

Control



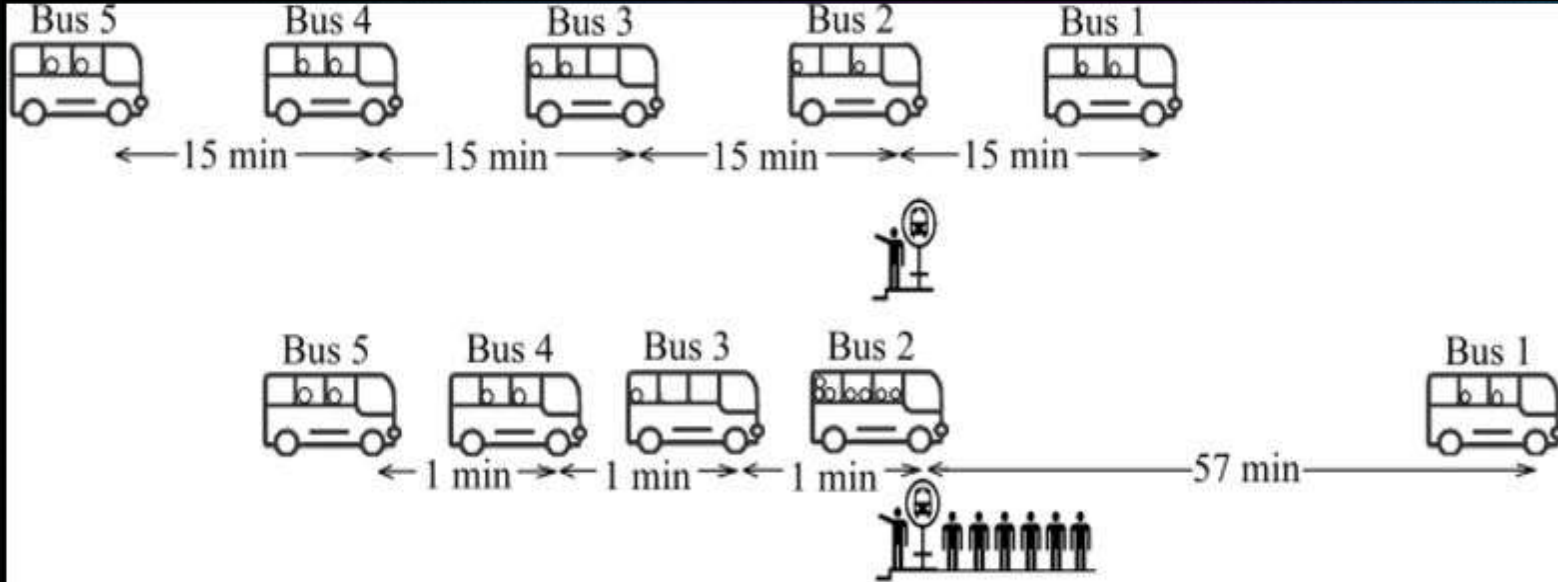
Assist

# AI for Bus Dispatching to Improve Reliability

*Joseph Rodriguez, Haris Koutsopoulos & Jinhua Zhao*



# Bus bunching



# Control strategy

**At the bus terminal,**

- 1. Hold the bus by 1 min, 2 mins, 3mins, ...**
- 2. Move departure time forward by 1 min, 2 mins, 3 mins,...**

Transit agencies experience major service reliability issues caused by driver shortage and driver absenteeism

TransitCenter |

# Bus Operators in Crisis

**The Steady Deterioration of One of Transit's Most Essential Jobs,  
and How Agencies Can Turn Things Around**

**CTA Service  
delivered: 84%**

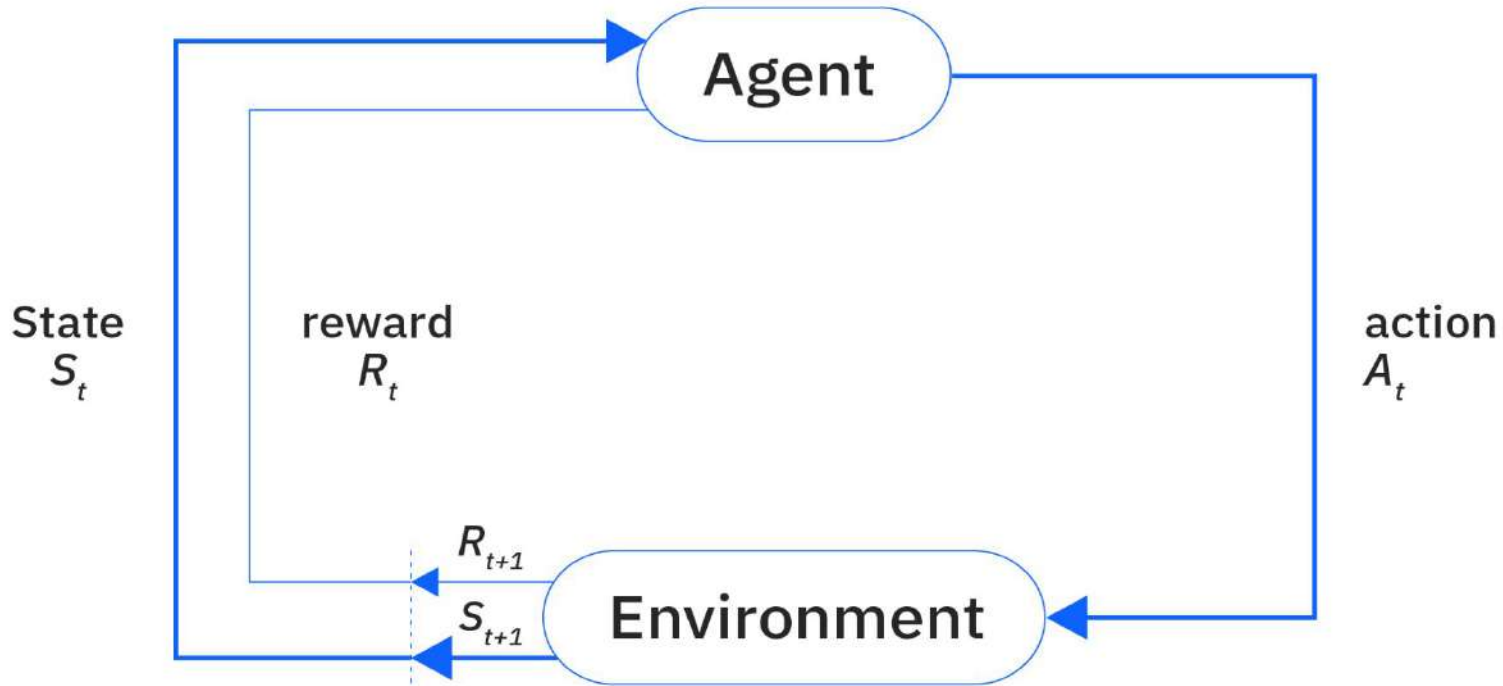
# Improve service reliability through bus control using reinforcement learning

## Robustness to uncertainties:

- Demand uncertainty: # of passengers
- Supply uncertainty: driver absenteeism

---

- Environment uncertainty: traffic conditions
- Operation uncertainty: non-compliance
- Information uncertainty: Incomplete or inaccurate



# Human in the loop #1: Bus route 81 (actual) vs. 20 (initial)

**Terminal:** Jefferson Park, serves 10 routes

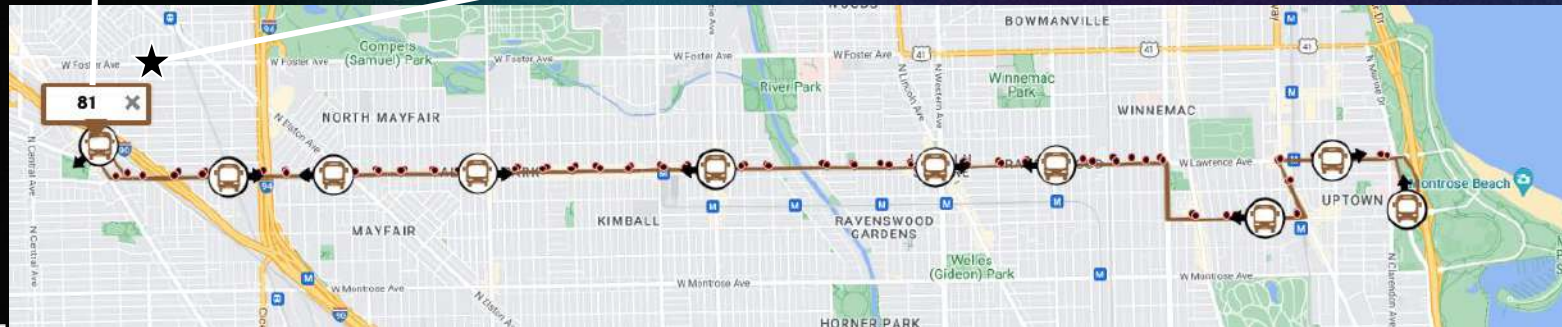


**Garage:** Forest Glen, North Chicago routes

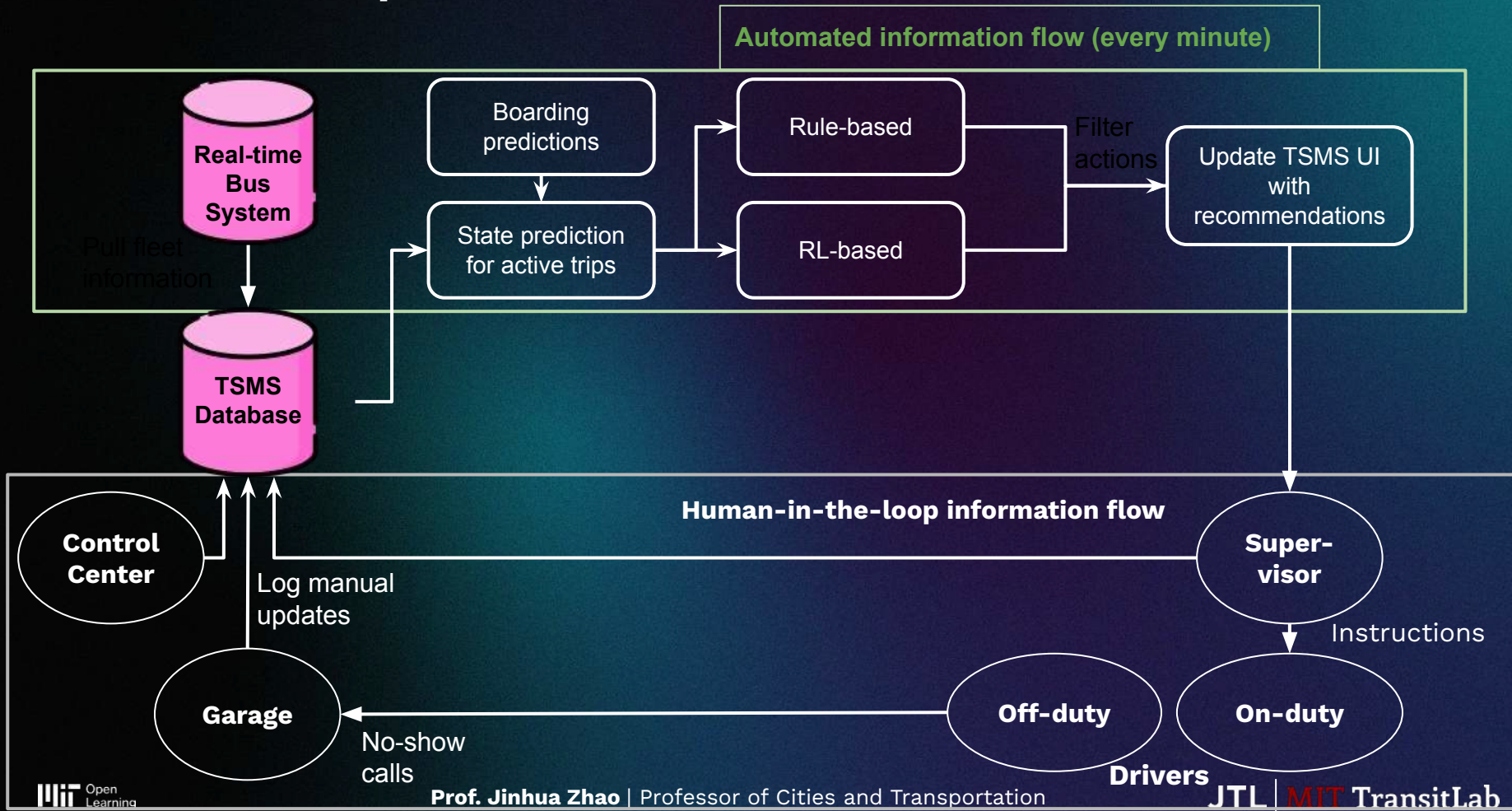


**Route:**

- 4 transfers to subway
- 50 stops each direction
- AM peak: 6-9 mins
- 15-25% missing trips
- 7,500 riders per weekday



# Human in the loop #2: AI recommendation → Human control execution



# Human in the loop #3: Field experiment in Chicago

Substantive support from the Chicago Transit Authority (CTA)

- President and CAA
- VP operation
- Garage supervisors
- Drivers



MIT developed the real-time control TSMS database, algorithm, and user interface

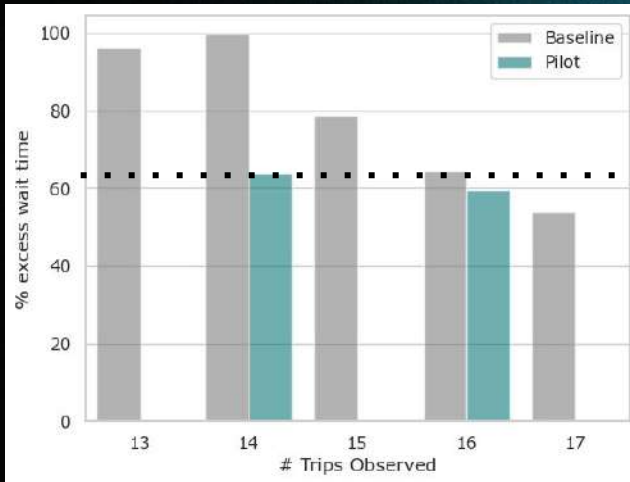
# Impacts

Excess Waiting Time: reduced >30%

# Impacts of Our Control Strategy: Performance Evaluation

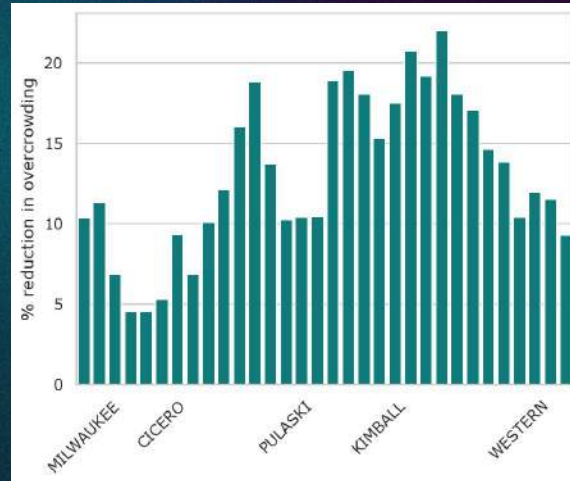
## Access wait time by # of bus trips in AM period (6-9am)

37% reduction,  
equivalent to adding 2 trips



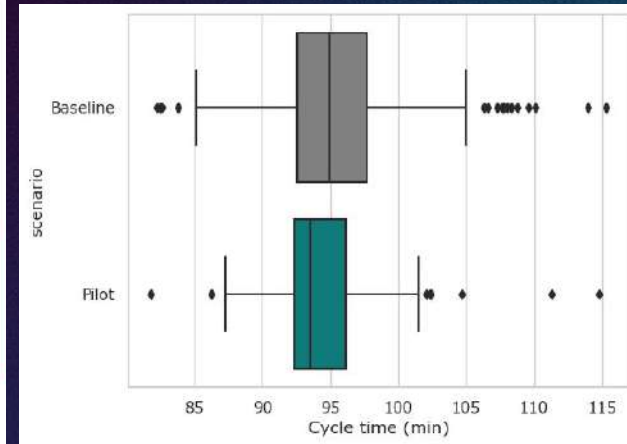
## Overcrowding

5-20% reduction



## 90<sup>th</sup> Cycle Times

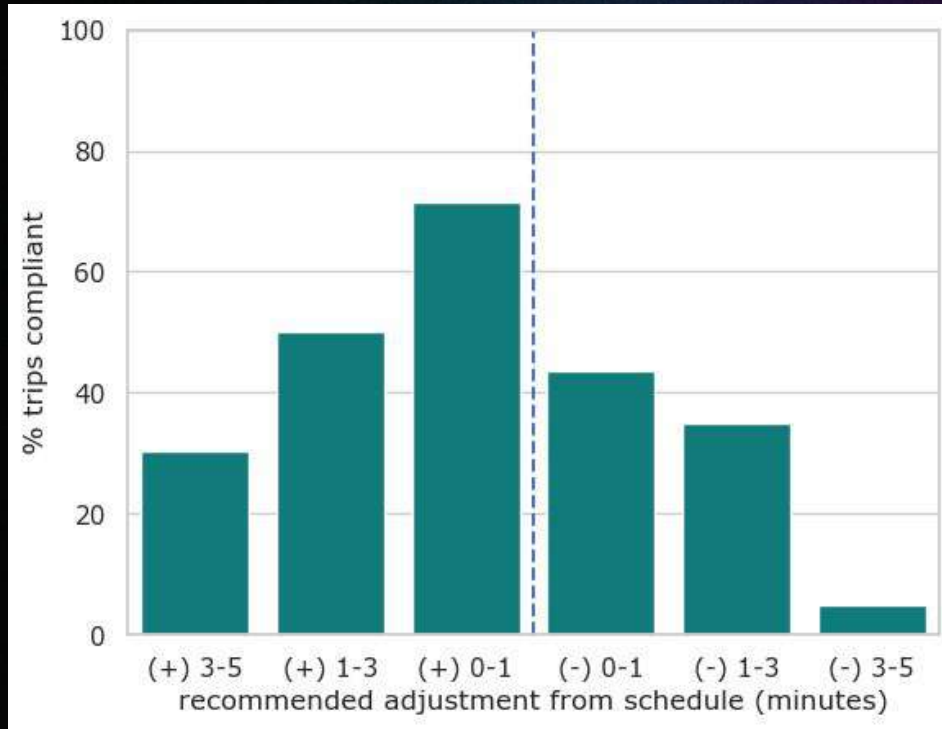
2.5 minutes reduction  
(2.4%) reduction



## Observation #3

Do drivers follow the AI guidance?

# Human in the loop #4: Do drivers follow A.I.?



Driver compliance decreases with the aggressiveness of the recommended control actions.

## Robustness to operational uncertainties

	Average excess wait time reduction (mins)		
Baseline (minutes)	Trips with information issues	Trips with non-compliant drivers	Trips with compliant drivers
3.6	Reduce 0.7 mins (19%)	Reduce 0.7 mins (19%)	Reduce 2.1 mins (58%)
	90th percentile headways reduction (mins)		
27.6	Reduce 7.2 mins (26%)	Reduce 5.8 mins (21%)	Reduce 11.2 mins (40%)

# Robustness to uncertainties: Driver compliance

	90th percentile headways	
Baseline (minutes)	Trips with non-compliant drivers	Trips with compliant drivers
27.6	21.8 (-21%)	16.4 (-40%)

## Knowledge Transfer: MIT → Chicago

Category	Pilot (BP2)	Full-scale (BP3)
Time period	1 week	1 month
Routes	81	All routes at the Jefferson Park terminal
Control actions	Terminal dispatch	Mid-route holding, dynamic interlining, short-turning
Implementation	RAs overseeing operations	CTA staff takes over the software

## Scaling Up: the US transit industry at large

The graphic features the MIT logo, Mobility Initiative, and Transit Lab logos at the top. Below them, it says "Introducing the first ever" followed by "Transit Research Consortium 2023". A map of the United States shows red dashed lines connecting seven major transit agencies: King County METRO, SFMTA, LA Metro, Chicago Transit Authority, MIT (circled in red), Massachusetts Bay Transportation Authority, MTA New York City Transit, and Washington Metropolitan Area Transit Authority. A red dashed box at the bottom of the map contains the text: "Collective Research Sharing among MIT and 7 of the Largest US Transportation Agencies".

# Observation #4

## Human in the loop

- Control level
- Garage level
- Organizational level



Create



Work

# AI for Generative Urban Design

*Qingyi Wang, Shenhao Wang, Mingyi He, Yuebing Liang & Jinhua Zhao*



# The New York High Line

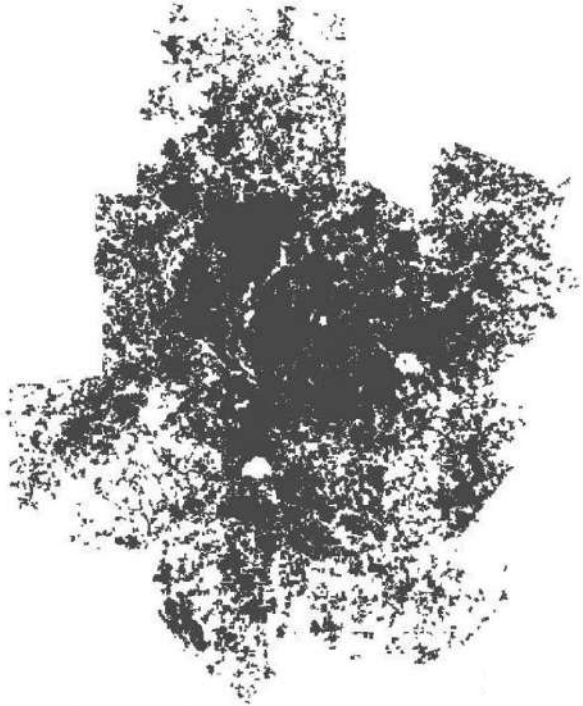


# Decarbonize Mobility: Urban structure is the starting point

**Atlanta, USA**

**Population: 2.5 million (1990)**

**Built-up Area: 4,280 sq. km.**



**Barcelona, Spain**

**Population: 2.8 million (1990)**

**Built-up Area: 162 sq. km.**



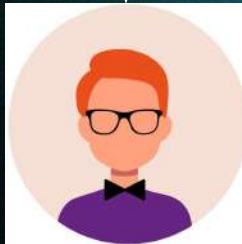
**(15 minute city)**

# Typical planning process

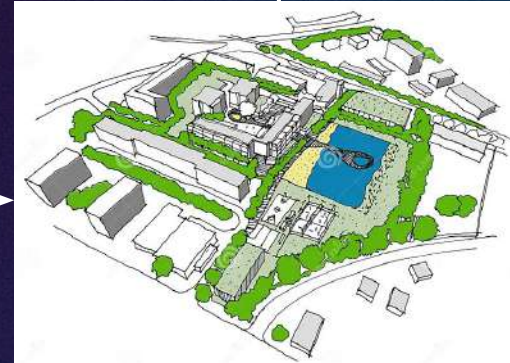
Zoning, Policy, Traffic,  
Social interactions,  
Natural Environment,  
Infrastructure, Public  
Engagement, .....

Feedback and Iterate (weeks and months)

Generate plans, layouts and  
visualizations



Urban Planner



[https://www.123rf.com/photo\\_16135751\\_sketch-of-a-new-development-urban-idea.html](https://www.123rf.com/photo_16135751_sketch-of-a-new-development-urban-idea.html)

# The human-machine planning collaboration

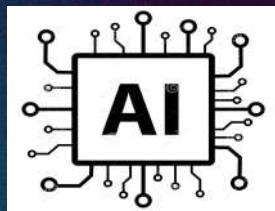
Zoning, Policy,  
Traffic, Social  
interactions...



Human Guidance

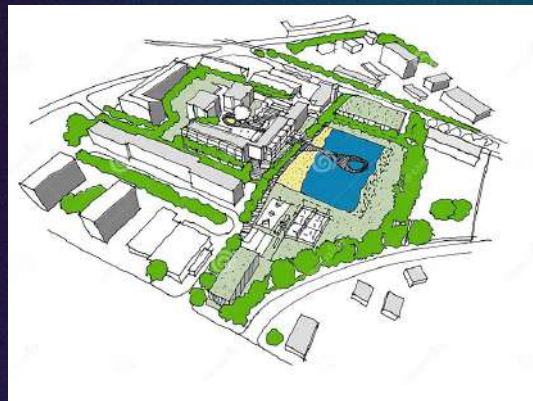


Give me a design in Chicago  
with 70% apartments, 10%  
park, 10% retail, and a lot of  
green space!



Generative AI

Generate  
Designs



Feedback and Iterate (minutes)

# What do planners want?

# An Example

# A piece of land 35 mile north of downtown Chicago



A target land use

- 35% residential
- 15% commercial
- 10% open parking

# What do planners want?

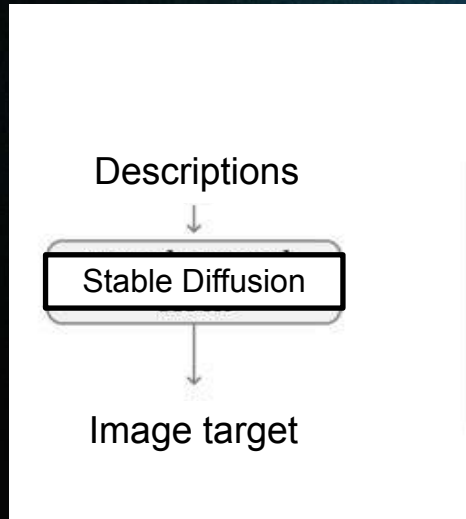
## Expectations of Generative Planning Tools

1. Having **control** over AI-generated land-use patterns
2. **Respecting** existing infrastructure and natural environment
3. Learn and apply **different urban textures**

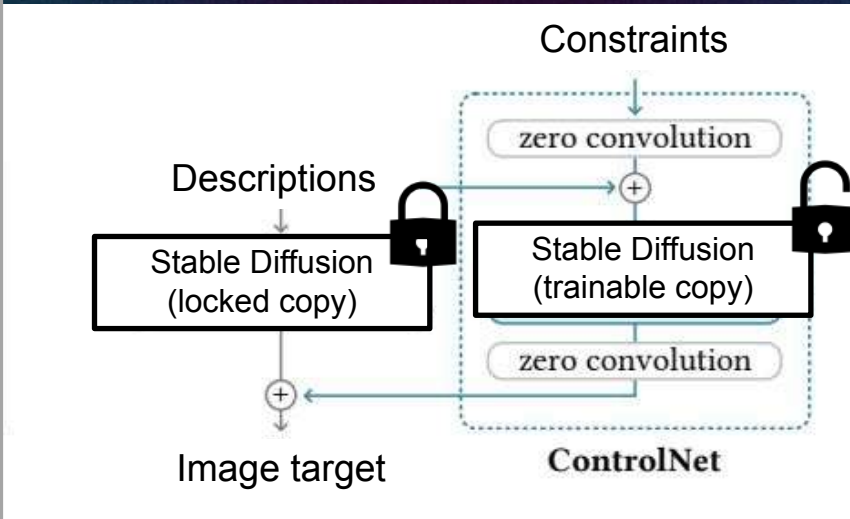
Can A.I. do it?

# Fine-tuning stable diffusion using ControlNet

## Stable Diffusion



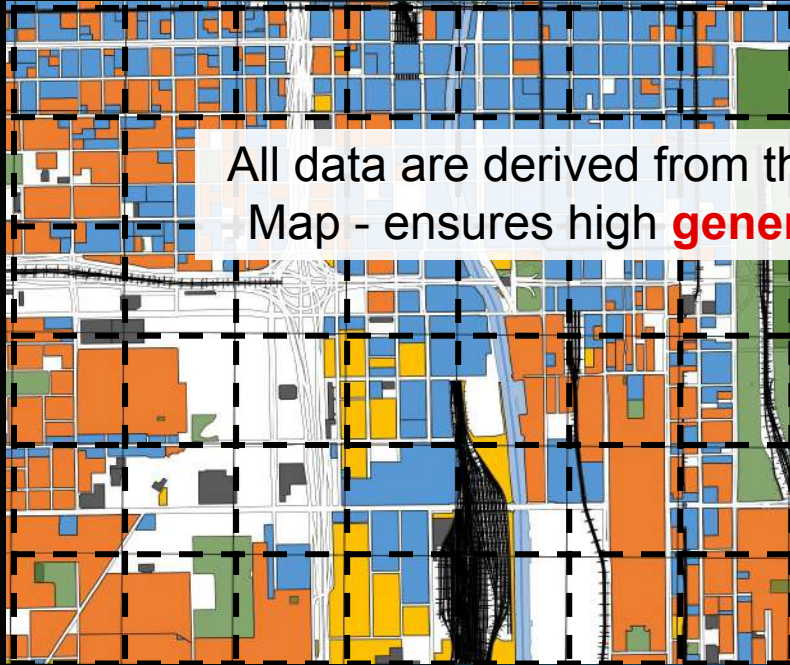
## Stable Diffusion with ControlNet



# Data: Open Street Map



Land use patterns



Existing infrastructure and natural environment



All data are derived from the publicly available Open Street Map - ensures high **generalizability** and **reproducibility**

Residential

Commercial

Park

Industrial

Open parking

Farmland

b

# A piece of land 35 mile north of downtown Chicago



- A target land use
- 35% residential
  - 15% commercial
  - 10% open parking



# A piece of land 35 mile north of downtown Chicago



- A target land use
- 65% industrial
  - 20% commercial
  - 5% open parking



# A piece of land 35 mile north of downtown Chicago



- A target land use
- 100% residential



# 1. Having **control** over AI-generated land-use patterns



35% residential  
15% commercial  
10% open parking



65% industrial,  
20% commercial,  
5% open parking



100% residential

## 2. Respecting **existing** infrastructure and natural environment

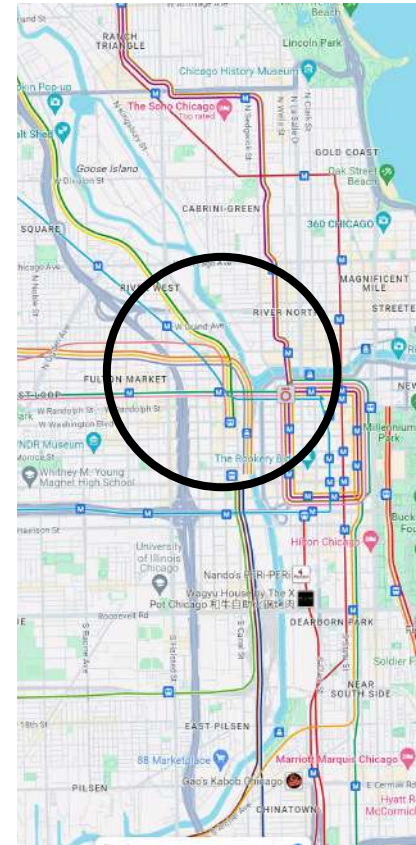
Descriptions

City in Chicago.  
35% residential  
20% commercial,  
5% open parking.  
Mix of houses and apartments.  
Dense building coverage.

Constraints



Generated Aerial Imagery

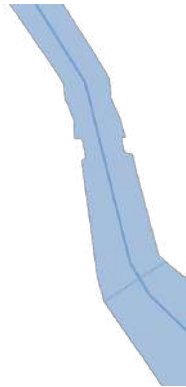


## 2. Respecting existing infrastructure and natural environment

Descriptions

City in Chicago.  
35% residential  
20% commercial,  
5% open parking.  
Mix of houses and apartments.  
Dense building coverage.

Constraints



Add waterways



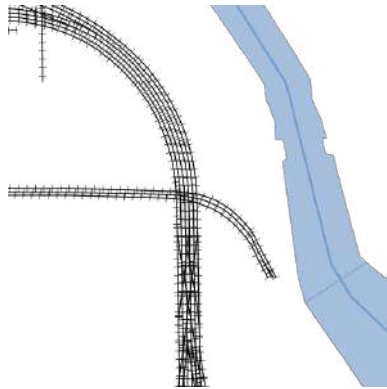
Generated Aerial Imagery

## 2. Respecting existing infrastructure and natural environment

Descriptions

City in Chicago.  
35% residential  
20% commercial,  
5% open parking.  
Mix of houses and apartments.  
Dense building coverage.

Constraints



Add railways



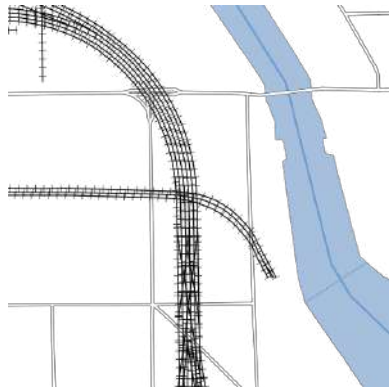
Generated Aerial Imagery

## 2. Respecting existing infrastructure and natural environment

Descriptions

City in Chicago.  
35% residential  
20% commercial,  
5% open parking.  
Mix of houses and apartments.  
Dense building coverage.

Constraints



Add roads



Generated Aerial Imagery



True Satellite Image

### 3. Learn and apply **different urban textures**

Descriptions

City in {city name}.

40% residential

15% industrial,

15% commercial,

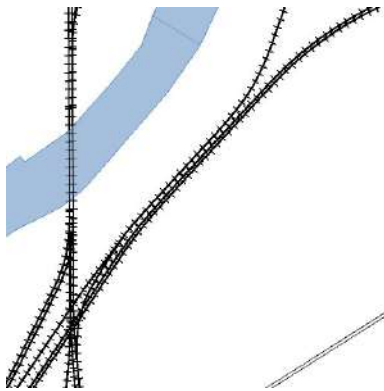
10% park,

5% open parking.

All houses.

Medium building coverage.

Constraints



Chicago



Dallas



Los Angeles

# What do planners want?

## Expectations of Generative Planning Tools

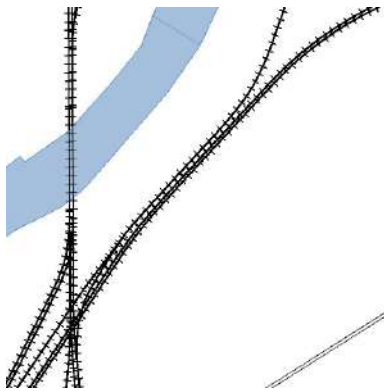
1. Having **control** over AI-generated land-use patterns
2. **Respecting** existing infrastructure and natural environment
3. Learn and apply **different urban textures**

# Bonus: Producing **alternate** designs for the same inputs

Descriptions

City in Chicago.  
40% residential  
15% industrial,  
15% commercial,  
10% park,  
5% open parking.  
All houses.  
Medium building coverage.

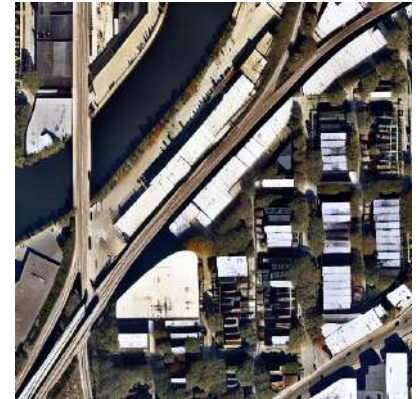
Constraints



Design 1



Design 2



Design 3

How do planners and AI work together?

# Generative Urban Design

## A Stepwise Approach Integrating Human Expertise with Diffusion Models

**Mingyi He**

([mingyihe@mit.edu](mailto:mingyihe@mit.edu))

Massachusetts Institute of Technology

He, M., Liang, Y., Wang, S., Zheng, Y., Wang, Q., Zhuang, D., Tian, L., & Zhao, J. (2025). *Human-guided urban form generation using multimodal diffusion models*. **Building and Environment**, Article

113892. <https://doi.org/10.1016/j.buildenv.2025.113892>

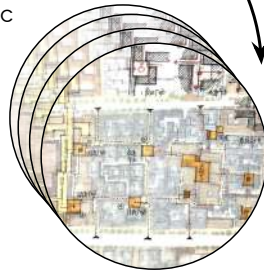


# Background / Urban Design

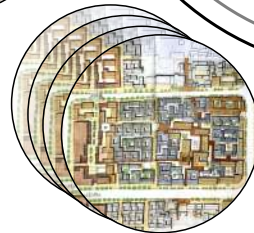
Urban design is

**iterative, labor-intensive, time-consuming.**

**Design Objectives**  
+  
**Existing Constraints**  
infrastructure, natural environment, and sociodemographic



**Design Drafts**



**Design Schemes**

**Adjusted Design Strategies**  
+  
**Objectives**

**Public Engagement**

**Communicate with a multitude of stakeholders for feedback and iteration,**

including local government, community residents, urban planners, and real estate developers

**Design Proposal**

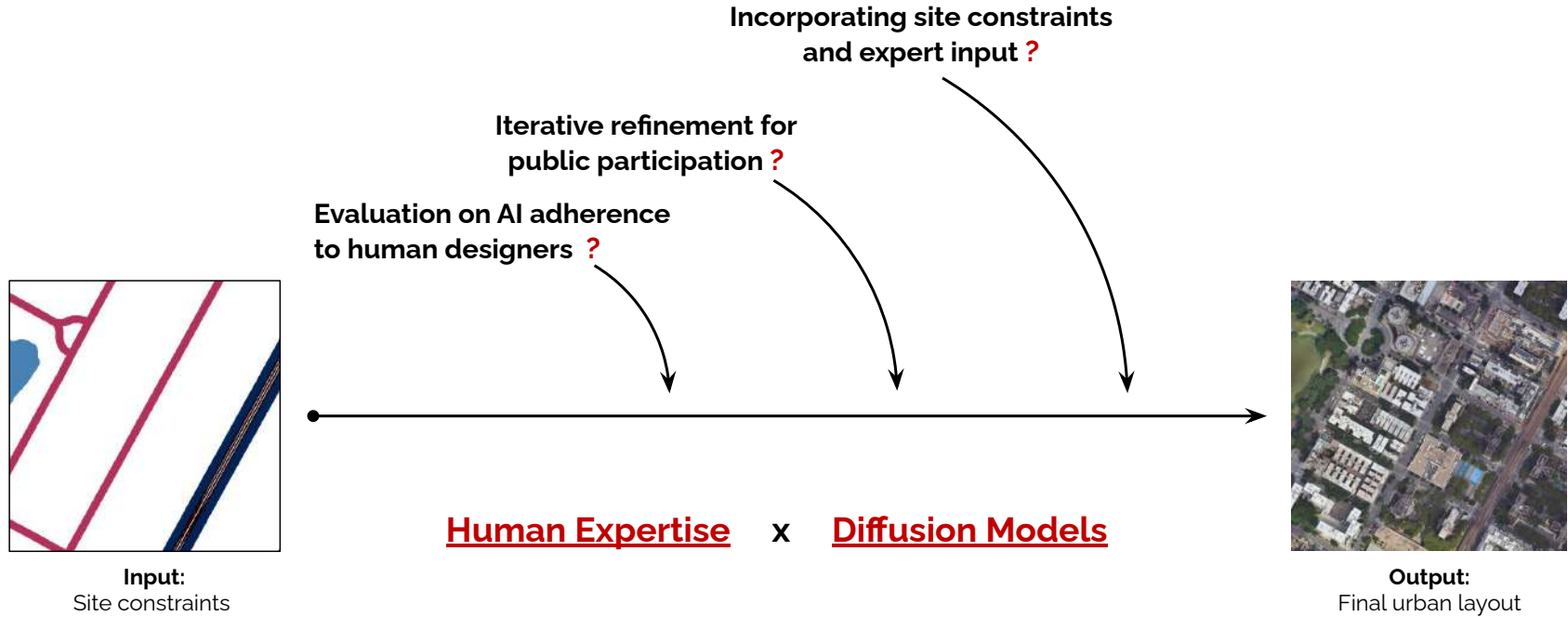
**Visualize urban landscapes,**  
which provide intuitive perspectives to the non-experts

**Final Design Proposal**

**Urban Design Process**

## Research Gap / **Generative Urban Design**

GenAI has potential to **enhance design efficiency, support participatory design...**  
Existing approaches **fall short of addressing real-world design complexities...**



# Research Framework / Human-Centered AI for Urban Design

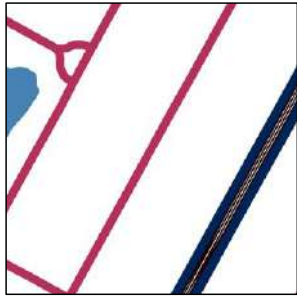
## A Stepwise Approach Integrating Human Expertise with Diffusion Models

### Human-in-the-loop Design Process:

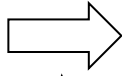
Integrates human expertise with AI capabilities.

Human designers maintain control over the design process, enabling iterative adjustments and refinements at each stage.

#### Stage 1 - Road Network and Land Use Planning



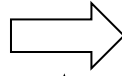
**Input 1**  
Site constraints (skeleton roads + railway + water)



Human Instructions



**Output 1 / Input 2**  
Cohesive road network and land use plan

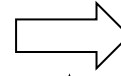


Human Instructions

#### Stage 2 - Building Layout Planning



**Output 2 / Input 3**  
Spatial arrangement of buildings



Human Instructions

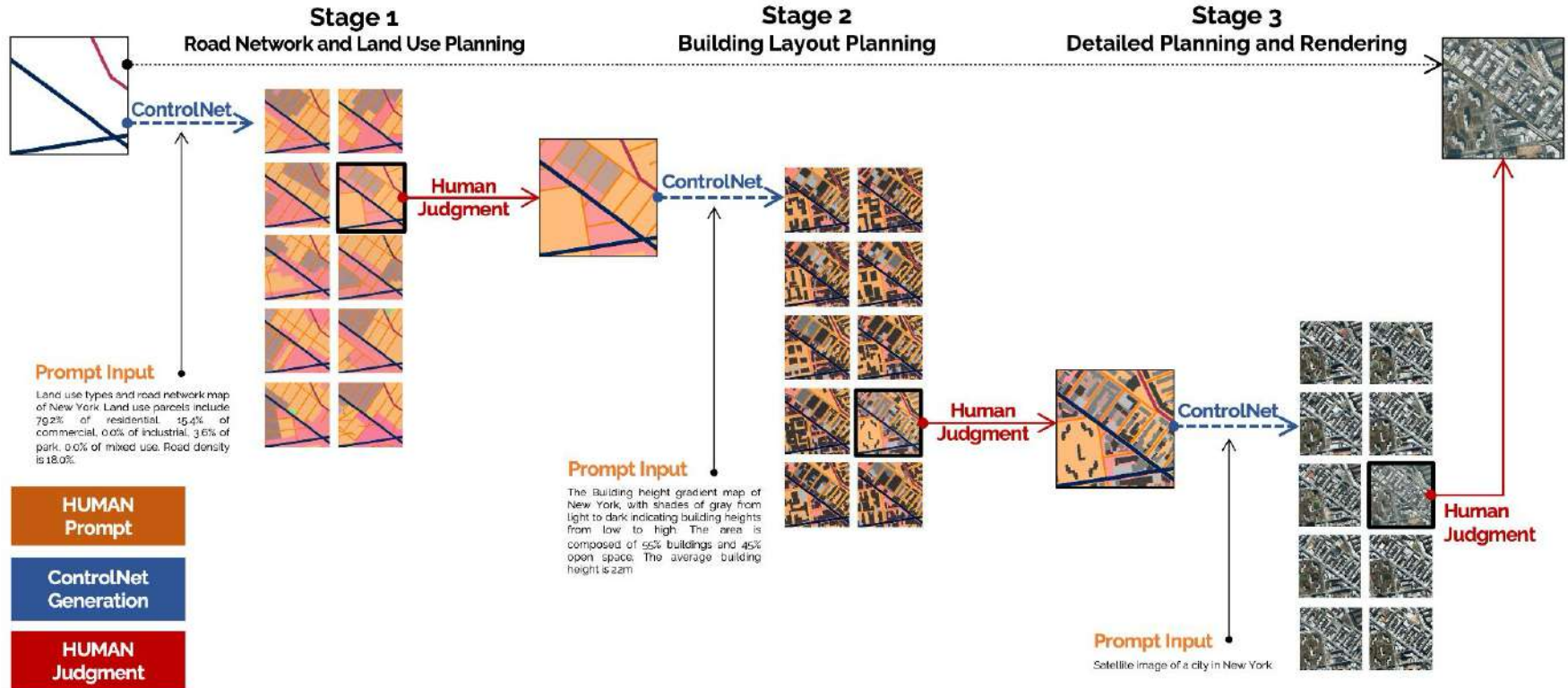
#### Stage 3 - Detailed Planning and Rendering



**Output 3**  
Final urban layout (e.g., satellite image)

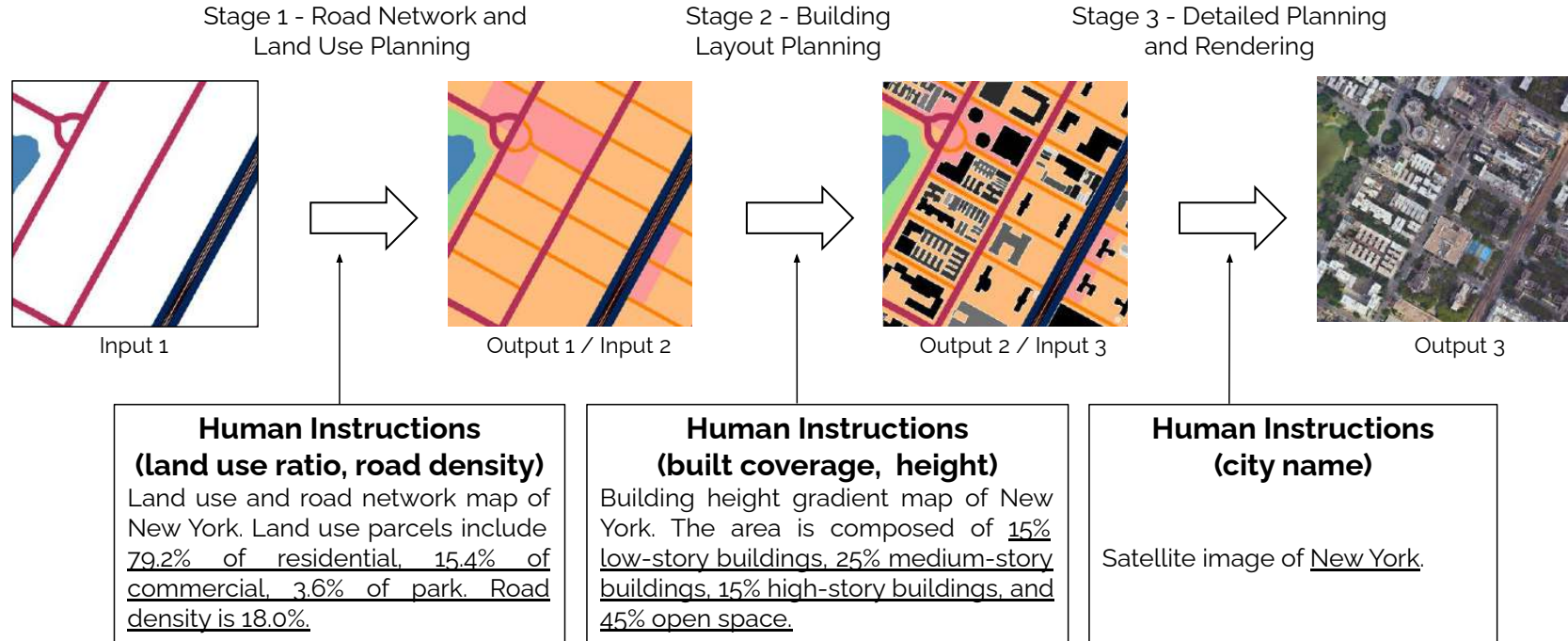
# Research Framework / Human-Centered AI for Urban Design

## A Stepwise Approach Integrating Human Expertise with Diffusion Models



# Research Framework / Human-Centered AI for Urban Design

## A Stepwise Approach Integrating Human Expertise with Diffusion Models



# Method / Data Processing

## Study Area:

- New York City
- Chicago

## Data Source:

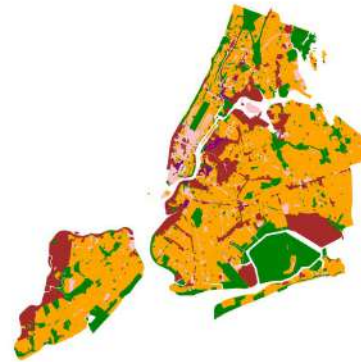
- OpenStreetMap (OSM)
- City official data hub
- Mapbox



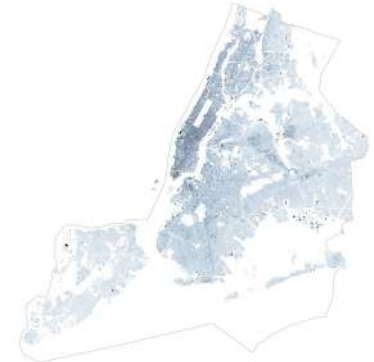
Basic data



Road network data

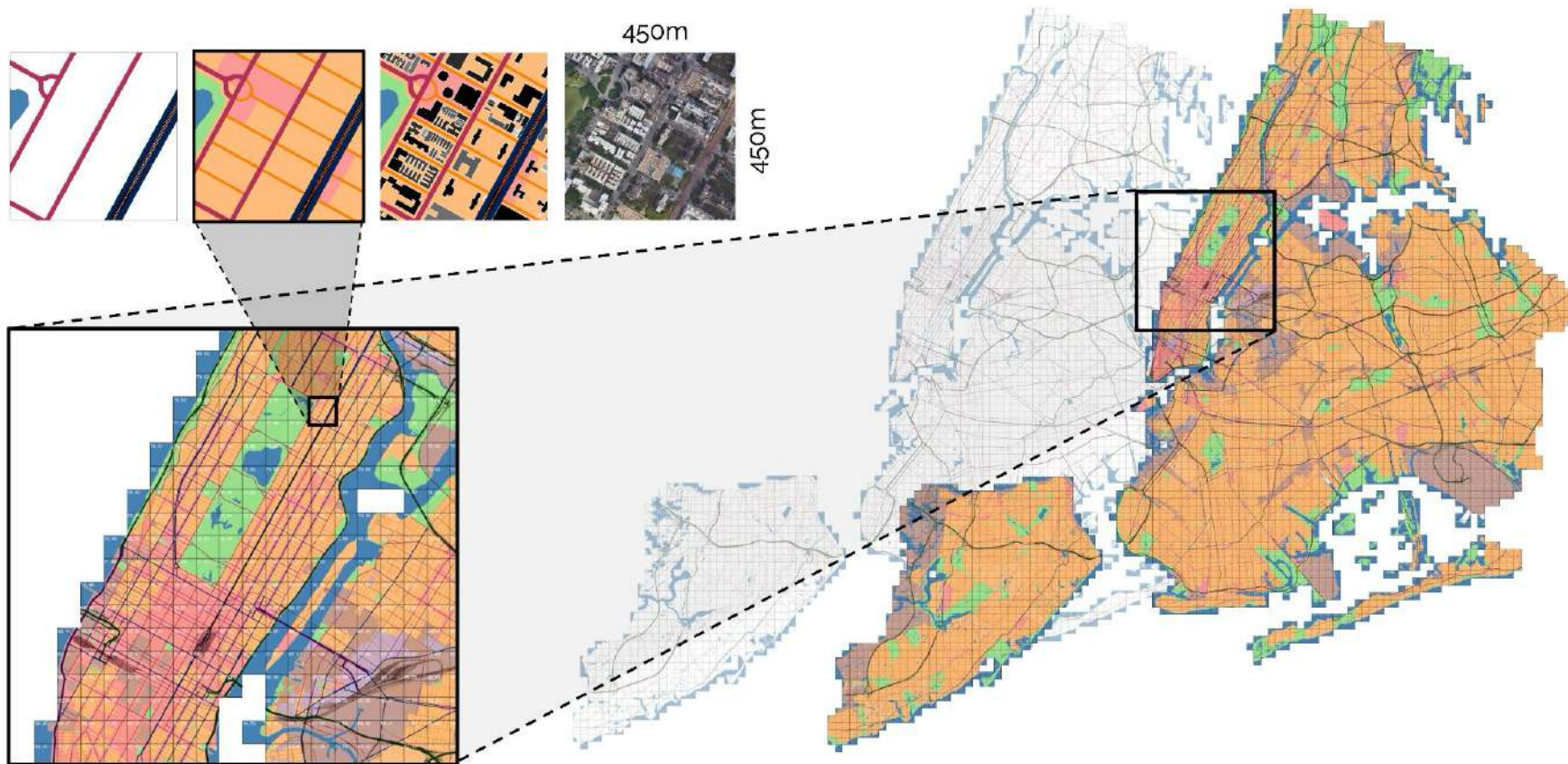


Land use data



Building height data

## Method / Data Processing

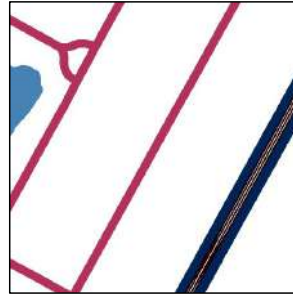


## Method / Feature Extraction



Urban layout

Feature  
Extraction



Design Constraints

+



Design Descriptions (Features)

Residential: 79.2%  
Commercial: 15.4%  
Manufacturing: 0.0%  
Park: 3.6%  
Mixed use: 0.0%

Road density: 18.0%

Building Coverage area: 28.1%  
Average building height: 79.25  
Building volume density: 2224.73%

### Prompt - Stage 1 - Road Network and Land Use Planning

Land use types and road network map of New York.

Land use parcels include 79.2% of residential, 15.4% of commercial, 0.0% of industrial, 3.6% of park, 0.0% of mixed use.

Road density is 18.0%.

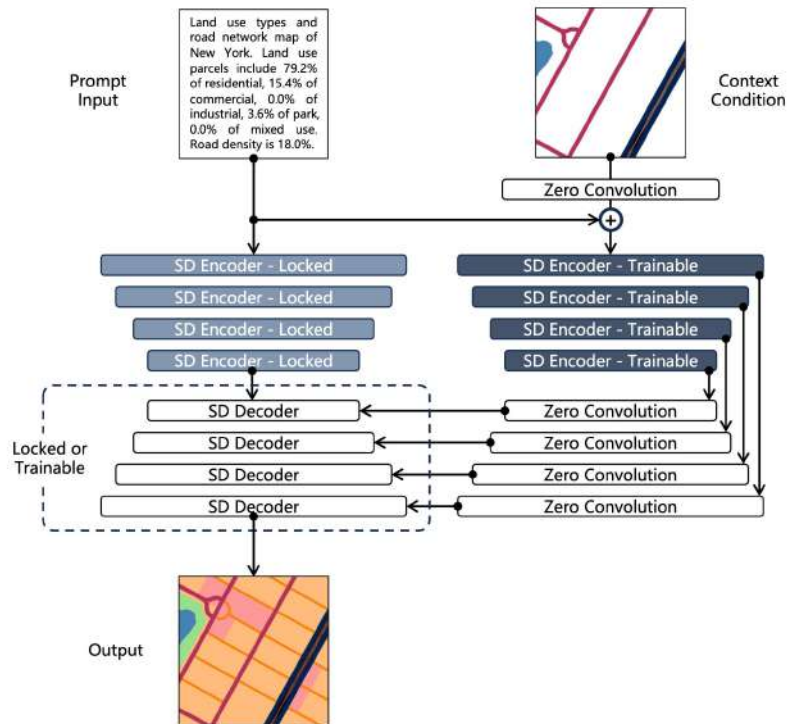
### Prompt - Stage 2 - Building Layout Planning

Building height gradient map of New York.

Buildings include 15% of low-story, 25% of medium-story, 15% of high-story, and 45% open space.

## Method / Model Training: **ControlNet**

A neural network structure designed to **enhance control** over diffusion models



**Diffusion models** exhibit **better scalability, parallelization, more stable training, and higher fidelity images**, comparing to GAN.

**ControlNet framework** shows promising generalization capability **on in-context learning and controllable image editing**.

# Method / **Model Evaluation**

## Baseline model

1. **Pix2Pix** (image-only inputs)
2. **Metric-enhanced Pix2Pix** (image + metrics inputs)
3. **LoRA-SD** (Stable Diffusion fine-tuned with Low-Rank Adaptation, text-only inputs, see Appendix)
4. **ChatGPT-5** (see Appendix)

## Evaluation

### 1. **Visual Fidelity**

Measures the quality by comparing their distribution similarity to real images.  
**FID (Fréchet Inception Distance)** between real and generated urban images.

### 2. **Instruction Compliance**

Whether the model accurately follows instructions regarding land use ratios, road density, building density, building height.  
**Performance Metrics (RMSE, MAE, MAPE, R<sup>2</sup>)** of Proportions of land use types, building categories, and road density.

### 3. **Output Diversity**

Whether the model generates diverse design diagrams conditioned on the same site constraints and human instructions.

# Experiments / 1. Visual Fidelity

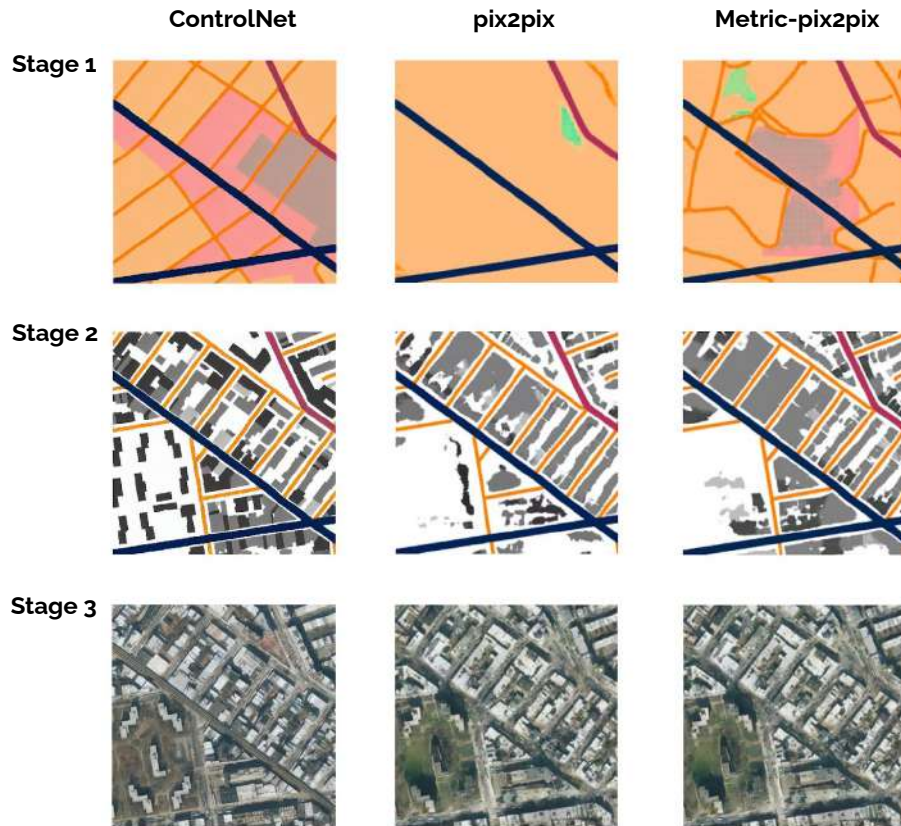
Lower FID means higher fidelity.

ControlNet clearly outperforms Pix2pix in terms of image fidelity.

ControlNet generates images with greater clarity and more precise edge definition.

Table 1: Fidelity performance (FID score) comparison of generated images

Stage	City	ControlNet	Pix2Pix	Metric-Pix2Pix
Stage 1	NYC	<b>30.68</b>	<u>127.78</u>	178.33
	Chicago	<b>34.28</b>	131.35	<u>77.78</u>
Stage 2	NYC	<b>12.84</b>	29.87	<u>20.34</u>
	Chicago	<b>17.88</b>	<u>23.76</u>	23.95
Stage 3	NYC	<b>53.50</b>	<u>56.51</u>	<u>56.51</u>
	Chicago	<b>35.40</b>	<u>44.33</u>	<u>44.33</u>



## Experiments / 2. Instruction Compliance

### Stage 1 Road network and land use planning

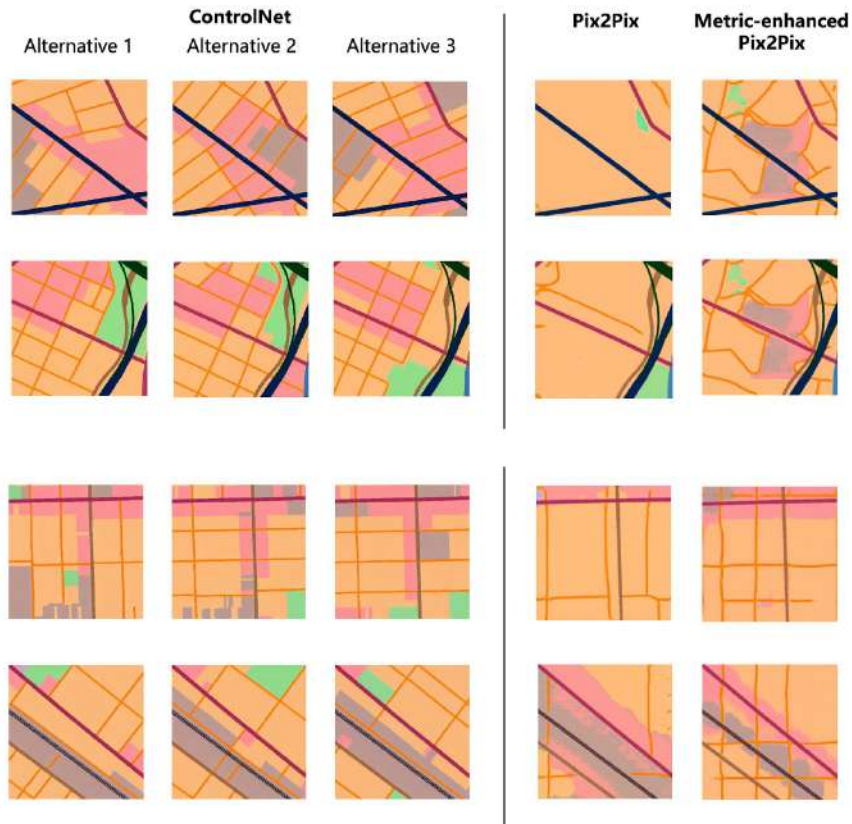
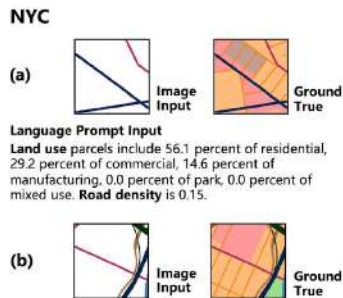
Pix2pix struggles to generate roads and land use layouts following human instructions.

ControlNet generates land use and road layouts, reflecting human intent.

Table 2

Quantitative evaluation of instruction compliance for stage 1 (mean  $\pm$  std).

Model	Metric	Road density		Land use	
		NYC	Chicago	NYC	Chicago
ControlNet	RMSE	0.02 $\pm$ 0.00	0.02 $\pm$ 0.00	0.06 $\pm$ 0.05	0.06 $\pm$ 0.05
	MAE	0.01 $\pm$ 0.01	0.01 $\pm$ 0.01	0.04 $\pm$ 0.04	0.04 $\pm$ 0.04
	R <sup>2</sup>	0.92 $\pm$ 0.01	0.91 $\pm$ 0.01	0.84 $\pm$ 0.27	0.78 $\pm$ 0.49
Pix2Pix	RMSE	0.06 $\pm$ 0.00	0.03 $\pm$ 0.00	0.20 $\pm$ 0.13	0.16 $\pm$ 0.10
	MAE	0.05 $\pm$ 0.04	0.03 $\pm$ 0.02	0.14 $\pm$ 0.09	0.12 $\pm$ 0.08
	R <sup>2</sup>	0.07 $\pm$ 0.12	0.58 $\pm$ 0.06	-0.71 $\pm$ 2.19	-0.12 $\pm$ 1.51
Metric-Pix2Pix	RMSE	0.03 $\pm$ 0.00	0.02 $\pm$ 0.00	0.10 $\pm$ 0.06	0.08 $\pm$ 0.05
	MAE	0.02 $\pm$ 0.02	0.02 $\pm$ 0.01	0.07 $\pm$ 0.04	0.06 $\pm$ 0.04
	R <sup>2</sup>	0.80 $\pm$ 0.03	0.83 $\pm$ 0.03	0.33 $\pm$ 1.89	0.68 $\pm$ 0.47
LoRA-SD	RMSE	0.15 $\pm$ 0.00	0.19 $\pm$ 0.01	0.22 $\pm$ 0.09	0.25 $\pm$ 0.08
	MAE	0.13 $\pm$ 0.08	0.16 $\pm$ 0.10	0.16 $\pm$ 0.07	0.18 $\pm$ 0.06
	R <sup>2</sup>	-4.25 $\pm$ 0.62	-11.23 $\pm$ 1.56	-0.22 $\pm$ 0.96	-0.64 $\pm$ 2.51



Residential

Commercial

Manufacturing

Park

Mixed use

## Experiments / 2. Instruction Compliance

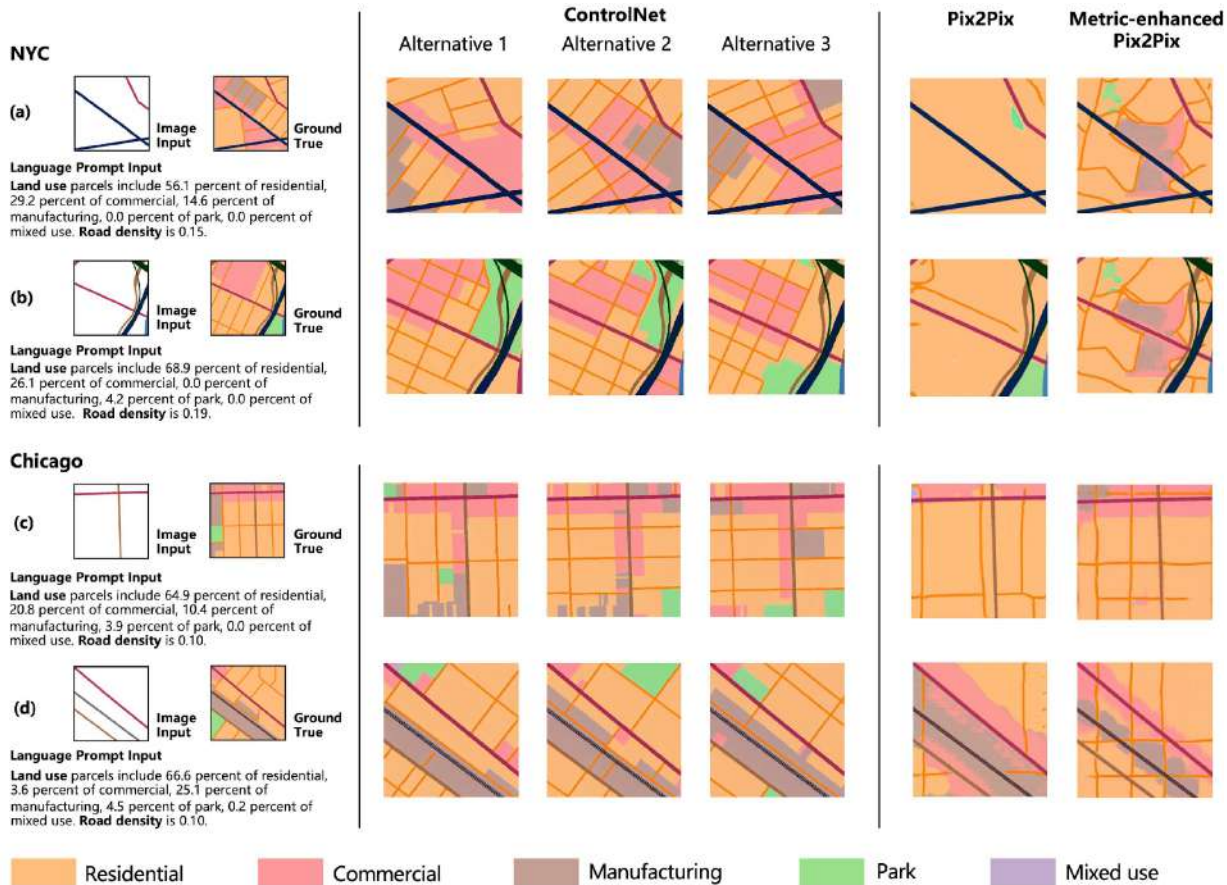
### Stage 1 Road network and land use planning

Pix2pix struggles to generate roads and land use layouts following human instructions.

**ControlNet generates land use and road layouts, reflecting human intent.**

**ControlNet shows spatial awareness:**

- Captures interactions between **land use types and transportation infrastructure**.
- **Commercial** areas align with high-order roads.
- **Residential** zones cluster along lower-order streets.



## Experiments / 2. Instruction Compliance

### Stage 2 Building layout planning

Pix2pix struggles to generate building layouts following human instructions.

**ControlNet generates building layouts and height, reflecting human intent.**

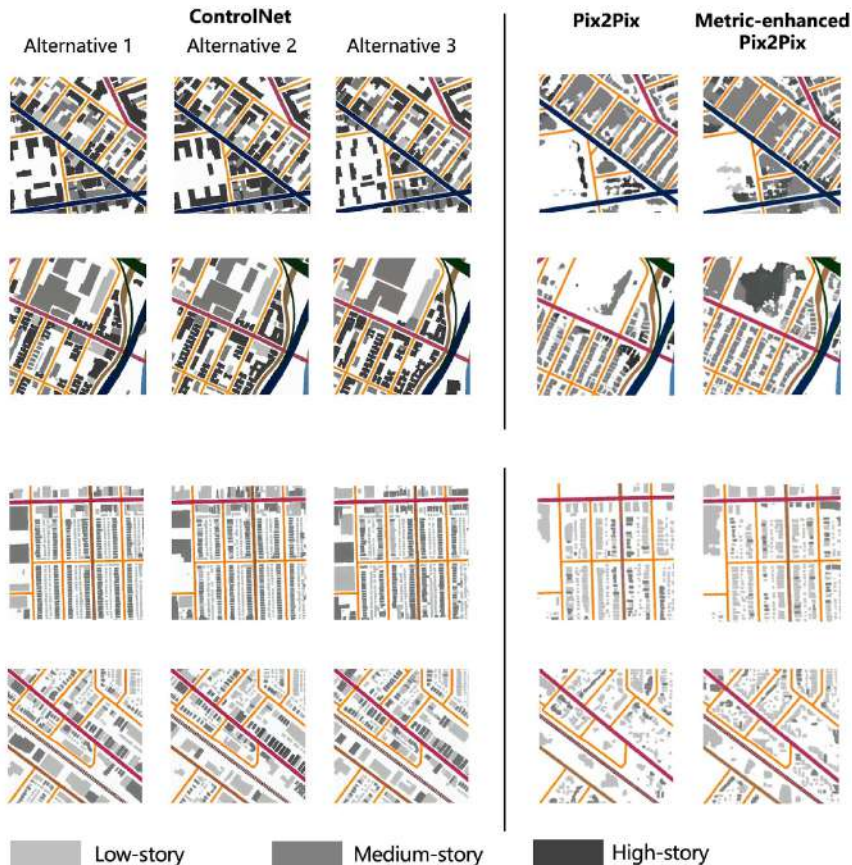
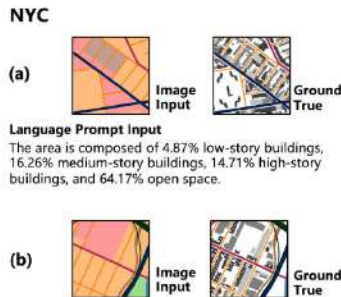


Table 3

Quantitative evaluation of instruction compliance for stage 2 (mean  $\pm$  std).

Model	Metric	Building height		Open space	
		NYC	Chicago	NYC	Chicago
ControlNet	RMSE	<b>0.03 <math>\pm</math> 0.00</b>	<b>0.04 <math>\pm</math> 0.00</b>	<b>0.05 <math>\pm</math> 0.00</b>	<b>0.05 <math>\pm</math> 0.00</b>
	MAE	<b>0.02 <math>\pm</math> 0.00</b>	<b>0.02 <math>\pm</math> 0.00</b>	<b>0.03 <math>\pm</math> 0.04</b>	<b>0.04 <math>\pm</math> 0.03</b>
	$R^2$	<b>0.87 <math>\pm</math> 0.01</b>	0.81 $\pm$ 0.04	0.91 $\pm$ 0.01	<b>0.90 <math>\pm</math> 0.02</b>
Pix2Pix	RMSE	0.06 $\pm$ 0.00	0.06 $\pm$ 0.00	0.07 $\pm$ 0.00	0.08 $\pm$ 0.00
	MAE	0.03 $\pm$ 0.00	0.04 $\pm$ 0.00	0.05 $\pm$ 0.05	0.05 $\pm$ 0.05
	$R^2$	0.62 $\pm$ 0.03	0.49 $\pm$ 0.04	0.81 $\pm$ 0.03	0.77 $\pm$ 0.04
Metric-Pix2Pix	RMSE	0.04 $\pm$ 0.00	<b>0.04 <math>\pm</math> 0.00</b>	<b>0.05 <math>\pm</math> 0.00</b>	0.06 $\pm$ 0.02
	MAE	0.03 $\pm$ 0.00	<b>0.02 <math>\pm</math> 0.00</b>	<b>0.03 <math>\pm</math> 0.03</b>	<b>0.03 <math>\pm</math> 0.06</b>
	$R^2$	0.79 $\pm$ 0.02	<b>0.84 <math>\pm</math> 0.02</b>	<b>0.92 <math>\pm</math> 0.01</b>	0.84 $\pm$ 0.10
LoRA-SD	RMSE	0.16 $\pm$ 0.00	0.07 $\pm$ 0.00	0.37 $\pm$ 0.01	0.26 $\pm$ 0.01
	MAE	0.11 $\pm$ 0.00	0.05 $\pm$ 0.00	0.34 $\pm$ 0.16	0.22 $\pm$ 0.14
	$R^2$	-1.83 $\pm$ 0.27	0.33 $\pm$ 0.04	-4.00 $\pm$ 0.36	-1.65 $\pm$ 0.28

## Experiments / 2. Instruction Compliance

### Stage 2 Building layout planning

Pix2pix struggles to generate building layouts following human instructions.

**ControlNet generates building layouts and height, reflecting human intent.**

**ControlNet shows spatial coherence:**

- Aligns buildings along streets, forming **continuous street edges**.
- Baseline models produce fragmented or misaligned layouts.
- Captures realistic land use patterns:**
  - Larger buildings in **commercial and industrial areas**.
  - Smaller ones in **residential zones**.

NYC



Language Prompt Input

The area is composed of 4.87% low-story buildings, 16.26% medium-story buildings, 14.71% high-story buildings, and 64.17% open space.



Language Prompt Input

The area is composed of 6.67% low-story buildings, 13.32% medium-story buildings, 10.48% high-story buildings, and 69.53% open space.

Chicago



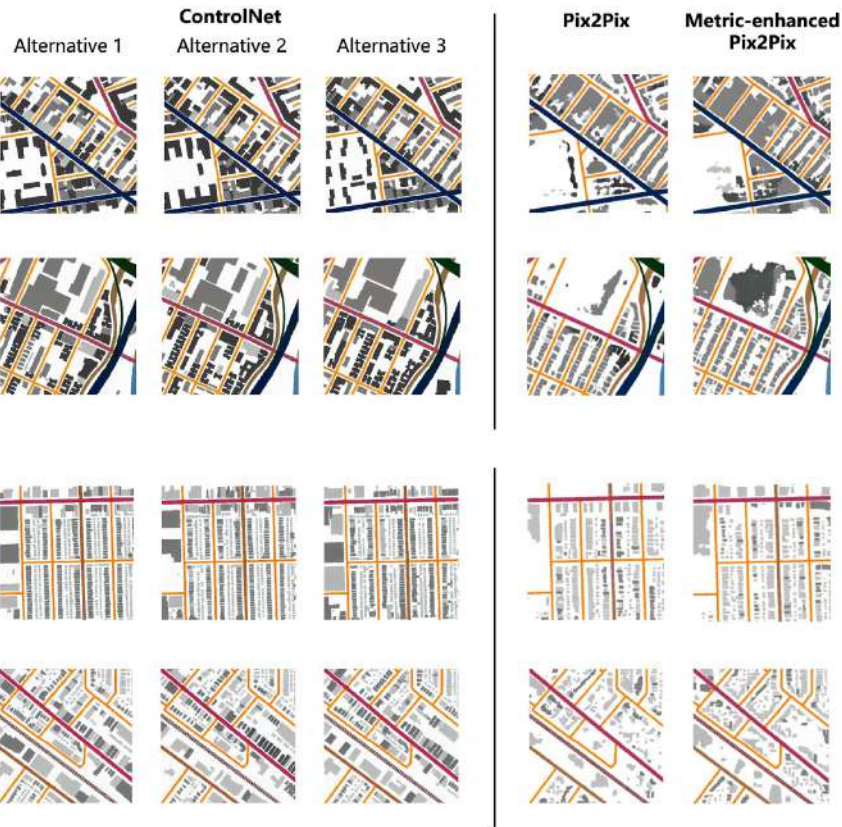
Language Prompt Input

The area is composed of 18.99% low-story buildings, 15.11% medium-story buildings, 0.00% high-story buildings, and 65.90% open space.



Language Prompt Input

The area is composed of 18.57% low-story buildings, 8.08% medium-story buildings, 0.00% high-story buildings, and 73.35% open space.



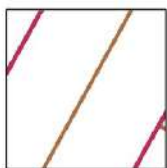
Low-story

Medium-story

High-story

## Experiments / 3. Output Diversity

### Stage 1 Road Network and Land Use Planning

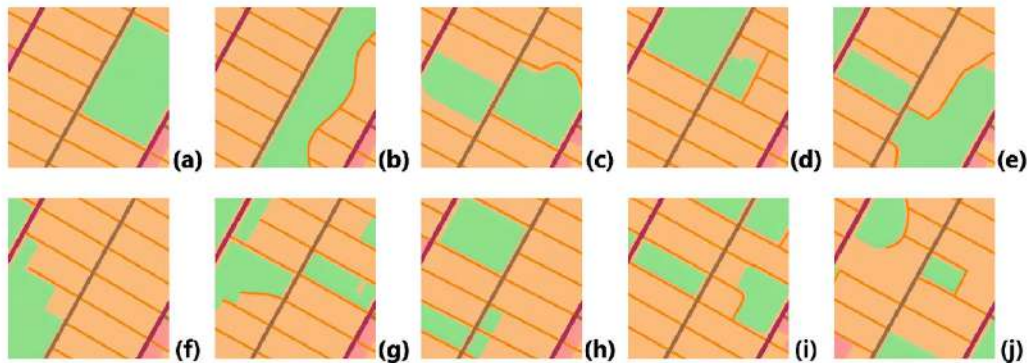


Input Image



Ground True

Same land use composition, diverse spatial layouts—from centralized parks to dispersed green spaces enhancing accessibility.



### Stage 2 Building Layout Planning

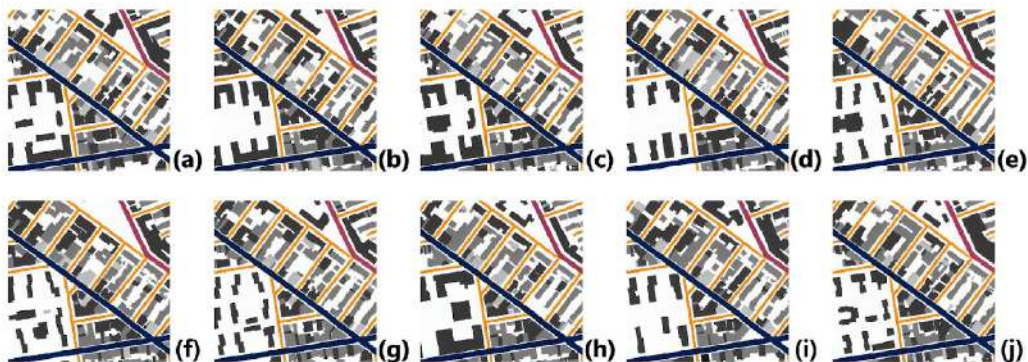


Input Image



Ground True

Same building height proportion, diverse residential layouts—from linear rows to courtyard and clustered forms.



# Experiments / 3. Output Diversity

## Stage 3 Detailed Planning and Rendering



Satellite image of a city in New York.



ControlNet strictly follows the constraint of the input image while generating diverse detailed planning, such as landscape design, rooftop design, etc.

## Experiments / 4. User Study

### Experiment Design

- Experts rated real or generated images (1–10 scale), seeing only one per case.
- Evaluations covered **site alignment** and **stage-specific design quality** across three stages.

### Results

Table 5: User study results across three stages (mean  $\pm$  std scores).

	Stage 1			Stage 2		Stage 3	
	Site constr.	Land use	Road network	Site constr.	Building layout	Site constr.	Realism
Real	8.25 $\pm$ 2.19	7.37 $\pm$ 2.21	7.08 $\pm$ 2.35	7.72 $\pm$ 2.28	6.45 $\pm$ 2.33	8.45 $\pm$ 1.79	8.24 $\pm$ 1.82
Generated	7.78 $\pm$ 2.26	6.99 $\pm$ 2.38	7.21 $\pm$ 2.33	8.05 $\pm$ 1.92	6.52 $\pm$ 2.41	8.12 $\pm$ 1.88	7.77 $\pm$ 1.91
Difference	-0.47	-0.38	+0.13	+0.33	+0.07	-0.33	-0.47
<i>t</i> -stat	1.41	1.12	-0.36	-1.05	-0.19	1.18	1.64
<i>p</i> -value	0.16	0.26	0.72	0.30	0.85	0.24	0.10

- Generated images **scored closely to** real ones in all stages.
- Slight advantages in **road network rationality** and **site constraint alignment**.
- **No statistically significant differences**  $\rightarrow$  **perceptually close to real images**.

## Experiments / 5. Stepwise v.s. End-to-End

Table 6: Comparison of stepwise and end-to-end framework

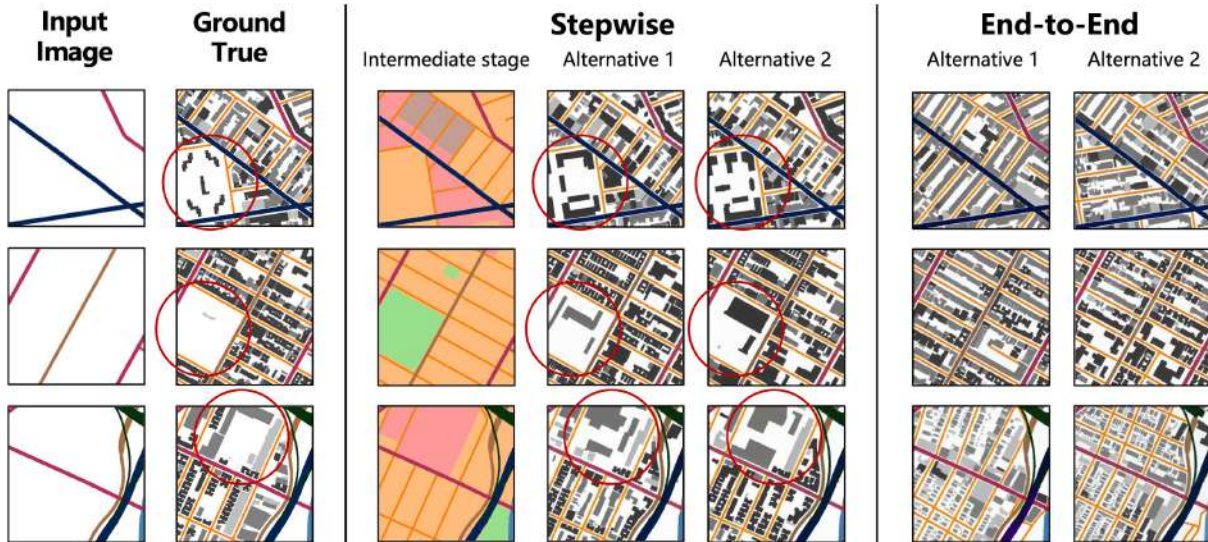
	Stepwise			End-to-end		
FID	12.84			32.38		
	RMSE	MAE	R <sup>2</sup>	RMSE	MAE	R <sup>2</sup>
Road density	0.02	0.01	0.92	0.05	0.04	0.44
Building Height	0.03	0.02	0.87	0.04	0.02	0.78
Open Space	0.05	0.03	0.91	0.12	0.09	0.48

**Stepwise:** Site Constraints -> Land Use Layout -> Building Layout

- produces **clearer and more functionally coherent layouts**.

**End-to-End:** Site Constraints -> Building Layout

- generates **repetitive, uniform patterns that lack diversity**.



**The stepwise approach proves superior to end-to-end:**

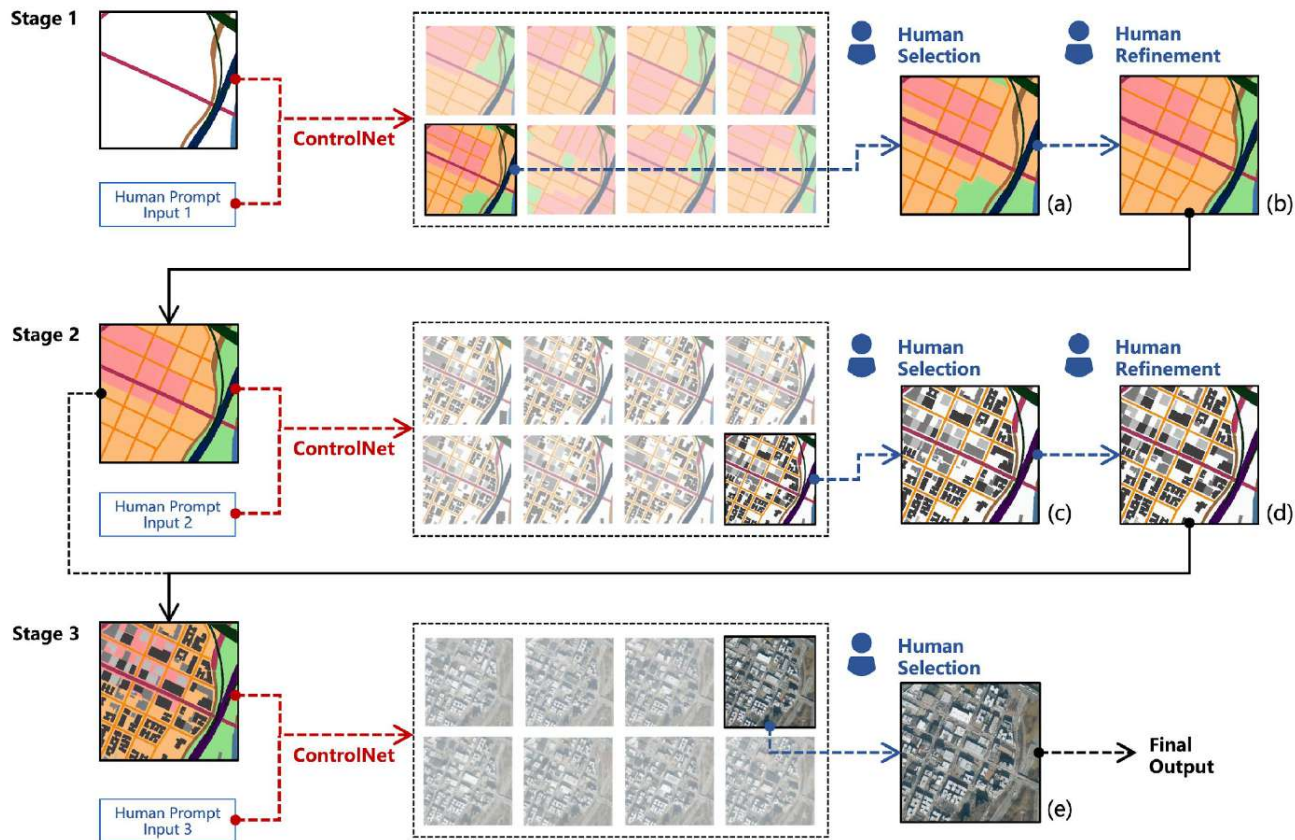
- **Alignment with instructions:** better follows human guidance through intermediate decision points.
- **Output quality:** generates more realistic and functionally coherent urban forms.
- **Better human control:** enables meaningful review and refinement at each stage.

## Experiments / 6. Human-in-the-loop Generation Case

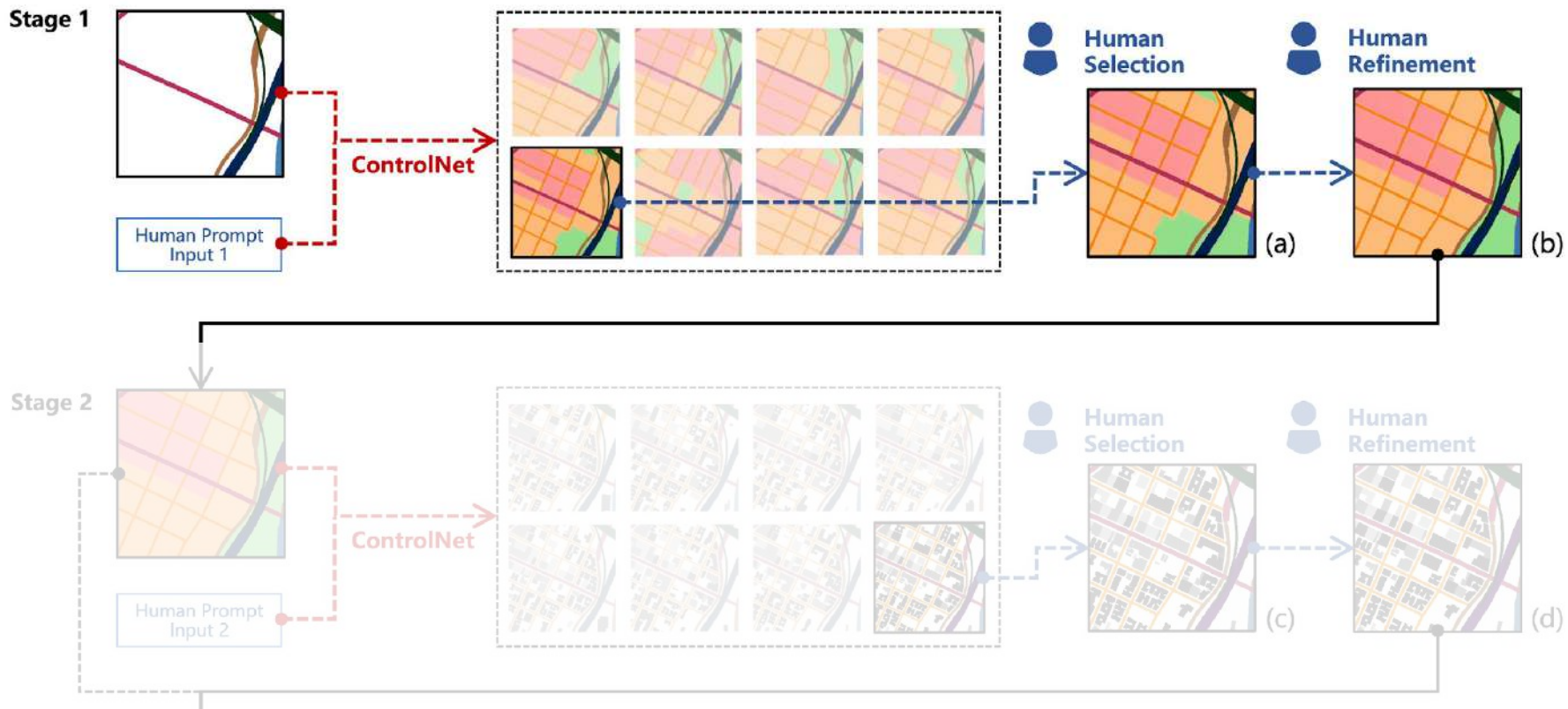
Demonstrates **iterative collaboration** between diffusion models and human expertise.

**Three-stage process** for urban form generation and refinement.

- Human Prompt
- Human Selection
- Human Refinement



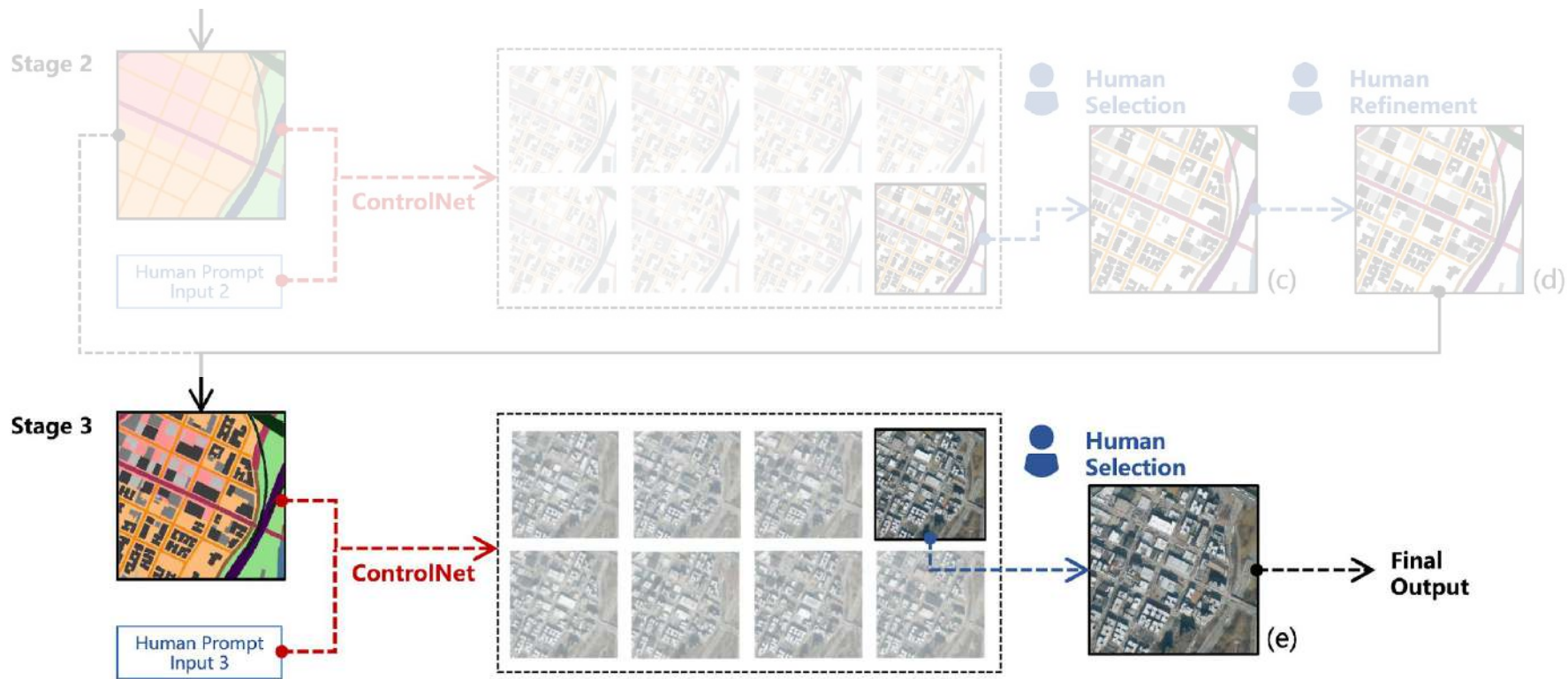
## Experiments / 6. Human-in-the-loop Generation Case



## Experiments / 6. Human-in-the-loop Generation Case



## Experiments / 6. Human-in-the-loop Generation Case

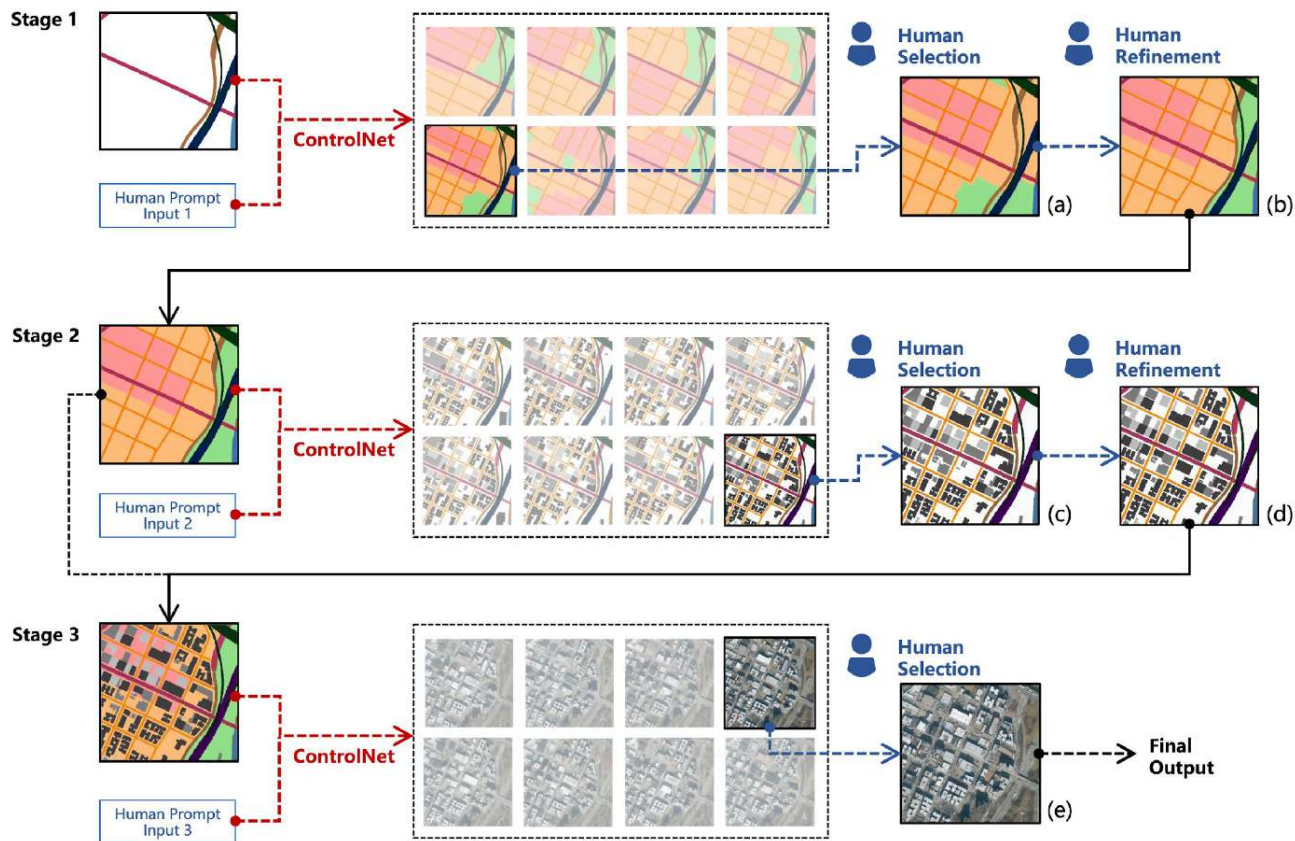


## Experiments / 6. Human-in-the-loop Generation Case

Demonstrates **iterative collaboration** between diffusion models and human expertise.

**Three-stage process** for urban form generation and refinement.

- Human Prompt
- Human Selection
- Human Refinement



# Main Findings

## 1. Model Effectiveness

- ControlNet understands designer prompts and generates effective urban layouts, **outperforming the baseline model in both image fidelity and prompt alignment**.

## 2. Stepwise Control Advantage

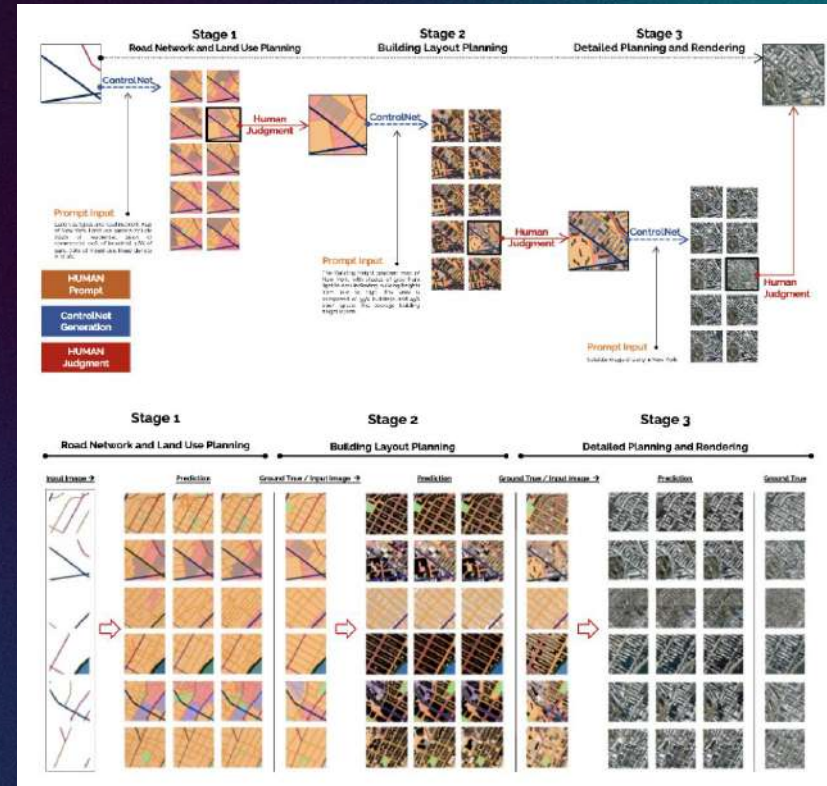
- Stepwise generation is superior to end-to-end in terms of **human control, output quality, and alignment with instructions**.
- An Example of Neuro-symbolic AI.

## 3. Collaborative Design

- Human-Centered AI** for Urban Design enhances diversity and human control, promoting efficiency in urban design and facilitating public communication and iterative refinement.

## 4. Design Pattern Discovery

- Through cross-city comparisons and the application of city model transfer, the model uncovers **hidden urban design patterns**.



What will future planners do?

# 1. Immediacy changes the dialog

*Planners have to be more responsive*

## 2. A.I. enables the public

*Planning is a tool to allocate power*

### 3. Essence of planners' job

Design creator → Design critique

Design creator → Design communicator

Design creator → Design negotiator

# AI for Future of Mobility

What is success and what defines the future

## AI for City Planning

Human In the Loop, Power Balance and Future Role of City Planners

## AI for Autonomous Vehicles

human agency and AV deployment

What's essential differences between AV and human driven cars?

Human Agency

# Spectrum of human vs machine decision making

High level:  
single vs. pooling;  
destination choice;  
ownership choice

Mid level:  
route choice;  
departure time  
choice

Low level:  
Acceleration,  
deceleration, change  
lanes

## “The Agency Frontier

— where human preference and  
machine intelligence meet.”

Fully  
human

A variety of middle  
points

Fully  
machine

Spectrum of human vs machine decision making

Operation scenarios

Emergency  
(large scale)

Emergency  
(small scale)

Accidents

Events

Normal

“The Command Matrix  
— the right actor in charge for  
each scenario.”

consumers

operators

car makers

authorities

Decision-making parties

# AI and Human Agency

## Incentive design framework

- motivate each decision party to acquire or release the control
- private purposes (pay for it) vs. public purposes (to be paid for it)

## Public policy

- Different layers of property rights
- A richer set of policy tools

# Three ways to change behavior

- Price
- Rules
- Norms

Direct control

# Conclusion

# Two Roles of Models

# Communicative planning

# Behavior + Computation

## Behavioral Thinking

- Emotional
- Social
- Perceptual



## Transportation Technology

- Electrification
- Automation
- Connectivity
- Sharing

---

## Computational Foundation

- Representation
- Explanation
- Prediction
- Control
- Creation

Representation

Prediction

Explanation

Control

Creation

Communication

What is success?