PURDUE UNIVERSITY
Department of Computer Science

EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO
2024

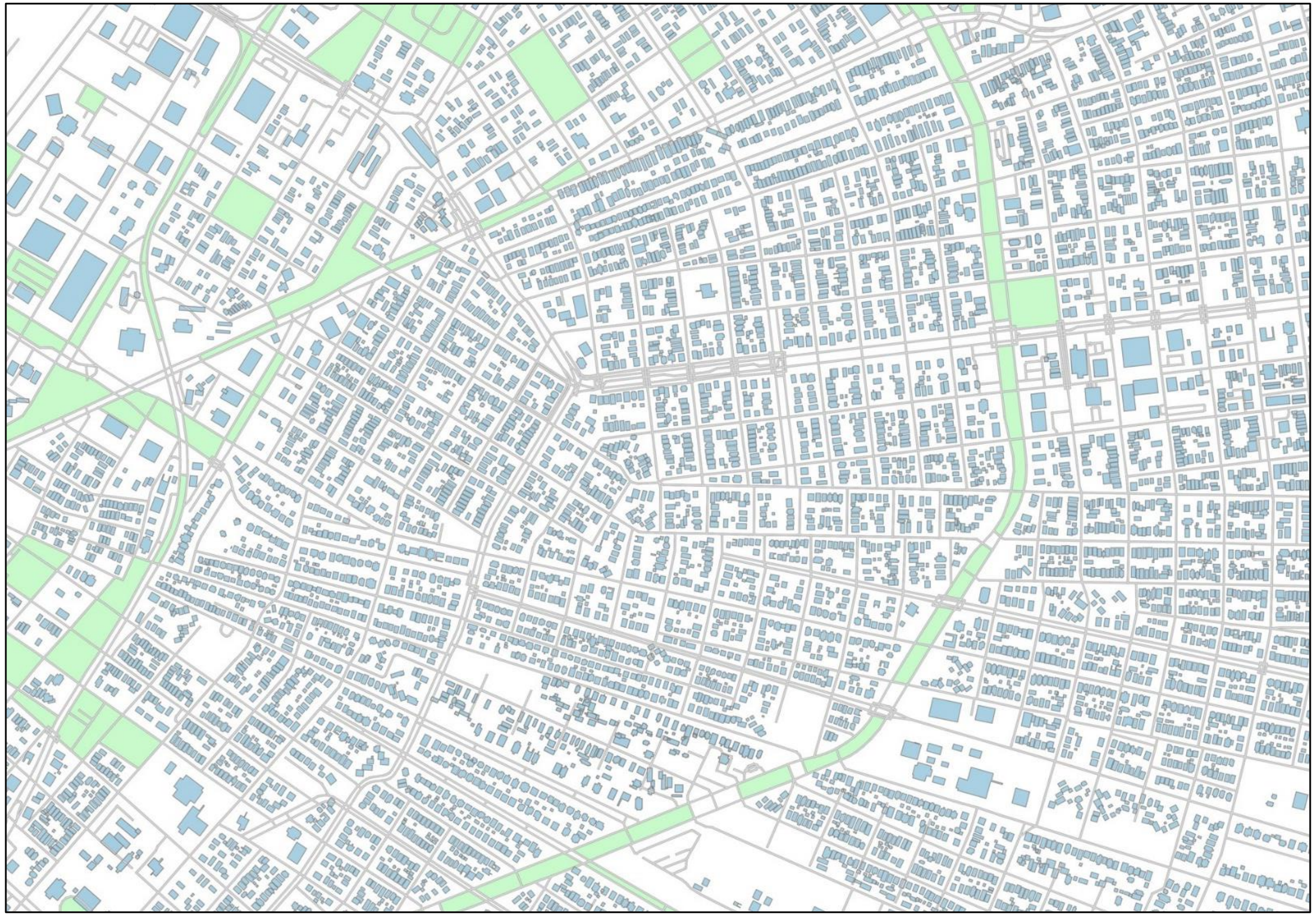
COHO: Context-Sensitive City-Scale Hierarchical Urban Layout Generation

Liu He, Daniel Aliaga

Project Website:

Author:

Goal: large-scale urban layout generation



e.g., New Orleans

PURDUE UNIVERSITY | Department of Computer Science

ECCV EUROPEAN CONFERENCE ON COMPUTER VISION MILANO

Example Results: 330 Cities in North America

Observations:

(1) **Representation:** Cities, communities, blocks, buildings, and roads are arbitrarily shaped and with complex topology.

(2) **Context:** Buildings, city blocks, and communities are built considering their neighboring structures and not in isolation.

(3) **Prioritization:** The stylistic and semantic importance, or priority, of city blocks and buildings varies.

Realistic, Context-Sensitive, Scalable

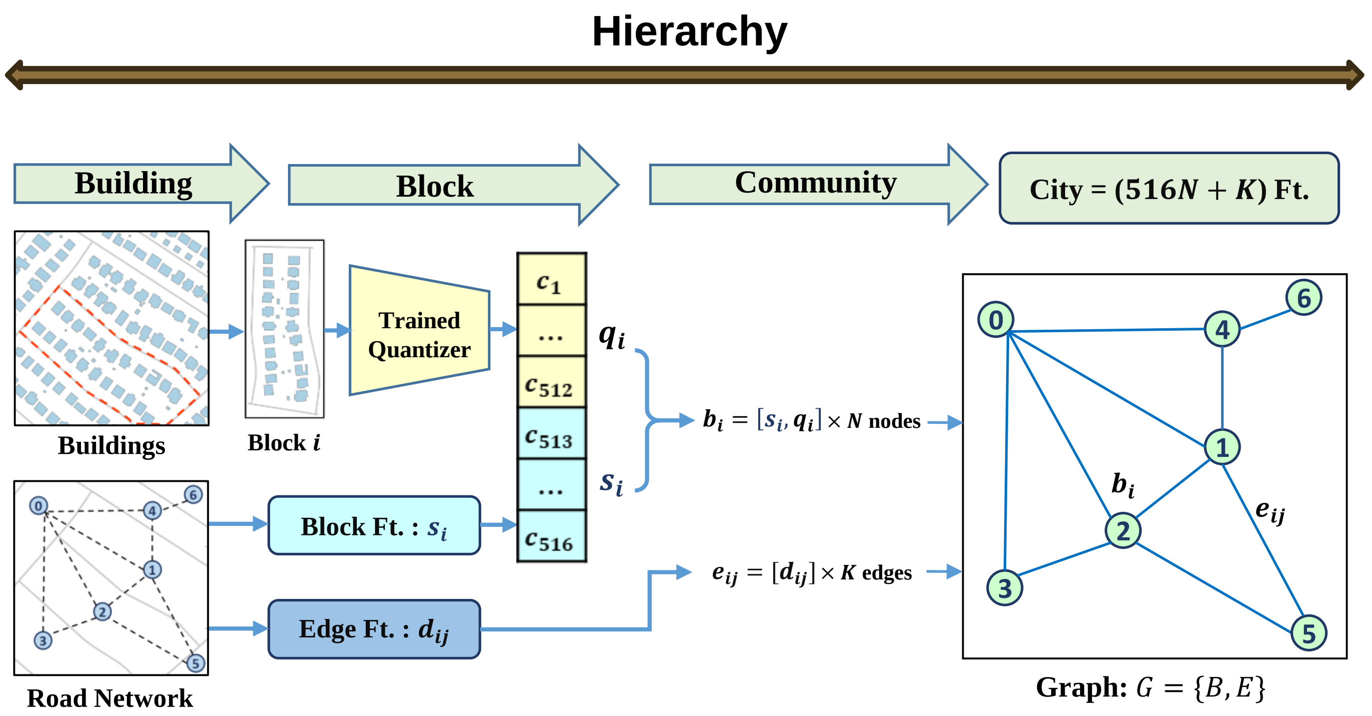
Approach:

(1) A canonical **hierarchical graph representation** for an entire city.

(2) A self-supervised Graph-based Masked AutoEncoder (GMAE) for **contextual sensitivity learning**.

(3) A **priority-based scheduled** iterative sampling for generation starting with any percentage of prior ([0, 100%]).

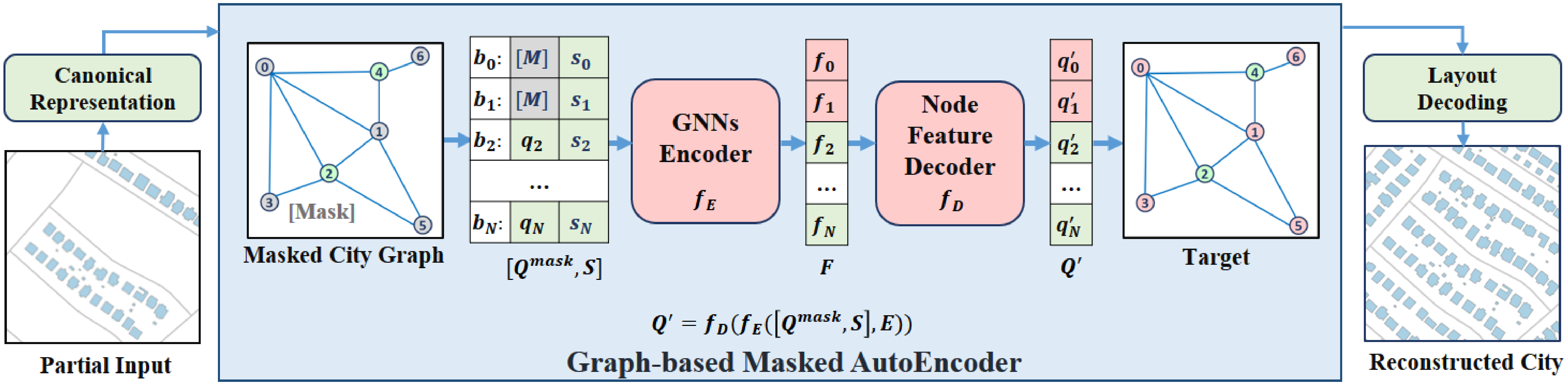
1. Canonical Representation for Arbitrary Layouts



The entire graph G corresponds to a city. Sub-graphs map to communities. Each node *b* represent a city block. Further, each node encodes its building layouts, shapes, and heights as a quantized feature by well-trained quantizer.

Example: a city with ***N*** nodes and ***K*** edges is represented by a total of (***516N + K***) variables.

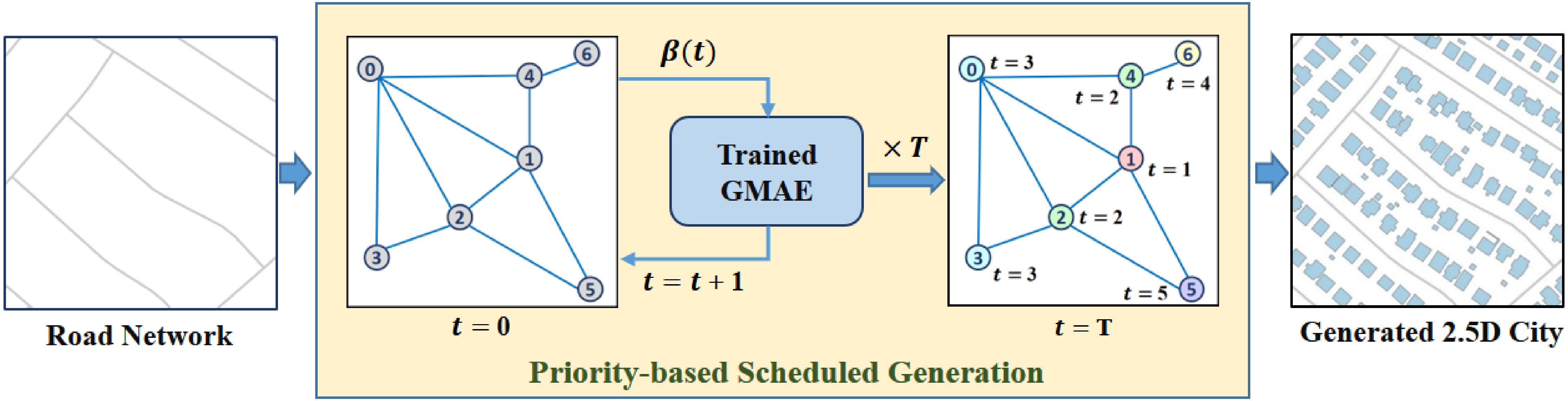
2. GMAE Enabling Context-Sensitive Learning



GMAE learns the mutual relations between sub-graphs and smaller components. It captures the context relationship within the components of the graph *G.*

Training: self-supervised with dynamic masking ratios [0.5, 1.0] using **17.5M buildings** and **1M city blocks** spanning **330 cities** containing thousands of communities.

3. Priority-based Iterative Generation



Priority-based iterative generation produces city blocks with high confidence first and the rest afterwards. This hold for any percentage of starting prior. It aims to keep the best fidelity and diversity.

Sampling speed: In each iteration, the subset $\beta(t) = 1 - \cos(t/T)$ nodes are accepted, producing a full graph after T iterations.

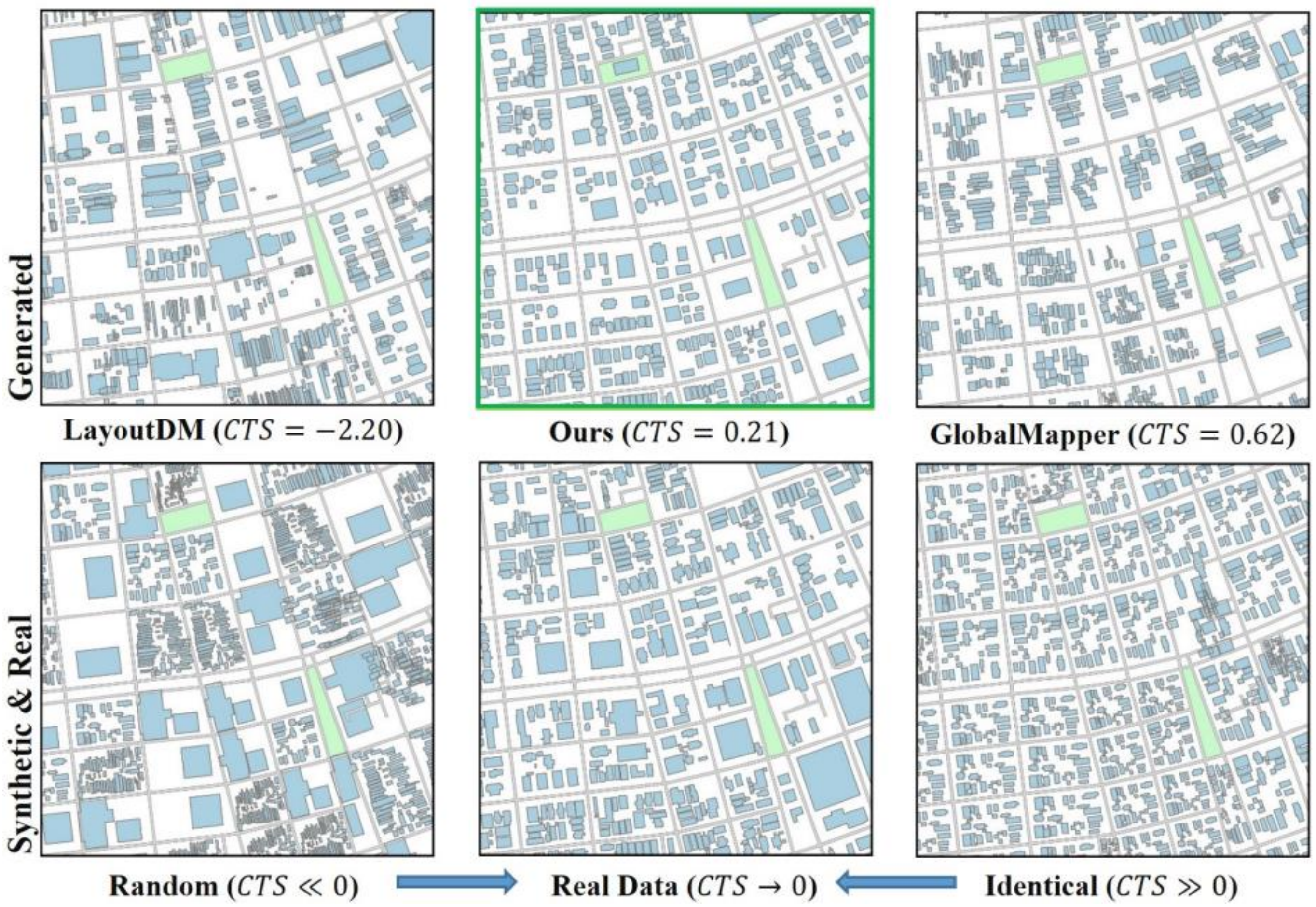
Results: Context Score

Context Score (CTS) measures contextual diversity:

- **CTS < 0 is over-diversity**
- **CTS > 0 is over-similar**
- **CTS = 0 is same as ground-truth**

$$CT = \frac{1}{\|N(i)\|} \sum_{j}^{N(i)} LayoutSim(b_i, b_j), \quad N(i) = \{j \mid (i, j) \in E\}$$

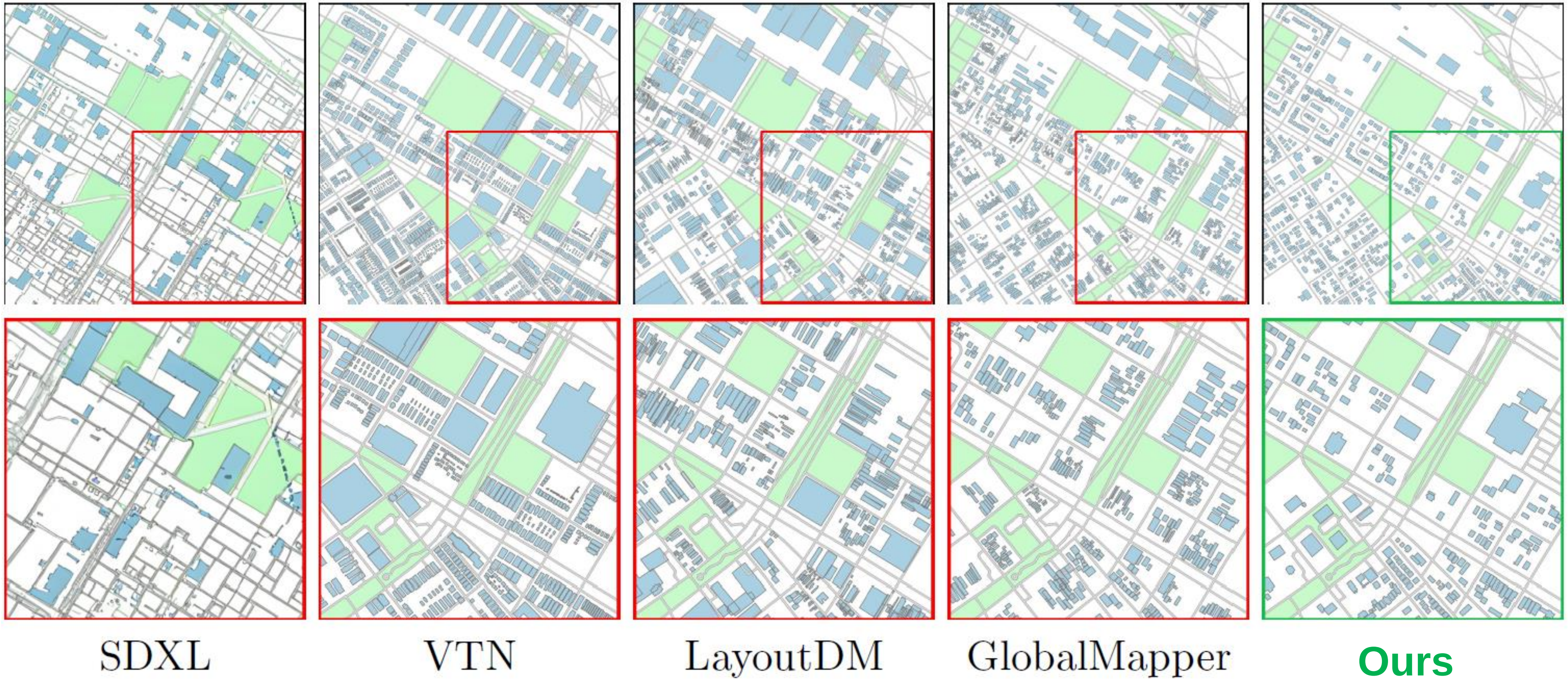
$$CTS = CT_{gen} - CT_{real}$$



PURDUE UNIVERSITY Department of Computer Science

ECCV EUROPEAN CONFERENCE ON COMPUTER VISION MILANO

Results: Qualitative Comparisons



Our results improve on large-scale support, realism, diversity, and contextual harmonization.

Results: Quantitative Comparisons

Table 1: Quantitative Comparisons. All methods are compared to the same real urban layouts (except SDXL [50] which cannot take-in a road network). Best values are in bold, second best values are underlined. Our method outperforms other existing methods in all but the overlap metric. See text for an explanation of the metrics.

Method	$CTS_{x\rightarrow 0}$	WD-5D↓	WD-CO↓	Overlap↓	O-Blk↓	FID↓	KID↓	LPIPS↓
SDXL [50]	-	-	-	-	-	120.24	0.079	0.48
VTN [3]	-1.14	3.18	5.81	**1.24**	7.35	69.14	0.047	<u>0.32</u>
LayoutDM [31]	-2.20	<u>2.92</u>	12.50	4.56	1.72	66.77	0.040	0.39
GlobalMP [21]	<u>0.62</u>	4.77	<u>4.14</u>	2.52	<u>0.68</u>	<u>49.55</u>	<u>0.024</u>	0.34
Ours	**0.21**	**2.28**	**1.91**	<u>1.27</u>	**0.42**	**23.63**	**0.005**	**0.20**

Our results improve on context sensitivity, fidelity, quality.

Applications: Socio-Economic Metric Prediction

Table 4: GMAE-based Prediction. Our GMAE combined with a conventional classifier (e.g. SVM or XGBoost) can be used to predict whether a set of city blocks corresponds to an advantaged and disadvantaged economic/social/environmental group.

Abbrv.	Metric Full Name	Best Acc. %
DSF_PFS	Diesel particulate matter exposure (percentile)	**89.76**
EBF_PFS	Energy burden (percentile)	**84.26**
LMI_PFS	Low median household income as a percent of area median income (percentile)	**83.62**
LLEF_PFS	Low life expectancy (percentile)	**80.30**
EBLR_PFS	Expected building loss rate (Natural Hazards Risk Index) (percentile)	**80.13**
LPF_PFS	Percent pre-1960s housing (lead paint indicator) (percentile)	**79.83**
HBF_PFS	Housing burden (percent) (percentile)	**76.82**
EPLR_PFS	Expected population loss rate (Natural Hazards Risk Index) (percentile)	**75.82**
P100_PFS	Percent of individuals $<$ 100% Federal Poverty Line (percentile)	**75.46**
FLD_PFS	Share of properties at risk of flood in 30 years (percentile)	**75.46**
TF_PFS	Traffic proximity and volume (percentile)	**75.15**

PURDUE UNIVERSITY | Department of Computer Science

Applications: Customized Urban Layout Editing

New Orleans

Future Work

1. Large-scale photo-realistic multi-view scene synthesis.
2. Digital Twin, City-scale 3D modeling.
3. Synthetic data generation in autonomous driving and world model.

PURDUE UNIVERSITY
Department of Computer Science

ECCV
EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO 2024

THANK YOU!

Project Website:

Author: