Controllable Road Marking Generation

Zhiyu Joey Cai, Yufan Zhang, Ruichen Tan, Zengxiang Lei, Satish Ukkusuri

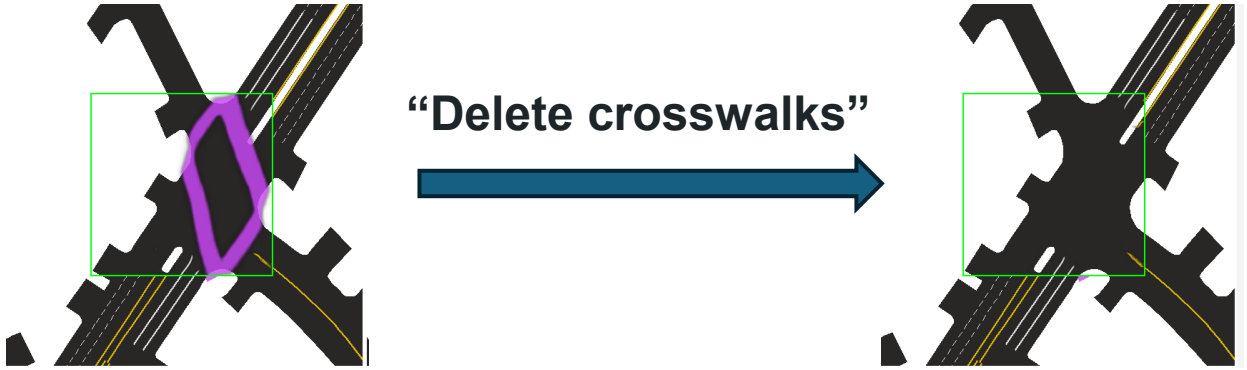
Purdue University

Background & Motivation

Road markings determine the behavior of road users.

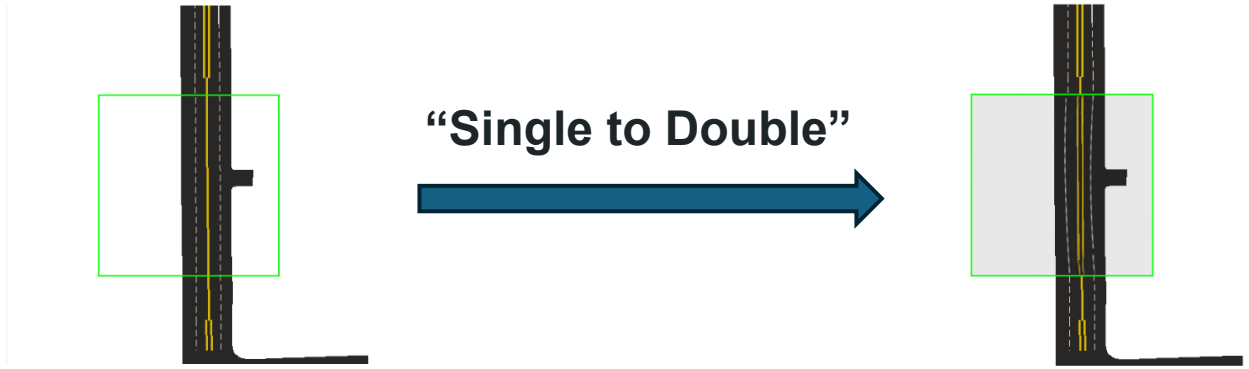
Why vary the markings?

Simulation needs different layouts.
Manual map edits are hard to scale.



The missing capability

Go beyond one observed map:
generate plausible alternatives
that follow a user’s request.



Can AI learn to design road markings from context and traffic intent?

Task: Generate the markings in target area

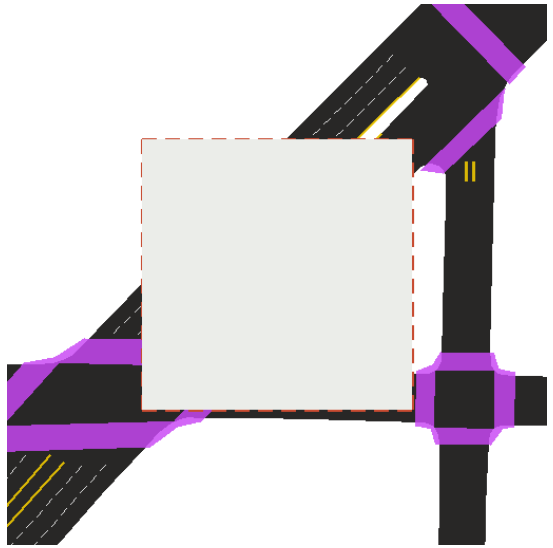
Given: 100 m tile · 50 m center

Drivable area



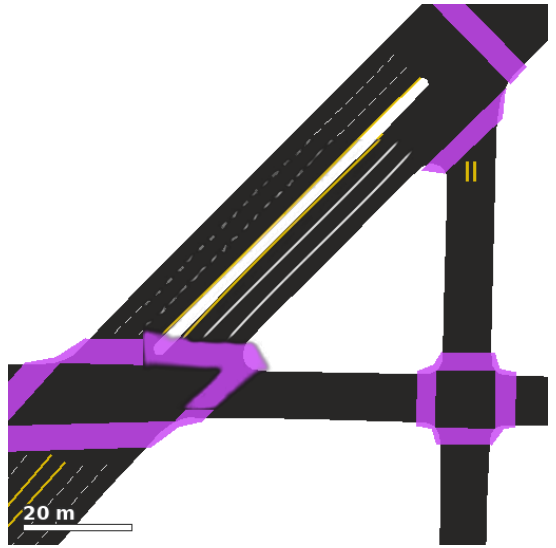
+

Optional outer-ring context



→

Generated markings



Text prompt “two-way road, solid dividers, with crosswalks”

3 channels

Lane divider

Road divider

Crosswalk

Target output: Appropriate lane dividers, road dividers, and pedestrian crossings from scratch

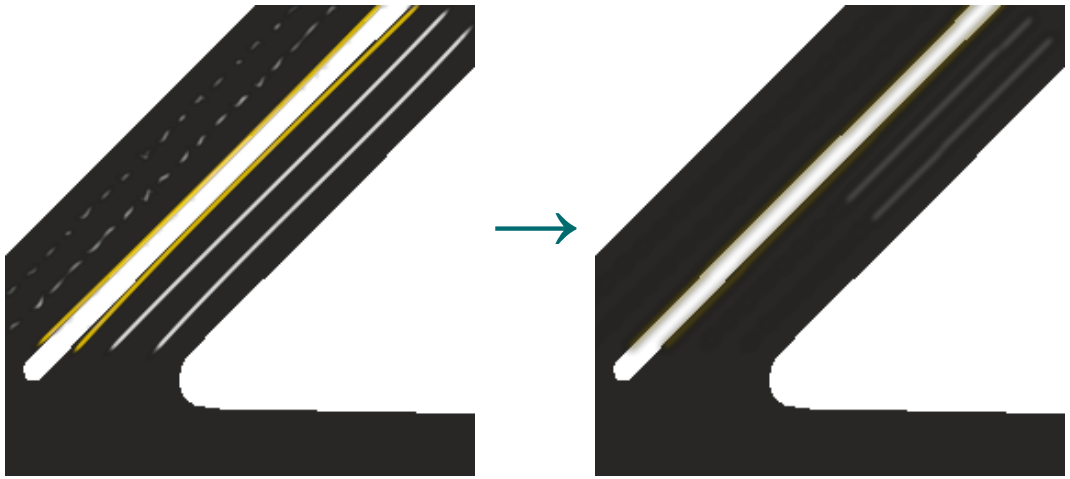
Challenges & Our solutions

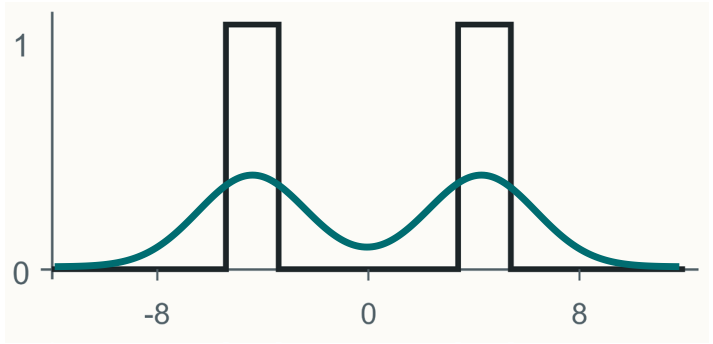
Thin dividers are sparse, and conditions are complex.

Blur the training targets

Sharp target

Blurred target





Conditional rectified-flow DiT

Road geometry + marking context

Frozen CLIP text embeddings

↓

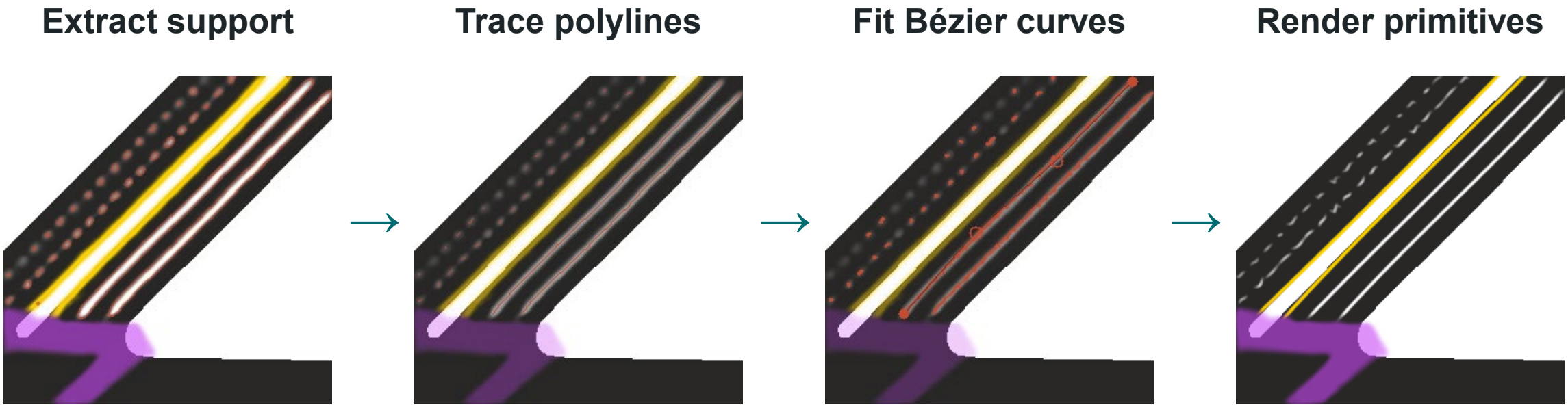
Rectified-flow DiT

Geometry anchors the layout.
Text steers marking attributes.

Topology-aware supervision encourages divider continuity

Challenges & Our solutions

Structured Gaussian Render: a training-free refinement step



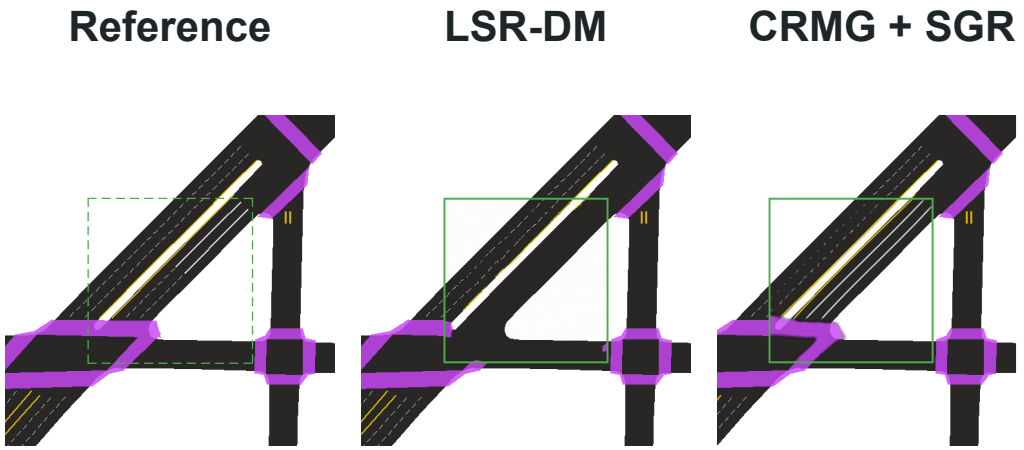
Steps: ①fit curves to the generated layout ②Render thin dividers

Refined channels: lane divider, road divider
Crosswalks are kept since they are raster masks

Results

Argoverse 2[1]: 4,597 test tiles

Method	Buffered F1	clDice
LSR-DM (adapted) [2]	38.8	24.6
CRMG + CNN deblur	62.7	46.3
CRMG + SGR	**76.8**	**50.2**



Green box: generated center region

+14.1 F1 points: SGR vs. CNN deblur

Waymo: 88.0 F1 / 66.2 clDice. Separately trained and evaluated

[1] Wilson, Benjamin, et al. "Argoverse 2: Next generation datasets for self-driving perception and forecasting." *arXiv preprint arXiv:2301.00493* (2023).
[2] Ruiz, Antonio, et al. "Lane Graph Extraction from Aerial Imagery via Lane Segmentation Refinement with Diffusion Models." *Remote Sensing* 17.16 (2025): 2845.

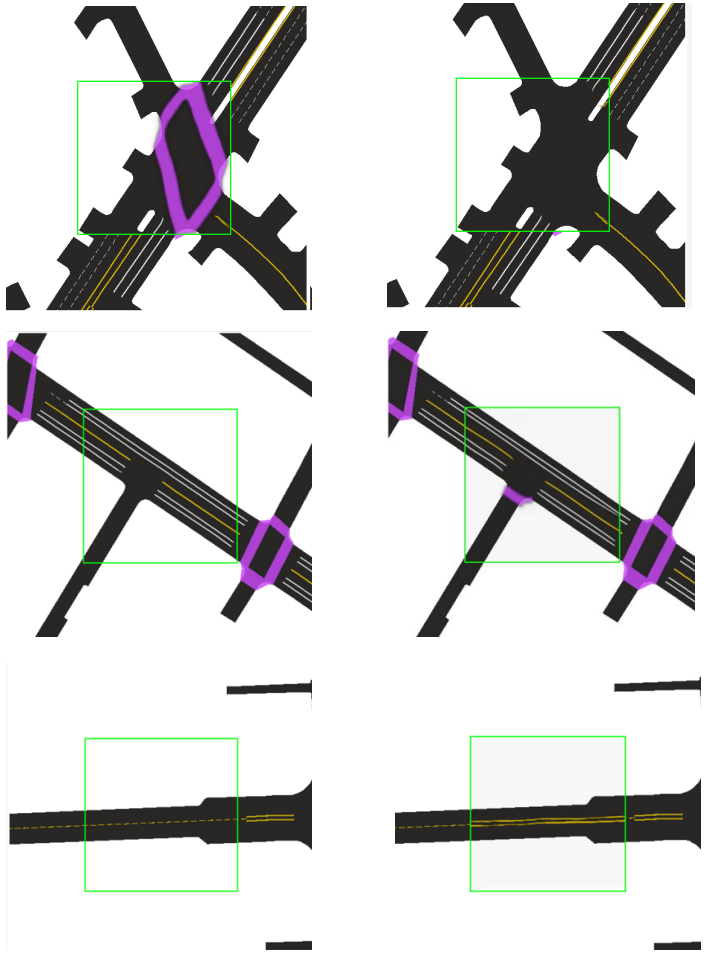
Results - Text controllability

Change only the prompt; keep geometry, context, and seed fixed.

Delete crosswalks

Add crosswalks

Single to double



Edit	Success	False act.
Add crosswalk	49.6%	9.2%
Delete crosswalk	75.1%	19.7%
Single → double	31.0%	7.4%

Divider success uses a skeleton-length proxy.

75.1% success for crosswalk deletion in eligible tiles.

What our work contributes

Task + benchmark A benchmark for geometry- and text-conditioned markings

Method

- Rectified diffusion model learns layouts
- Training-free module recovers sharp edge geometry

Evidence 76.8 F1 on AV2, plus measured prompt-driven edits

Toward controllable road-map variation for simulation

Paper

Code

Thank you

github.com/RU0804/CRMG