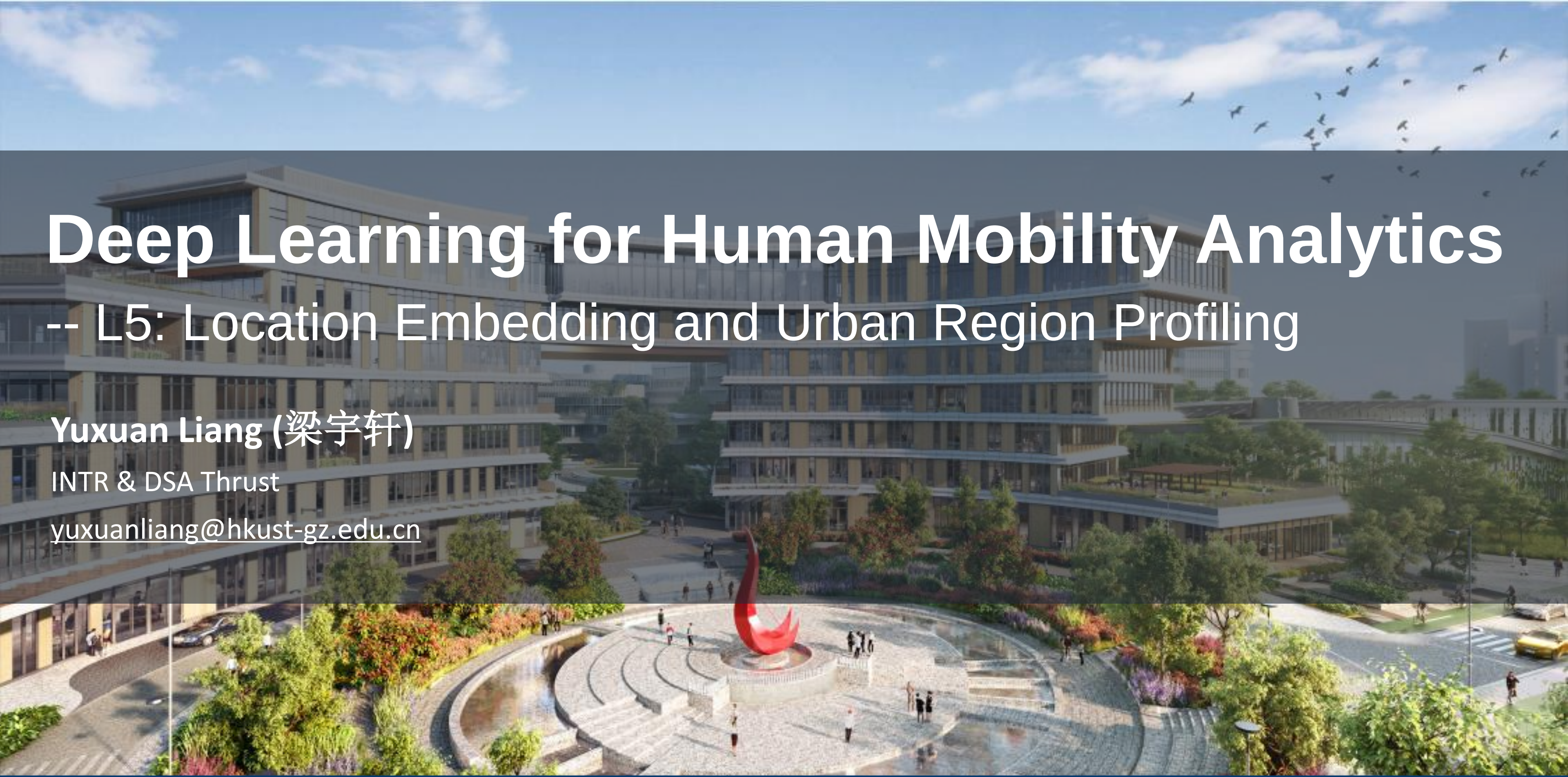


Deep Learning for Human Mobility Analytics

-- L5: Location Embedding and Urban Region Profiling

Yuxuan Liang (梁宇轩)

INTR & DSA Thrust

yuxuanliang@hkust-gz.edu.cn

Notes

- Submit 1-page project plan to yuxuanliang@hkust-gz.edu.cn

	Date	Lecture/Tutorial Topics	Pre	Remarks
1	2 Sep	L1: Overview and Introduction		
2	9 Sep	L2: Tutorials on Deep Learning	✓	Introduce foundations in deep learning techniques, such as CNN, RNN, GNN, etc.
13	16 Sep	Transfer to **an online course for project discussion** (each team 20 mins) in **early November**. Details will be informed in late October.		Mid-Autumn Festival
3	23 Sep	L3: Spatio-Temporal Data Sensing and Management	✓	To form team (2 people)
4	30 Sep	L4: Cross-Domain Data Fusion and Multimodal Learning	✓	
5	14 Oct	L5: Location Embedding and Urban Region Profiling	✓	**Meeting 1**: Submit 1-page plan for project (by **14 Oct @ 23:59**)
6	21 Oct	L6: Learning Spatio-Temporal Trajectory Data	✓	

1-Page Plan Samples

Personalized Vehicle Energy Consumption Prediction and Fuel-Saving Route Recommendation under Multi-View and Multi-Modal Approach

Introduction: With the increasing consumption of energy and environmental problems, energy conservation and emission reduction have become the focus of people's attention. For the transportation industry, developing a method that can predict driving energy consumption and recommend fuel-saving routes is of great significance. In daily car driving, there are three main factors that affect driving energy consumption, including path features, environmental features, and driver's driving characteristics. Among them, path features include the presence or absence of trajectories and POIs; environmental features include factors such as weather, wind direction, shade, temperature, engine conditions, and air conditioning energy consumption that affect ground driving, while driver's driving characteristics refer to the user's driving behavior on the recommended road, such as speed, sudden acceleration, deceleration, braking, and clutching. While it affects the final energy consumption, it is also influenced by path features and environmental features, whereas path and environmental features are obviously not influenced by driver's driving characteristics. Therefore, simply concatenating these three features in one model is inaccurate. A personalized multi-view energy consumption prediction framework should be proposed. Additionally, to improve the energy prediction model, we can use satellite maps as a supplement to the modalities to complete path features, such as road surface conditions, road quality, and whether there are shades on the road. At the same time, the supplement of path and environmental features will directly improve the energy consumption prediction model, and also improve the driver's driving feature prediction model, achieving a win-win effect.

Key points:

1. **Personal driving style prediction model based on user's historical spatial-temporal trajectories**

First, a personal driving style prediction model is trained based on the user's historical spatial-temporal trajectories. The model aims to predict the user's driving behavior on the recommended road, such as speed, sudden acceleration, deceleration, braking, and clutching. The model training strategy is as follows: Input features: time, location, historical trajectory, environmental features (such as regional weather and temperature), and road features. Output results: driving status (such as speed, acceleration, and braking). *Baseline: (https://dl.acm.org/doi/10.1145/3580305.3599767).*

2. **Multi-modal and multi-view road feature extraction**

To obtain comprehensive road features, this study adopts a multi-modal and multi-view approach. Specifically, we extract the following three types of features:

- POI features: including the distribution of surrounding points of interest.
- Basic information features: including road speed limits, lane numbers, and other basic attributes.
- Satellite map features: including potential features such as vegetation, road conditions, and surrounding environment.

3. **Driving energy consumption prediction model based on historical data**

Combining the above features, this study trained a driving energy consumption prediction model based on historical data. The model training strategy is as follows: • Input features: driving behavior, road features, trajectory information, and environmental information. • Output results: energy consumption. *Baseline:(https://link.springer.com/article/10.1007/s11277-018-5495-x).*

4. **Personalized energy consumption prediction model optimization**

To further improve the performance of the energy consumption prediction model, this study uses the Transformer framework and fine-tunes individual users. This helps to improve the model's adaptability to different user driving characteristics.

5. **Fuel-saving route recommendation based on energy consumption model**

Finally, this study conducted low-energy-consumption route recommendation based on the energy consumption prediction model. By incorporating energy consumption factors into path planning, this method can achieve more energy-efficient driving route recommendations. We also compared the proposed method with existing models to verify its performance advantages.

1-Page Plan Samples

Urban Traffic Prediction from Spatio-Temporal Data Using LLMs

Introduction

The increasing urbanization and population growth in cities have led to an increase in traffic congestion, which is a major challenge for urban planners and transportation authorities. In recent years, the availability of spatio-temporal data and the development of large language models have opened up new opportunities for urban traffic prediction. With the popularity of LLMs, people witnessed the potinetial resasoning and generating ability of foundation models in various tasks. Most recent work focus on how to represent spatio-temporal data and design specific netowrk architecture to solve prediction tasks, such as RNNs, CNNs and GCNs.*We aim to try leveraging different modality information in traffic with LLMs, by making specific prompts to finetune LLM, in order to achieve future flow prediction.*

Methodology

We propose a framework that utilizes large language models for traffic flow prediction, and the structure of the framework is shown in Fig. 1. We borrowed the idea of ST-ResNet to categorize historical traffic flow features into temporal closeness, period, and trend. These three types of features are first tokenized and then concatenated with external factors to contruct the prompts. The fine-tuned LLMs then uses the prompts to predict the future traffic flow tokens. These tokens are finally converted to inflow/outflow matrixes through a decoder model.

In our framework, the *Tokenizer* and *Decoder* can be a two-layer MLPs, and these two submodules can be trained in advance following the image reconstruction pipeline. The prompts can be designed as follow:

```
Prompts:

External factors: {Date: 10:00 AM, 2023.10.12; City: BeiJing; Week: Thursday; Weather: Sunny; Holidays: None; Policy: None}
Historical spatio-temporal tokens: {[clossness tokens], [period tokens], [trend tokens]}

Output:
target tokens: [...]
```

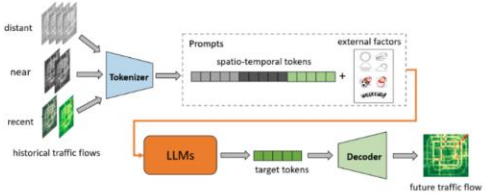


Fig. 1. The architecture of the proposed framework.

Experiment Settings

We follow the dataset settings in ST-ResNet, but only use TaxiBJ datasets. The data preprocessing also as the same in ST-ResNet, which use the Min-Max normalization method to scale the data into the range [-1, 1]. We partition a city into an $I * J$ grid map based on the longitude and latitude where a grid denotes a region, and also introduce the inflow/outflow matrixes, just as that in ST-ResNet.

1-Page Plan Samples

Predicting Urban EV Charging Demand Using Trajectory Data

Introduction

The rise of electric vehicles (EVs) in urban landscapes has increased the need for optimal charging infrastructure. Managing the demand for EV charging stations requires an understanding of spatio-temporal dynamics. Traditional approaches based on statistical methods such as Gaussian processes provide some insight, but may fail to capture complex spatial and temporal correlations. Conversely, modern learning-based models, especially neural networks, provide deeper insights into these correlations, but often fall short in providing uncertainty estimates. In our project, inspired by Spatio-Temporal Graph Neural Processes (STGNP), we aim to integrate the strengths of both traditional and learning-based methods. We'll use trajectory data not only to make accurate predictions, but also to account for the inherent uncertainties. This balanced approach will pave the way for effective and efficient EV infrastructure planning, ensuring both accuracy and reliability.

Literature Review

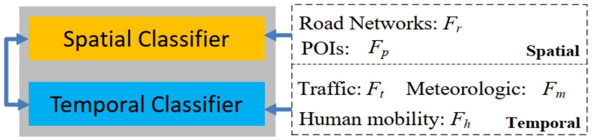
No.	Title	Dataset	Method	Conclusion	Remark
1	Short-term electric vehicle charging demand prediction: A deep learning approach (Link)	**trajectory dataset** containing over 76,000 private EVs in Beijing in January 2018,contained a large number of **charging events** at over 1,000 charging stations	**LSTM**	The **LSTM model outperforms** the two traditional time series prediction models, namely Auto-Regressive Moving Average model (ARIMA), and Multiple Layer Perceptron model (MLP)	Predicted the EV charging demand at the station level for the **short term** next few hours e.g., 1-5 h; dataset confidential.
2	Charging demand prediction in Beijing based on real-world electric vehicle data (Link)	Real-world operation data from 25,489 EVs within three months in Beijing	EV user clustering through **ML**, a trip-chain **simulation** model is constructed	A two-step driving, charging fragment and users clustering model is proposed for comprehensive profile analysis of EV user; A trip-chain-simulation-based charging demand prediction model is proposed to support CS planning	**Combining** simulation (main) with machine learning; dataset confidential.
3	Sensing the Pulse of Urban Refueling Behavior: A Perspective from Taxi Mobility (Link)	4-month trajectories of 32,476 taxicabs, 689 gas stations, and the self-reported refueling details of 8,326 online users	The time estimated part: **context-aware tensor factorization (CATF)**, a factorization model that considers a variety of contextual factors (e.g., price, brand, and weather condition)	Propose **a complete data-driven system for**: detection of individual refueling events (REs) from which refueling preference can be analyzed; estimates of gas station wait times from which recommendations can be made; an indication of overall fuel demand from which macroscale economic decisions can be made, and a spatial, temporal, and economic view of urban refueling characteristics	**Contextual information** is considered by its method, to solve data sparsity problem; **Various data** is used to identify station visits; **Refuel behaviour is somehow similar** to EV charging, only that the time is longer.
4	Electric vehicle charging in smart grid: A spatial-temporal simulation method (Link)	America national household travel survey (NHTS).	Monte Carlo sampling method, Dijkstra's shortest path method, speed-flow model, fuzzy theory	A spatial-temporal simulation method based on the **vehicle-transportation-grid trajectory**, verified by a transportation and distribution coupled test case. Predicted the charging. Two typical days, "workday" and "holiday", were simulated and compared under different EV penetration levels, different trip chain ratio, for verification.	Constructed the vehicle-transportation-grid trajectory for prediction. The prediction is static and did not provide uncertainty estimates.
5	Graph Neural Processes for Spatio-Temporal Extrapolation (Link)	-	-	-	In the last section, it mentioned that the method could be considered for ST data forecasting, and we will further explore it.

Datasets

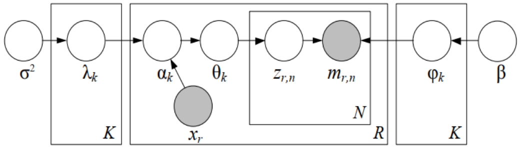
1. GPS trajectory data: TaxiBJ (Link), TaxiBJ21 (Link).
2. Charging demand data (under investigation)

Revisit: Cross-Domain Data Fusion

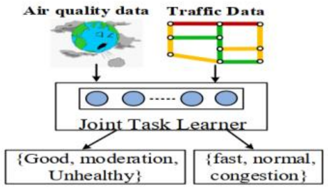
- Stage-based data fusion
- Feature-level-based data fusion
  - Feature concatenation + regularization
  - DNN-based
- Semantic meaning-based fusion
  - Multiple-view-based: like co-training
  - Similarity-based: Coupled matrix factorization
  - PGM-based
  - Transfer learning-based



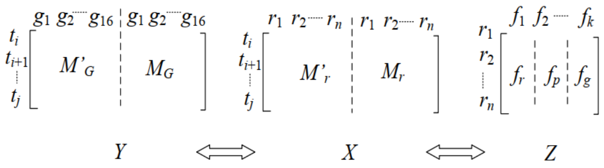
Multi-view learning (Co-training)



Pro. dependency-based (Topic Models)

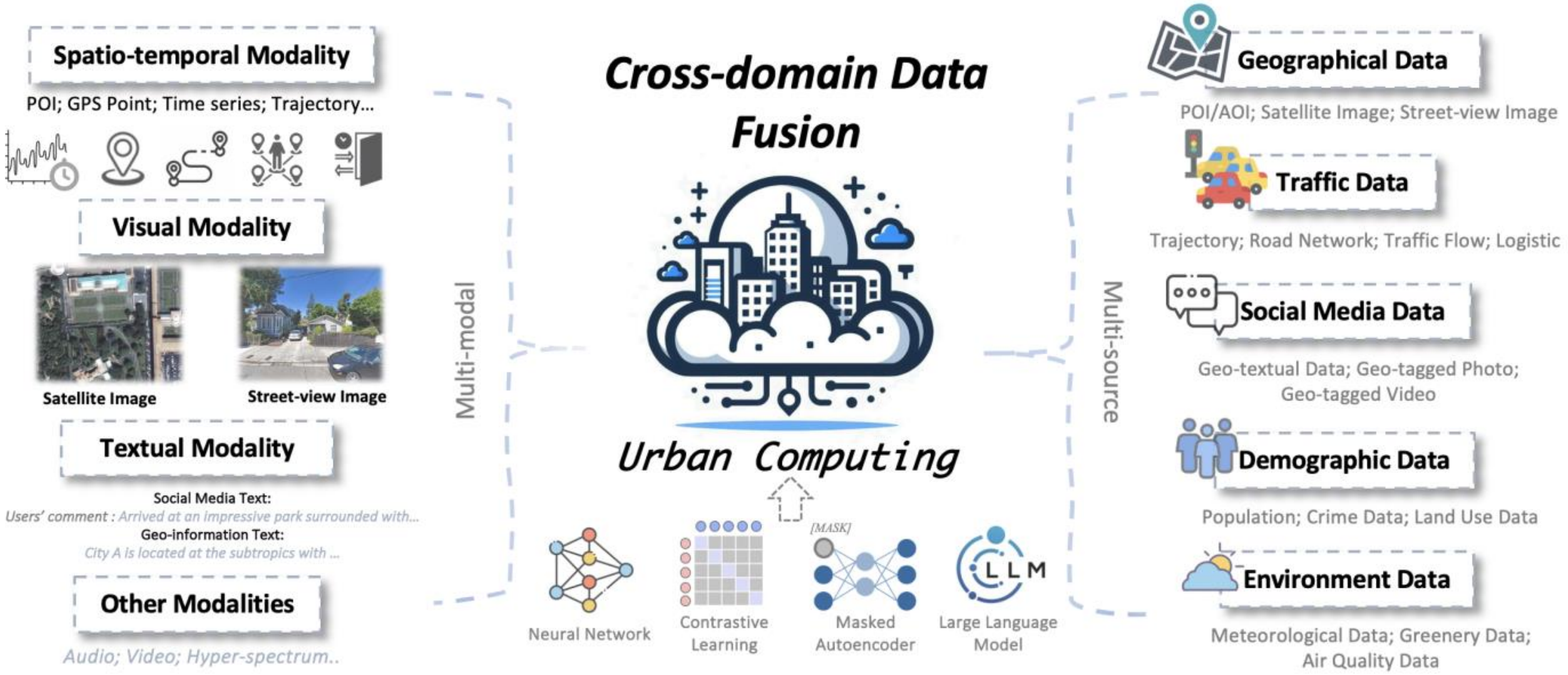


Transfer Learning-based

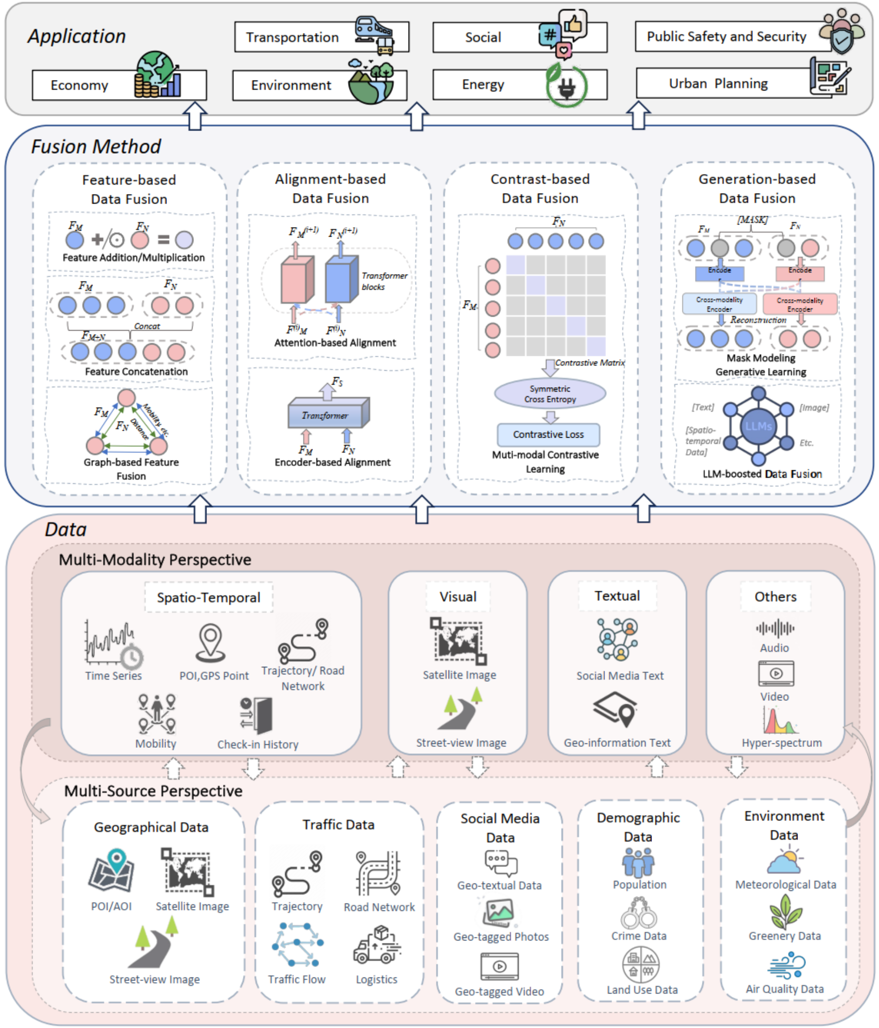


Similarity-based (matrix factorization)

Revisit: Deep Learning-based Fusion



X. Zou, Y. Yan, **Y. Liang*** et al., Deep Learning for Cross-Domain Data Fusion in Urban Computing:Taxonomy, Advances, and Outlook, ongoing

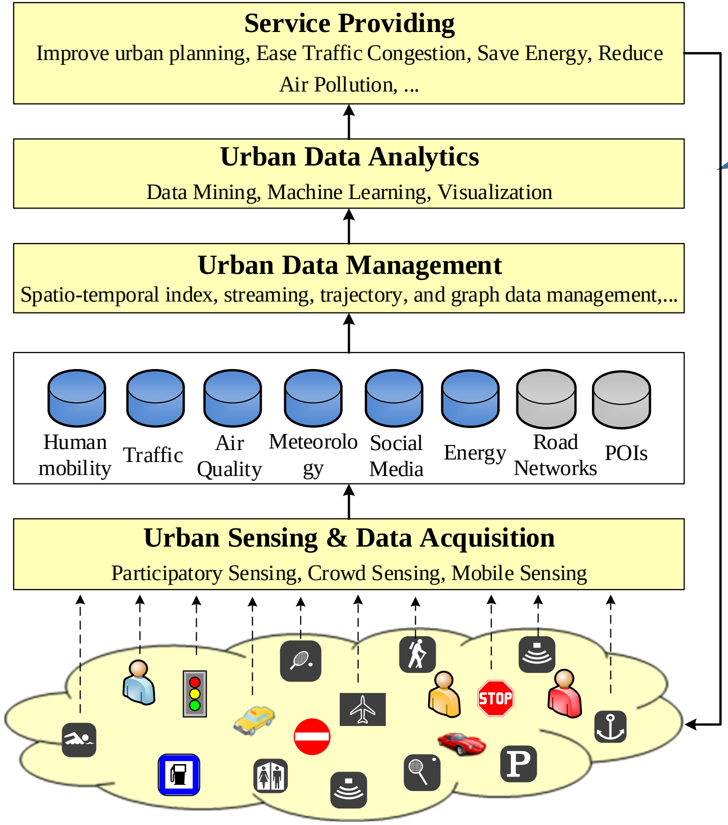

Application
Transportation
Social
Public Safety and Security
Economy
Environment
Energy
Urban Planning
Fusion Method
Feature-based Data Fusion
Feature Addition/Multiplication
Concat
Feature Concatenation
Graph-based Feature Fusion
Alignment-based Data Fusion
Transformer blocks
Attention-based Alignment
Transformer
Encoder-based Alignment
Contrast-based Data Fusion
Contrastive Matrix
Symmetric Cross Entropy
Contrastive Loss
Muti-modal Contrastive Learning
Generation-based Data Fusion
[MASK]
Encode
Encode
Cross-modality Encoder
Cross-modality Encoder
Reconstruction
Mask Modeling Generative Learning
[Text]
[Image]
[Spatio-temporal Data]
Etc.
LLMs
LLM-boosted Data Fusion
Data
Multi-Modality Perspective
Spatio-Temporal
Time Series
POI,GPS Point
Trajectory/ Road Network
Mobility
Check-in History
Visual
Satellite Image
Street-view Image
Textual
Social Media Text
Geo-information Text
Others
Audio
Video
Hyper-spectrum
Multi-Source Perspective
Geographical Data
POI/AOI
Satellite Image
Street-view Image
Traffic Data
Trajectory
Road Network
Traffic Flow
Logistics
Social Media Data
Geo-textual Data
Geo-tagged Photos
Geo-tagged Video
Demographic Data
Population
Crime Data
Land Use Data
Environment Data
Meteorological Data
Greenery Data
Air Quality Data

Objectives of this Course

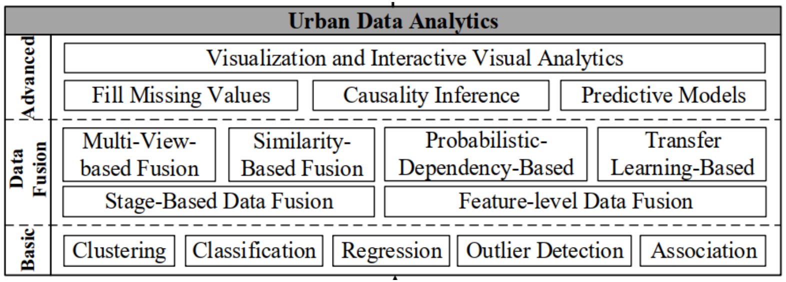
To introduce

- Explicit learning of location embeddings
  - Task-specific supervised learning
  - Self-supervised learning
- Implicit learning of location embeddings
  - Low-rank matrix decomposition
  - Meta learning approaches

3rd Stage: Urban Data Analytics

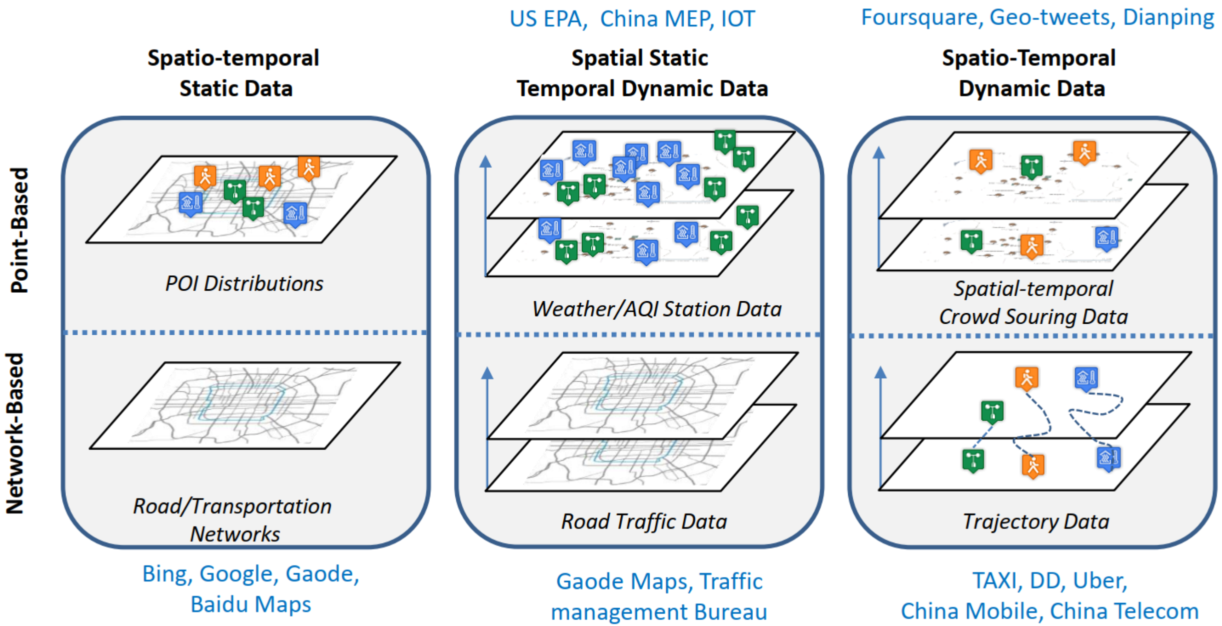


- Texts and images → spatio-temporal data
- A single data source → cross-domain data sources
- Separate data mining algorithms → ML + data management
- Visual and interactive data analytics



Spatio-Temporal Big Data

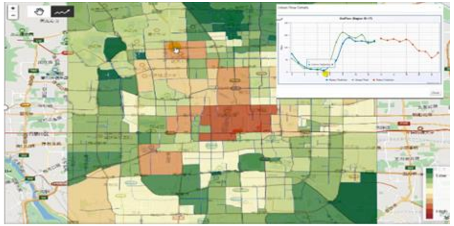
- Data Structures
- Spatio-Temporal (ST) Properties

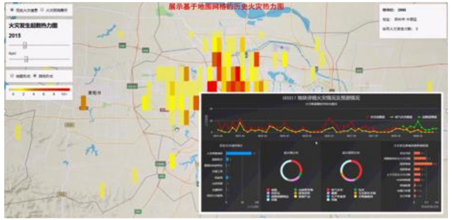


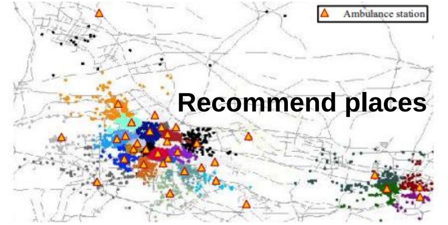
Spatio-Temporal AI

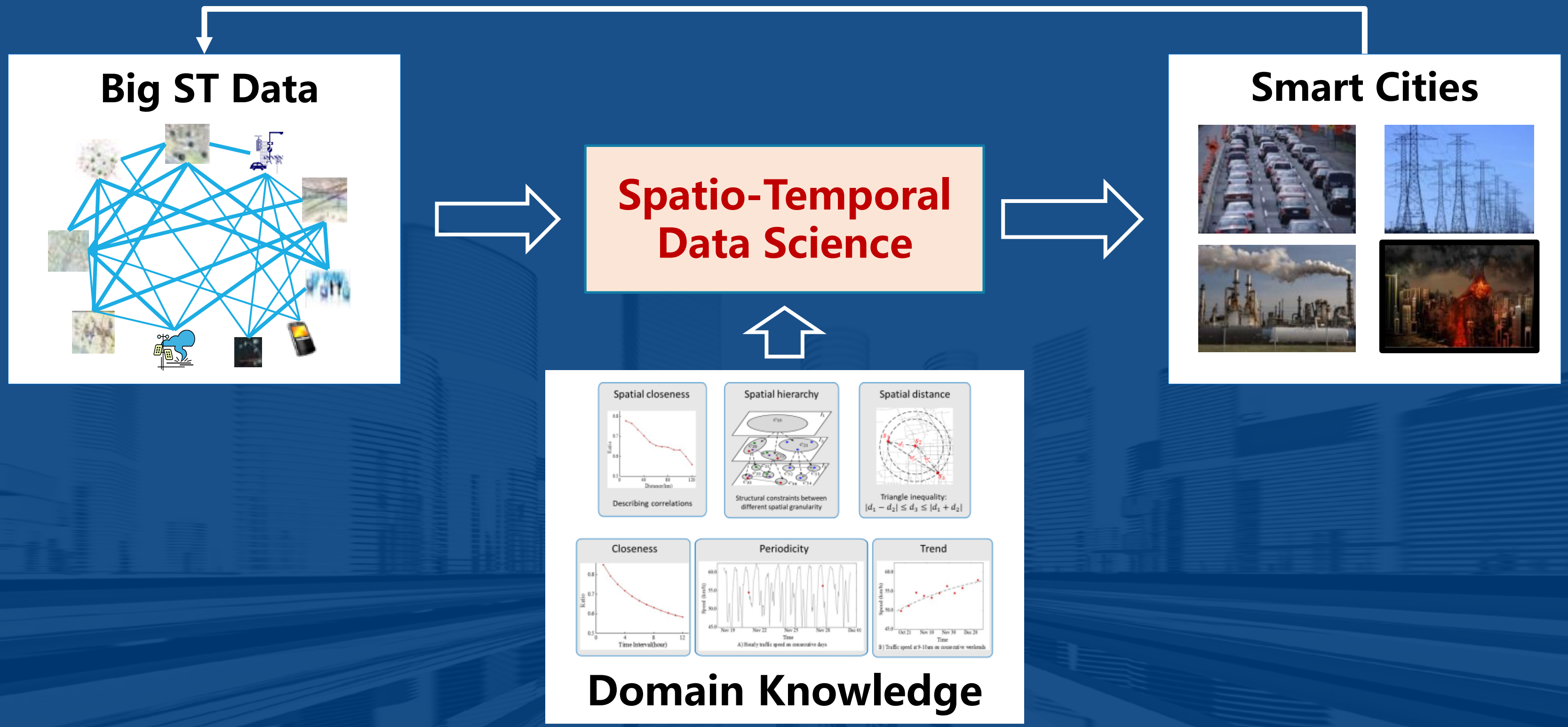
General AI

ST data science (oriented to ST data)








Big ST Data
Spatio-Temporal
Data Science
Smart Cities
Spatial closeness
Describing correlations
Spatial hierarchy
Structural constraints between
different spatial granularity
Spatial distance
Triangle inequality:
$|d_1 - d_2| \le d_3 \le |d_1 + d_2|$
Closeness
Time Interval(hour)
Periodicity
Trend
Domain Knowledge

Modeling Pipeline of Spatio-Temporal AI

Raw ST Data → Data Transformation → **Standard ST Dataset** → Model Selection → **Spatio-Temporal AI Model**

POI

RN

⋮

Traj

A variety of data in ST domains

ST Dataset

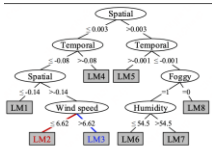
ST Trajectory

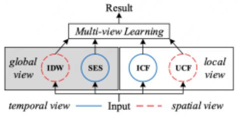
ST Raster Data

ST Graph Data

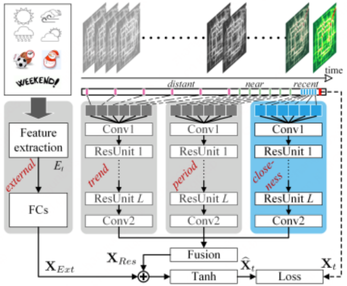
ST Point Data

Feature Engineering + ML



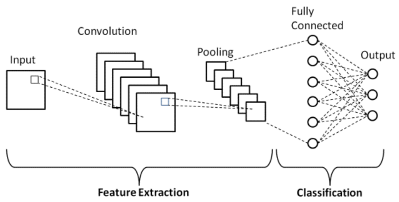


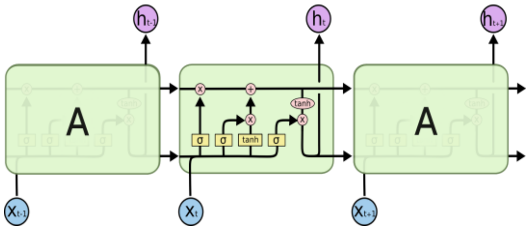
Spatio-Temporal Neural Networks

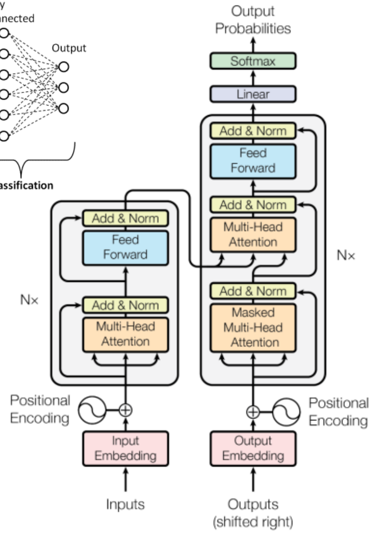


Deep Spatio-Temporal Neural Networks

- Advantages beyond traditional ML methods
  - Larger capacity
  - Automatic feature extraction
  - Friendly to cross-domain data fusion, e.g., considering external factors
- Basic building blocks
  - Multi-layer perceptron (MLP)
  - Convolutional neural networks (CNN)
  - Recurrent neural networks (RNN)
  - Attention models, e.g., Transformers
  - Graph neural networks (GNNs)







Modeling Pipeline of Spatio-Temporal AI

Raw ST Data → Data Transformation → **Standard ST Dataset** → Model Selection → **Spatio-Temporal AI Model** → Training & deploy → **Real-world Application**

POI

RN

⋮

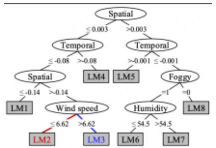
Traj

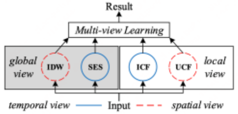
A variety of data in ST domains

ST Dataset

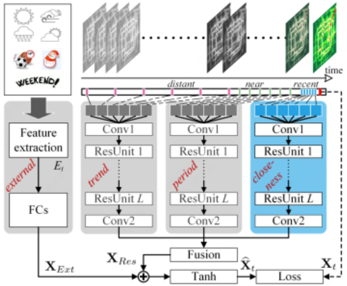
- ST Trajectory
- ST Raster Data
- ST Graph Data
- ST Point Data

Feature Engineering + ML





Spatio-Temporal Neural Networks

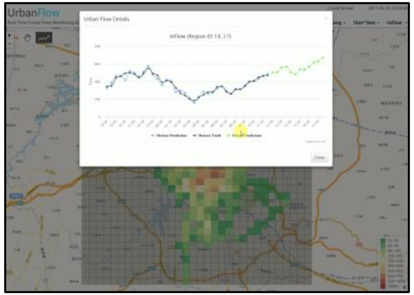


Tasks

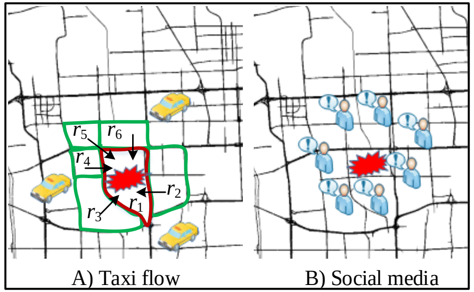
- Prediction
- Anomaly
- Interpolation
- Classification
- Recommendation
- Clustering
- Scheduling
- ⋮

Popular Tasks in Spatio-Temporal AI

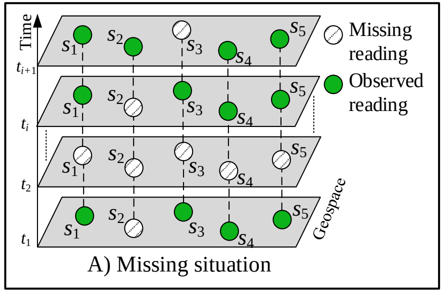
ST Prediction



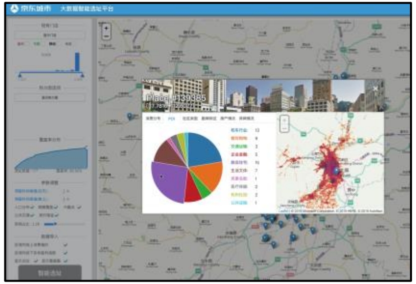
Anomaly Detection



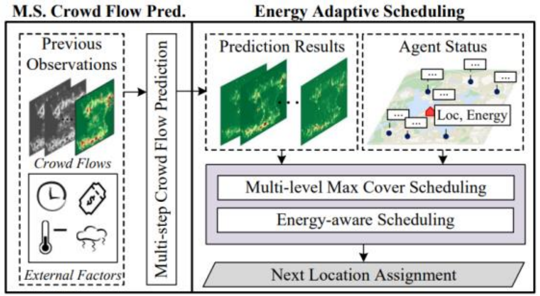
ST Interpolation



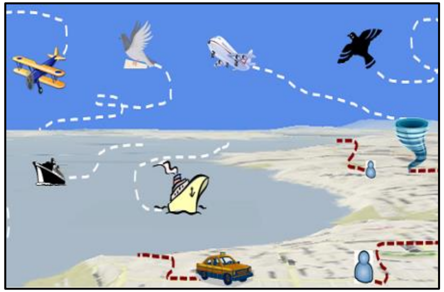
ST Recommendation



Scheduling

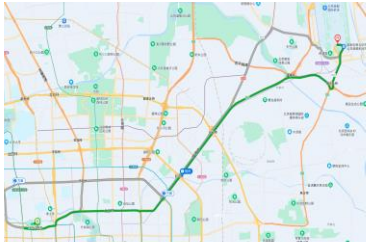


Classification



Our Methodologies & Applications

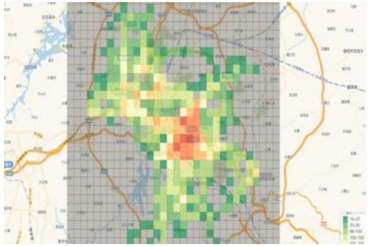
Modeling ST Trajectory



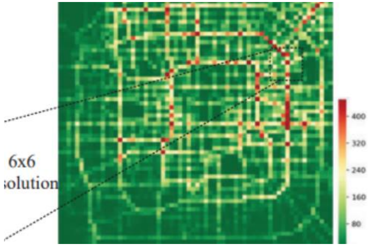
TrajODE [IJCAI’21]

TrajFormer [CIKM’22]

Modeling ST Rasters

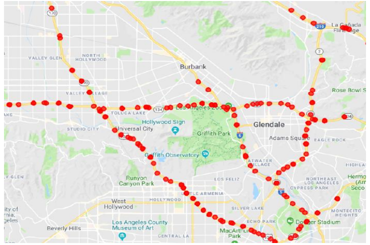


STRN [WWW’21]

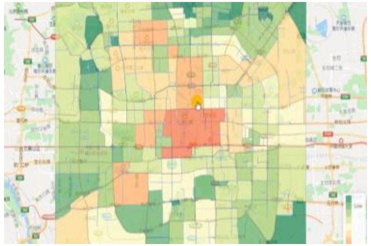


UrbanFM [KDD’19]

Modeling ST Graphs

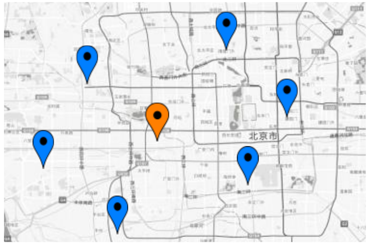


ST-MetaNet [KDD’19]

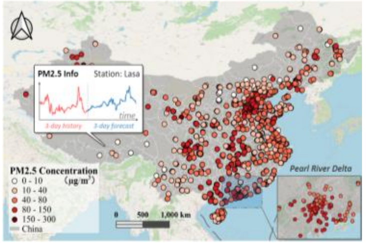


MixRNN [TKDE’22]

Modeling ST Series



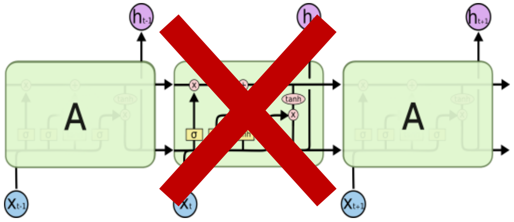
GeoMAN [IJCAI’18]



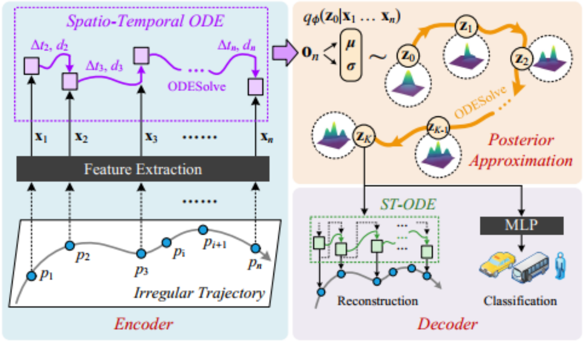
AirFormer [AAAI’23]

Our Exploration on ST Trajectories

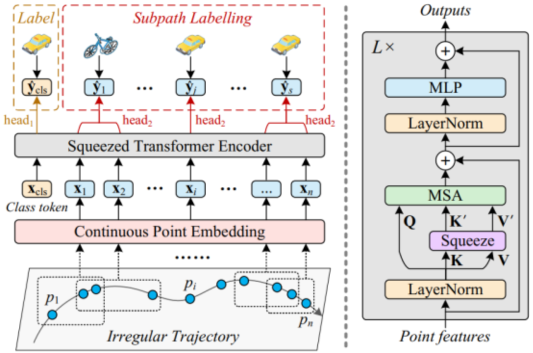
- Existing AI approaches for CV/NLP are NOT always good choices for modeling trajectories
- Capturing the irregularity of trajectories is of great importance to trajectory modeling
- We demonstrate how to encode the domain knowledge (i.e., irregularity) into existing AI methods, including RNNs and Transformers



Classic RNNs



Continuous trajectory modeling

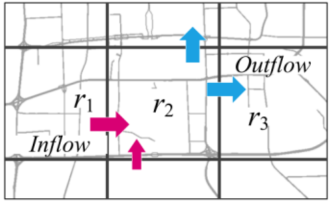


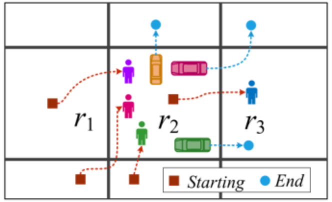
Efficient trajectory modeling

Our Exploration on ST Raster Data

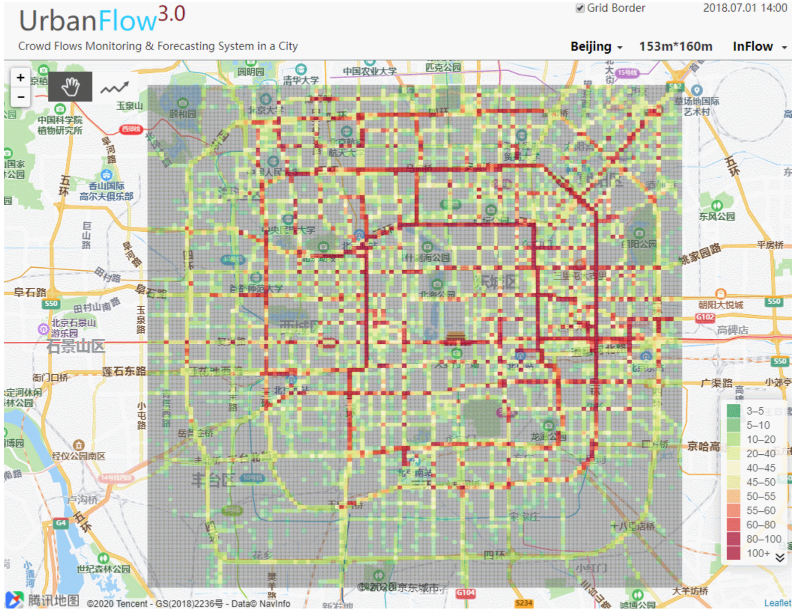
An Example: Grid-based citywide crowd flow prediction

- Predicting the inflow/outflow of every region in several hours





Our UrbanFlow system

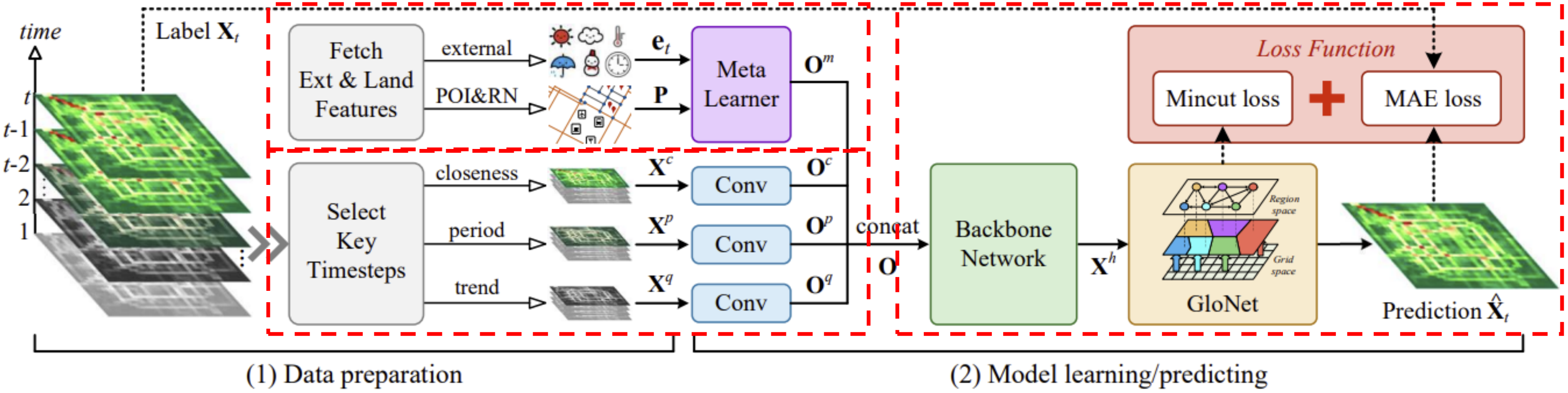


- It can provide insights to
  - Traffic control, risk assessment
- Challenges
  - Complex ST dependencies
  - External factor influence

Y. Liang, R. Zimmermann et al., Fine-grained Urban Flow Prediction. WWW 2021.

Sample Solution: STRN Framework

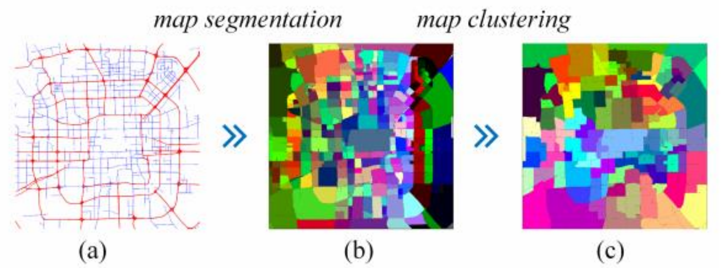
- Spatio-Temporal Relation Network (STRN)
  - Modeling temporal properties: closeness, periodicity, trend
  - Learning the impact of external factors
  - Learning local and global spatial dependencies

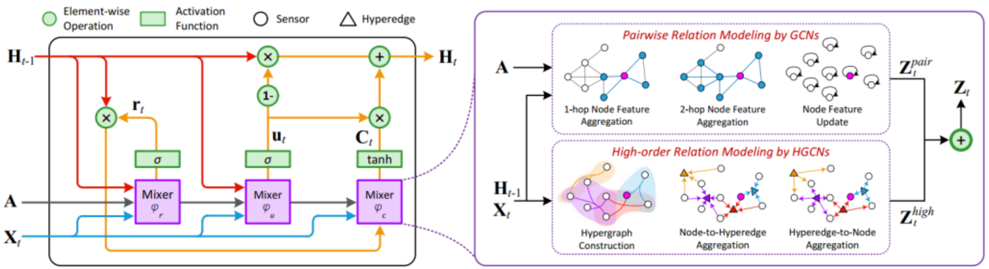


<u>Y. Liang</u>, R. Zimmermann et al., Fine-grained Urban Flow Prediction. WWW 2021.

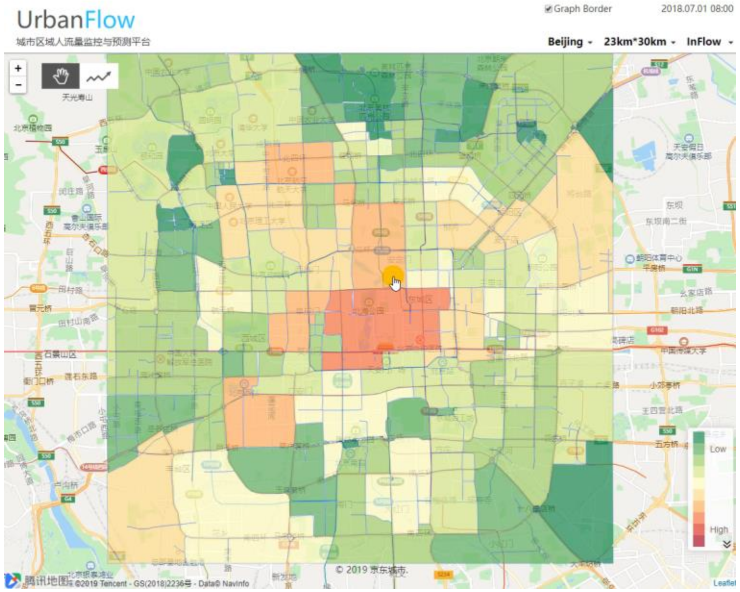
Our Exploration on ST Graphs

- Prior works mainly focused on predicting the crowd flows in regular gridded regions
- We can also predict on irregular regions





(a) Details of MixRNN cell (b) Illustration of Mixer block



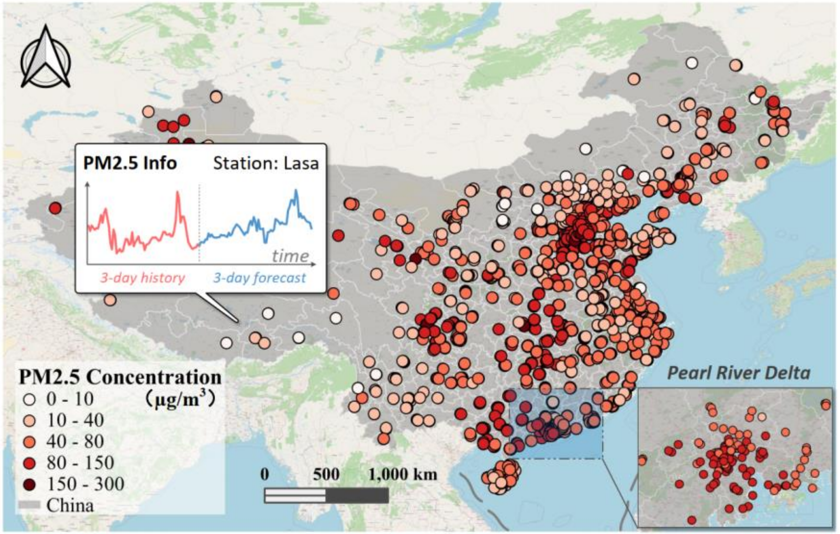
Y. Liang, R. Zimmermann et al., Mixed-Order Relation-Aware Recurrent Neural Networks for Spatio-Temporal Forecasting. TKDE 2022.

Our Exploration on ST Series

An Example: Nationwide Air Quality Prediction in China

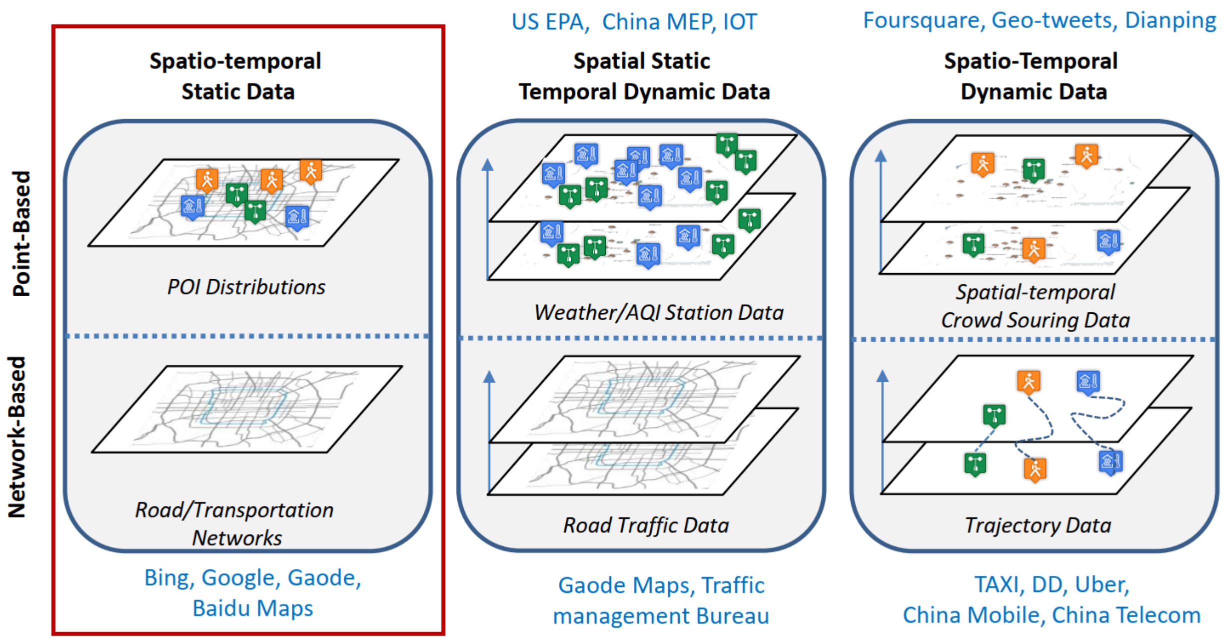
- We present the first attempt to *collectively* predict air quality in the Chinese mainland with an **unprecedented fine spatial granularity**, covering 1,000+ stations.

- Benefits of nationwide prediction
- To capture dynamic spatial correlations
  - Using **self-attention mechanism**
  - Challenge: quadratic complexity w.r.t #locations



Y. Liang, R. Zimmermann et al., AirFormer: Predicting Nationwide Air Quality in China with Transformers. AAAI 2023.

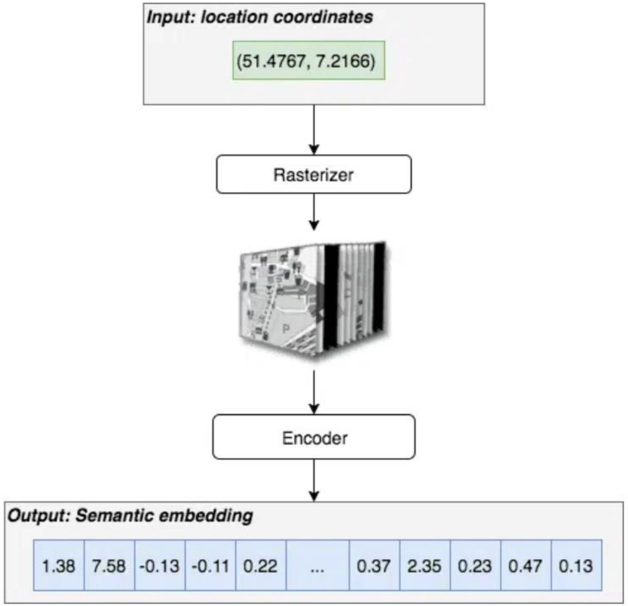
Learning Location Embeddings



What is Location Embedding?

- Location embedding refers to the process of representing a location within a city as a set of features or attributes that capture its physical, social, and cultural characteristics.

- Applications
  - predicting traffic patterns
  - identifying areas at risk of crime
  - recommending personalized services to users



How Location Embedding is Related to Mobility?

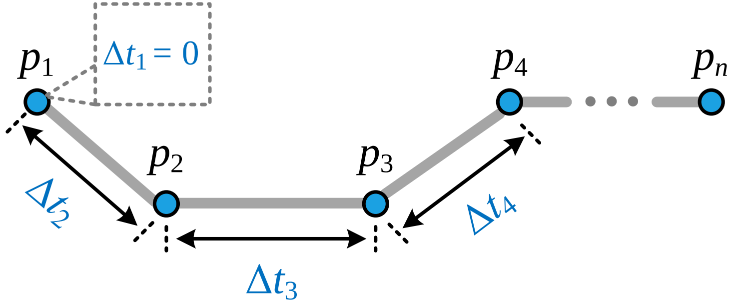
- A human trajectory is a sequence derived from a moving object in geographical spaces, formulated by a series of chronologically ordered points

Timestamp

$$T = p_1 \to p_2 \to \cdots \to p_n, \quad p_i = (\underline{a_i, b_i}, \boxed{t_i})$$

Location (latitude & longitude)

	Latitude,	Longitude,	Time
p_1:	Lat1,	Lng1,	t_1
p_2:	Lat2,	Lng2,	t_2
	……	……	……
p_n:	Latn,	Lngn,	t_n

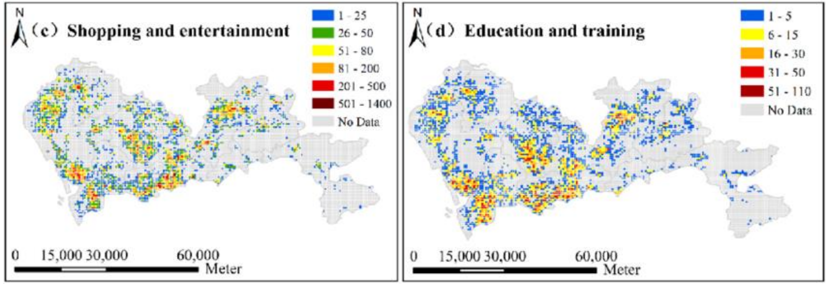


Which Kinds of Data Can be Used?

Satellite images

Street-view images

Road networks



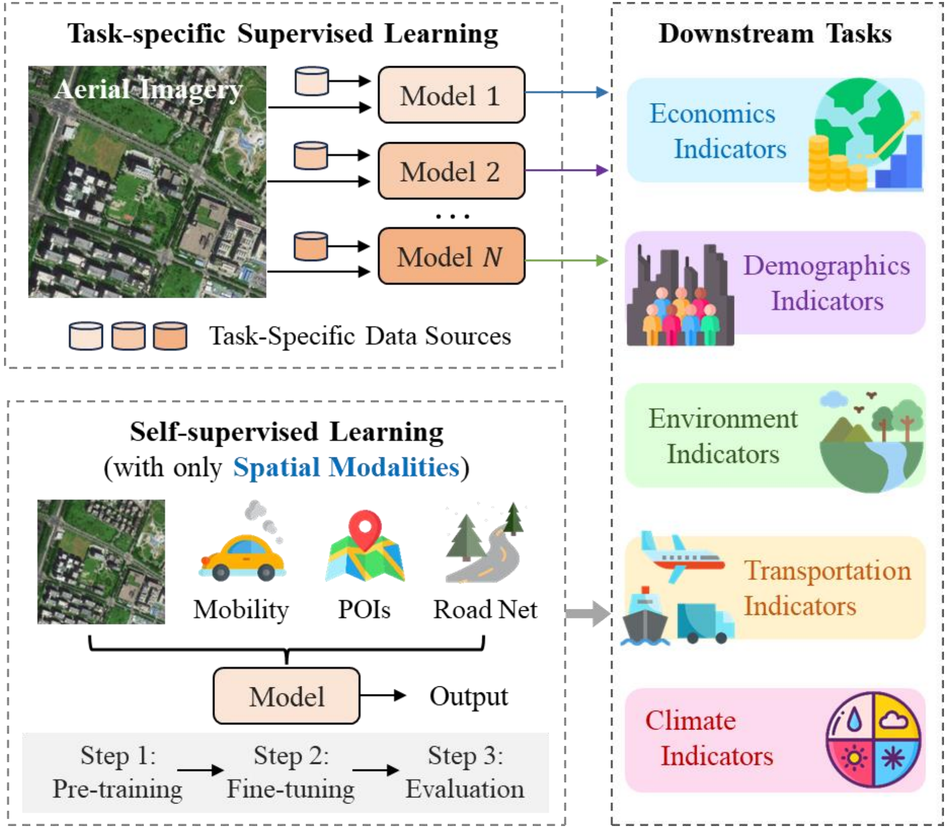
POI information

Mobility data

Outline

- Learning embeddings for locations
  - Explicit learning
    - Task-specific supervised learning
    - Self-supervised learning
  - Implicit learning
    - Low-rank matrix decomposition
    - Meta learning approaches

Explicit Learning

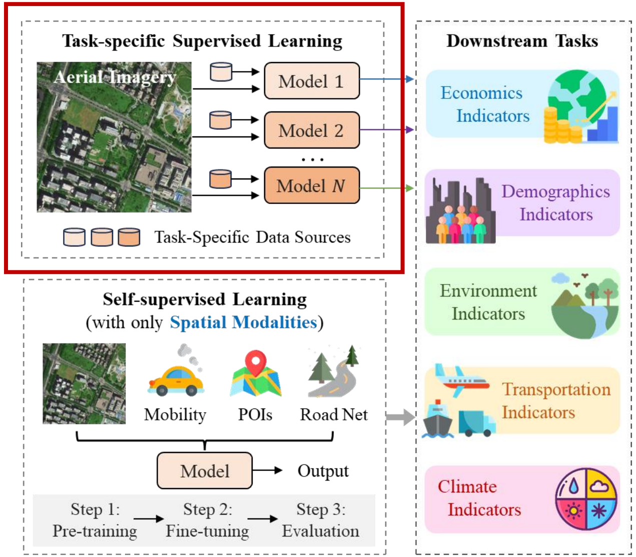


Outline

- Learning embeddings for locations
  - Explicit learning
    - Task-specific supervised learning
    - Self-supervised learning
  - Implicit learning
    - Low-rank matrix decomposition
    - Meta learning approaches

Task-Specific Supervised Learning

- It acquires urban region representations through supervised training using data sources (e.g., satellite imagery) specific to particular tasks
  - Examples: poverty levels, crop yields, population, land cover and commercial activeness
- Drawback
  - Hard to generalize across different tasks
  - Require a large number of training samples



Case 1: Predicting Poverty from Satellite Images

- A multi-task CNN for simultaneously predicting the material of roof, source of lighting and source of drinking water from satellite images

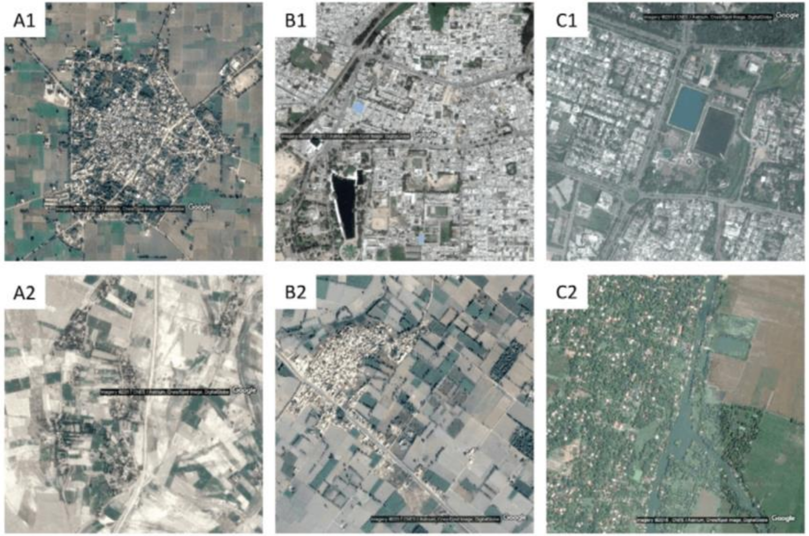
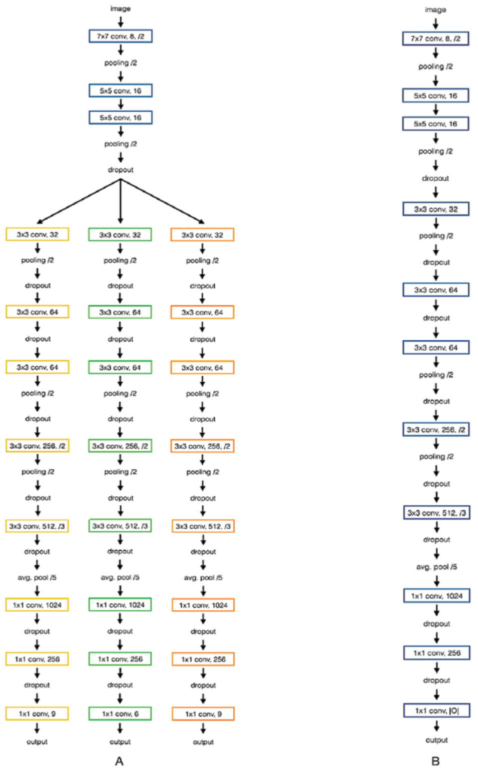


Figure 1: Regions with (A1) concrete roofs and (A2) thatch roofs, (B1) 100% electricity and (B2) 0% electricity for lighting, and (C1) 85.9% households with tap water and (C2) 99.1% with river/canal as drinking water source. Distinct visual features in satellite imagery can be associated with the presence or lack of economic development.

Pandey et al., Multi-Task Deep Learning for Predicting Poverty From Satellite Images. IAAI 2018.

Case 1: Predicting Poverty from Satellite Images

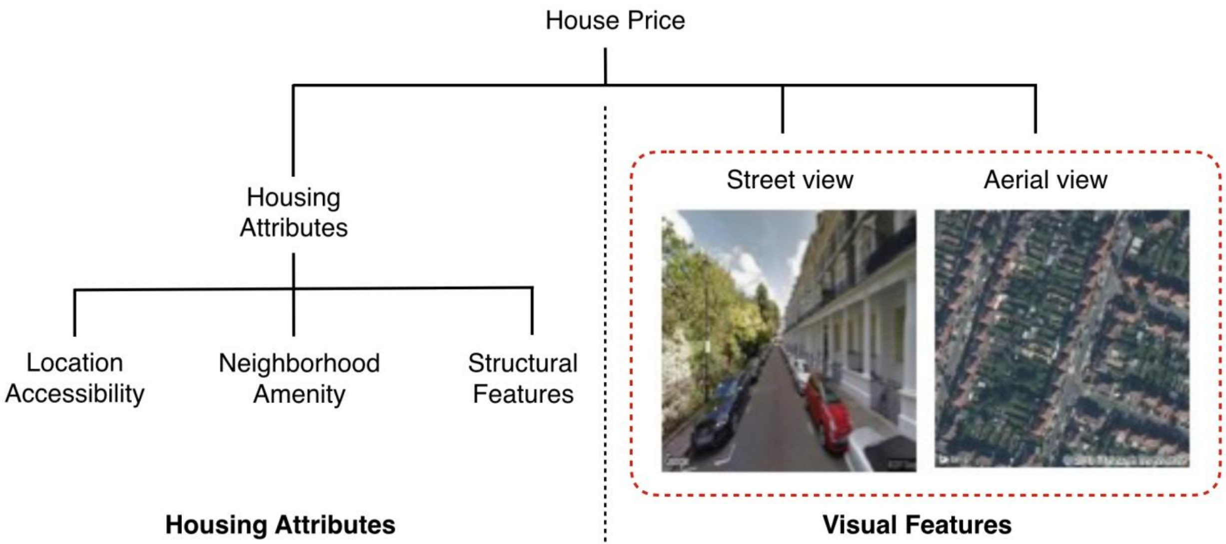
- Firstly, they engineer a multi-task fully convolutional deep network for simultaneously predicting the material of roof, source of lighting and source of drinking water from satellite images.
- Secondly, they use the predicted developmental statistics to estimate poverty.



Pandey et al., Multi-Task Deep Learning for Predicting Poverty From Satellite Images. IAAI 2018.

Case 2: Estimating House Price from Multi-View Images

- The authors show that street image and satellite image data can capture these urban qualities and improve the estimation of house prices



Law et al., Take a Look Around: Using Street View and Satellite Images to Estimate House Prices . TIST 2019.

Case 2: Estimating House Price from Multi-View Images

- Model architectures

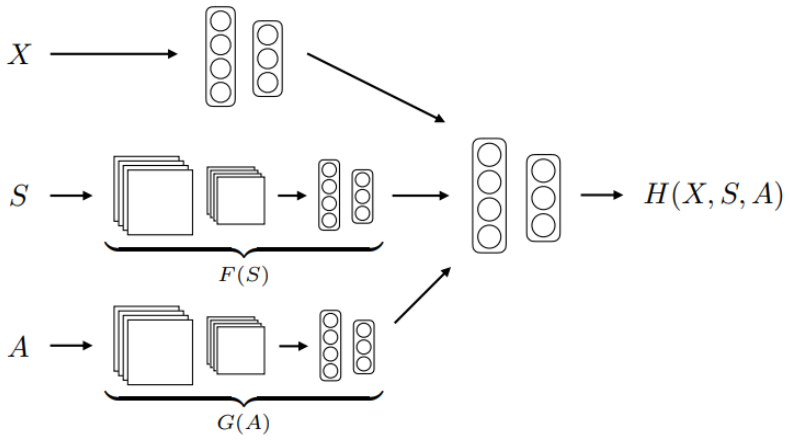


Fig. 6. Fully nonlinear model network structure.

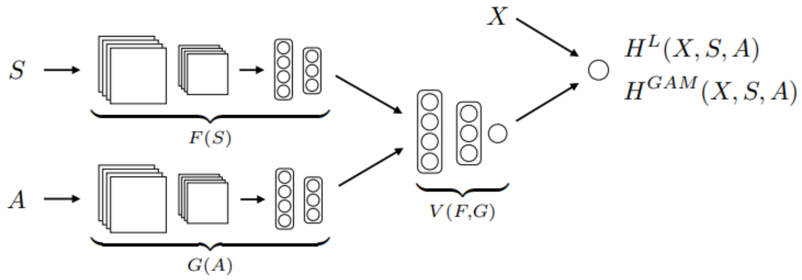


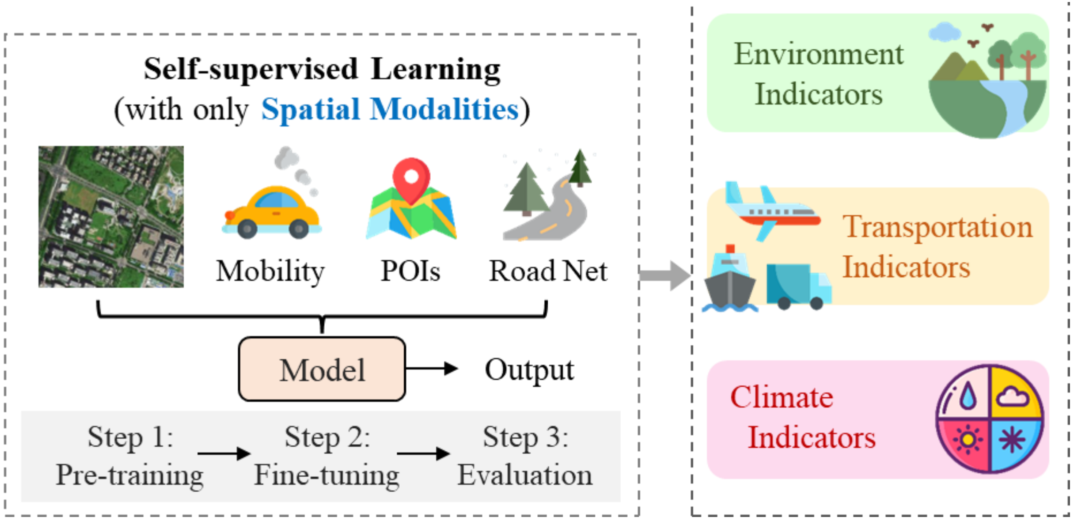
Fig. 7. Semi-interpretable model network structure.

Outline

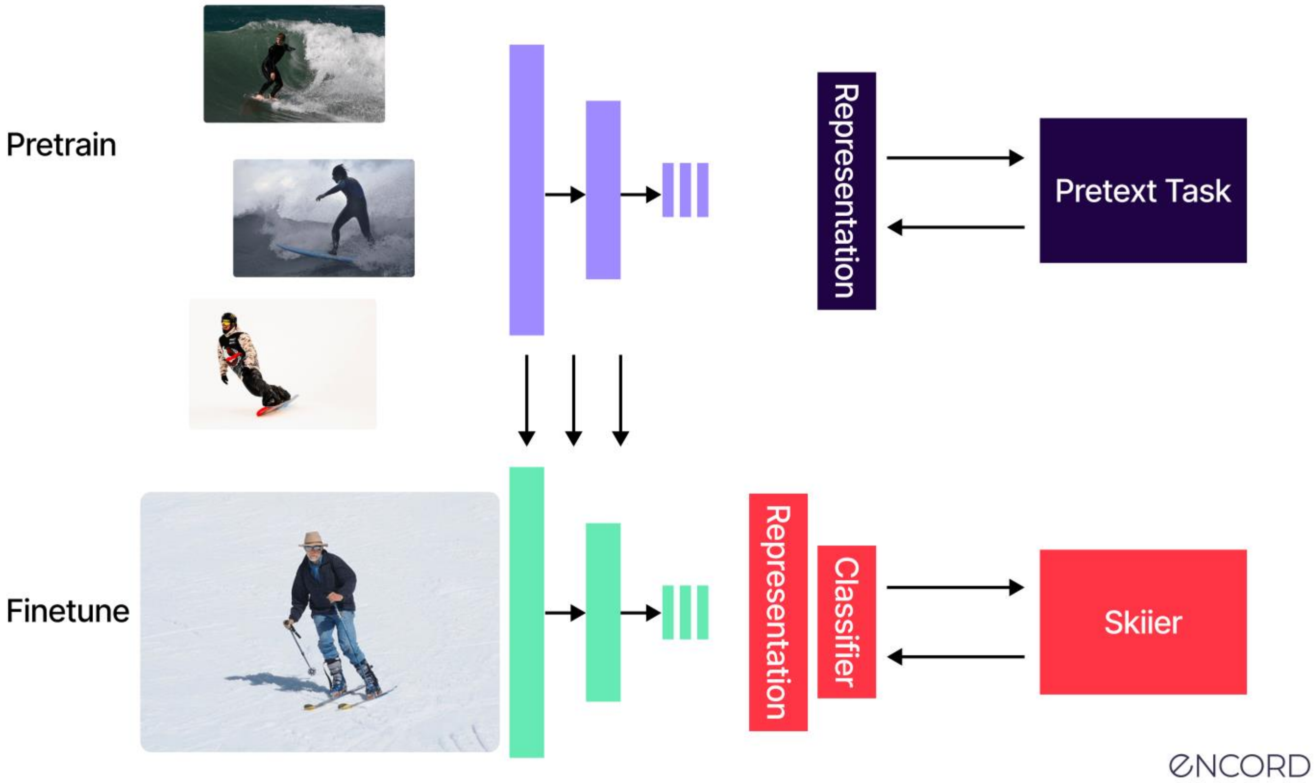
- Learning embeddings for locations
  - Explicit learning
    - Task-specific supervised learning
    - Self-supervised learning
  - Implicit learning
    - Low-rank matrix decomposition
    - Meta learning approaches

Self-Supervised Learning (SSL)

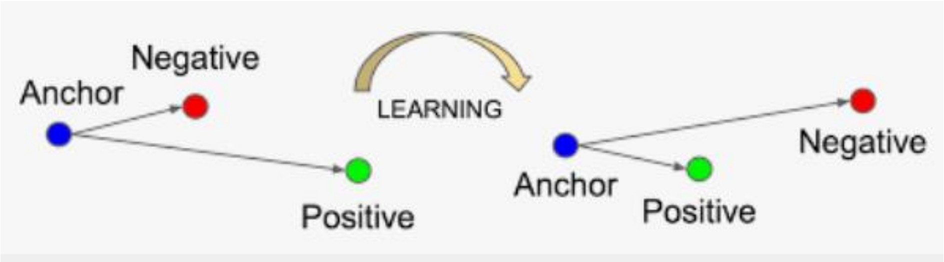
- SSL integrates diverse auxiliary spatial modalities to generate comprehensive urban representations
- These representations boast wide applicability, readily generalizing across numerous urban indicator tasks,

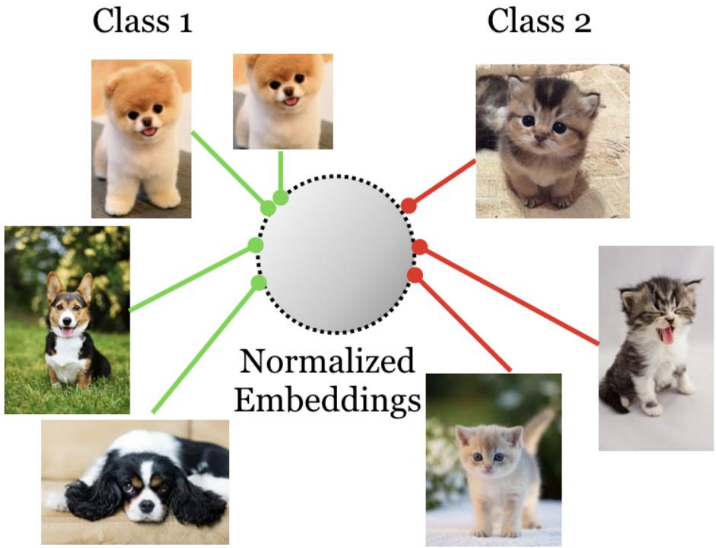


Self-Supervised Learning (SSL)



Contrastive Learning

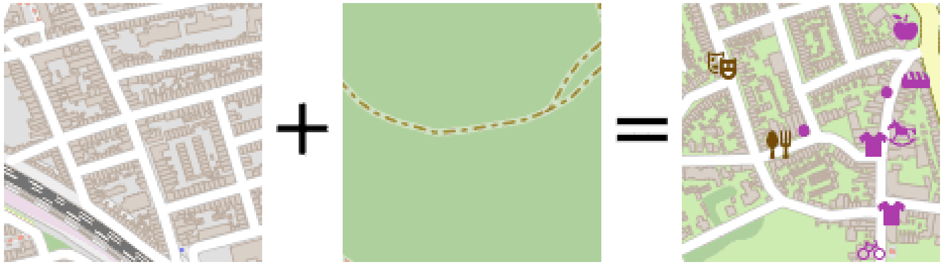


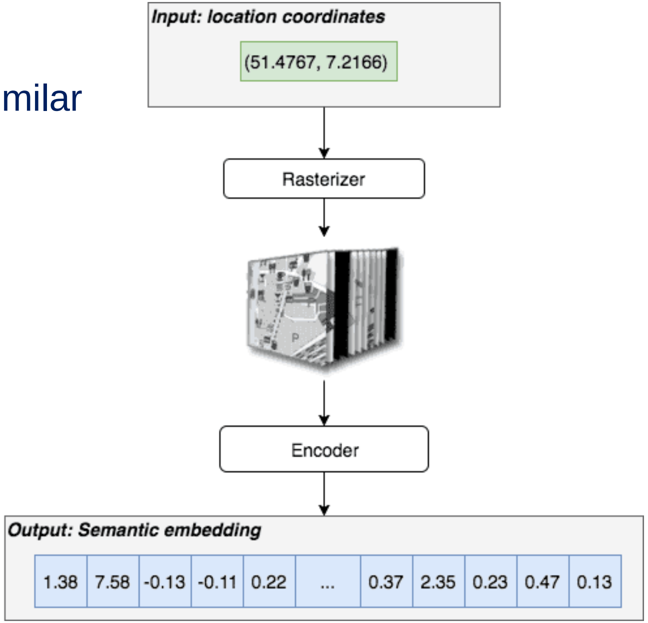


Supervised Contrastive

Loc2Vec

- Location → Vector: given a location with coordinates, we aim to generate a vector embedding to represent it and support downstream tasks
  - It should preserve location functionality
  - Tobler's first law of geography: close things are usually more similar

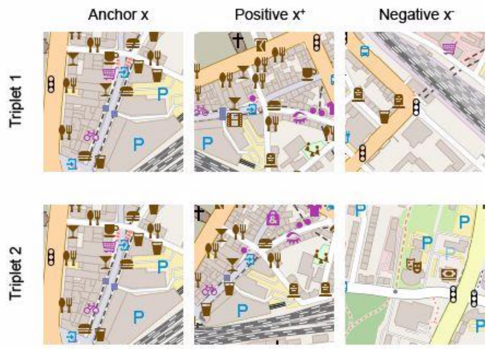


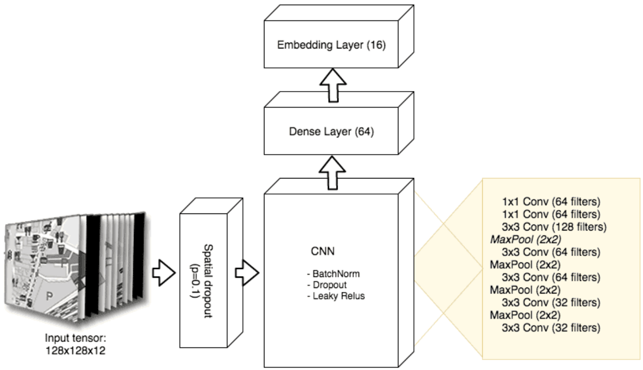


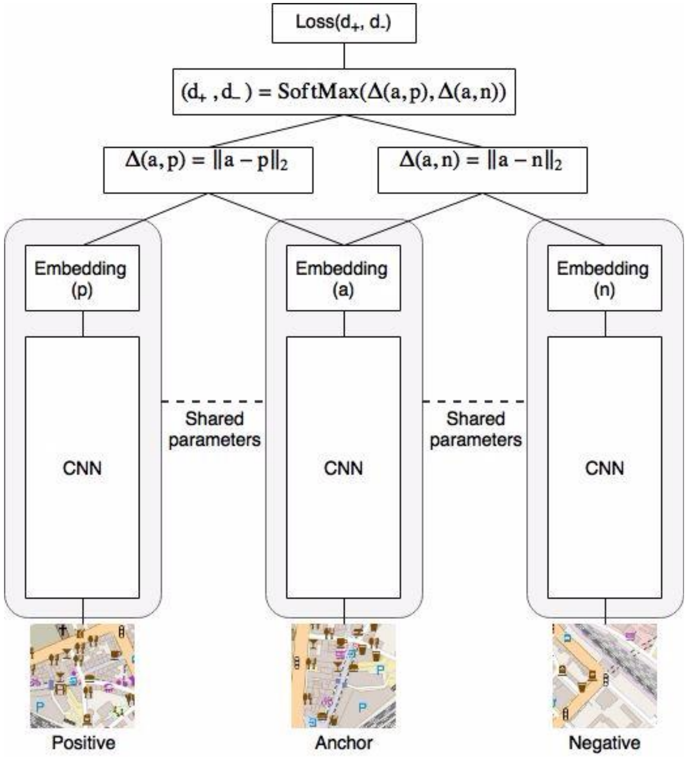
https://sentiance.com/loc2vec-learning-location-embeddings-w-triplet-loss-networks

Self-Supervised Learning

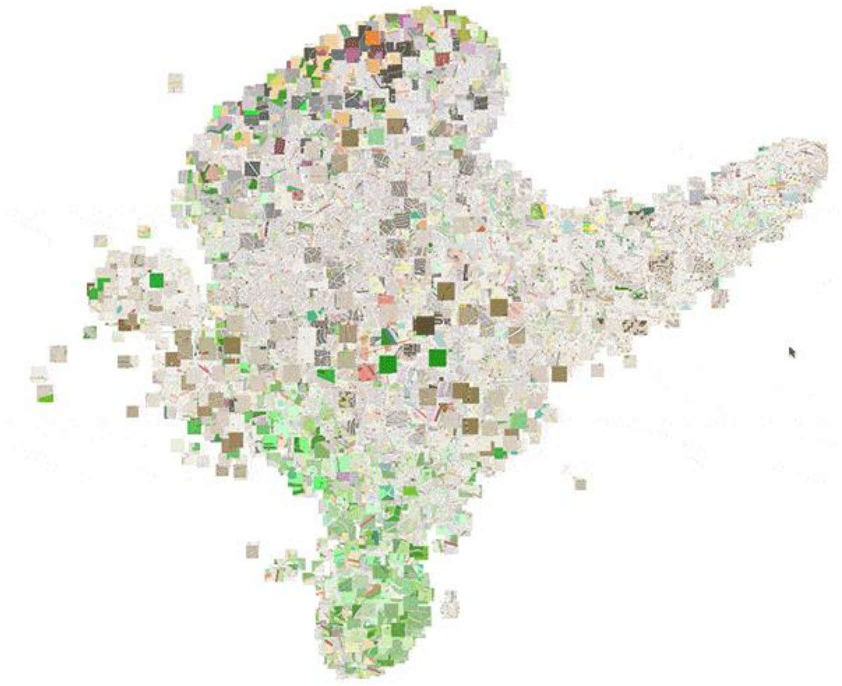
- Training through a Triplet Loss

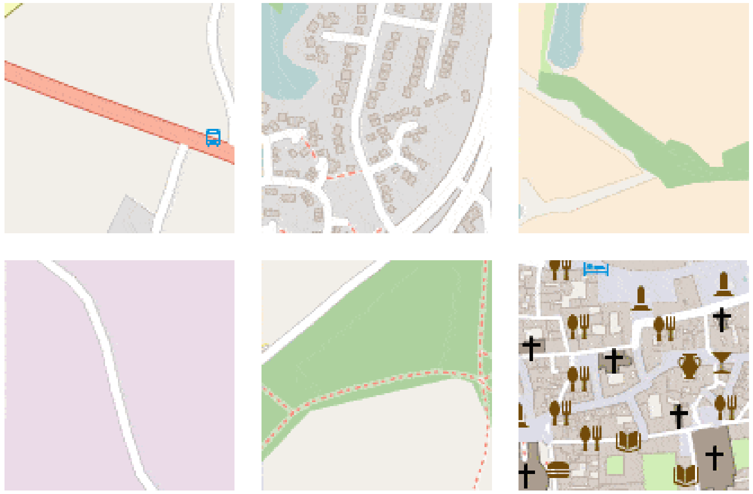






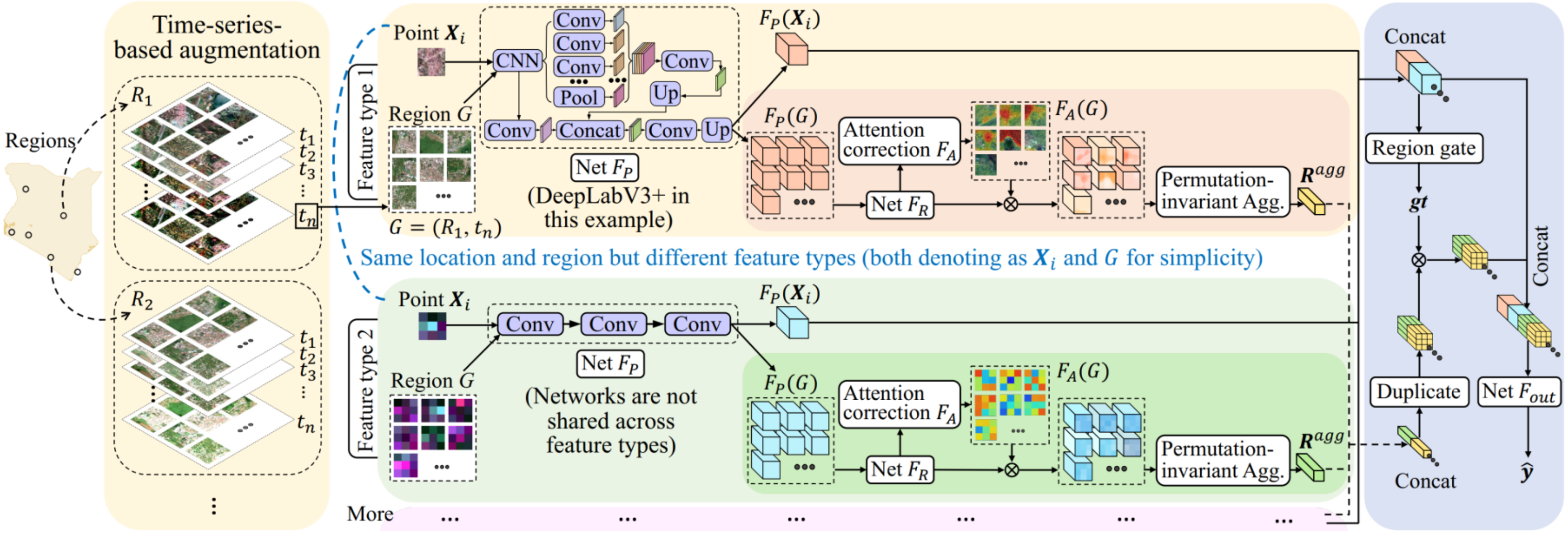
Visualization





Random walk over 6 locations

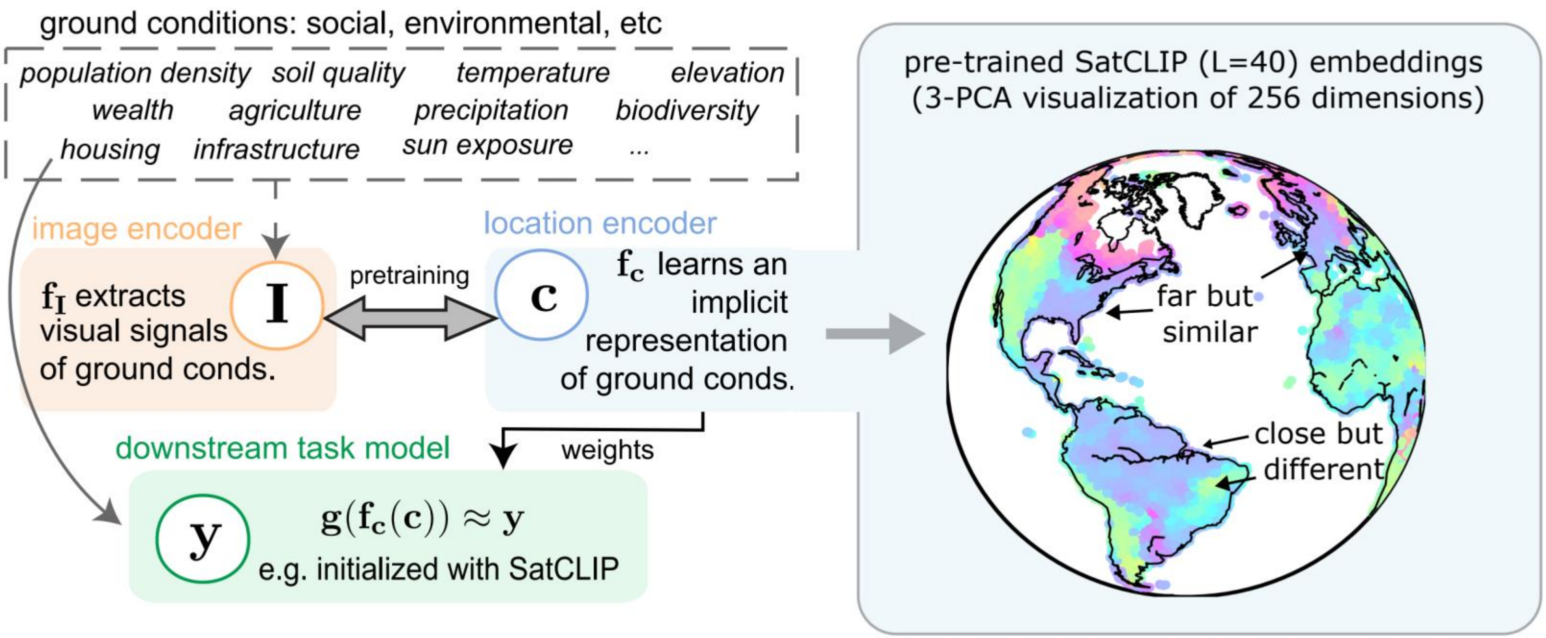
Predicting Poverty from Satellite Images via SSL



Li et al., Point-to-Region Co-learning for Poverty Mapping at High Resolution Using Satellite Imagery. AAAI 2023.

SatCLIP

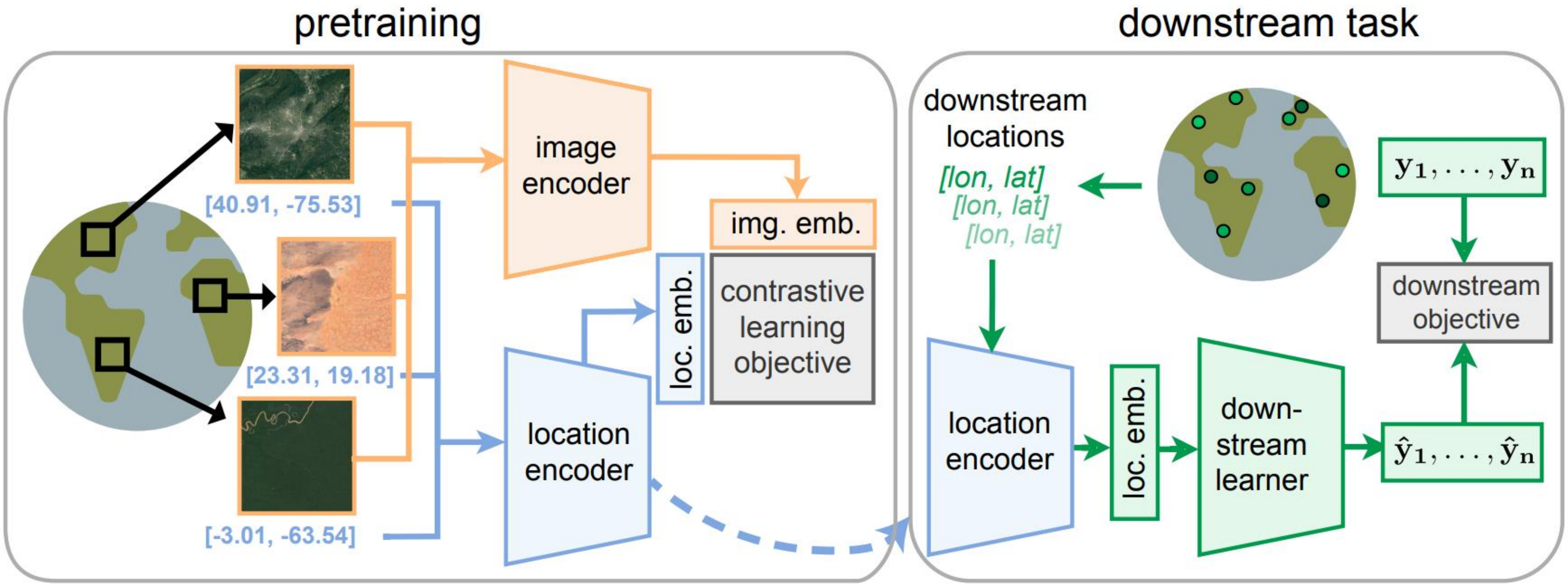
- Objective: transform a geo-coordinate into vector representations



Klemmer et al., SatCLIP: Global, General-Purpose Location Embeddings with Satellite Imagery. arXiv 2024.

SatCLIP

- Based on Contrastive Language Image Pretraining (CLIP)

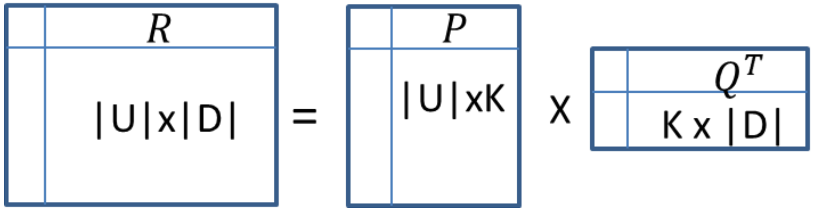


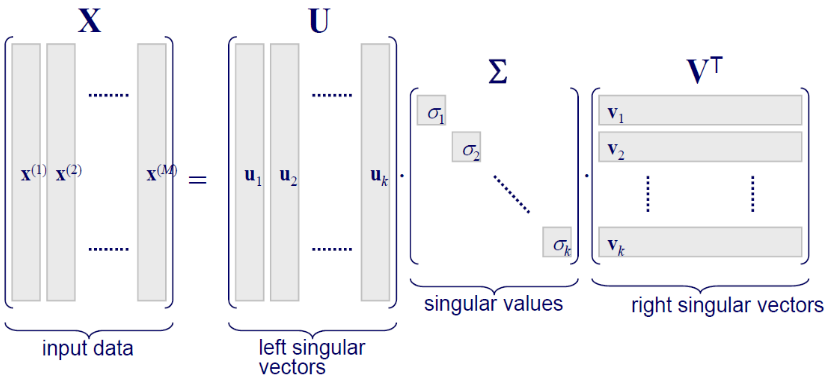
Outline

- Learning embeddings for locations
  - Explicit learning
    - Task-specific supervised learning
    - Self-supervised learning
  - Implicit learning
    - Low-rank matrix decomposition
    - Meta learning approaches

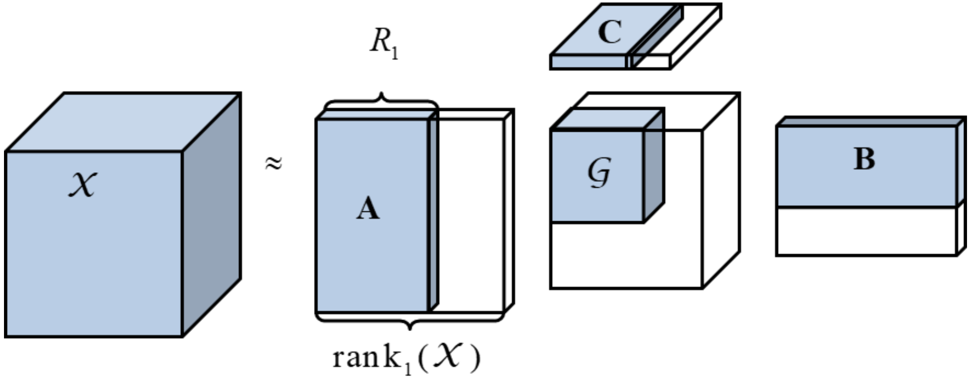
Low-Rank Matrix Decomposition

- Matrix Factorization



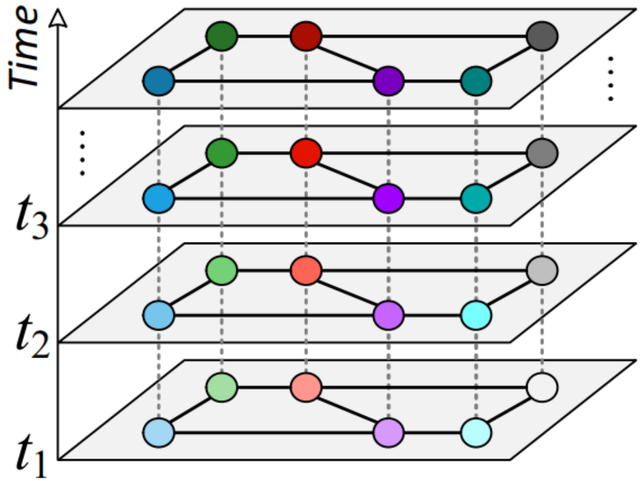


- Tucker Decomposition



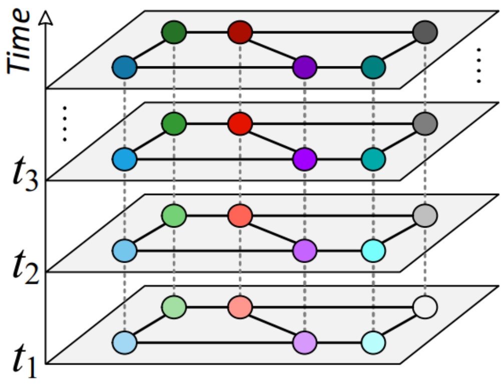
Definition of ST Graphs

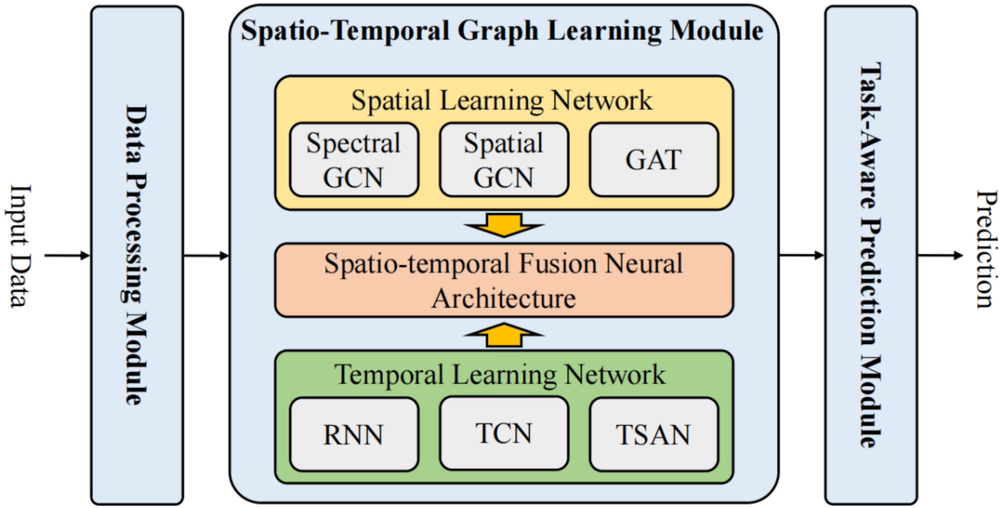
- There are numerous sensors deployed in the physical world
- Properties
  - Each sensor has a unique geospatial location
  - Constantly reporting time series readings
  - With structural correlation between readings
    - ➢ Usually represented as graphs
- Examples
  - Traffic speed/flow over road networks
  - Crowd flow in irregular urban regions



Spatio-Temporal Representation Learning

- Spatio-Temporal Graph Neural Networks (STGNNs) have emerged as the most popular approach for learning spatial and temporal dependencies

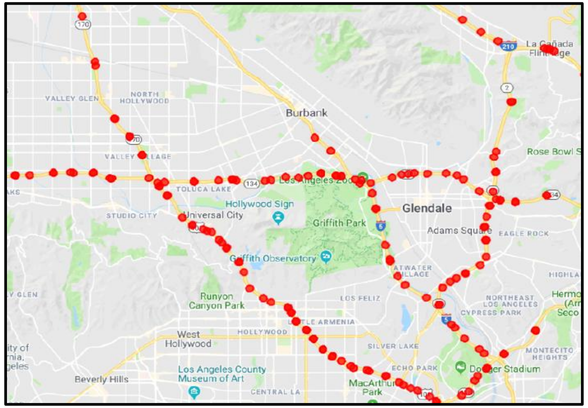




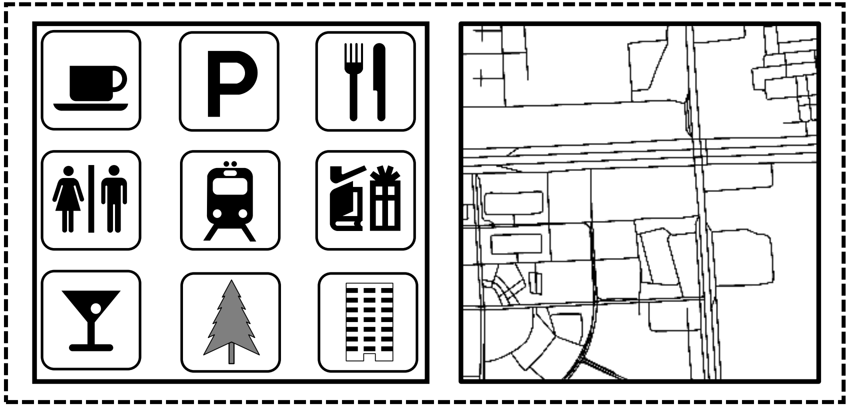
Application: Predicting ST Graphs over Road Networks

Traffic Forecasting over Road Networks

Traffic data



Geo-attributes



Predict urban traffic on each location at next time interval

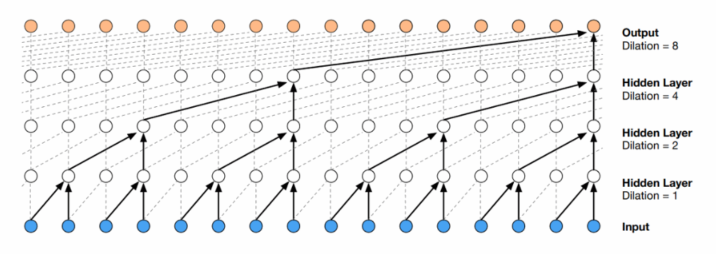
throughout a city by using ***historical traffic data*** and ***geo-attributes***

(e.g., points of interests and road networks)

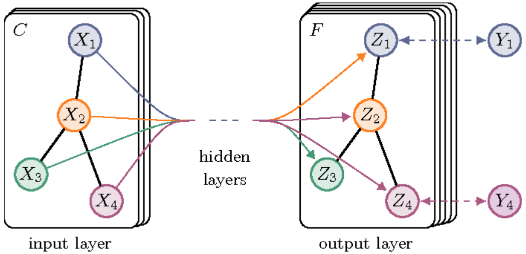
Graph WaveNet

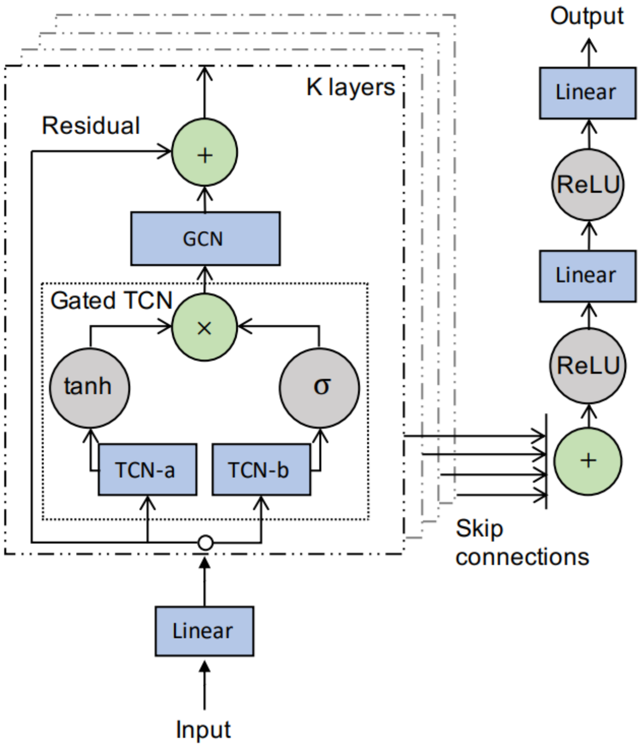
- Pipeline
  - Temporal learning via TCNs
  - Spatial learning via GNNs

TCNs



GNNs

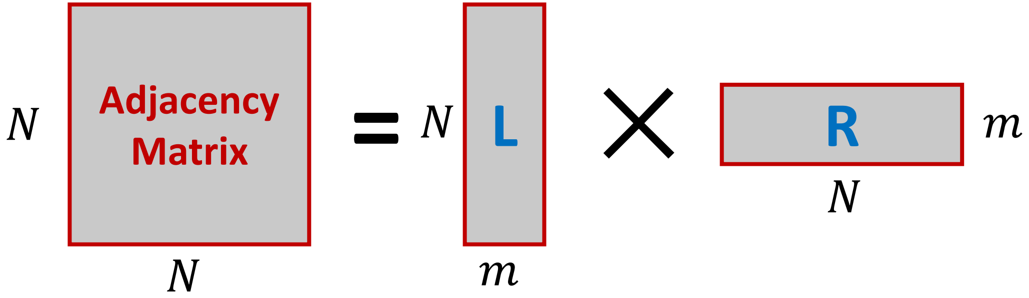




Wu et al., Graph WaveNet for Deep Spatial-Temporal Graph Modeling. IJCAI 2019.

Implicit Learning of Location Embeddings

- **Adaptive adjacency matrix** (inspired by RecSys)
  - Adaptively generated from two low-rank matrices
  - L and R can be viewed as location embeddings
  - The embeddings can reflect the similarity between different nodes in the latent space



(a) The heatmap of the learned self-adaptive adjacency matrix for the first 50 nodes.

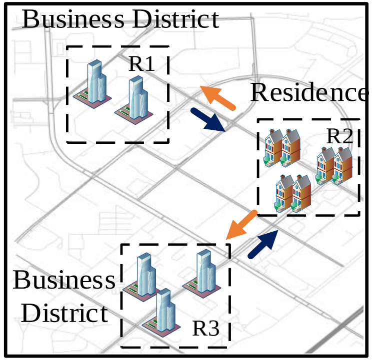
(b) The geographical location of a part of nodes marked on Google Maps.

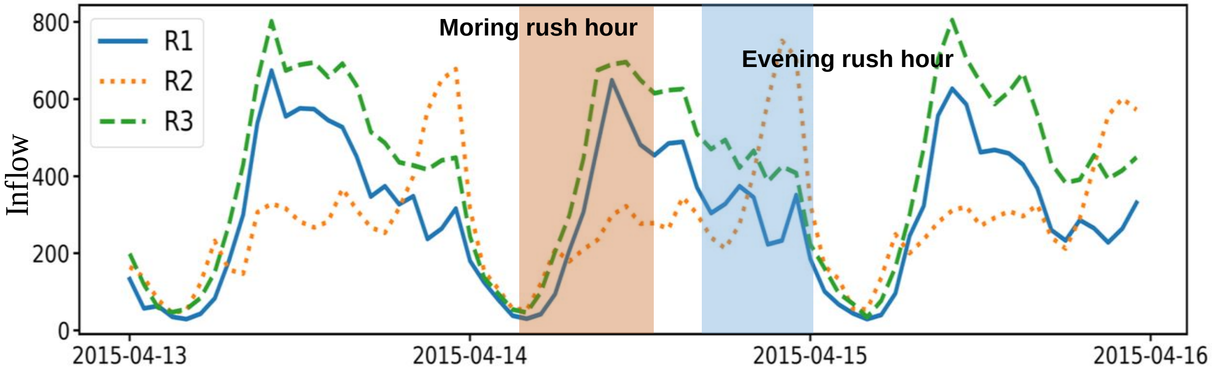
Outline

- Learning embeddings for locations
  - Explicit learning
    - ➢ Task-specific supervised learning
    - ➢ Self-supervised learning
  - Implicit learning
    - ➢ Low-rank matrix decomposition
    - ➢ Meta learning approaches

Challenge – Diverse ST Dependencies

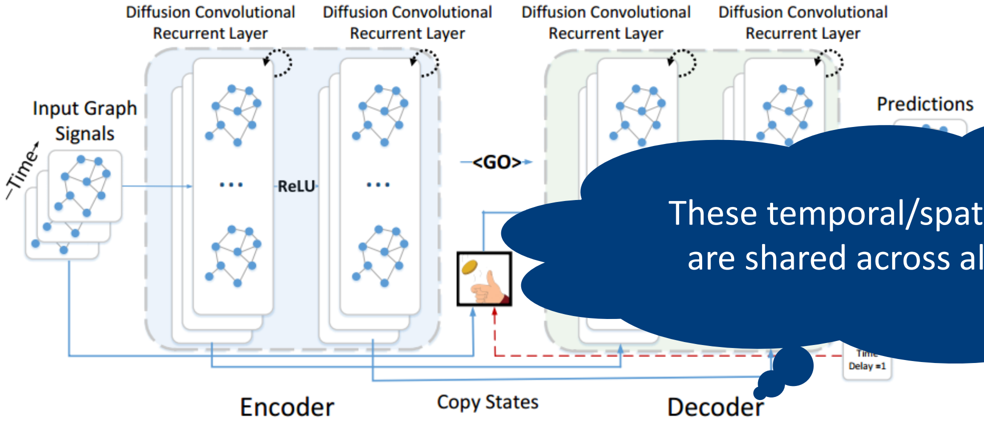
- However, spatial location characteristics and spatial interrelationships are diverse and often depend on **spatial attributes**
- Spatial locations with similar geo-attributes tend to have
  - Similar temporal properties
  - Similar spatial correlations

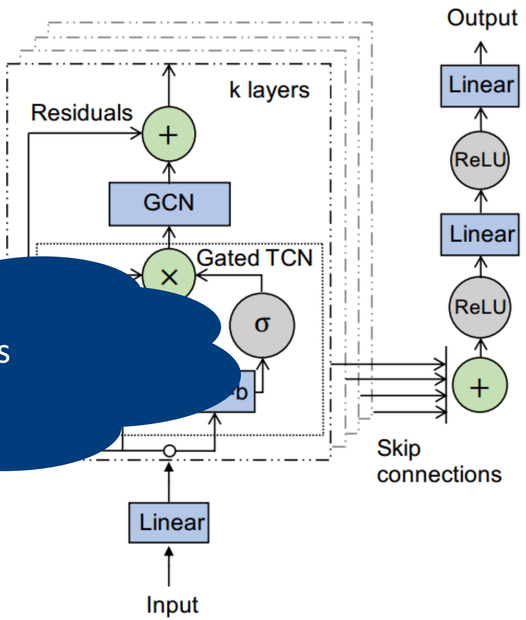




Existing Approaches

- Capturing ST dependencies is of great importance to traffic forecasting
  - Capturing spatial correlations by Graph Neural Networks (GNNs)
  - Learning temporal dependencies by RNNs or CNNs

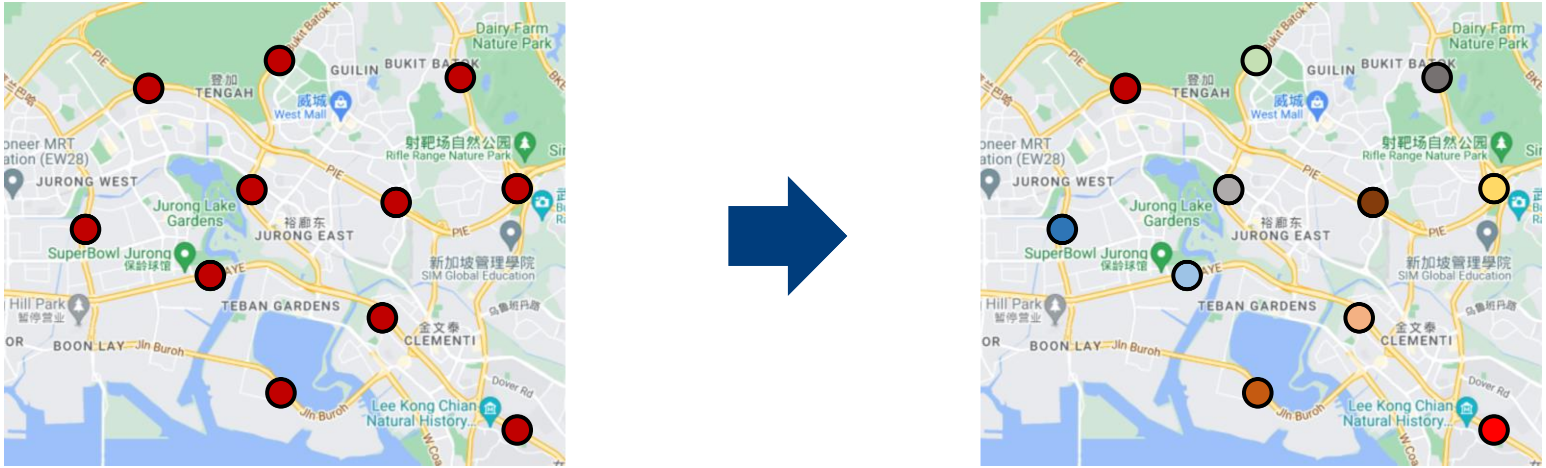




These temporal/spatial modules are shared across all locations

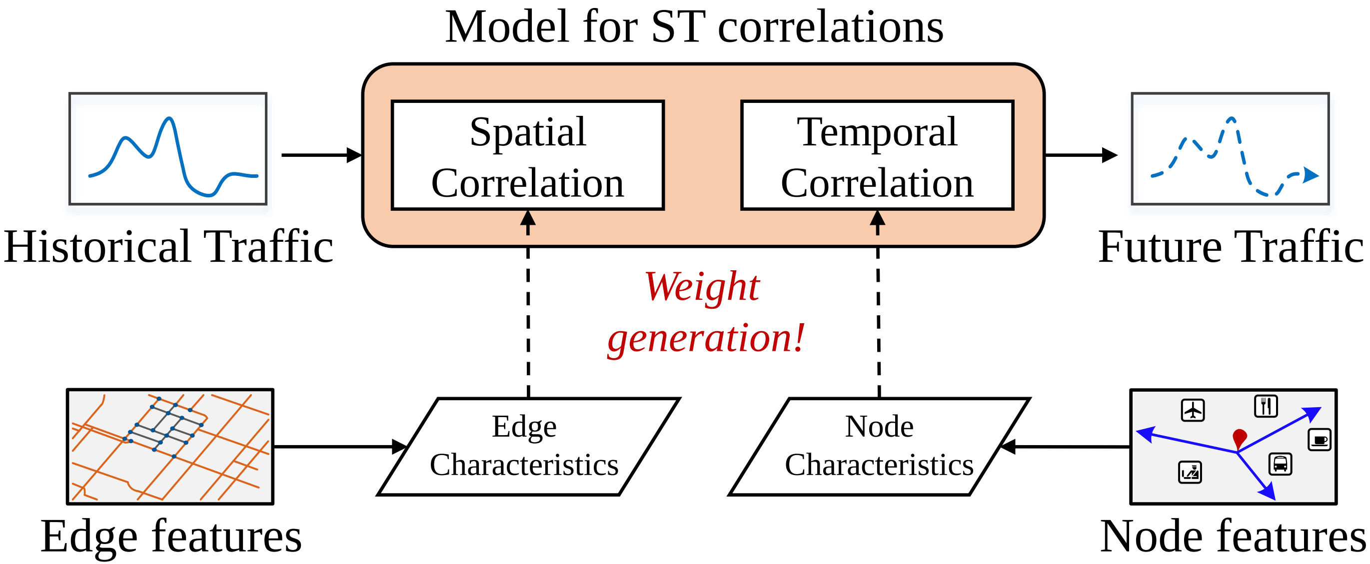
An Intuitive Idea

- Using non-shared modules to customize the model for each location
  - **Parameter explosion**: the parameter size will increase by #node times
  - Prone to overfitting, verified by empirical studies

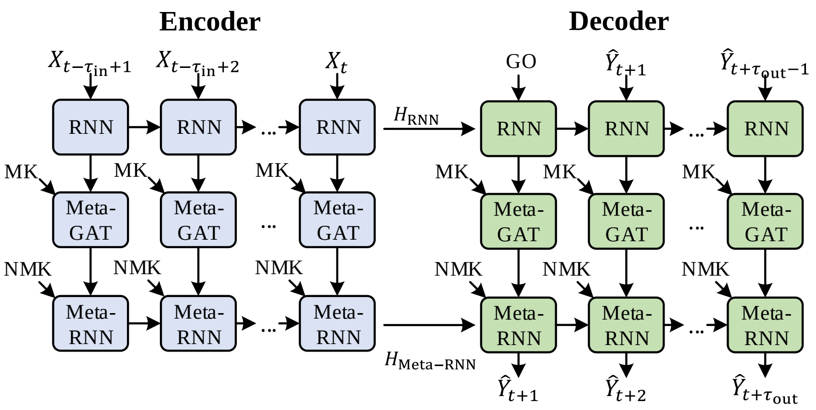


Our Intuition

- **Geospatial attributes** can reflect the characteristics of nodes and edges, and affect different kinds of spatio-temporal characteristics



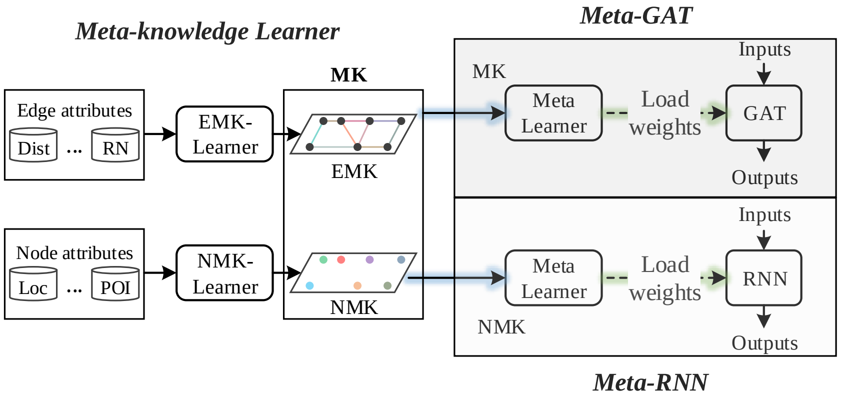
ST-MetaNet: Spatio-Temporal Meta Network



Recurrent Neural Network (RNN)
❑ Embedding the sequence of urban traffic.

Meta Graph Attention Network (Meta-GAT)
❑ Modeling diverse spatial correlations.

Meta Recurrent Neural Network (Meta-RNN)
❑ Modeling diverse temporal correlations.

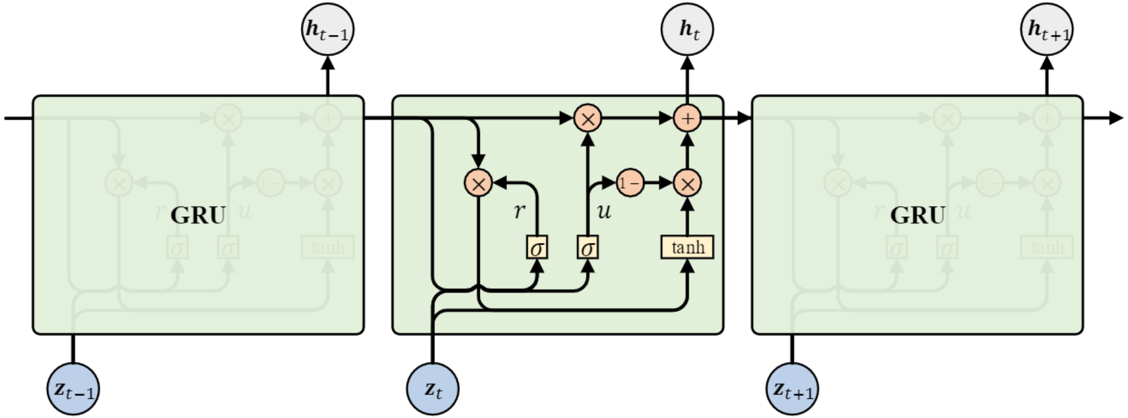


Meta-knowledge Learner
❑ Learning node & edge characteristics from geo-attributes.

Meta Learner
❑ Generating parameter weights in GAT and RNN.

Z. Pan, Y. Liang et al., Urban Traffic Prediction from Spatio-Temporal Data Using Deep Meta Learning. KDD 2019.

GRUs (a variant of RNNs)



$$u = \mathrm{sigmoid}\left(\mathbf{W}_{u,i}\boldsymbol{z}_{t,i} + \mathbf{U}_{u,i}\boldsymbol{h}_{t-1,i} + b_{u,i}\right),$$

$$r = \mathrm{sigmoid}\left(\mathbf{W}_{r,i}\boldsymbol{z}_{t,i} + \mathbf{U}_{r,i}\boldsymbol{h}_{t-1,i} + b_{r,i}\right),$$

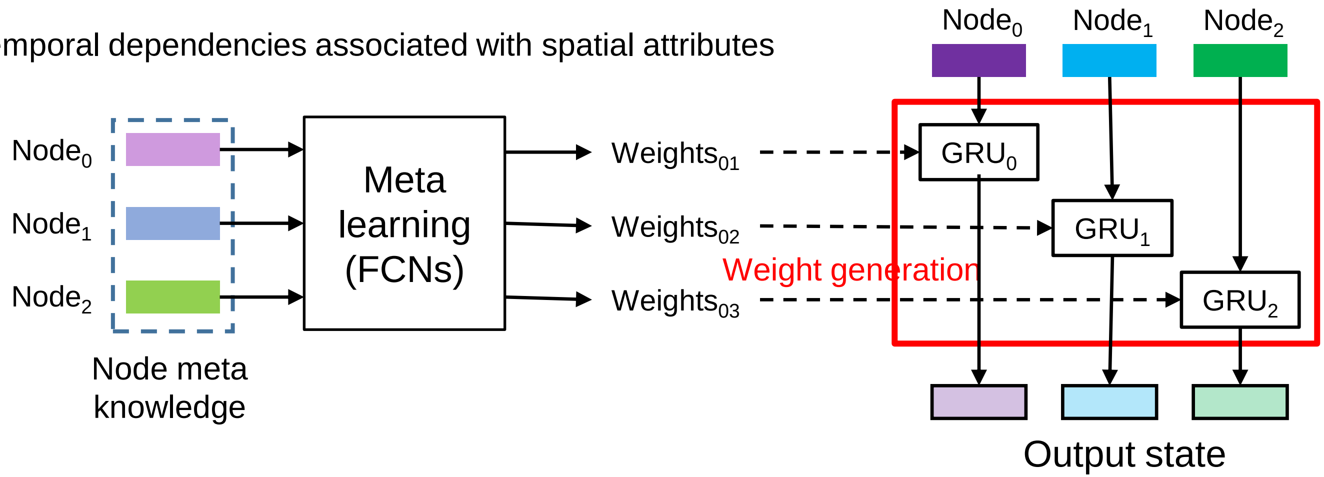
$$\boldsymbol{h}_{t,i} = u \circ \boldsymbol{h}_{t-1,i} + (1-u) \circ \tanh\left(\mathbf{W}_{h,i}\boldsymbol{z}_{t,i} + \mathbf{U}_{h,i}\left(r \circ \boldsymbol{h}_{t-1,i}\right) + b_{h,i}\right).$$

The parameters shared by all nodes for feature embedding

Meta-GRU

Meta-GRUs

- Modeling temporal dependencies associated with spatial attributes



$$u = \mathrm{sigmoid}\left(\boxed{\mathbf{W}_{u,i}}\boldsymbol{z}_{t,i} + \boxed{\mathbf{U}_{u,i}}\boldsymbol{h}_{t-1,i} + \boxed{b_{u,i}}\right),$$

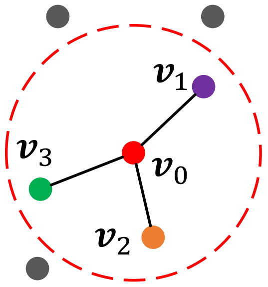
$$r = \mathrm{sigmoid}\left(\boxed{\mathbf{W}_{r,i}}\boldsymbol{z}_{t,i} + \boxed{\mathbf{U}_{r,i}}\boldsymbol{h}_{t-1,i} + \boxed{b_{r,i}}\right),$$

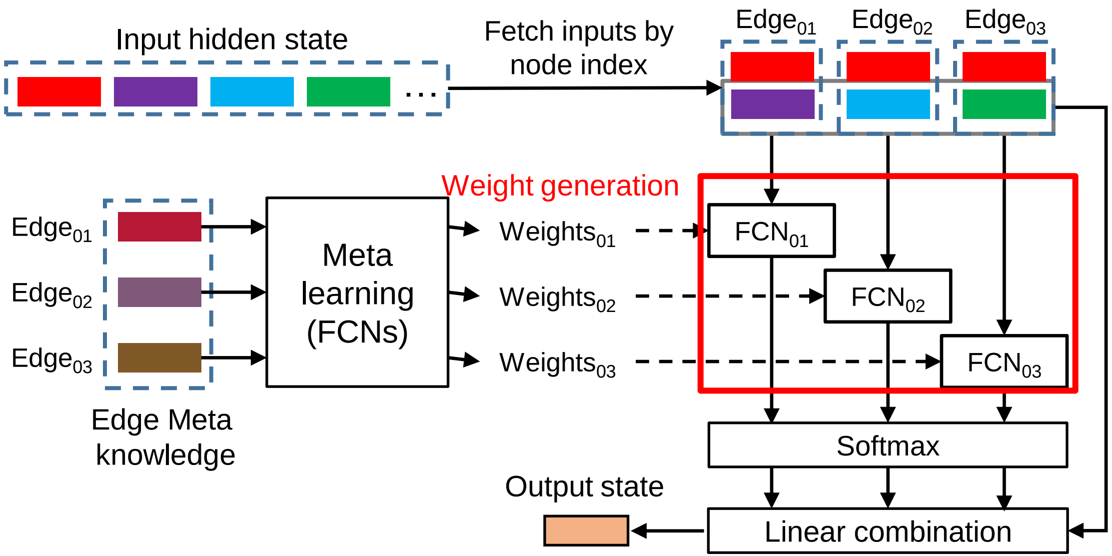
$$\boldsymbol{h}_{t,i} = u \circ \boldsymbol{h}_{t-1,i} + (1-u) \circ \tanh\left(\boxed{\mathbf{W}_{h,i}}\boldsymbol{z}_{t,i} + \boxed{\mathbf{U}_{h,i}}\left(r \circ \boldsymbol{h}_{t-1,i}\right) + \boxed{b_{h,i}}\right).$$

Meta Graph Attention Network (Meta-GAT)

Meta-GAT

- Modeling spatial correlation associated with spatial attributes





Computing attention weights

$$\boldsymbol{w}_{ij} = \mathrm{LeakyReLU}\left(\boxed{\mathbf{W}_{ij}}\left[\boldsymbol{h}_i \parallel \boldsymbol{h}_j\right] + \boxed{b_{ij}}\right)$$

Extracting edge-related meta knowledge

$$\mathrm{MK}_{ij} = \mathrm{NMK}(\boldsymbol{v}_i) \parallel \mathrm{NMK}(\boldsymbol{v}_j) \parallel \mathrm{EMK}(\boldsymbol{e}_{ij})$$

Network weight generation

$$\mathbf{W}_{ij} = \mathrm{G}_{\mathbf{W}}(\mathrm{MK}_{ij}), \quad b_{ij} = \mathrm{G}_b(\mathrm{MK}_{ij})$$

Normalization by softmax

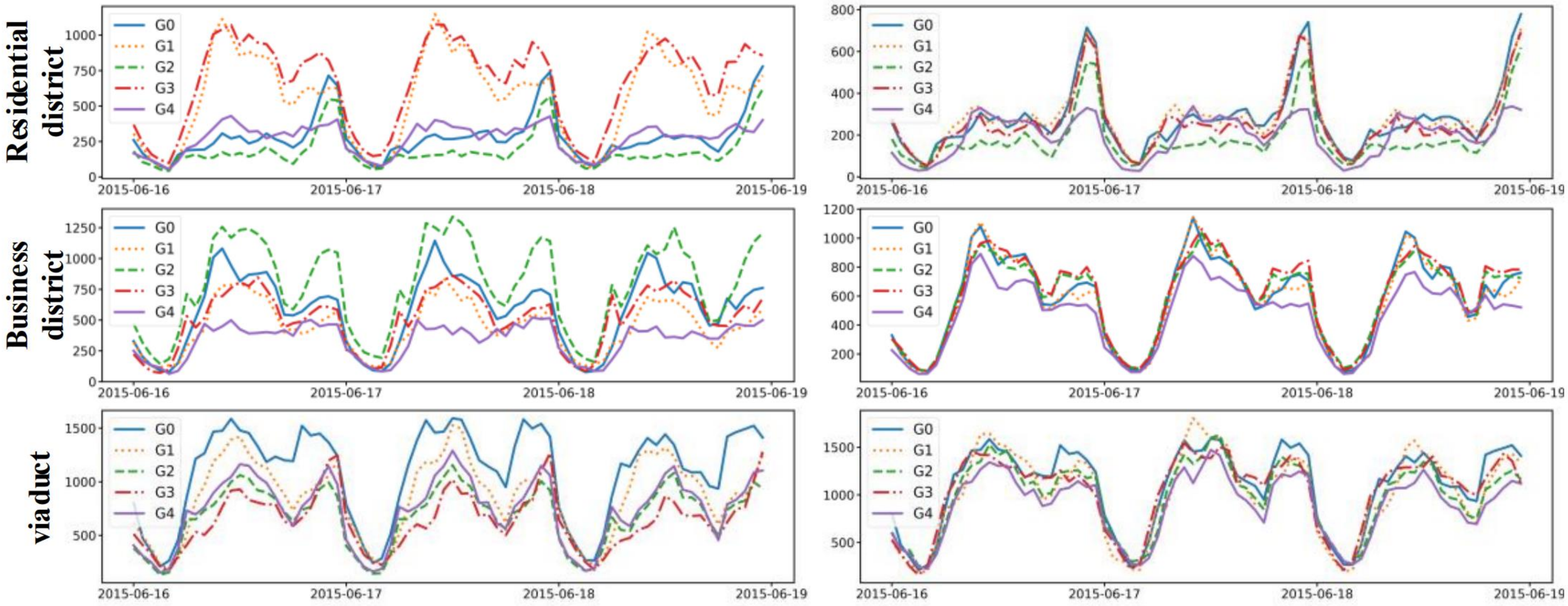
$$(1-\lambda_i)\boldsymbol{h}_i + \lambda_i \mathrm{ReLU}\left(\sum_j \frac{\exp(\boldsymbol{w}_{ij})}{\sum_k \exp(\boldsymbol{w}_{ik})} \boldsymbol{h}_j\right)$$

Model Comparison

- We evaluate our model on two different tasks, predicting taxi flow in Beijing and forecasting traffic speed on highways of Los Angeles.
- Compared to SOTA, our method achieves lower error while using fewer parameters

Models			HA	ARIMA	GBRT	Seq2Seq	GAT-SeqSeq	SOTA	**ST-MetaNet**
Taxi flow	overall	MAE	26.2	40.0	28.8	21.3 ± 0.06	18.3 ± 0.13	18.7 ± 0.53	**16.9** ± 0.13
		RMSE	56.5	86.8	60.9	42.6 ± 0.14	35.6 ± 0.23	36.1 ± 0.59	**34.0** ± 0.25
	1h	MAE	26.2	27.1	22.3	17.8 ± 0.05	16.3 ± 0.12	16.8 ± 0.50	**15.0** ± 0.14
		RMSE	56.5	58.3	47.7	35.1 ± 0.07	31.9 ± 0.21	31.9 ± 0.69	**29.9** ± 0.08
	2h	MAE	26.2	41.2	29.8	22.0 ± 0.06	18.7 ± 0.12	18.9 ± 0.57	**17.3** ± 0.14
		RMSE	56.5	77.0	62.6	43.6 ± 0.16	36.3 ± 0.20	36.4 ± 0.71	**34.7** ± 0.25
	3h	MAE	26.2	51.8	34.2	24.2 ± 0.09	19.9 ± 0.14	20.3 ± 0.52	**18.4** ± 0.10
		RMSE	56.5	108.0	70.3	48.1 ± 0.20	38.4 ± 0.30	39.5 ± 0.46	**37.1** ± 0.41
	# params		-	-	-	333k	407k	445k	268k
Traffic Speed	overall	MAE	4.79	4.03	3.85	3.55 ± 0.01	3.28 ± 0.00	3.10 ± 0.01	**3.05** ± 0.02
		RMSE	8.72	7.94	7.48	7.27 ± 0.01	6.66 ± 0.01	6.31 ± 0.03	**6.25** ± 0.02
	15min	MAE	4.79	3.27	3.16	2.98 ± 0.01	2.83 ± 0.01	2.75 ± 0.01	**2.68** ± 0.02
		RMSE	8.72	6.14	6.05	5.88 ± 0.01	5.47 ± 0.01	5.33 ± 0.02	**5.15** ± 0.02
	30min	MAE	4.79	3.99	3.85	3.57 ± 0.01	3.31 ± 0.00	3.14 ± 0.01	**3.09** ± 0.03
		RMSE	8.72	7.78	7.50	7.26 ± 0.01	6.68 ± 0.00	6.45 ± 0.04	**6.28** ± 0.02
	60min	MAE	4.79	5.18	4.85	4.38 ± 0.01	3.93 ± 0.01	**3.60** ± 0.02	**3.60** ± 0.04
		RMSE	8.72	10.10	9.08	8.88 ± 0.02	8.03 ± 0.02	7.65 ± 0.06	**7.52** ± 0.01
	# params		-	-	-	81k	113k	373k	85k

Case Study on Learned Location Embedding



(a) Seq2seq integrated with graph attention

(b) The proposed model

Presented Papers

0	UrbanCross: Enhancing Satellite Image-Text Retrieval with Cross-Domain Adaptation	Link	ACM MM	2024	HKUST(GZ)	Yongzi Yu
1	Beyond the First Law of Geography: Learning Representations of Satellite Imagery by Leveraging Point-of-Interests	https://dl.acm.org/doi/10	WWW '22	2022	University of Helsinki	Yixuan Wang
2	When Urban Region Profiling Meets Large Language Models	Link	WWW	2024	HKUST(GZ)	Jiaxi Hu
3	Urban Region Profiling With Spatio-Temporal Graph Neural Network	.org/stamp/stamp.jsp?tp=	IEEE	2022	Dalian University of Technology	Tianyu Wei

Q&A

Thanks!

CityMind Lab