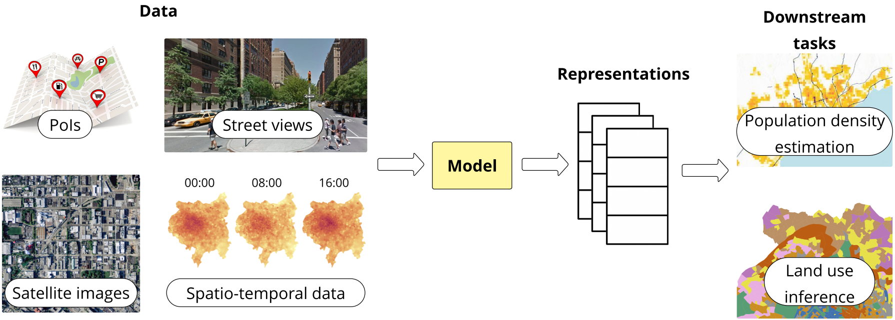
From Service-Specific Mobile Traffic data to Urban Region Representations

Giulio Loddi[1] Chiara Pugliese[2] Francesco Lettich[3]
Fabio Pinelli[1,3] Chiara Renso[3]

[1]IMT Lucca, Italy
[2]IIT-CNR / University of Pisa, Italy
[3]ISTI-CNR, Italy

What is Urban Representation Learning?

- **Urban representation learning** aims to compute meaningful and multi-purpose **embeddings** of urban areas, which capture the **characteristics of the urban landscape**.
- It can leverage a variety of **data sources**.
- The quality of the embeddings is evaluated through different **downstream tasks**

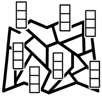


Service-specific mobile traffic data

People daily use different **mobile applications** (*e.g.*, Instagram, WhatsApp, Spotify).

This generates high-volume traffic, linked to specific **digital behaviors** (e.g., social networking, leisure activities, working)

The **rich spatio-temporal patterns** of **service-level traffic** offers a fine-grained lens on urban life.



We aim to demonstrate the potential of service-specific mobile traffic data to generate **high-quality urban region representations**.

Our Contributions

Our Contributions:

- We introduce the **first framework** for learning urban region embeddings from **service-specific mobile traffic data**[1].
- Our **two-step pipeline** combines a temporal convolutional autoencoder and a region-level aggregator trained with contrastive learning.
- We conduct extensive **evaluations on a real-world dataset** (NetMob23) across two downstream tasks and a qualitative clustering analysis.

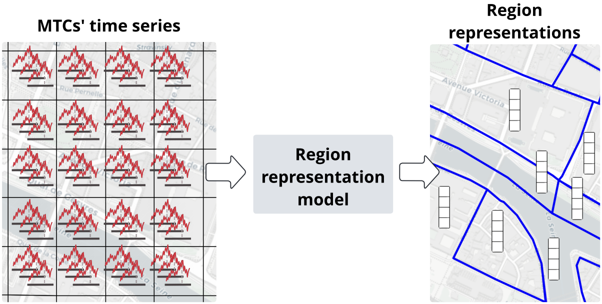
[1] G. Loddi et al., "Urban Region Embeddings from Service-Specific Mobile Traffic Data," 2025 26th IEEE International Conference on Mobile Data Management (MDM), IEEE, 2025.

Problem definition

Given:

- A city partitioned into Mobile Traffic Cells (**MTCs**), each with a multivariate time series of service-specific traffic.
- A **target tessellation** grouping MTCs into urban regions.

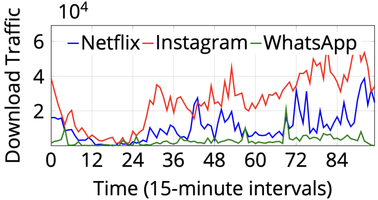
Goal: learn compact **region embeddings** that capture urban characteristics and account for both spatial context and temporal dynamics.

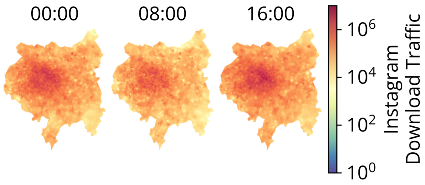


NetMob23 Dataset

We use the **NetMob23 dataset**[2], which captures mobile traffic activity in 20 French cities over approximately 2.5 months.

- The dataset contains **upload/download volumes** of **68 mobile applications** (e.g., WhatsApp, YouTube, Gmail), sampled every **15 minutes**.
- Traffic volumes are provided on a **100 m** × **100 m** MTC grid, with one multivariate time series per cell (one channel per app).





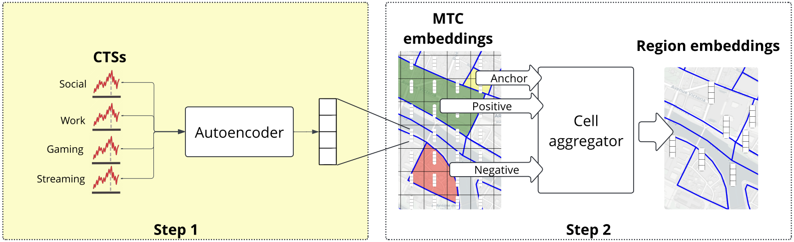
[2] released by Orange, to the participants of NetMob 2023 Data Challenge.

Method overview

Objective: (1) Learn compact representations of Mobile Traffic Cells (MTCs) from multivariate time series. (2) Aggregate MTC embeddings into region-level embeddings.

- Train a **TCN autoencoder** to produce a 44-dimensional embedding for each MTC.
- **Two aggregation strategies:**
  - Transformer encoder + average pooling
  - Learnable weighted sum

  Train using **triplet loss** (Tobler's First Law): pull nearby regions closer, separate distant ones.



Results

Land Use Inference

Model	Category	KL-div	L1	Cosine Similarity
Our method (transformer)	Gaming	0.410±0.007	0.578±0.010	0.869±0.003
Our method (weighted sum)	Gaming	0.395±0.003	0.571±0.007	0.878±0.017
RegionDCL	–	0.354±0.005	0.540±0.008	0.888±0.003
Our method (weighted sum)	All	0.328±0.019	0.503±0.016	0.904±0.008
Our method (transformer)	**All**	**0.313±0.017**	**0.485±0.016**	**0.906±0.008**

Population Density Estimation

Model	Category	MAE	RMSE	R^2
RegionDCL	–	3627.8±69.9	5098.7±77.7	0.4192±0.0197
Our method (transformer)	Gaming	2952.0±49.3	4423.0±98.4	0.5628±0.0213
Our method (weighted sum)	Gaming	2763.3±55.7	4224.4±95.5	0.6012±0.0204
Our method (transformer)	All	2316.0±86.1	3653.4±155.1	0.7013±0.0260
Our method (weighted sum)	**All**	**2243.8±80.1**	**3598.1±189.3**	**0.7099±0.0331**

Observations:

- Our method consistently **outperforms the baseline**.
- The transformer and weighted sum aggregators exhibit comparable performance across both tasks.
- Leveraging all service categories consistently outperforms using a single app type (e.g., Gaming), demonstrating **the value of a multi-service data**.

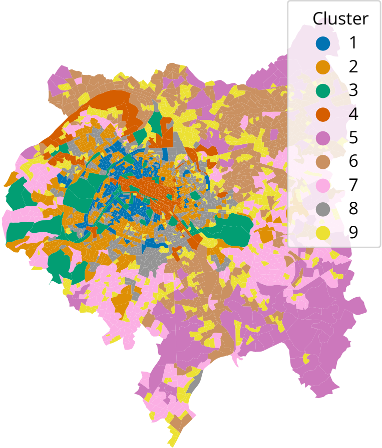
Clustering Qualitative Analysis

Goal: Assess whether region embeddings capture interpretable urban structure.

Method: hierarchical clustering (Ward's method, Euclidean distance) on region embeddings.

Findings:

- A clear center–periphery structure consistent with urban form.
- Semantically similar areas are grouped: e.g., green areas (e.g., parks), major landmarks (e.g., Eiffel Tower), train stations, and airports.
- Demonstrates that embeddings trained on mobile traffic alone reflect real-world urban patterns.



Conclusions and Future Work

Summary:

- We proposed a methodology to **learn urban region embeddings from service-specific mobile traffic data**.
- Our embeddings **outperform a strong multi-modal baseline** on land use and population density prediction tasks.
- Clustering analysis confirms that the **learned representations capture meaningful urban structure**.

Future Directions:

- Extend evaluations to **additional downstream tasks** (e.g., socio-economic indicators, mobility demand).
- Investigate **time-dynamical** region embeddings.

Thank you for your attention!

For questions feel free to contact me at:

giulio.loddi@imtlucca.it

Scan to access the GitHub repository