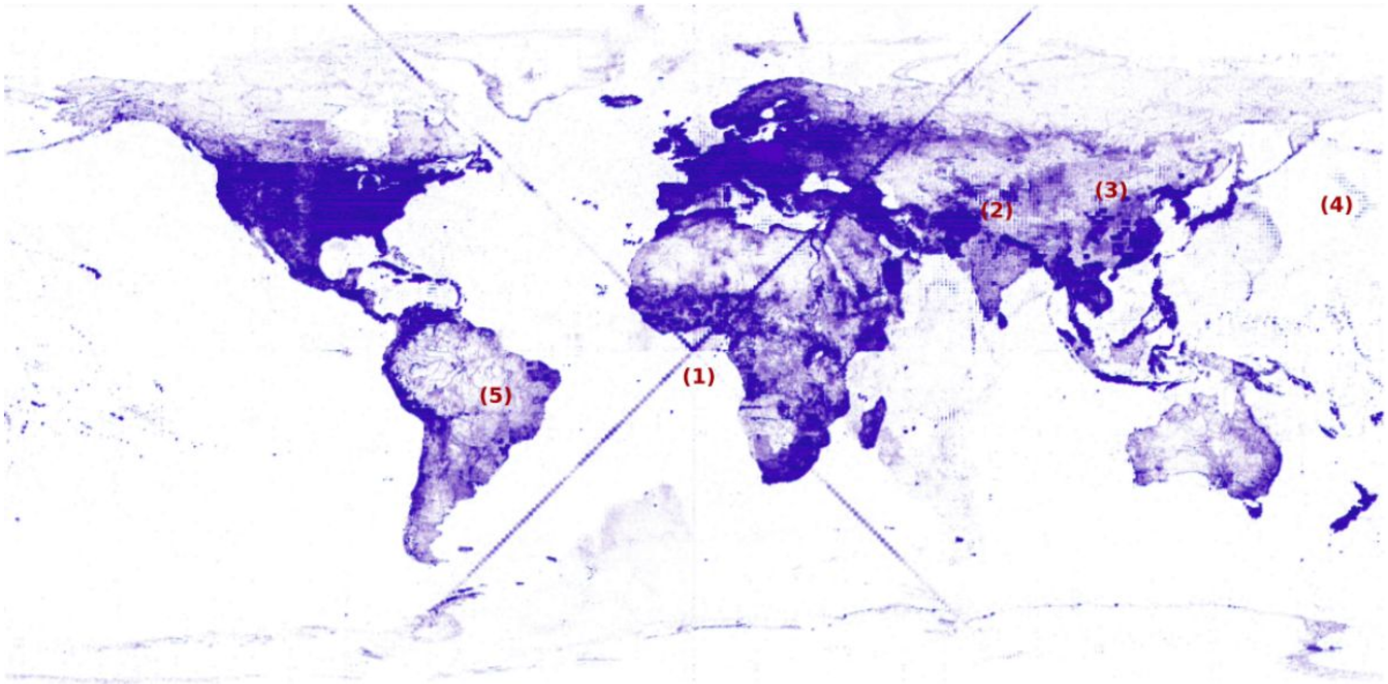
Prof. Waldo R. Tobler, UCSB Geography



Jean, N., et al.. **Tile2vec: Unsupervised representation learning for spatially distributed data**. AAAI 2019.

Why Spatial is Special?

Geographic bias and errors



Janowicz, K., et al.. **Moon landing or safari? a study of systematic errors and their causes in geographic linked data**. GIScience 2016.

Spatially-Explicit Model

A model is said to be **spatially explicit** when it differentiates behaviors and predictions according to spatial location

- **The invariance test**
  - spatially explicit models are not invariant under relocation
- **The representation test**
  - spatially explicit models include spatial representations in their implementations
- **The formulation test**
  - spatially explicit models include spatial concepts in their formula
- **The outcome test**
  - spatial structures of inputs and outcomes are different

Prof. Michael F. Goodchild
UCSB Geography

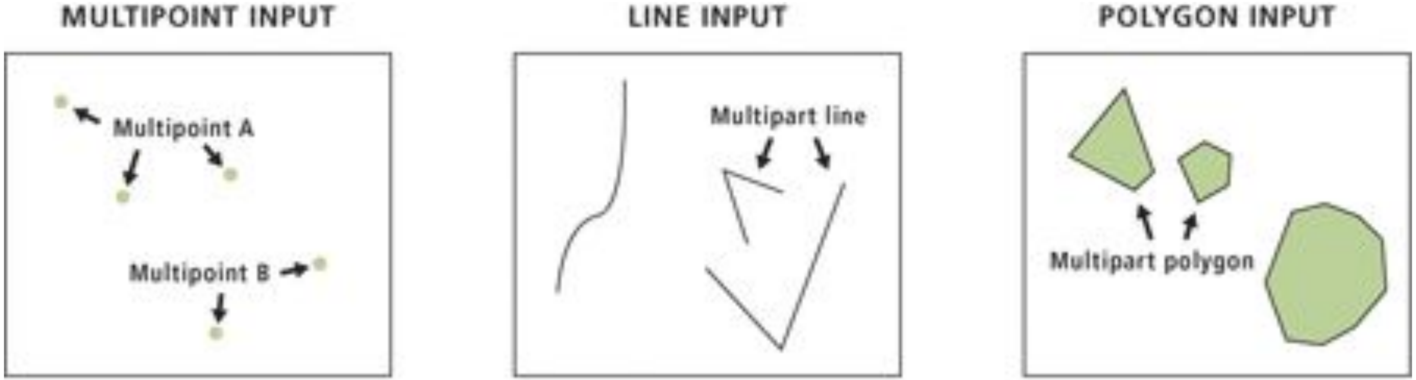
Spatially-Explicit Machine Learning Model

Spatially Explicit Machine Learning Model: Improve the performance of current state-of-the-art machine learning models by using **spatial thinking** and **spatial inductive bias** such as:

- spatial variability
- distance decay effect
- map projection

Represent Spatial Data into the Embedding Space

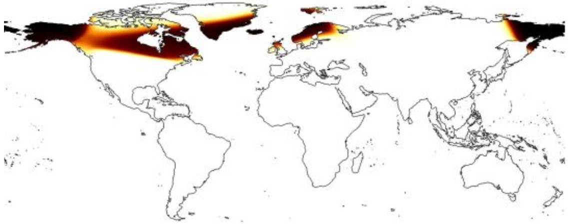
How to represent different types of spatial data into an embedding space?



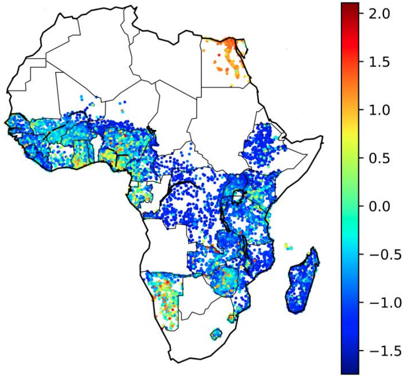
Various Geospatial Tasks

Climate Science:
Precipitation Prediction

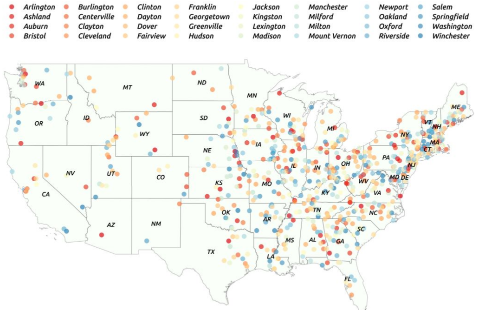
Ecology:
Species Distribution Modeling
(Mac Aodha et al., 2019; Mai et al., 2020)



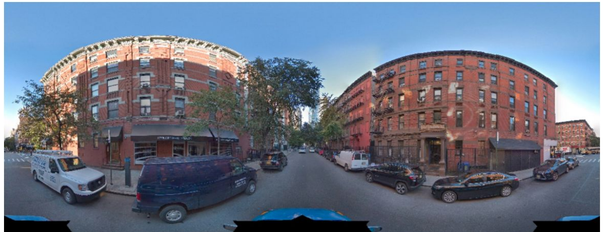
Economy:
Wealth Index Prediction
(Sheehan et al., 2019)



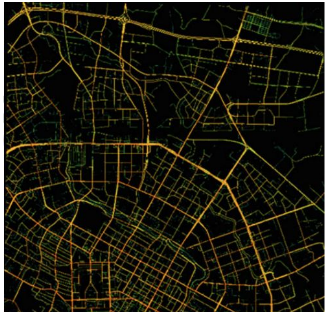
Geospatial Semantics:
Place Name Disambiguation
(Ju et al., 2016)



Smart City:
Indoor/outdoor Navigation
(Mehta et al., 2020)



Human Mobility:
Trajectory Prediction
(Gao et al., 2021)



Problem Statement

Distributed representation of point-features in space:

Given a set of points $\mathcal{P} = \{p_i\}$, i.e., Point Of Interests (POIs), in L-D space (L = 2,3), each point $p_i = (\mathbf{x}_i, \mathbf{v}_i)$ is associated with a location $\mathbf{x}_i$ and attributes $\mathbf{v}_i$ (i.e., POI feature such as type, name). We define function

$$f_{\mathcal{P},\theta}(\mathbf{x}) : \mathbb{R}^L \to \mathbb{R}^d \ (L \ll d)$$

which maps any coordinate x in space to a vector representation of d dimension

Unsupervised Location Encoding

1. Radial Basis Function (RBF)

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$$

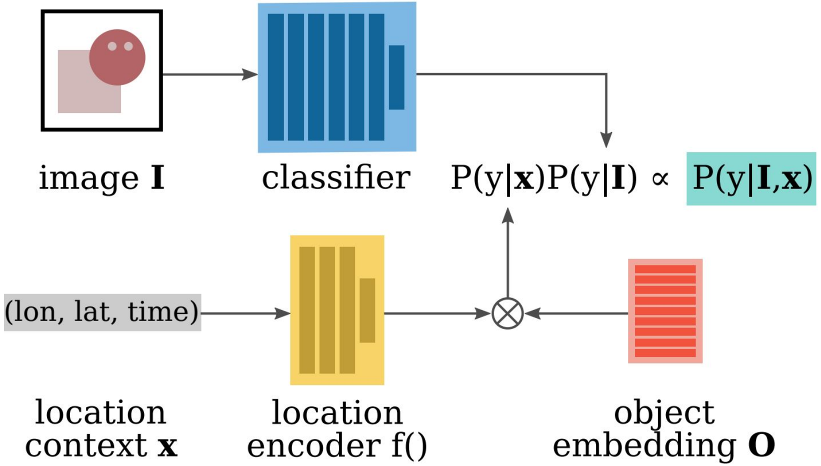
- choosing the correct scale is challenging
- Need to memorize the training samples

2. Tile-based approaches (Berg at al. 2014; Tang et al. 2015): discretize the study area into regular grids

- choosing the correct scale is challenging
- does not scale well in terms of memory

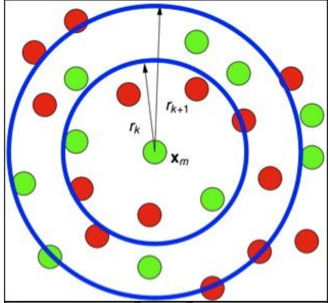
3. Directly feed the coordinates into a FFN (inductive single-scale location encoder)

- hard to capture fine grained distributions



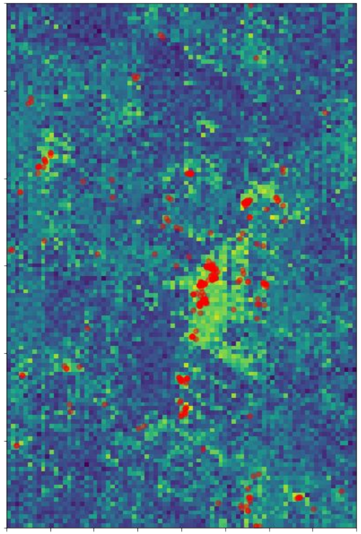
Geo-aware Image Classification (Mac Aodha et al., 2019)

The Key Challenge for Location Encoding

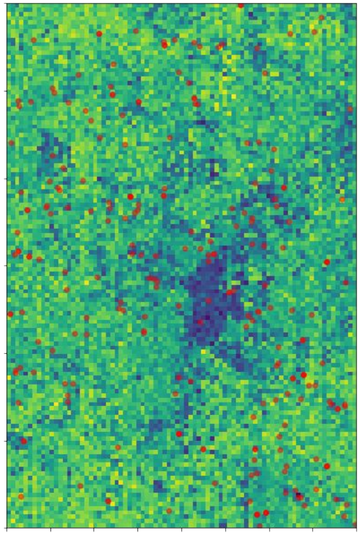


$$K(t) = A\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} I_{t(i,j)}/n^2$$

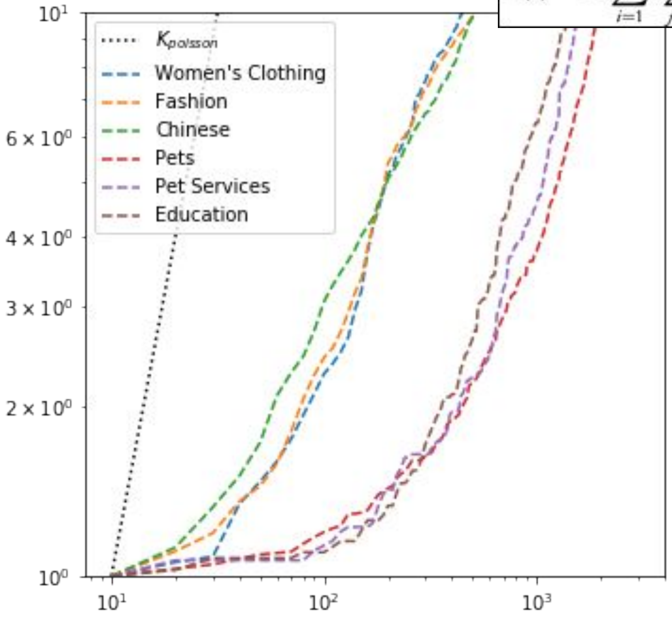
- Joint modeling distributions with very different characteristics
- => **multi-scale location representations**



Women's Clothing
(Clustered Distribution)



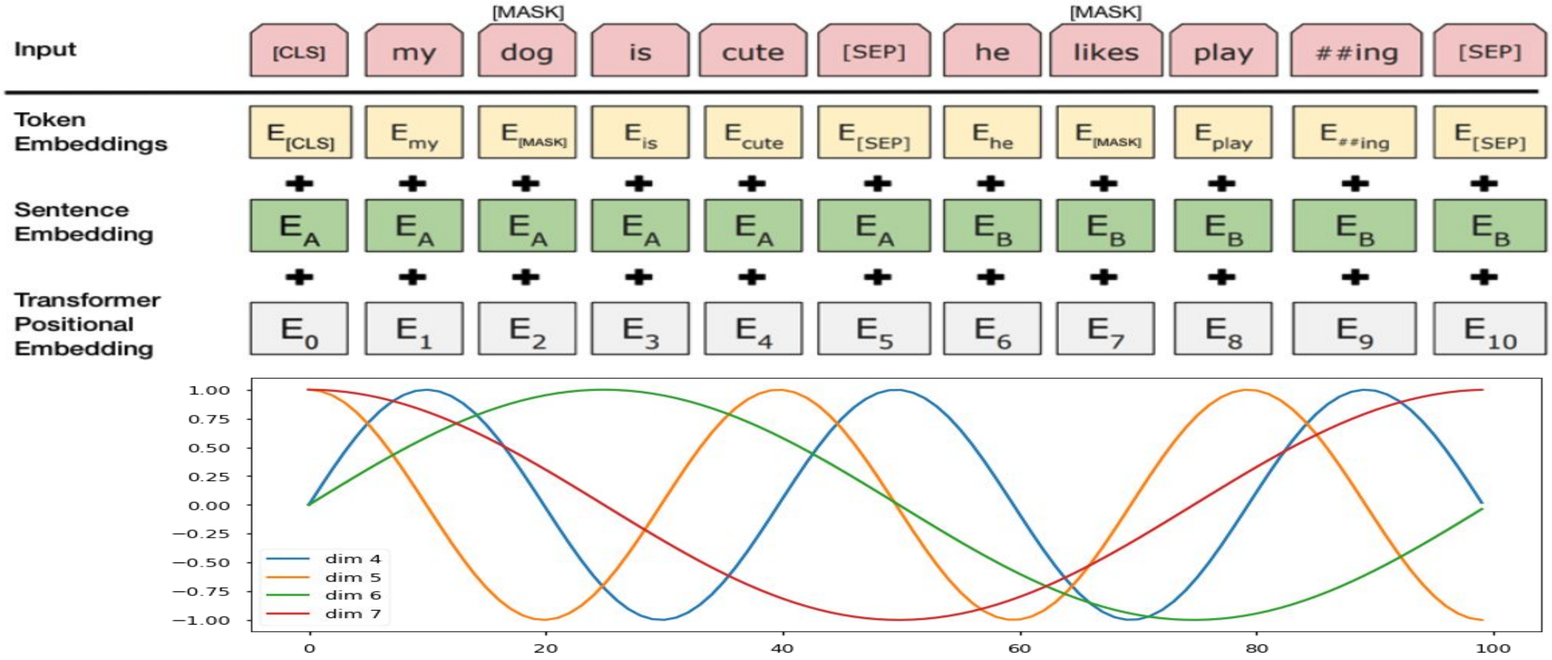
Education
(Even Distribution)



Renormalized Ripley's K
for different POI types

Unsupervised Text Encoding

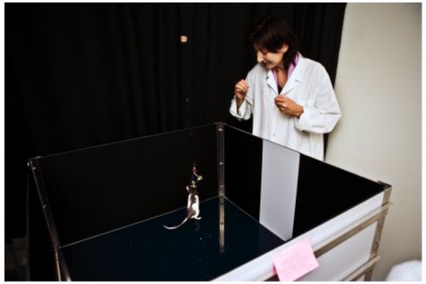
Position Encoding: encode word positions with sinusoid functions of different frequencies



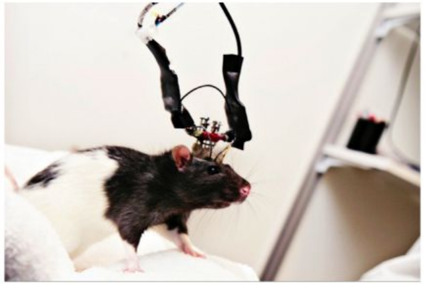
Transformer (Vaswani et al., 2017) BERT (Devlin et al., 2019)

Grid Cell Based Multi-Scale Location Encoding

By recording the **activities of rat neurons** during navigation within a square region, some neurons in its entorhinal cortex have a **hexagonal firing pattern**. (Hafting et al. 2005; Hafting et al. 2005; Fyhn et al. 2008; Yartsev et al. 2011; Killian et al. 2012)

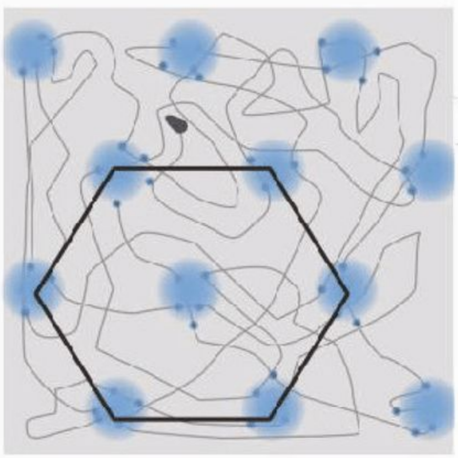


(a)



(b)

Grid cells in mammals provide a **multi-scale periodic representation** that functions as a metric for location encoding. (Banino et al., 2018)

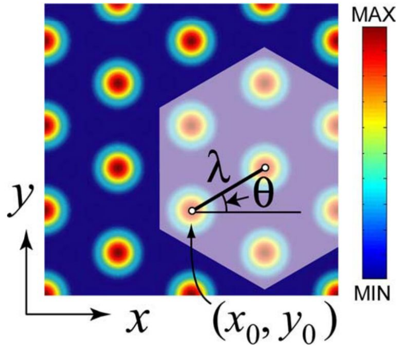


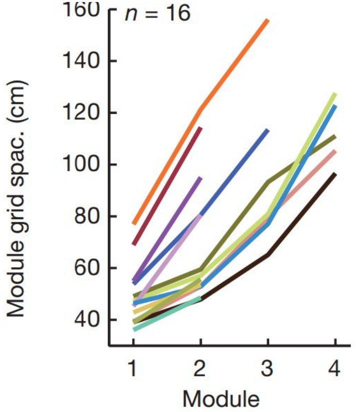
Grid Cell Based Multi-Scale Location Encoding

Grid cell representation can be simulated by summing **three cosine grating functions** oriented 60 degree apart (**a simple Fourier model of the hexagonal lattice**). (Blair et al. 2007)

$$G(\mathbf{r}) = g\left(\sum_{k=1}^{3} \cos(\boldsymbol{\omega}_k \cdot (\mathbf{r} - \mathbf{c}))\right),$$

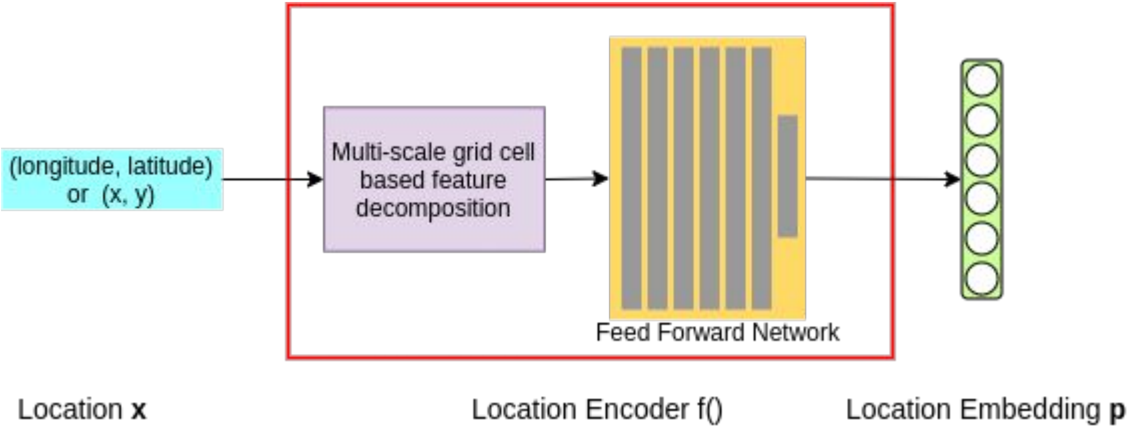
It is more likely that a 2D location **x** is represented by **a population of neurons**, i.e., grid cells, so that these grid cells form **a vector representation** of this location **x.** (Gao et al. 2019)





Mean grid spacing for all modules (M1–M4) in all animals (colour-coded)

Point Space Encoder: Space2Vec



Given a location x:

$$\mathbf{e}[\mathbf{x}] = Enc^{(x)}_{theory}(\mathbf{x}) = \mathbf{NN}(PE^{(t)}(\mathbf{x})) \quad (1)$$

$$\boxed{PE^{(t)}(\mathbf{x}) = [PE^{(t)}_0(\mathbf{x}); ...; PE^{(t)}_s(\mathbf{x}); ...; PE^{(t)}_{S-1}(\mathbf{x})]} \quad (2)$$

$$PE^{(t)}_s(\mathbf{x}) = [PE^{(t)}_{s,1}(\mathbf{x}); PE^{(t)}_{s,2}(\mathbf{x}); PE^{(t)}_{s,3}(\mathbf{x})] \quad (3)$$

$$\boxed{PE^{(t)}_{s,j}(\mathbf{x}) = [\cos(\frac{\langle \mathbf{x}, \mathbf{a}_j \rangle}{\lambda_{min} \cdot g^{s/(S-1)}}); \sin(\frac{\langle \mathbf{x}, \mathbf{a}_j \rangle}{\lambda_{min} \cdot g^{s/(S-1)}})] \forall j = 1, 2, 3;} \quad (4)$$

Point Feature Encoder

Point feature encoder $Enc^{(v)}()$ encodes such features $\mathbf{v}_i$ nto a feature embedding $\mathbf{e}[\mathbf{v}_i] \in \mathbb{R}^{d^{(v)}}$

$\mathbf{e}[\mathbf{v}_i]$

For example, if each point represents a POI with multiple POI types, the feature embedding can simply be the mean of each POI types' embeddings:

$$\mathbf{e}[\mathbf{v}_i] = \frac{1}{H} \sum_{h=1}^{H} \mathbf{t}_h^{(\gamma)}$$

$\mathbf{t}_h^{(\gamma)}$ indicates the hth POI type embedding of a POI pi with H POI types

POI classification - Location Modeling

Location Decoder $Dec_s()$: Directly reconstructs point feature embedding $\mathbf{e}[\mathbf{v}_i]$ given its space embedding $\mathbf{e}[\mathbf{x}_i]$

$$\mathbf{e}[\mathbf{v}_i]' = Dec_s(\mathbf{x}_i; \theta_{\mathrm{dec}_s}) = \mathbf{NN}_{\mathrm{dec}}(\mathbf{e}[\mathbf{x}_i])$$

For training we use inner product to compare the reconstructed feature embedding $\mathbf{e}[\mathbf{v}_i]'$ against the real feature embeddings $\mathbf{e}[\mathbf{v}_i]$ and other negative points

POI classification - Spatial Context Modeling

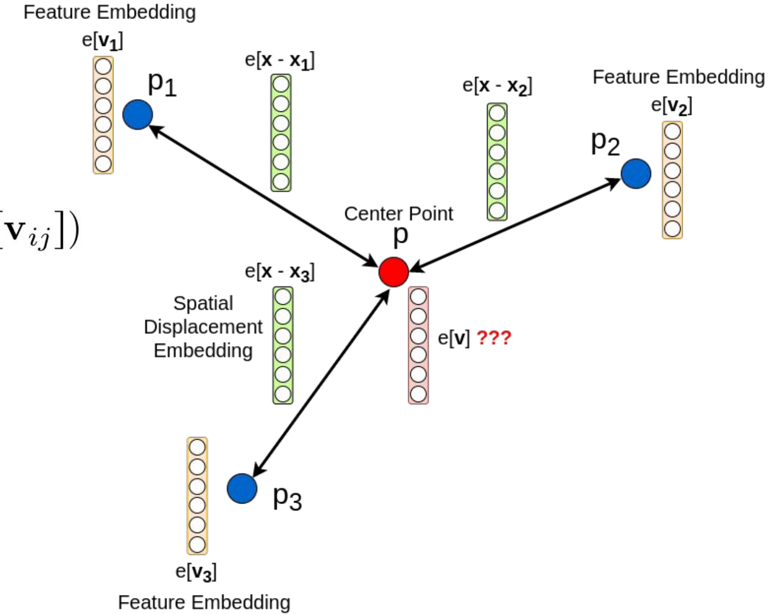
Spatial Context Decoder $Dec_c()$: reconstructs the feature embedding $\mathbf{e}[\mathbf{v}_i]$ of the center point p_i based on the space and feature embeddings $\{\mathbf{e}_{i1}, ..., \mathbf{e}_{ij}, ..., \mathbf{e}_{in}\}$ of n nearby points $\{p_{i1}, ..., p_{ij}, ..., p_{in}\}$

Space-Aware Graph Attention Network Model:

$$\mathbf{e}[\mathbf{v}_i]' = Dec_c(\mathbf{x}_i, \{\mathbf{e}_{i1}, ..., \mathbf{e}_{ij}, ..., \mathbf{e}_{in}\}; \theta_{\text{dec}_c}) = g(\frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{n}\alpha_{ijk}\mathbf{e}[\mathbf{v}_{ij}])$$

$$\alpha_{ijk} = \frac{exp(\sigma_{ijk})}{\sum_{o=1}^{n} exp(\sigma_{iok})}$$

$$\sigma_{ijk} = LeakyReLU(\mathbf{a}_k^T[\mathbf{e}[\mathbf{v}_i]_{init}; \mathbf{e}[\mathbf{v}_{ij}]; \boxed{\mathbf{e}[\mathbf{x}_i - \mathbf{x}_{ij}]}])$$



Unsupervised Training

The unsupervised learning task can simply be maximizing the log likelihood of observing the true point p_i at position $\mathbf{x}_i$ among all the points in P

$$\mathcal{L}_{\mathcal{P}}(\theta) = -\sum_{p_i \in \mathcal{P}} \log P(p_i | p_{i1}, ..., p_{ij}, ..., p_{in}) = -\sum_{p_i \in \mathcal{P}} \log \frac{\exp(\mathbf{e}[\mathbf{v}_i]^T \mathbf{e}[\mathbf{v}_i]')}{\sum_{p_o \in \mathcal{P}} \exp(\mathbf{e}[\mathbf{v}_o]^T \mathbf{e}[\mathbf{v}_i]')}$$

Negative Sampling:

$$\mathcal{L}'_{\mathcal{P}}(\theta) = -\sum_{p_i \in \mathcal{P}} \left(\log \sigma(\mathbf{e}[\mathbf{v}_i]^T \mathbf{e}[\mathbf{v}_i]') + \frac{1}{|\mathcal{N}_i|} \sum_{p_o \in \mathcal{N}_i} \log \sigma(-\mathbf{e}[\mathbf{v}_o]^T \mathbf{e}[\mathbf{v}_i]') \right)$$

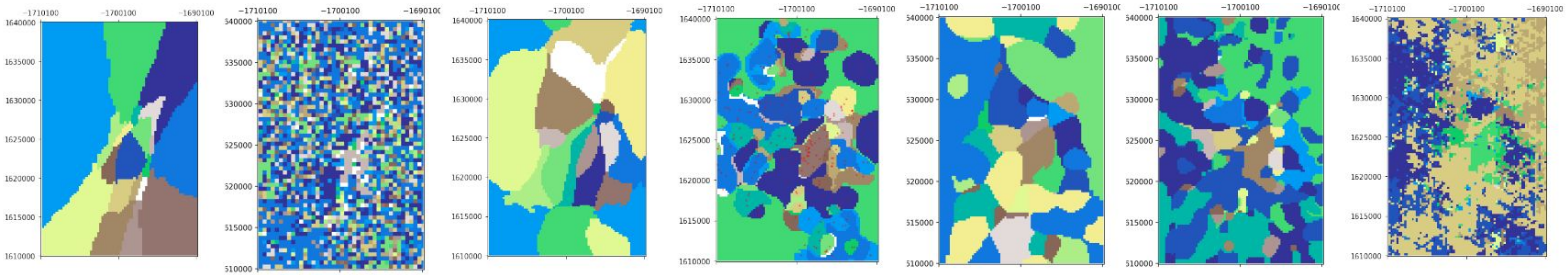
POI classification - Location Modeling Evaluation

Table 1: The evaluation results of different location models on the validation and test dataset.

	Train	Validation			Testing	
	NLL	NLL	MRR	HIT@5	MRR	HIT@5
$random$		-	0.052 (0.002)	4.8 (0.5)	0.051 (0.002)	5.0 (0.5)
$direct$	1.285	1.332	0.089 (0.001)	10.6 (0.2)	0.090 (0.001)	11.3 (0.2)
$tile$ (c=500)	1.118	1.261	0.123 (0.001)	16.8 (0.2)	0.120 (0.001)	17.1 (0.3)
$wrap$(h=3,o=512)	1.222	1.288	0.112 (0.001)	14.6 (0.1)	0.119 (0.001)	15.8 (0.2)
rbf (σ=1k)	1.209	1.279	0.115 (0.001)	15.2 (0.2)	0.123 (0.001)	16.8 (0.3)
$grid$ (λ_{min}=50)	1.156	1.258	0.128 (0.001)	18.1 (0.3)	0.139 (0.001)	**20.0** (0.2)
$hexa$ (λ_{min}=50)	1.230	1.297	0.107 (0.001)	14.0 (0.2)	0.105 (0.001)	14.5 (0.2)
$theorydiag$ (λ_{min}=50)	1.277	1.324	0.094 (0.001)	12.3 (0.3)	0.094 (0.002)	11.2 (0.3)
$theory$ (λ_{min}=1k)	1.207	1.281	0.123 (0.002)	16.3 (0.5)	0.121 (0.001)	16.2 (0.1)
$theory$ (λ_{min}=500)	1.188	1.269	0.132 (0.001)	17.6 (0.3)	0.129 (0.001)	17.7 (0.2)
$theory$ (λ_{min}=50)	1.098	1.249	**0.137** (0.002)	**19.4** (0.1)	**0.144** (0.001)	**20.0** (0.2)

Multi-scale Analysis of Location Modeling

POI Groups	Clustered ($r \leqslant 100m$)	Middle ($100m < r < 200m$)	Even ($r \geqslant 200m$)
direct	0.080 (-0.047)	0.108 (-0.030)	0.084 (-0.047)
wrap	0.106 (-0.021)	0.126 (-0.012)	0.122 (-0.009)
tile	0.108 (-0.019)	0.135 (-0.003)	0.111 (-0.020)
rbf	0.112 (-0.015)	0.136 (-0.002)	0.119 (-0.012)
theory	0.127 (-)	0.138 (-)	0.131 (-)
# POI	16,016	7,443	3,915
Root Types	Restaurants; Shopping; Food; Nightlife; Automotive; Active Life; Arts & Entertainment; Financial Services	Beauty & Spas; Health & Medical; Local Services; Hotels & Travel; Professional Services; Public Services & Government	Home Services; Event Planning & Services; Pets; Education



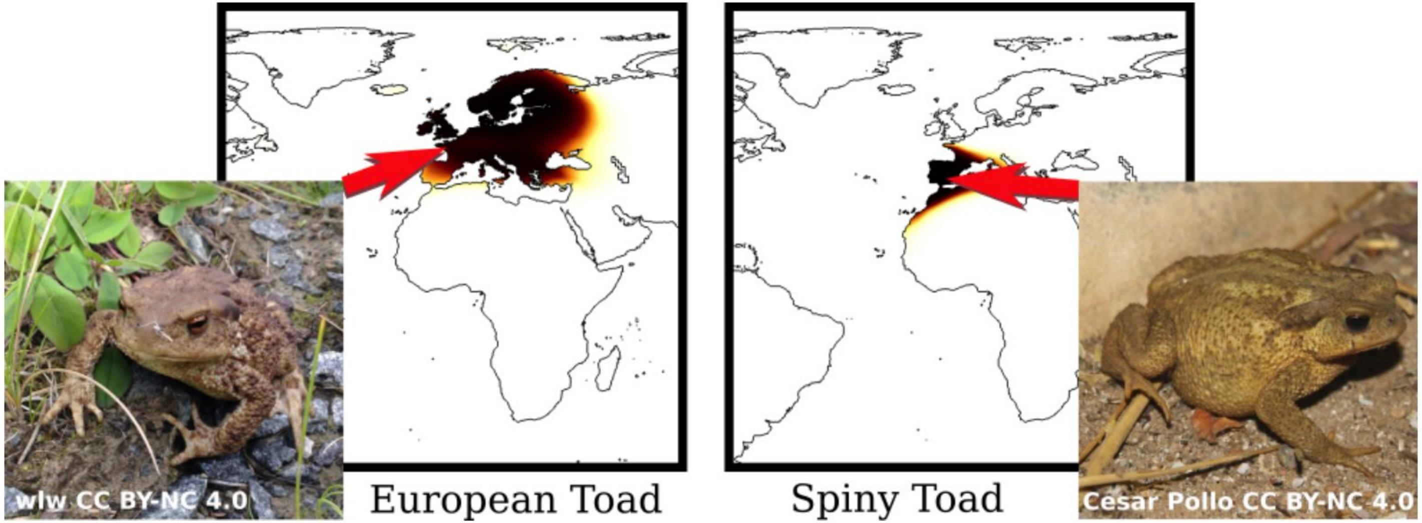
(a) *direct* (b) *tile* (c) *wrap* (d) *rbf* (σ=1k) (e) λ_{min}=1k (f) λ_{min}=500 (g) λ_{min}=50

Spatial Context Modeling Evaluation

Table 3: The evaluation results of different spatial context models on the validation and test dataset. All encoders contains a 1 hidden layer FFN. All grid cell encoders set λ_{min}=10, λ_{max}=10k.

	Train	Validation			Testing	
$Space2Vec$	NLL	NLL	MRR	HIT@5	MRR	HIT@5
$none$	1.163	1.297	0.159 (0.002)	22.4 (0.5)	0.167 (0.006)	23.4 (0.7)
$direct$	1.151	1.282	0.170 (0.002)	24.6 (0.4)	0.175 (0.003)	24.7 (0.5)
$polar$	1.157	1.283	0.176 (0.004)	25.4 (0.4)	0.178 (0.006)	24.9 (0.1)
$tile$ $(c = 50)$	1.163	1.298	0.173 (0.004)	24.0 (0.6)	0.173 (0.001)	23.4 (0.1)
$polar\_tile(S = 64)$	1.161	1.282	0.173 (0.003)	25.0 (0.1)	0.177 (0.001)	24.5 (0.3)
$wrap$ (h=2,o=512)	1.167	1.291	0.159 (0.001)	23.0 (0.1)	0.170 (0.001)	23.9 (0.2)
rbf $(\sigma = 50)$	1.160	1.281	**0.179** (0.002)	25.2 (0.6)	0.172 (0.001)	25.0 (0.1)
$scaled\_rbf$ (σ=40,β=0.1)	1.150	1.272	0.177 (0.002)	**25.7** (0.1)	0.181 (0.001)	25.3 (0.1)
$grid$(λ_{min}=10)	1.172	1.285	0.178 (0.004)	24.9 (0.5)	0.181 (0.001)	25.1 (0.3)
$hexa$ (λ_{min}=10)	1.156	1.289	0.173 (0.002)	24.0 (0.2)	0.183 (0.002)	25.3 (0.2)
$theorydiag$ $(\lambda_{min} = 10)$	1.156	1.287	0.168 (0.001)	24.1 (0.4)	0.174 (0.005)	24.9 (0.1)
$theory$(λ_{min}=200)	1.168	1.295	0.159 (0.001)	23.1 (0.2)	0.170 (0.001)	23.2 (0.2)
$theory$(λ_{min}=50)	1.157	1.275	0.171 (0.001)	24.2 (0.3)	0.173 (0.001)	24.8 (0.4)
$theory$(λ_{min}=10)	1.158	1.280	0.177 (0.003)	25.2 (0.3)	**0.185** (0.002)	**25.7** (0.3)

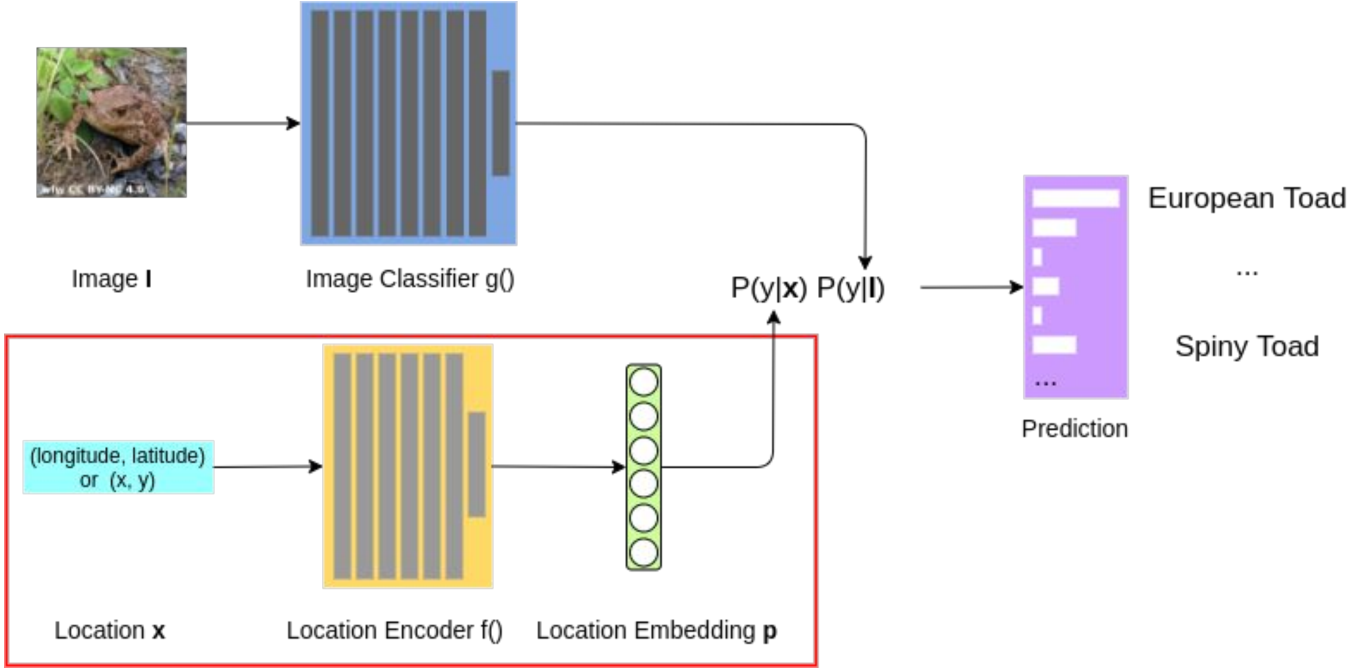
Geo-Aware Image Classification



Species with similar appearance information may have distinct geographic prior distributions. (Figure from Mac Aodha et al., 2019)

Geo-Aware Image Classification

Developing a **geo-aware image classification** model by **fusing** our Space2Vec location encoding model with the state-of-the-art image classification models such as Inception V3 (Szegedy et al. 2016).



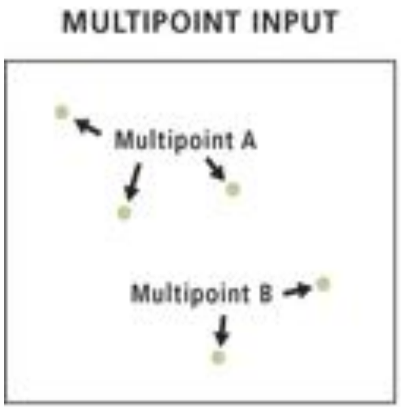
Geo-Aware Image Classification

	BirdSnap†	NABirds†
No Prior (i.e. uniform)	70.07	76.08
Nearest Neighbor (num)	77.76	79.99
Nearest Neighbor (spatial)	77.98	80.79
Adaptive Kernel (Berg et al., 2014)	78.65	81.11
tile (Tang et al., 2015) (location only)	77.19	79.58
wrap (Mac Aodha et al., 2019) (location only)	78.65	81.15
rbf (σ=1k)	78.56	81.13
grid (λ_{min}=0.0001, λ_{max}=360, S = 64)	**79.44**	81.28
theory (λ_{min}=0.0001, λ_{max}=360, S = 64)	79.35	**81.59**

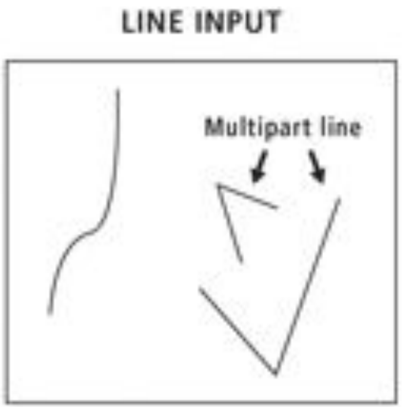
Conclusion for Space2Vec:

- We introduced an encoder-decoder framework as a general-purpose representation model for space inspired by **biological grid cells' multi-scale periodic representations**.
- We show the effectiveness of Space2Vec on two tasks: **POI classification** and **geo-aware image classification**.
- Our analysis reveals that it is the **ability to integrate representations of different scales** that makes the grid cell models outperform other baselines on these two tasks

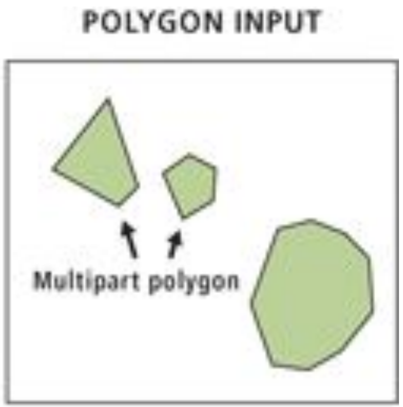
Some Interesting Work



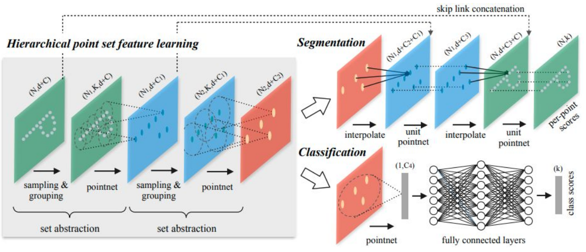
Point



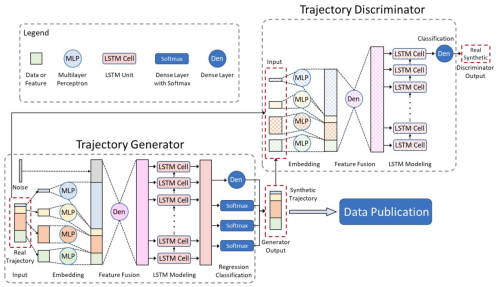
Line



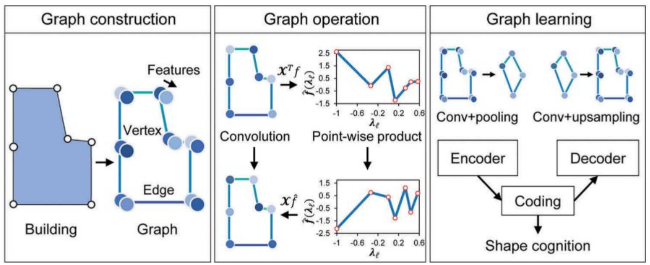
Polygon



PointNet++ (Qi et al., 2017)



LSTM-TrajGAN (Rao et al., 2020)



GCAE (Yan et al., 2021)

Reference

1. **Gengchen Mai**, Krzysztof Janowicz, Ling Cai, Rui Zhu, Blake Regalia, Bo Yan, Meilin Shi, Ni Lao. SE-KGE: A Location-Aware Knowledge Graph Embedding Model for Geographic Question Answering and Spatial Semantic Lifting. *Transactions in GIS*. DOI:10.1111/TGIS.12629 [arxiv paper]
2. **Gengchen Mai**, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao. Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells, In: *Proceedings of International Conference on Learning Representations (ICLR) 2020*, Apr. 26 - 30, 2020, Addis Ababa, ETHIOPIA . [OpenReview paper] [arxiv paper] [code] [video] [slides] * **Spotlight Paper**
3. **Gengchen Mai**, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao. Contextual Graph Attention for Answering Logical Queries over Incomplete Knowledge Graphs, In: *Proceedings of K-CAP 2019*, Nov. 19 - 21, 2019, Marina del Rey, CA, USA. [arxiv]


Thank you!

Acknowledgement

Grants:

University of California, Santa Barbara

SCHMIDT FAMILY FOUNDATION

AI for Earth

NSF

esri

THE SCIENCE OF WHERE

Industry Support:

THE MOONSHOT FACTORY

Alphabet

mosaic

esri

THE SCIENCE OF WHERE

My advisor:

- Prof. Krzysztof Janowicz

Lab Members:

- Dr. Bo Yan
- Dr. Rui Zhu
- Dr. Blake Regalia
- Ling Cai
- Meilin Shi
- ...

Collaborators:

- Dr. Ni Lao, SayMosaic Inc.
- Dr. Wenyun Zuo, Stanford University
- Yao Xuan, UCSB

Reference

1. **Gengchen Mai**, Krzysztof Janowicz, Ling Cai, Rui Zhu, Blake Regalia, Bo Yan, Meilin Shi, Ni Lao. SE-KGE: A Location-Aware Knowledge Graph Embedding Model for Geographic Question Answering and Spatial Semantic Lifting. *Transactions in GIS*. DOI:10.1111/TGIS.12629 [arxiv paper]
2. **Gengchen Mai**, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao. Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells, In: *Proceedings of International Conference on Learning Representations (ICLR) 2020*, Apr. 26 - 30, 2020, Addis Ababa, ETHIOPIA . [OpenReview paper] [arxiv paper] [code] [video] [slides] *** Spotlight Paper**
3. **Gengchen Mai**, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao. Contextual Graph Attention for Answering Logical Queries over Incomplete Knowledge Graphs, In: *Proceedings of K-CAP 2019*, Nov. 19 - 21, 2019, Marina del Rey, CA, USA. [arxiv]
4. **Gengchen Mai**, Bo Yan, Krzysztof Janowicz, Rui Zhu. Relaxing Unanswerable Geographic Questions Using A Spatially Explicit Knowledge Graph Embedding Model, In: *Proceedings of AGILE 2019*, June 17 - 20, 2019, Limassol, Cyprus. *** 1st Best Full Paper Award**
5. Bo Yan, Krzysztof Janowicz, **Gengchen Mai**, Rui Zhu. A Spatially-Explicit Reinforcement Learning Model for Geographic Knowledge Graph Summarization. *Transactions in GIS*, 23(2019), 620-640. DOI:10.1111/tgis.12547
6. Will Hamilton, Payal Bajaj, Marinka Zitnik, Dan Jurafsky, and Jure Leskovec. Embedding logical queries on knowledge graphs. In Advances in Neural Information Processing Systems, pp. 2026-2037. 2018.
7. Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. Knowledge Graph Embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering* 29, no. 12 (2017): 2724-2743.