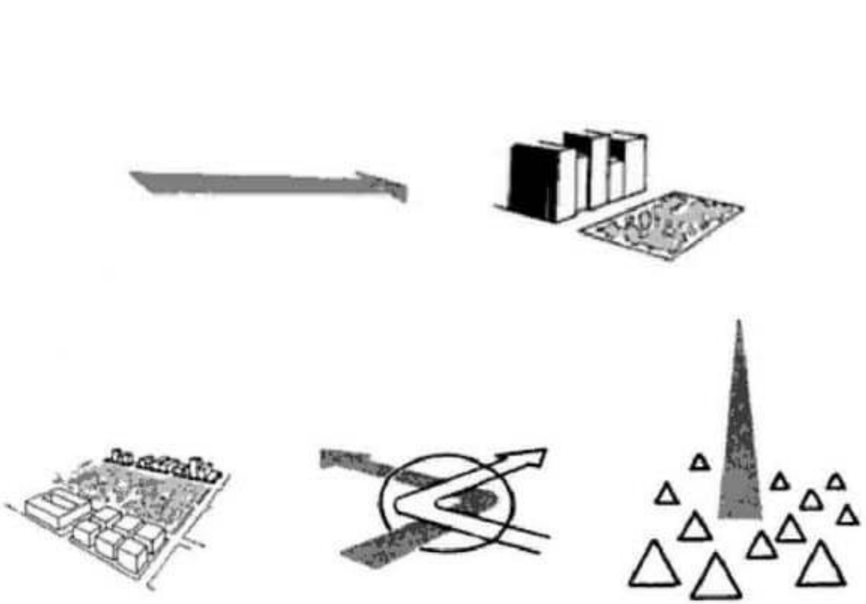
senseable city lab.

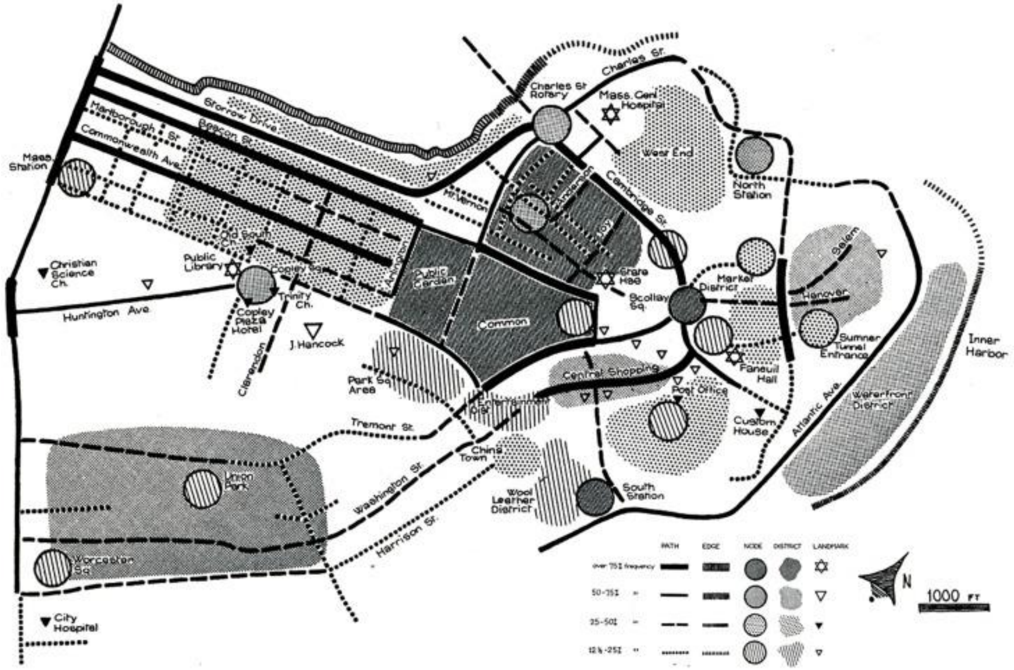
Prior Work

KEVIN LYNCH, ALLAN JACOBS, WILLIAM WHYTE

KEVIN LYNCH

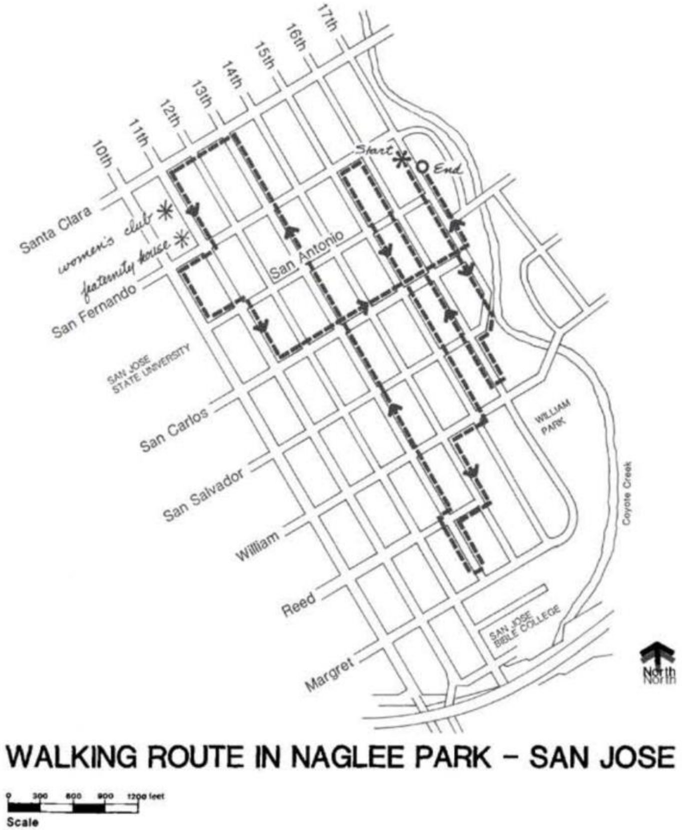


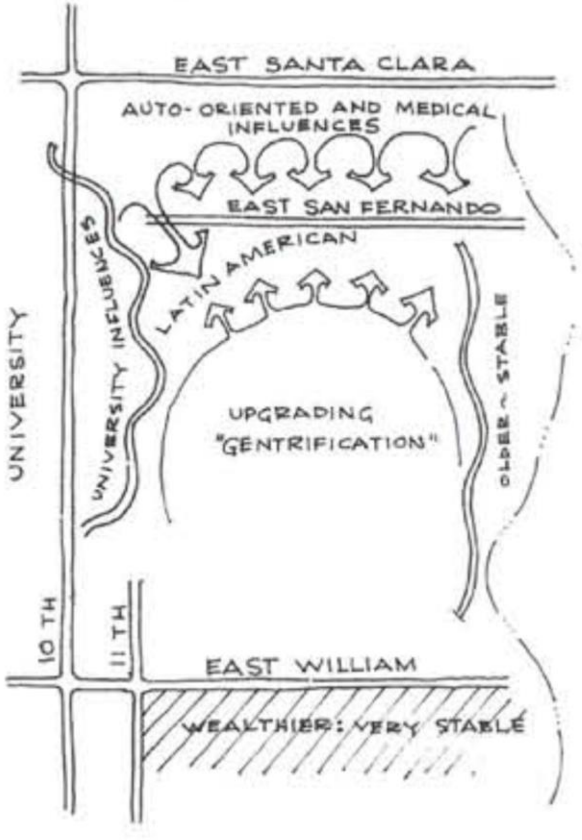
Perceptible urban elements



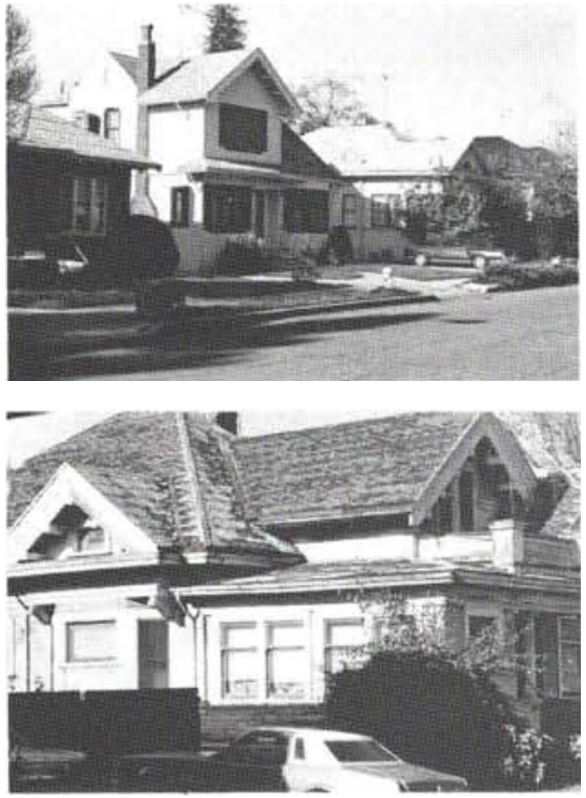
Perceived urban image in Boston community

ALLAN JACOBS





Observing and Interpreting Naglee Park



Visual Cues of Gentrification

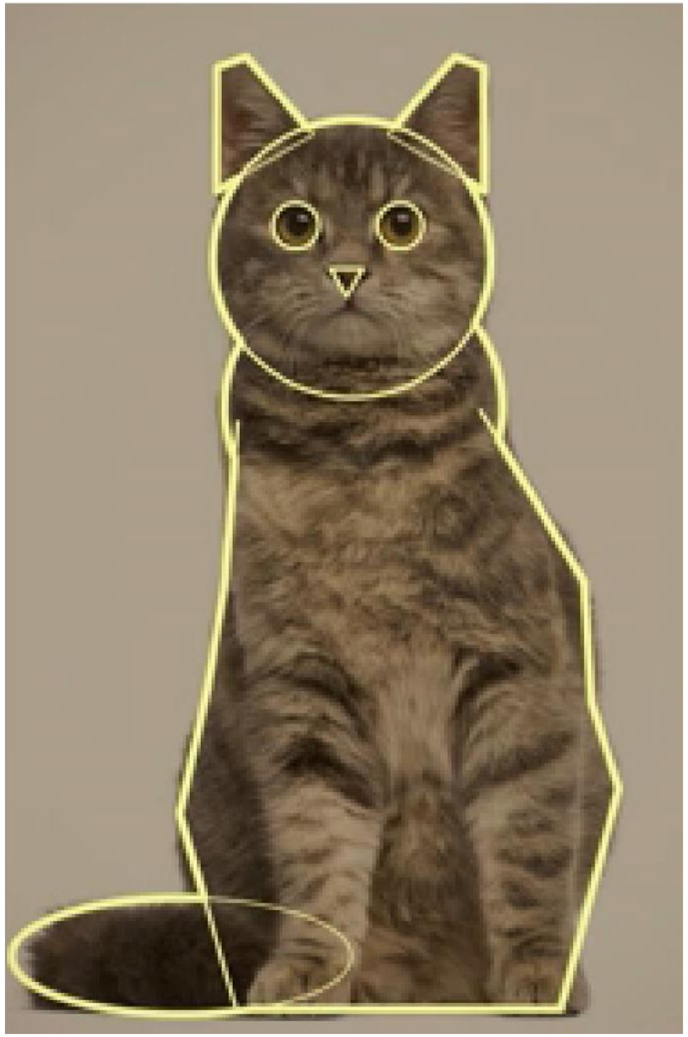
Digital turning point

THE CAT PROBLEM

FEI-FEI LI

How come a toddler can identify a cat and computers cannot?

FEI-FEI LI



FEI-FEI LI

Dataset of images + labels

Machine learning to train classifier (training dataset)

Evaluate classifier on new images (testing dataset)

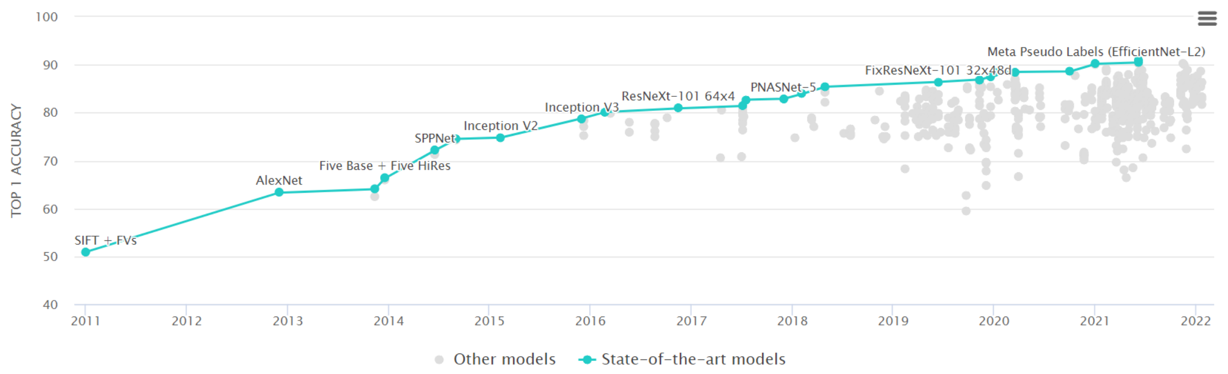
ImageNet

- 14 million images annotated into 20,000 object categories
- Benchmark to compare models

Image Classification on ImageNet

Leaderboard   Dataset

View: Top 1 Accuracy   by: Date   for: All models



Filter: ImageNet-1k only | Transformer | ResNet | CNN | EfficientNet | JFT-300M | ImageNet-22k | ResNeXt | MLP | PatchConvnet | JFT-3B | Conv+Transformer | MViT | Mixer | untagged | Hardware Burden | Operations per network pass | Robustness reports

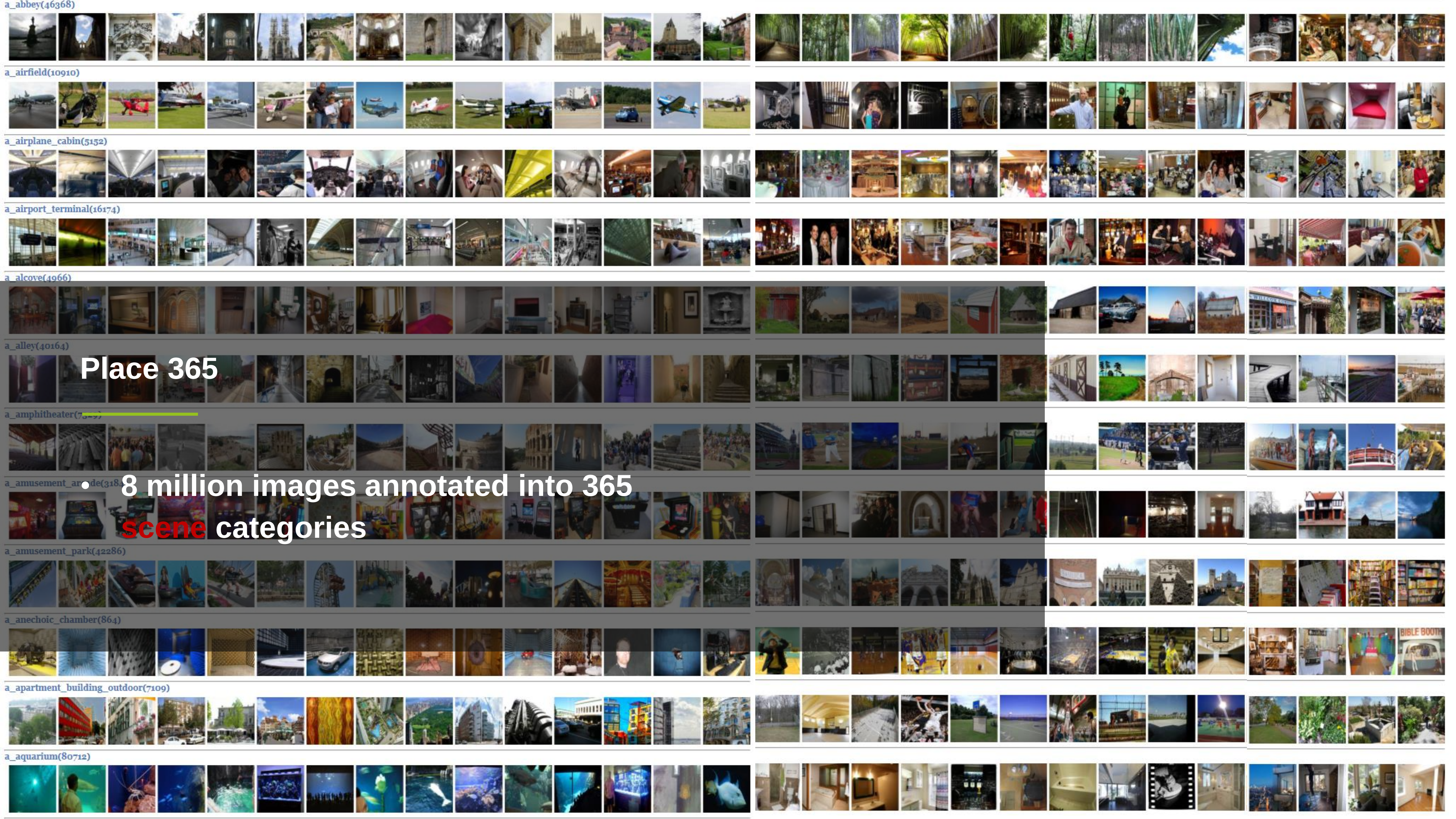
Edit Leaderboard

Rank	Model	Top 1 Accuracy	Top 5 Accuracy	Number of params	Extra Training Data	Paper	Code	Result	Year	Tags
1	**CoAtNet-7**	90.88%		2440M	✓	CoAtNet: Marrying Convolution and Attention for All Data Sizes			2021	Conv+Transformer, JFT-3B
2	**ViT-G/14**	90.45%		1843M	✓	Scaling Vision Transformers			2021	Transformer, JFT-3B
3	**CoAtNet-6**	90.45%		1470M	✓	CoAtNet: Marrying Convolution and Attention for All Data Sizes			2021	Conv+Transformer, JFT-3B
4	**ViT-MoE-15B** (Every-2)	90.35%		14700M	✓	Scaling Vision with Sparse Mixture of Experts			2021	Transformer, JFT-3B
5	**Meta Pseudo Labels** (EfficientNet-L2)	90.2%	98.8%	480M	✓	Meta Pseudo Labels			2021	EfficientNet, JFT-300M

https://paperswithcode.com/sota/image-classification-on-imagenet

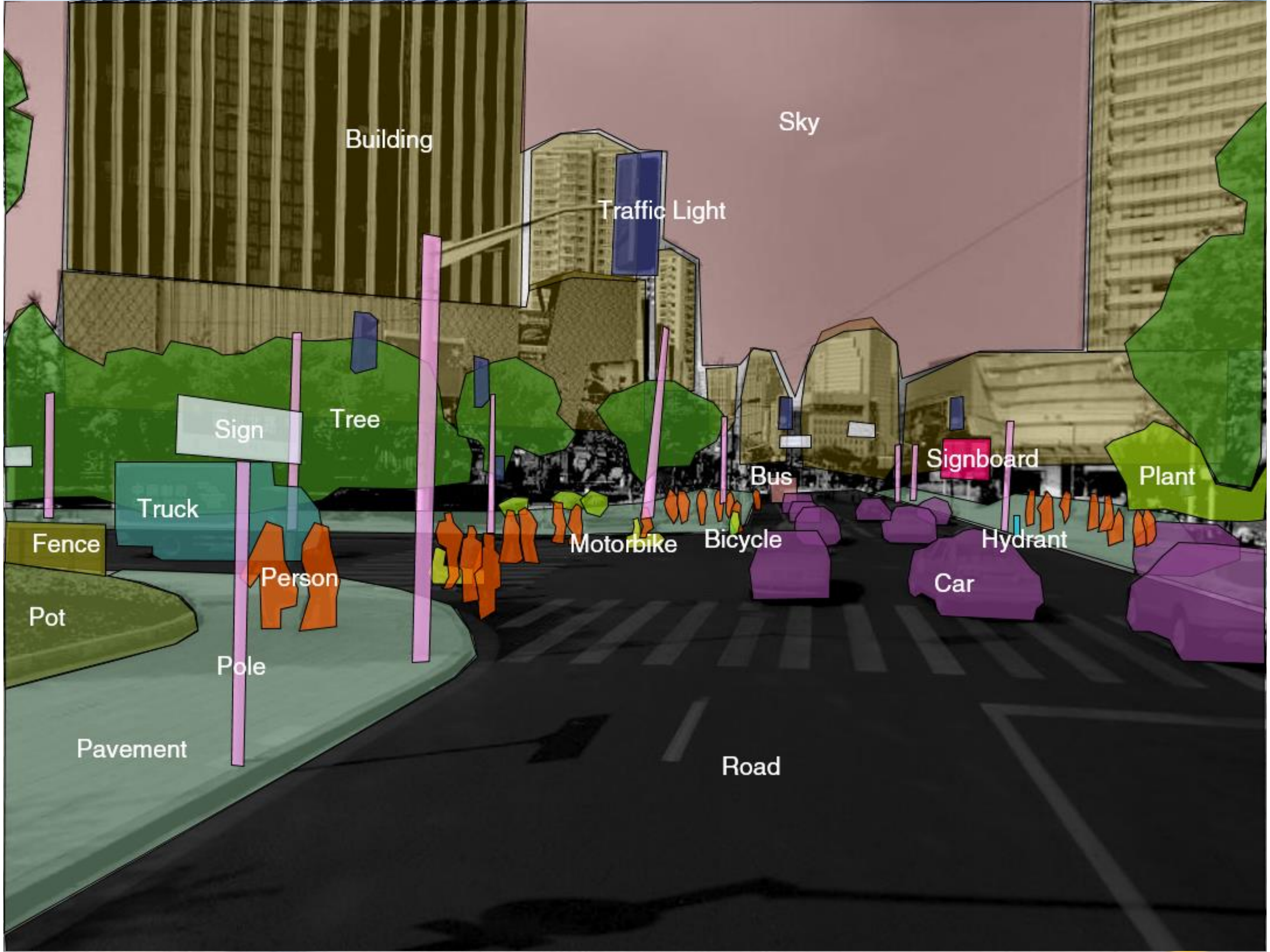
Place 365

- 8 million images annotated into 365 scene categories



ADE20K

- 25,574 images annotated at pixel-level into 150 categories


Building
Sky
Traffic Light
Sign
Tree
Signboard
Plant
Bus
Truck
Fence
Motorbike
Bicycle
Hydrant
Person
Car
Pot
Pole
Pavement
Road

SCL Visual AI

PASSIVE, ACTIVE,

CROWDSOURCED

Passive Data Collection

Treepedia

ENVIRONMENT, COMPUTER VISION


100 Massachusetts Ave
Boston, Massachusetts
Google, Inc.
Street View - Aug 2017
Room&Board
MUJI
MUJI
無印良品
Trader Joe's
90
Sheraton Boston
Google
Image capture: Aug 2017
© 2018 Google
United States
Terms
Report a problem

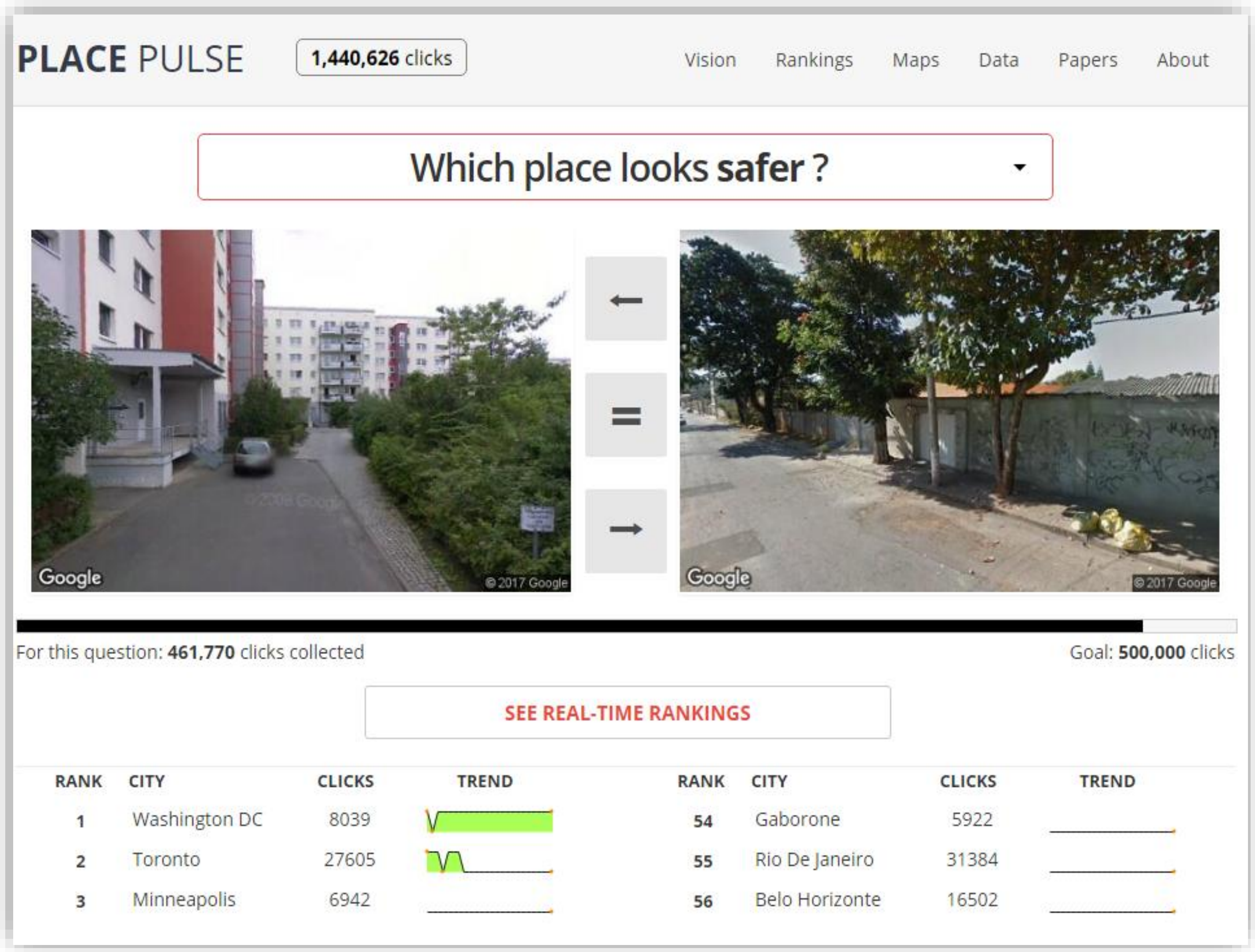
AI Perception Map

BUILT ENVIRONMENT, HUMAN PERCEPTION, DEEP LEARNING

Photo by Jonas Hoss on Unsplash

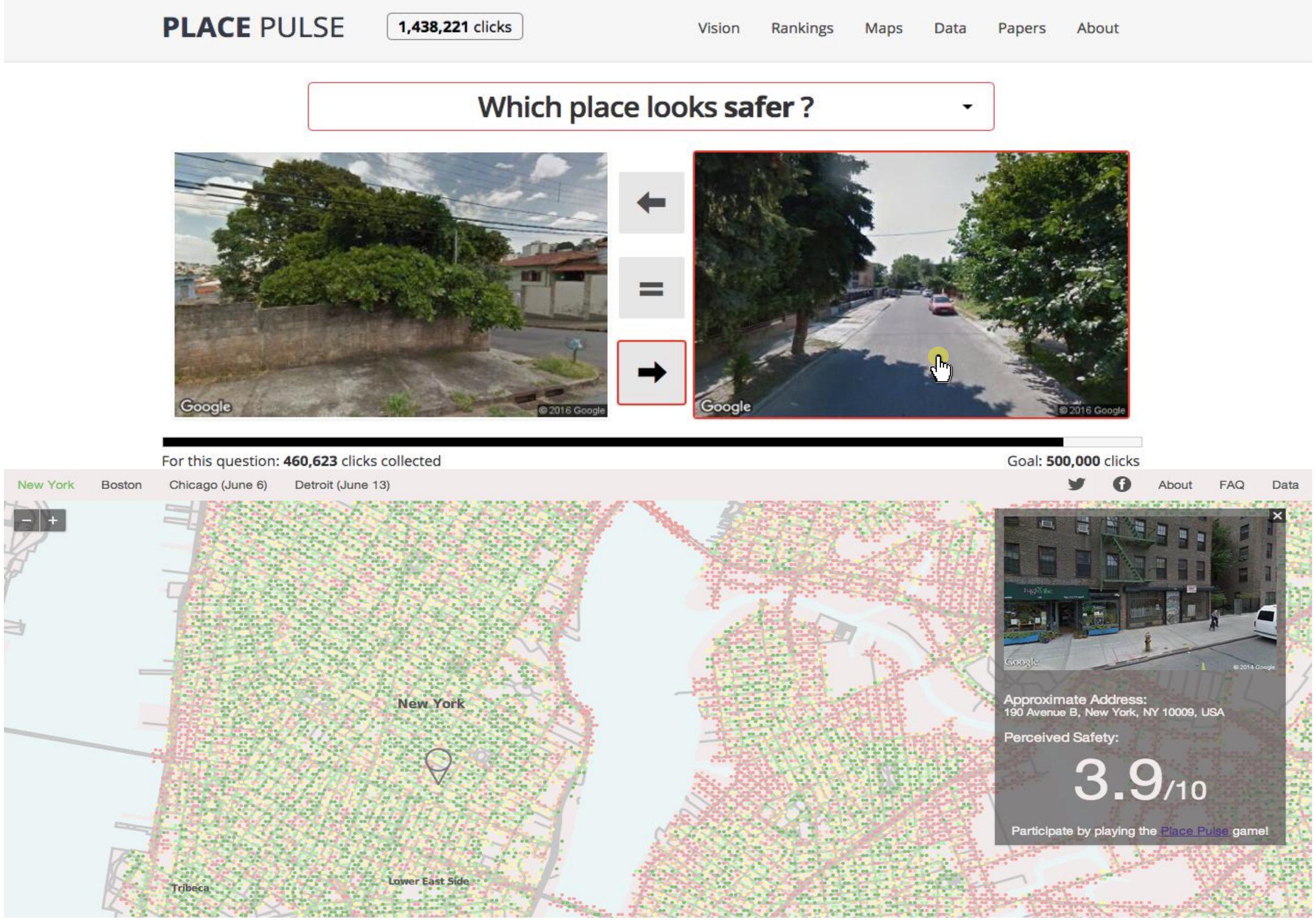

Photo by Evgeniy Grozev on Pexels

Urban Perception Preference Dataset



- 1,170,000 pairwise comparisons provided by
  81,630 online volunteers
- 110,988 images from
  56 cities, 28 countries, 6 continents
- Six perceptual dimensions:
  Safety, Lively, Boring,
  Wealthy, Depressing, Beautiful.

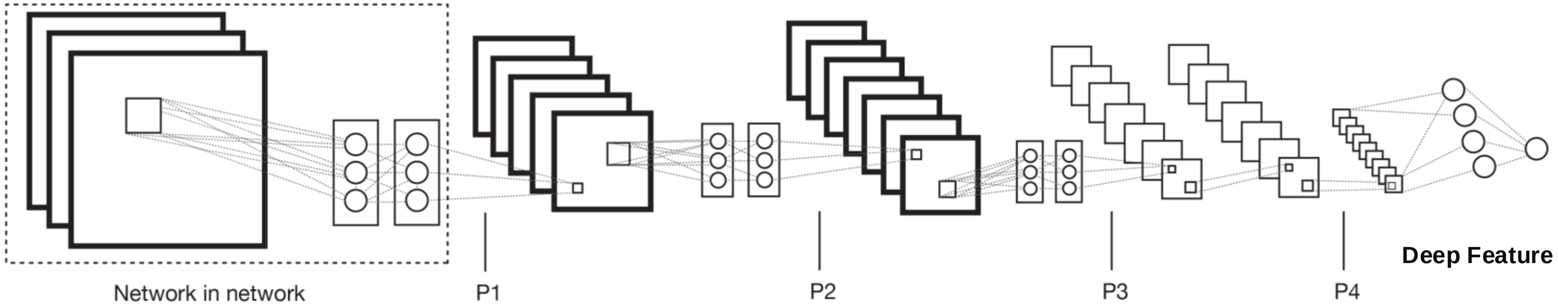
Urban Perception Preference Dataset



Evaluating Perceptual Preference using Deep Learning

Street view image

Deep Convolutional Neural Network



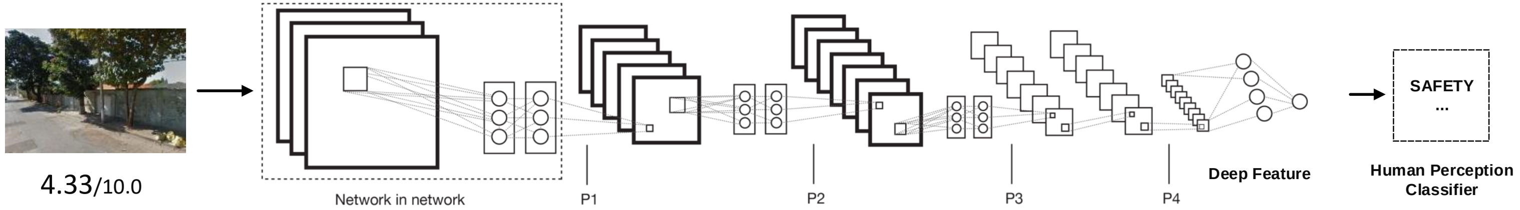
Classifier

SAFETY
...

Human Perception Classifier

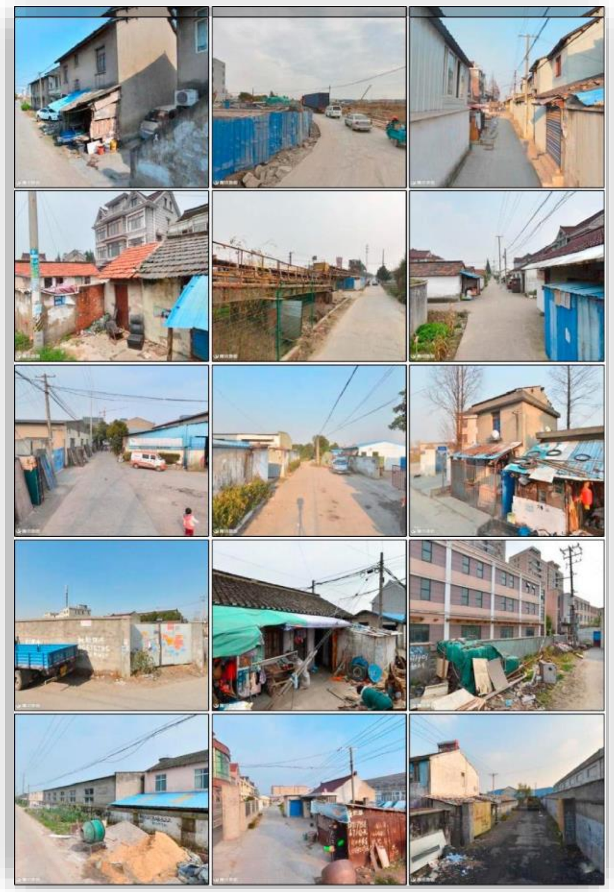
Zhang, Fan, Bolei Zhou, Liu Liu, Yu Liu, Helene H. Fung, Hui Lin, and Carlo Ratti. "Measuring human perceptions of a large-scale urban region using machine learning." *Landscape and Urban Planning* 180 (2018): 148-160.

Applying the model to new images

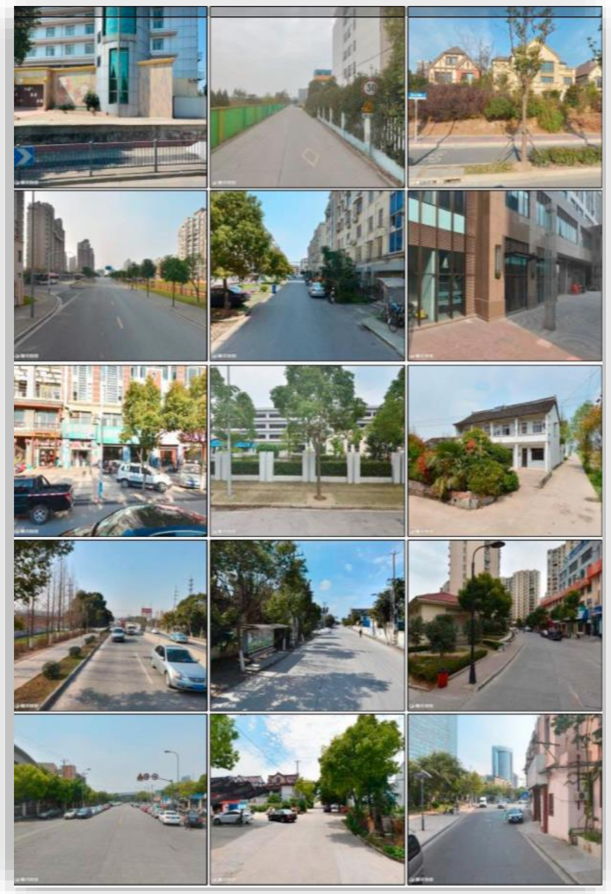


Evaluating Street View Images in China using Deep Learning

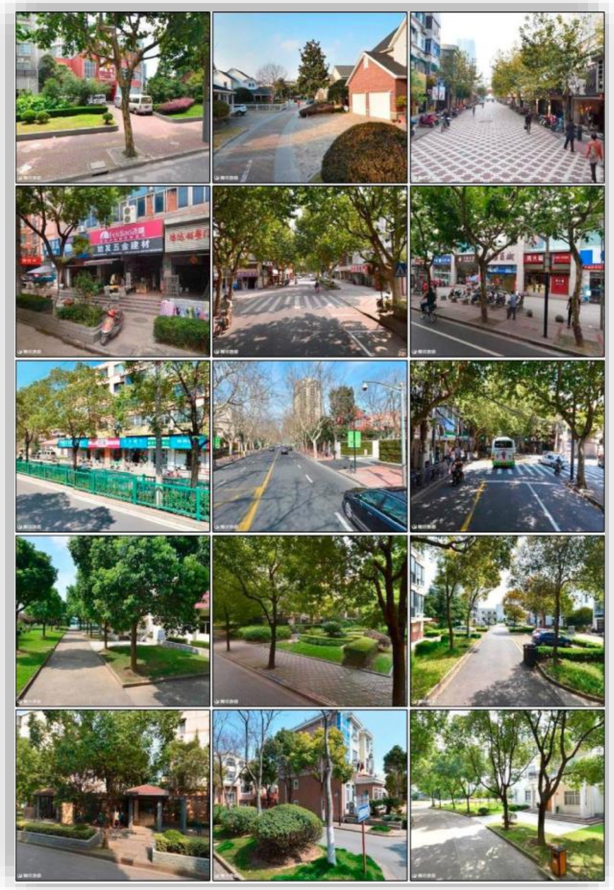
- 250,000 images of Beijing
- 140,000 images of Shanghai



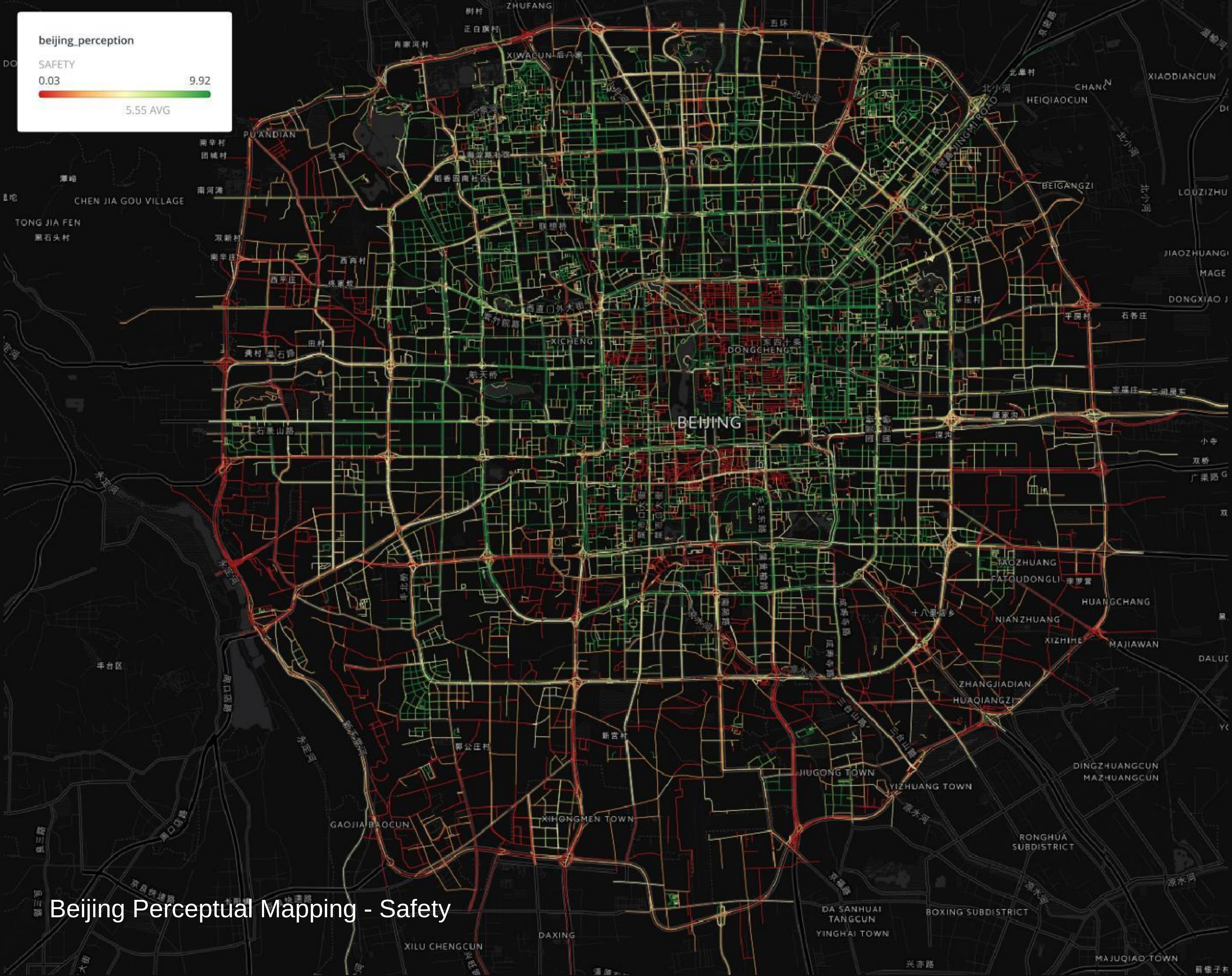
(Q safe <= 3)



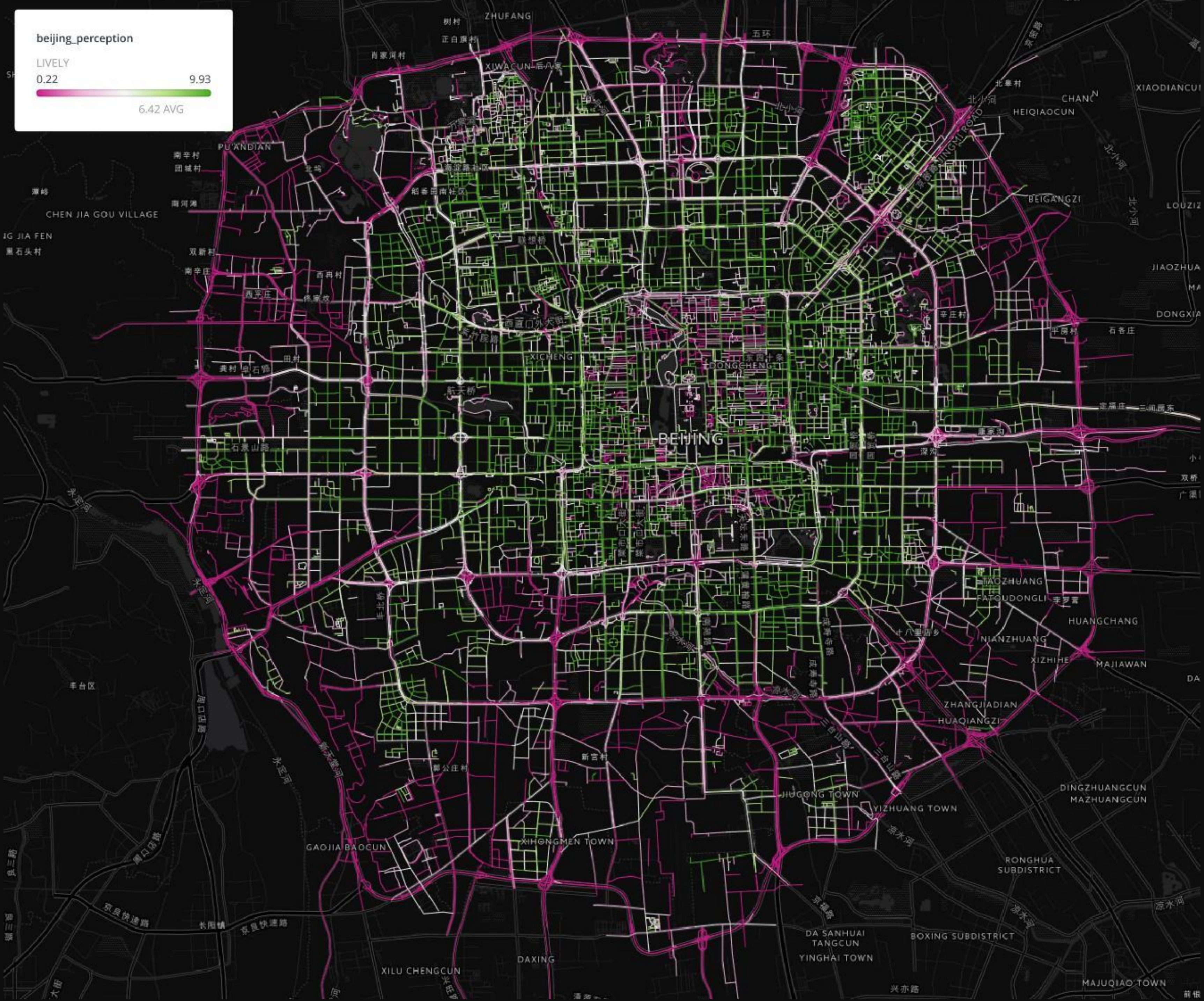
(3 < Q safe <= 7)



(Q safe <= 3)



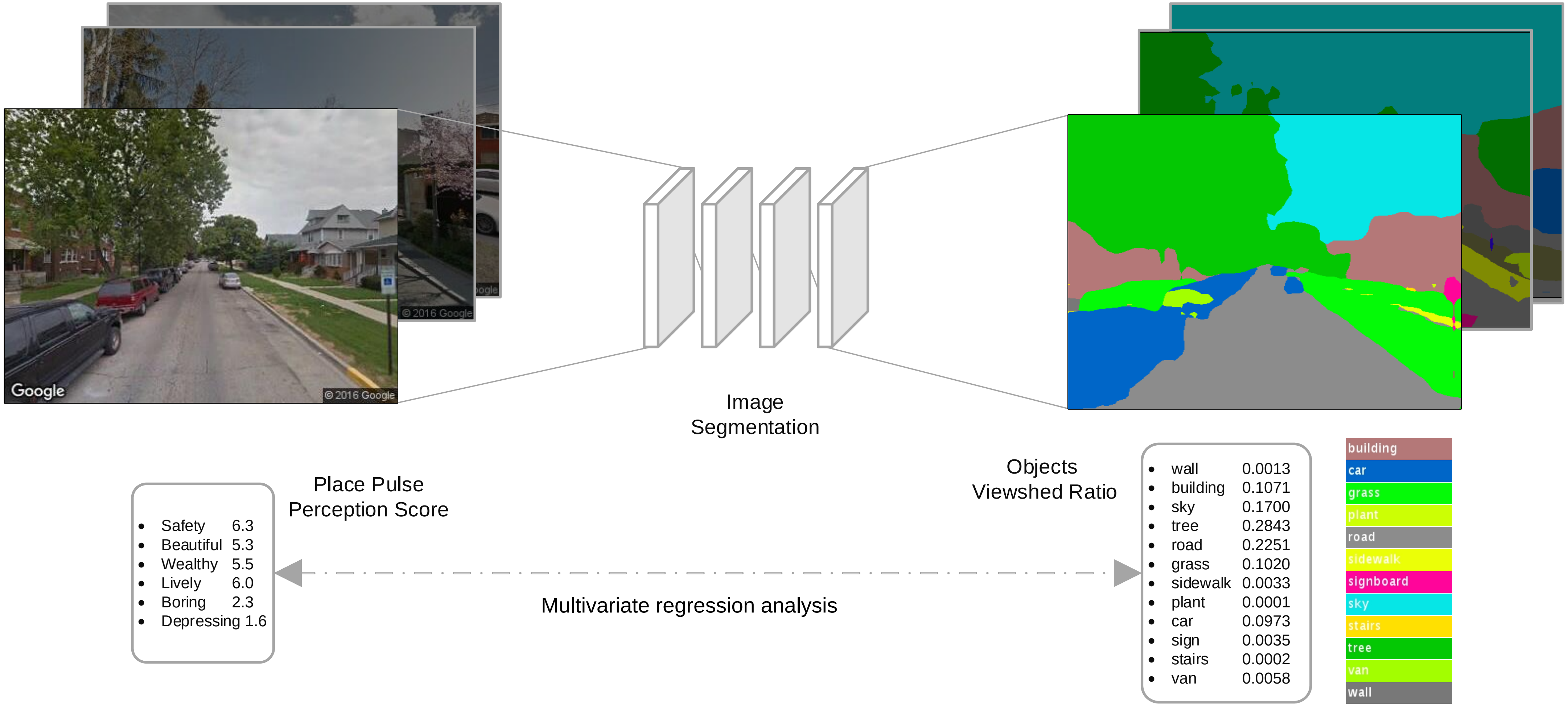
Beijing Perceptual Mapping - Safety


beijing_perception
LIVELY
0.22
9.93
6.42 AVG
ZHUFANG
XIWANGCUN
PU'ANDIAN
CHEN JIA GOU VILLAGE
HEIQIAOCUN
BEIGANGZI
XICHENG
DONGCHENG
BEIJING
FATOUDONGLI
HUANGCHANG
NIANZHUANG
XIZHIHE
MAJIAWAN
ZHANGJIADIAN
HUAQIANGZI
JIUGONG TOWN
YIZHUANG TOWN
DINGZHUANGCUN
MAZHUANGCUN
RONGHUA
SUBDISTRICT
GAOJIA BAOCUN
XIHONGMEN TOWN
DA SANHUAI
TANGCUN
YINGHAI TOWN
BOXING SUBDISTRICT
DAXING
XILU CHENGCUN
MAJUQIAO TOWN

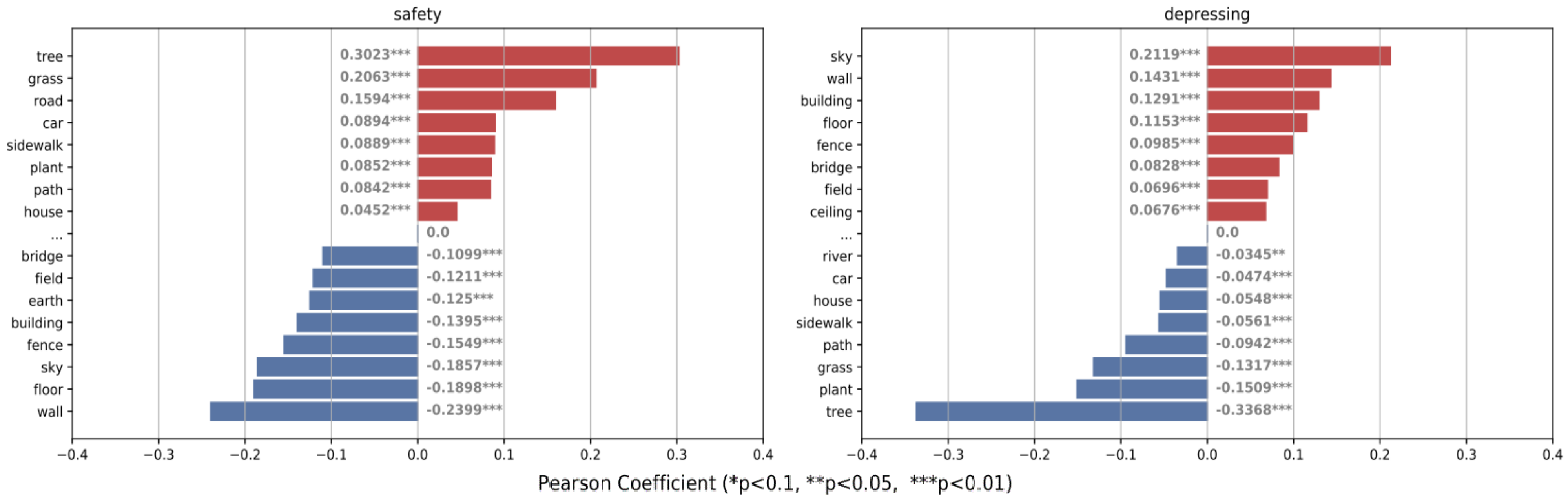
Shanghai Perceptual Mapping - Safety

Shanghai Perceptual Mapping - Lively

What visual elements are associated with perceptions?



What visual elements are associated with perceptions?



Fan Zhang, Bolei Zhou, Liu Liu, Yu Liu, Helene H. Fung, Hui Lin, and Carlo Ratti. **Measuring human perceptions of a large-scale urban region using machine learning**. *Landscape and Urban Planning*, 180:148–160, 2018

Results:

- Greenery, vehicles, etc. are associated with positive perceptions
- Wall, buildings, large open sky, etc. are associated with negative perceptions

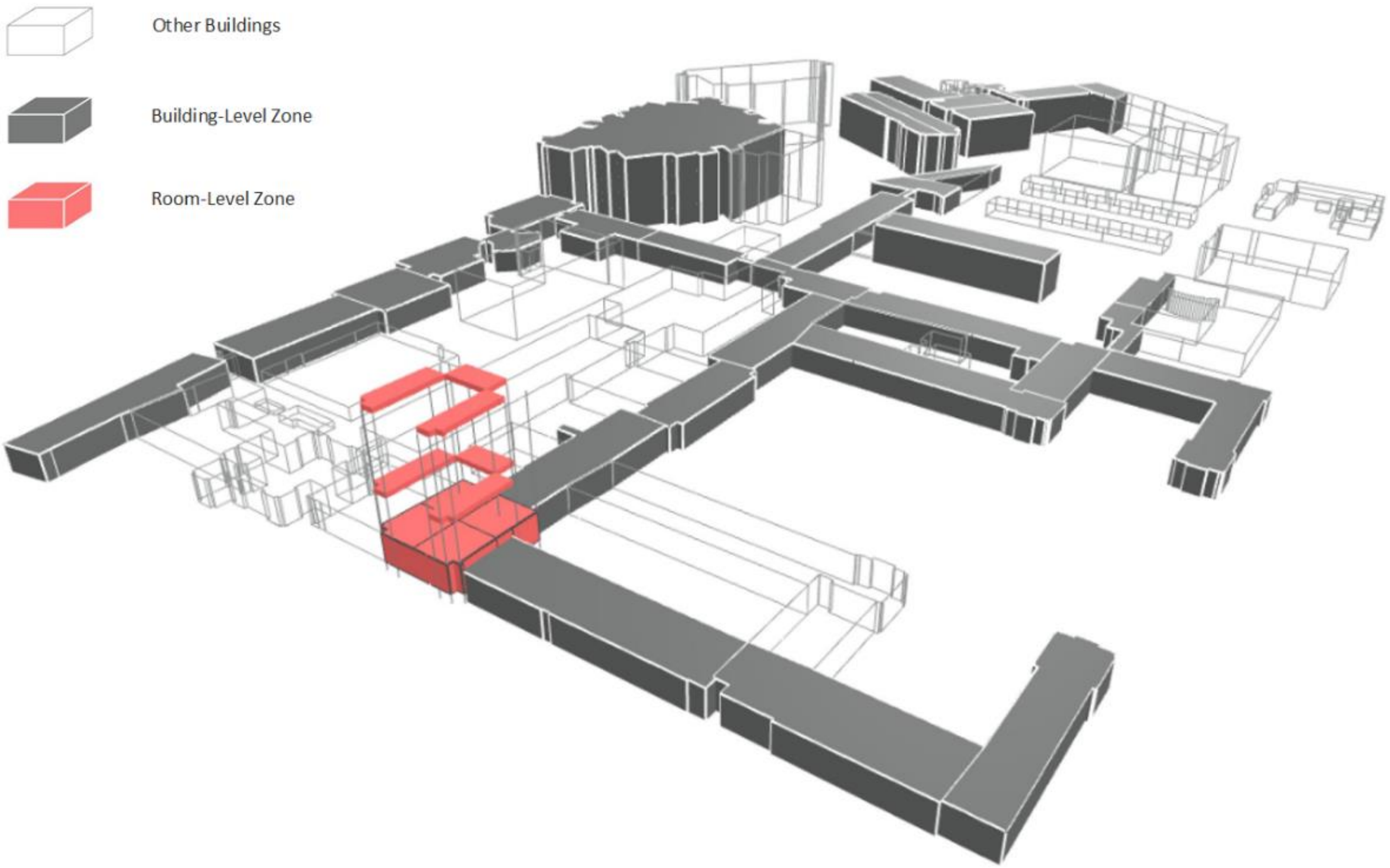
Active Data Collection

Infinite Corridor

NAVIGATION, COMPUTER VISION


EXIT
Building 56
Main Campus


MIT Campus



Data Collection

GoPro

Building 1_C0

Building 1_C1

Building 2_C0

Building 2_C1

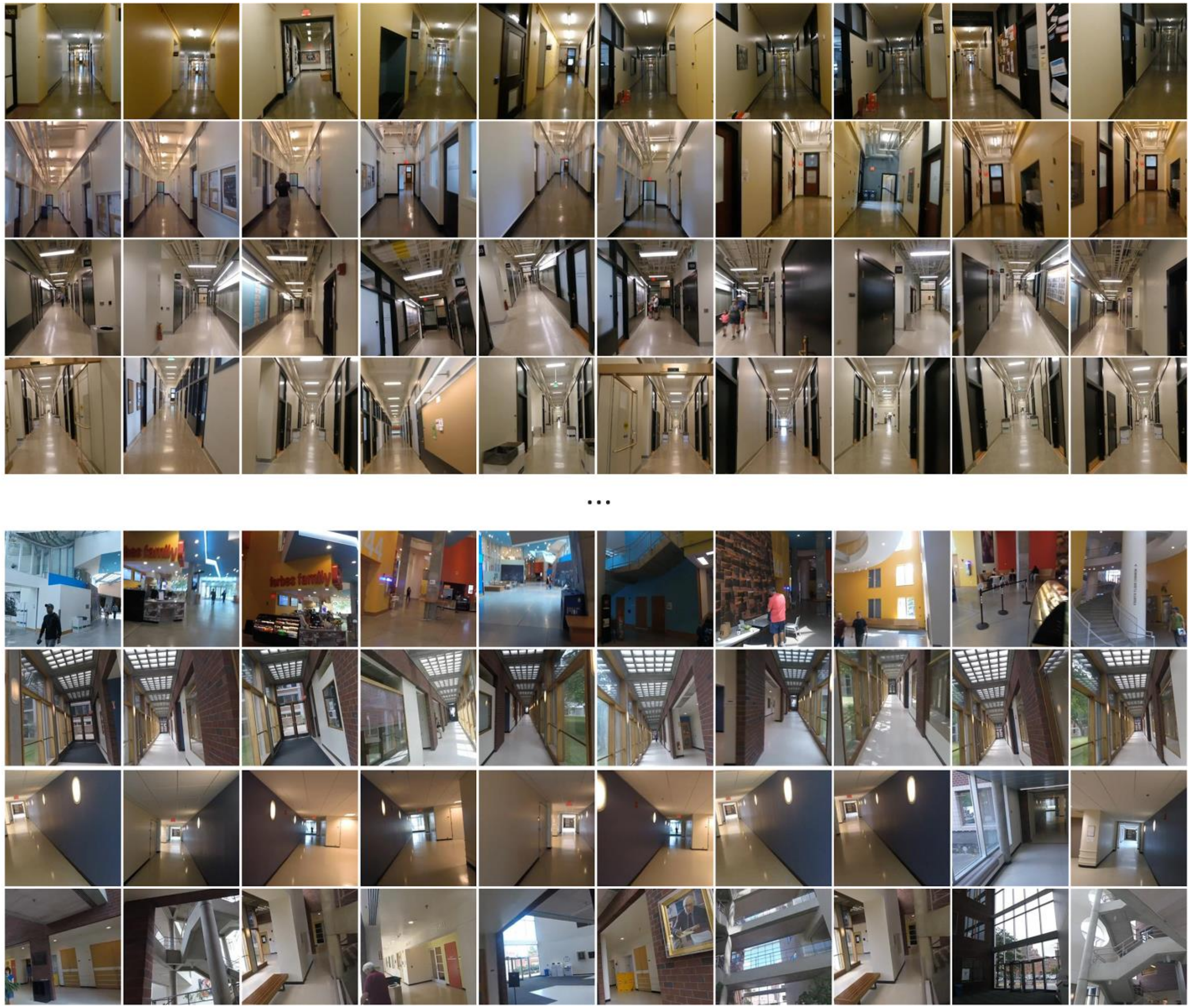
...

Building 32_L0

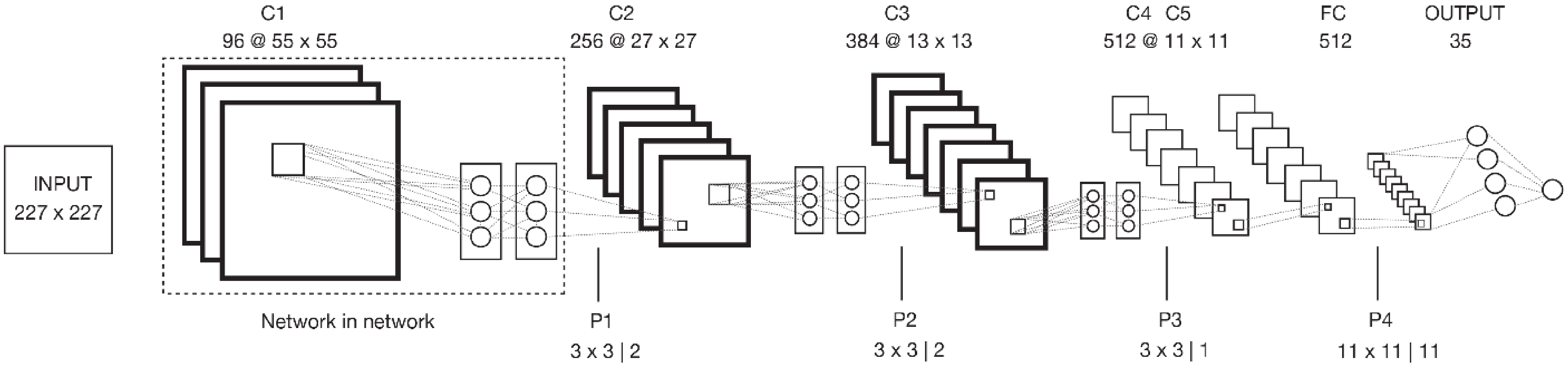
Building E25_C0

Building E25_C1

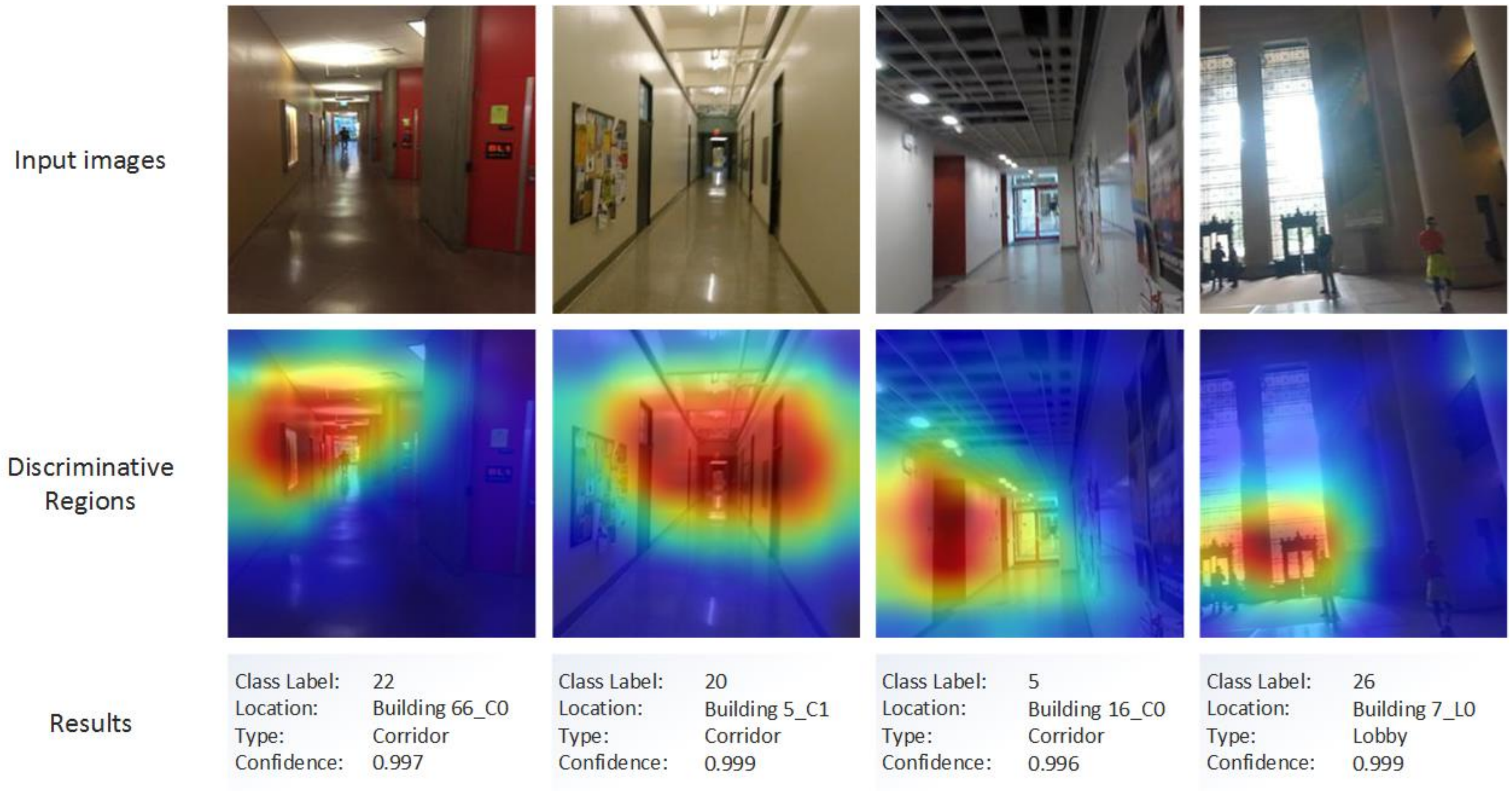
Building E25_L2



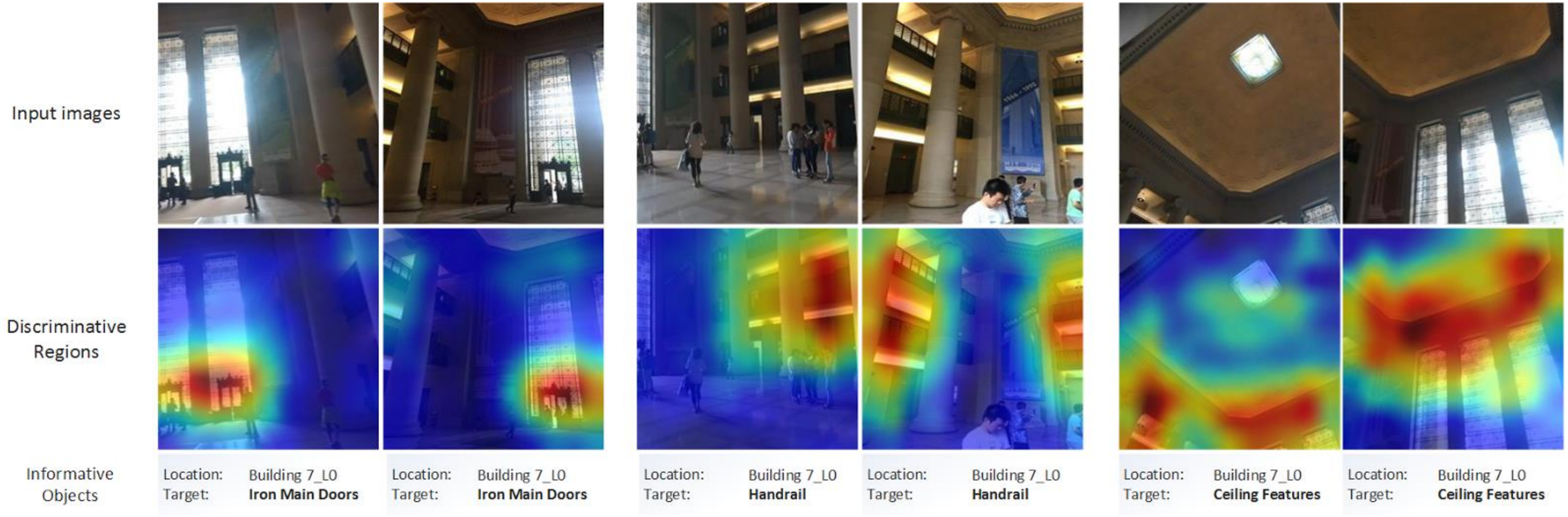
Deep Convolutional Neural Network



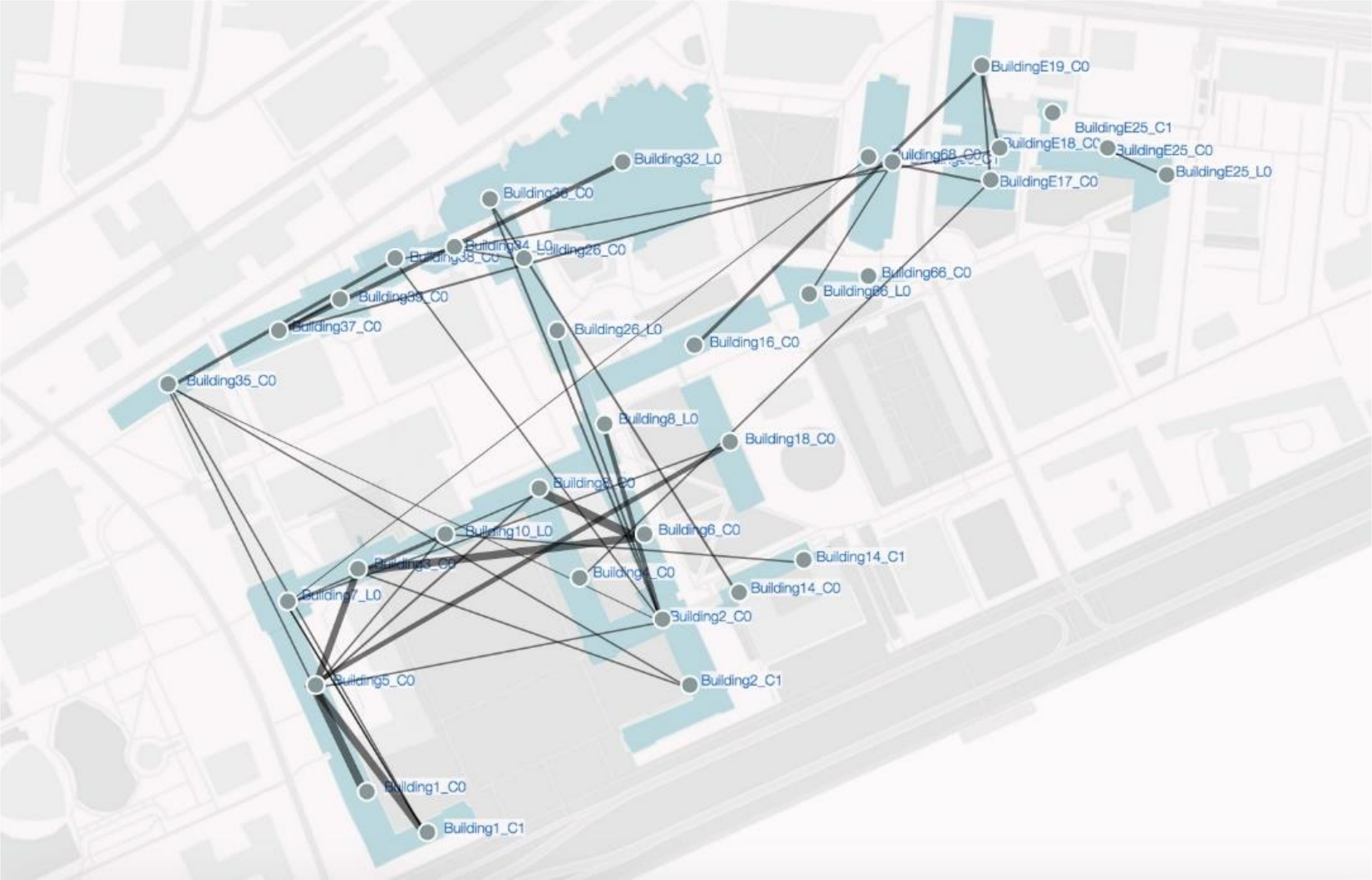
What features did the model learn?



What features did the model learn?



Spatial Analysis – Building Similarity



Spatial Analysis – Building Distinctiveness

What makes the building look different?

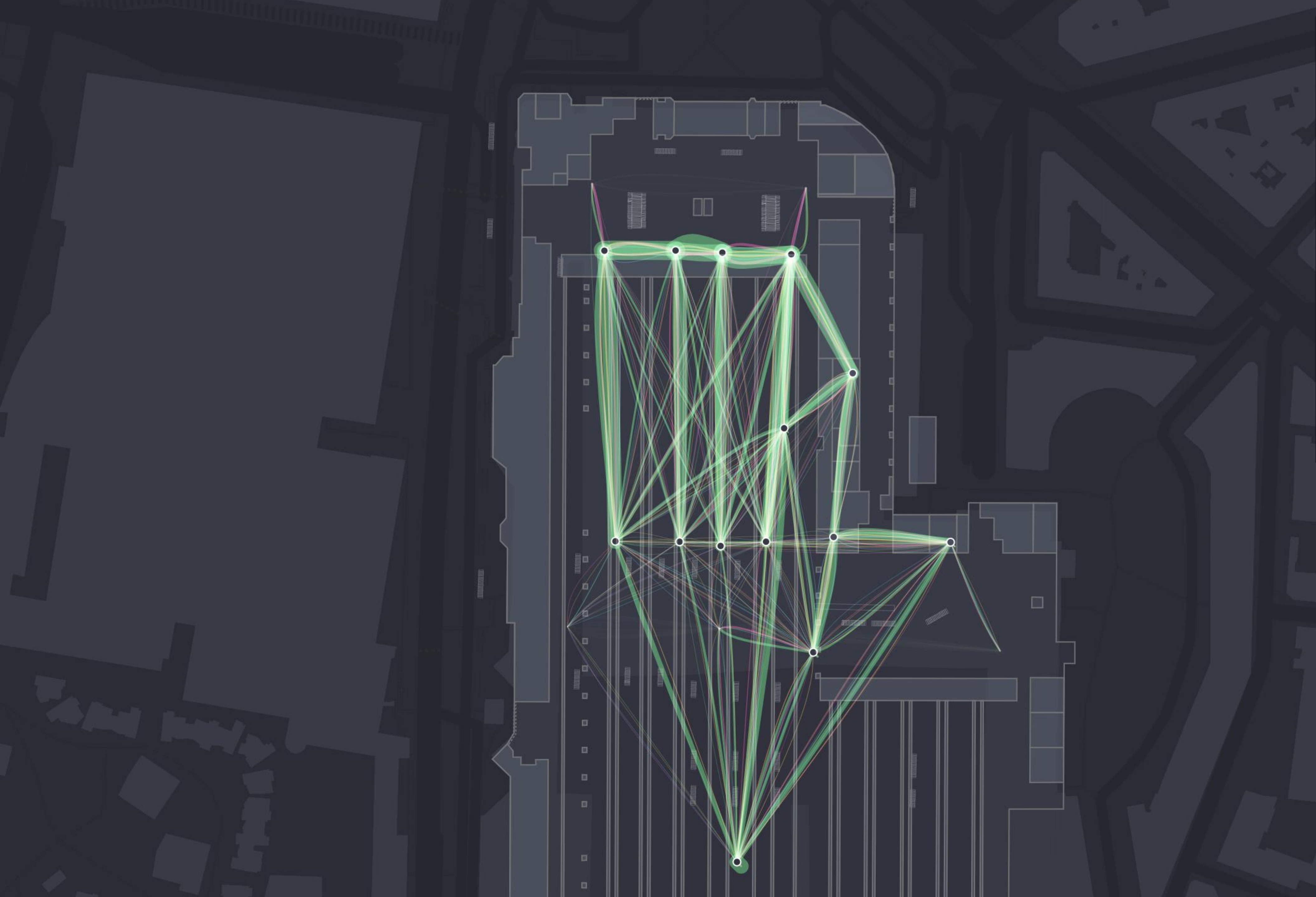
Building 34_L0 | Building 14_C1 | Building 68_C1 | Building 8_L1 | Building 4_C0

AI Station

MOBILITY, SPATIAL PERCEPTION

This is Gare de Lyon,
one of the largest
train stations in France.

100,500,000
PASSENGERS/YEAR



Show map of:

Ground level

Level -1

ALL USERS

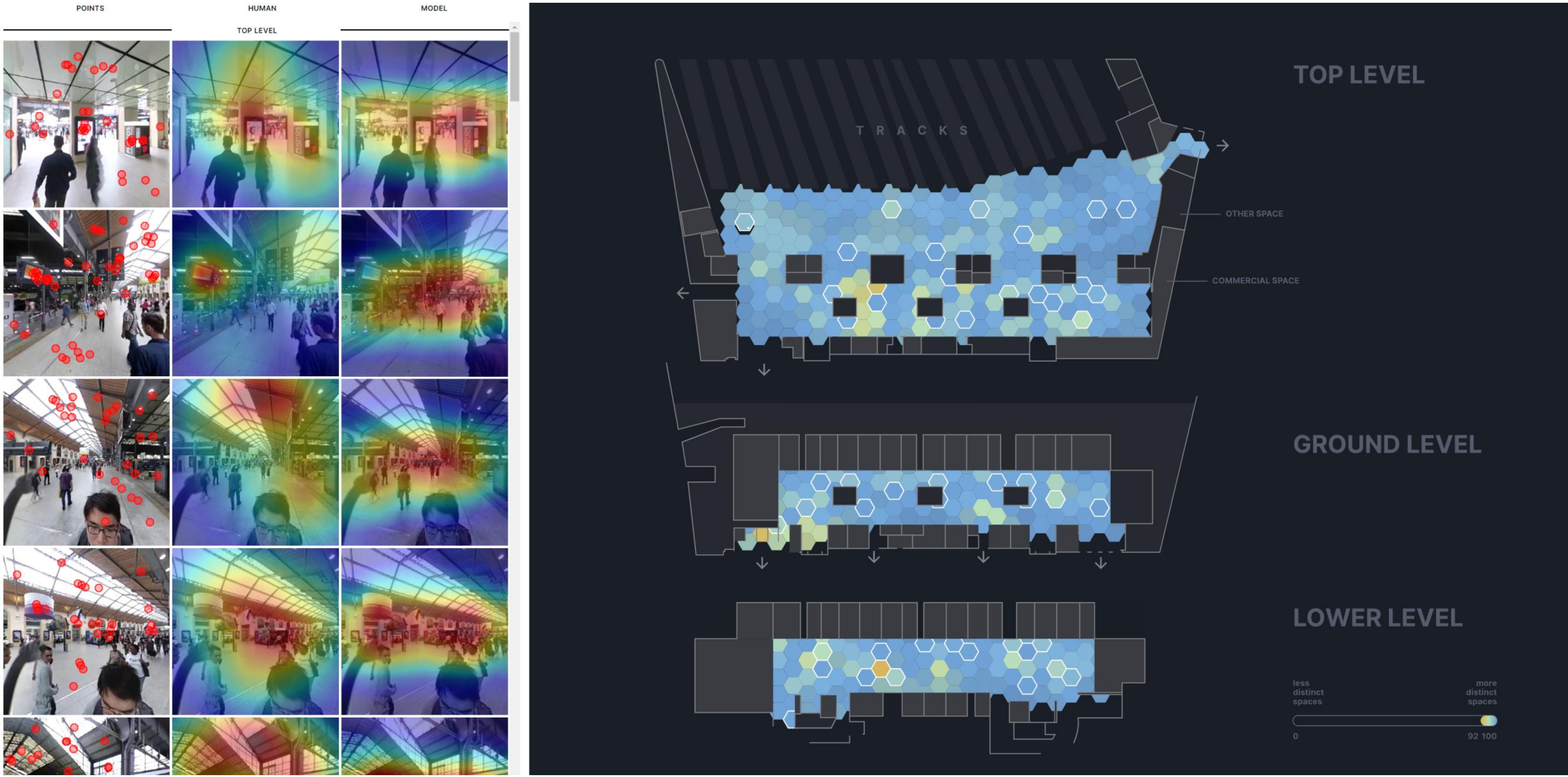
57,500 people

Density of trajectories

1 14,000 28,000

Mouse over the circles to see the zones connected to that area of the station


Google Earth



POINTS
HUMAN
MODEL
TOP LEVEL
T R A C K S
TOP LEVEL
OTHER SPACE
COMMERCIAL SPACE
GROUND LEVEL
LOWER LEVEL
less distinct spaces
more distinct spaces
0
92 100

Smart Curbs

HUMAN DYNAMICS, COMPUTER VISION

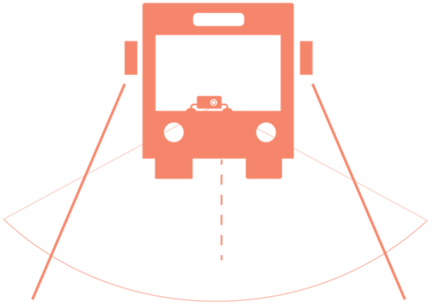
Curbs are the urban asset of tomorrow.

Curbs are the urban asset of tomorrow.

Leverage *Artificial Intelligence* to understand how curbs are used in *real-time,* to better manage them.

traffic light
traffic light
pole
pole
pole
sign
pole
traffic light
traffic light
traffic light
traffic light
car
bus
bicycle
bicycle
person
person
person
person
person
person
van
person
person
person
person
person
handbag
suitcase
handbag
bicycle

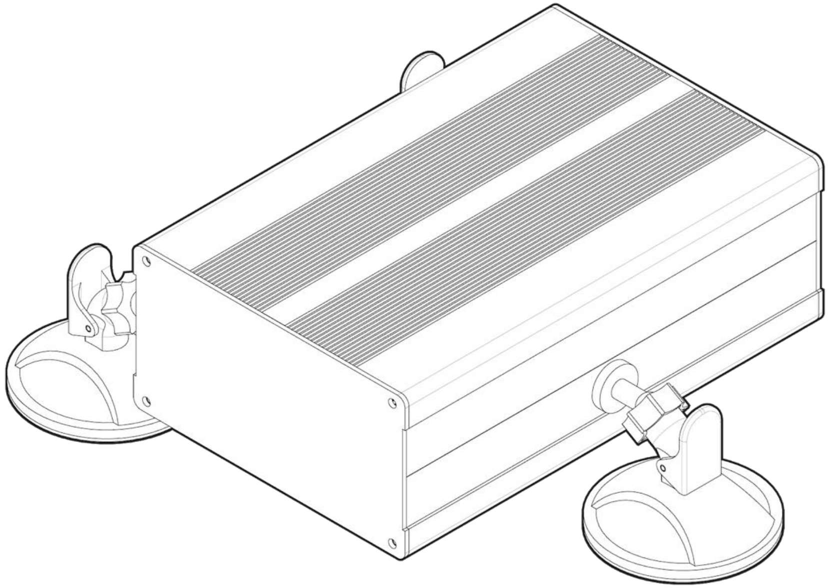
Measuring Human Activity

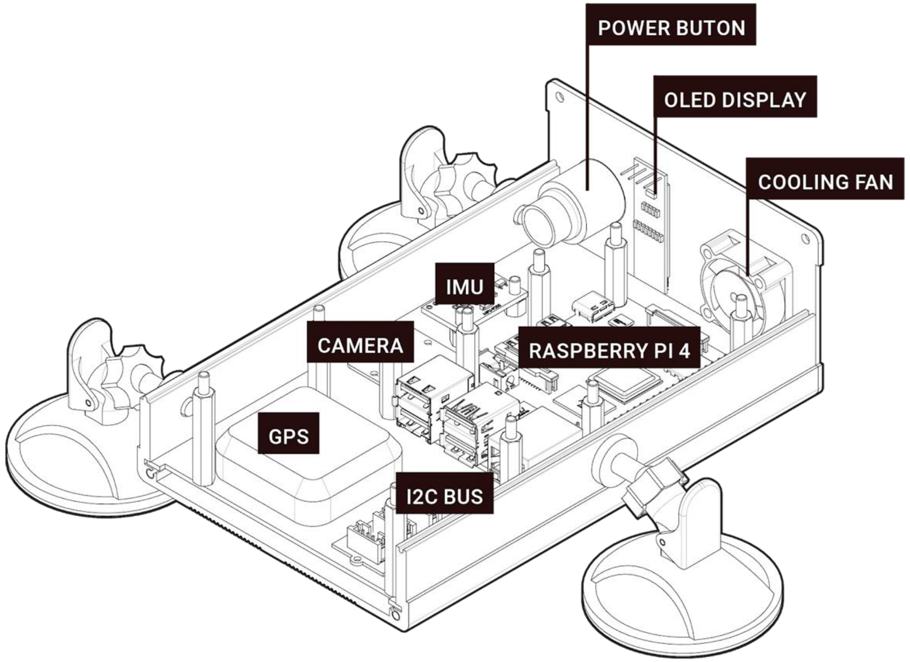


Approach
equip RATP buses with camera-based system

1. Computer Vision Model
2. Hardware Design

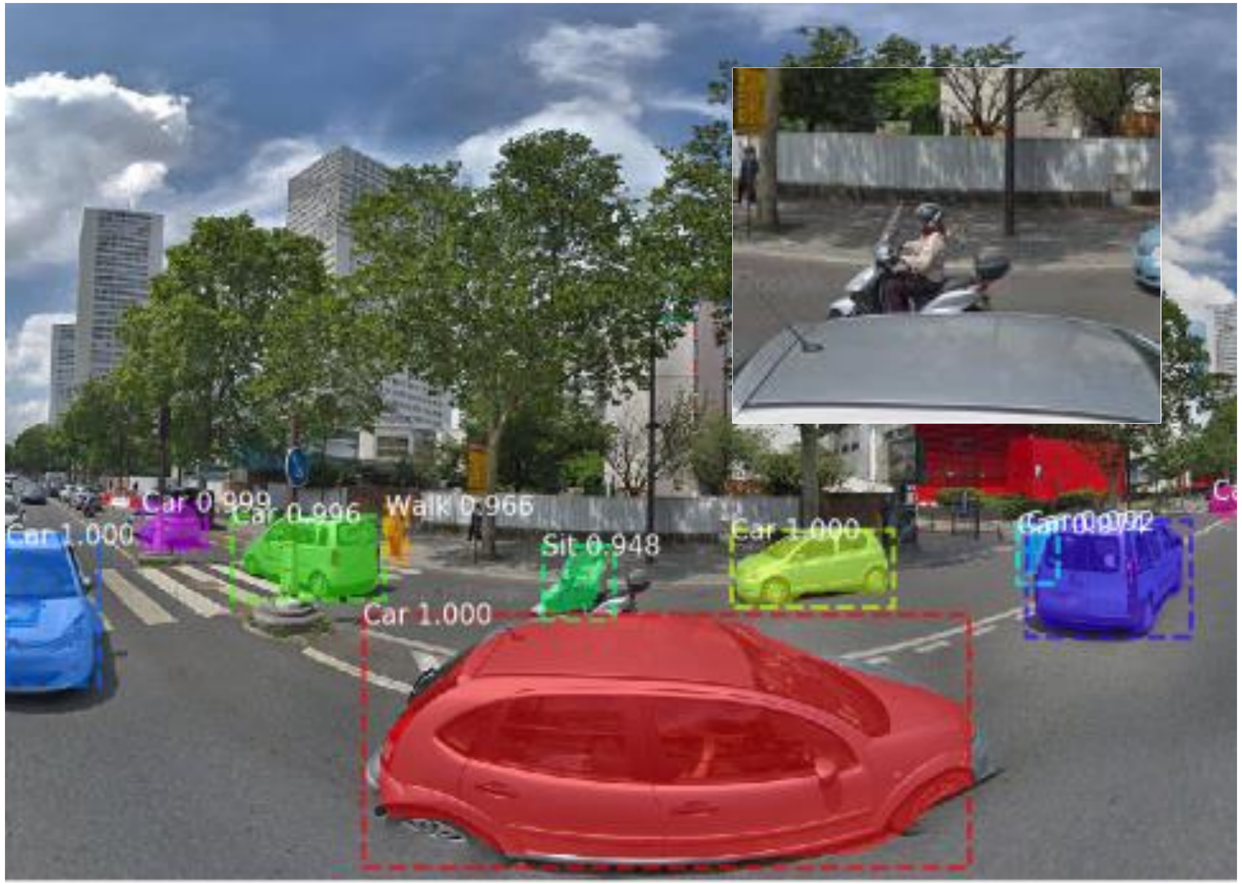
Hardware Design





Testing AI Model in Paris Street - Example

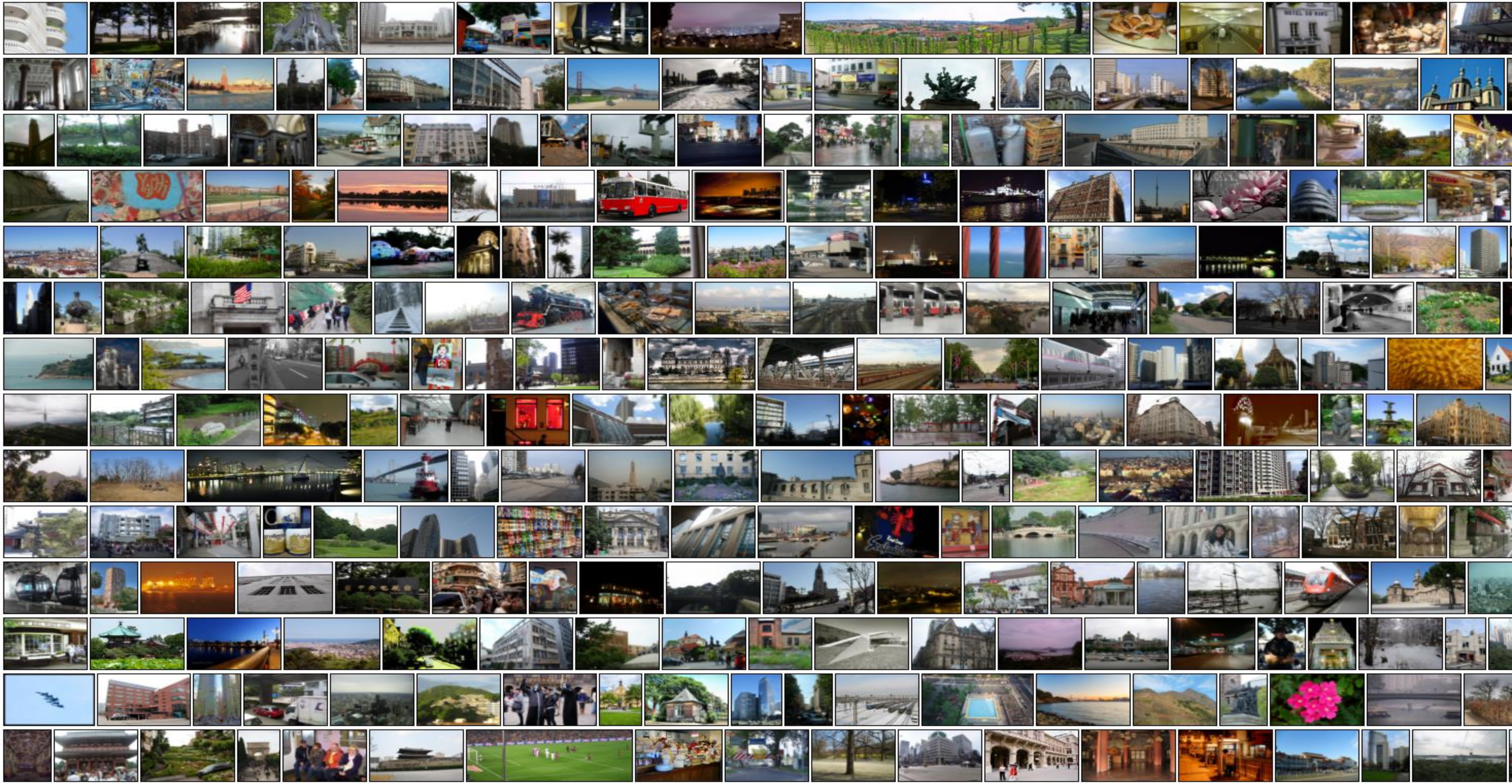
Testing AI Model in Paris Street - Example

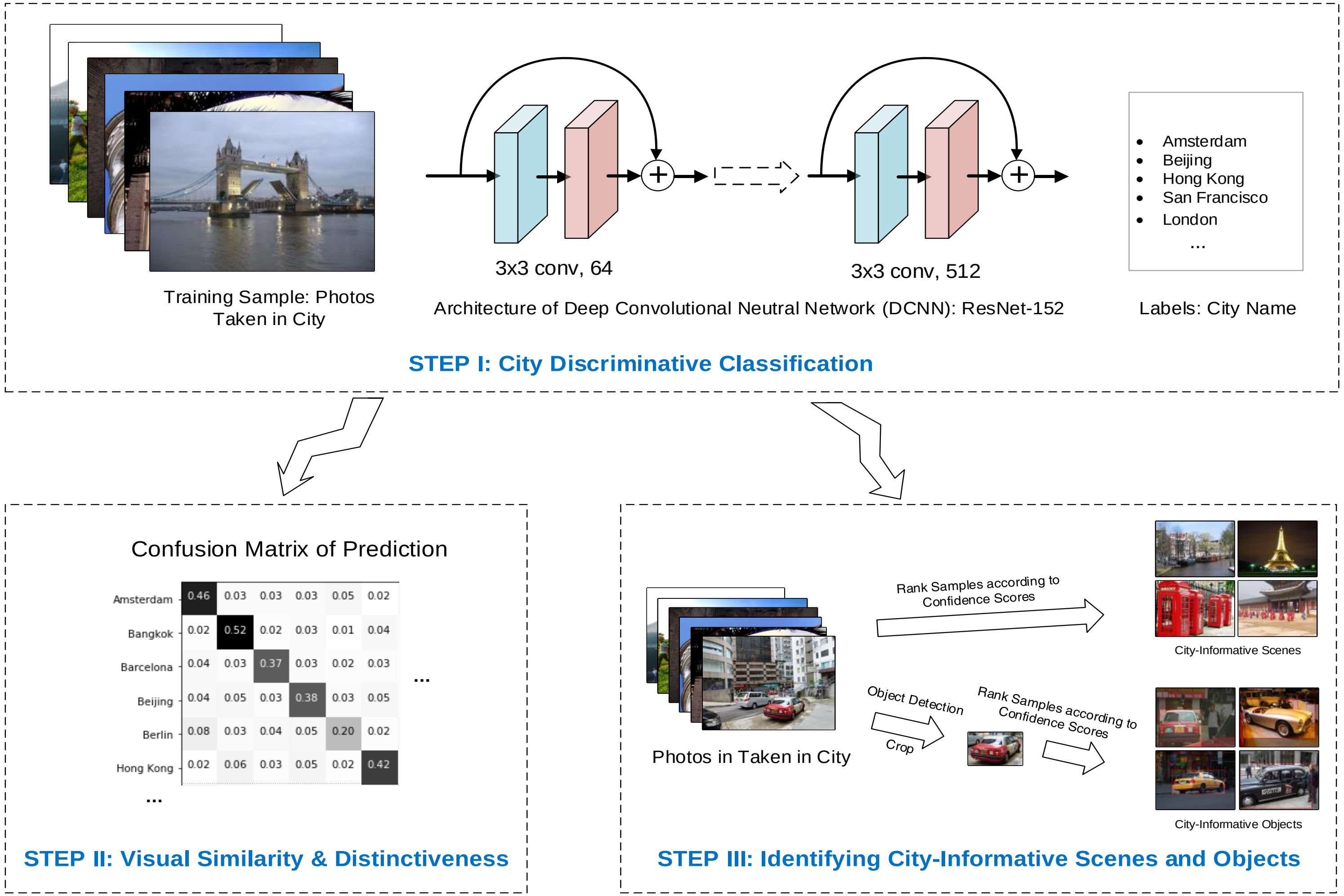


Crowdsourced Images

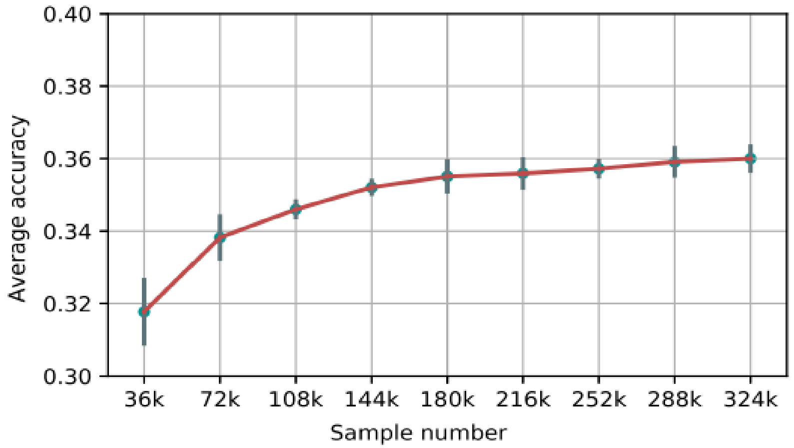
[In]Distinct Cities

TOURISM, VISUAL INFORMATICS



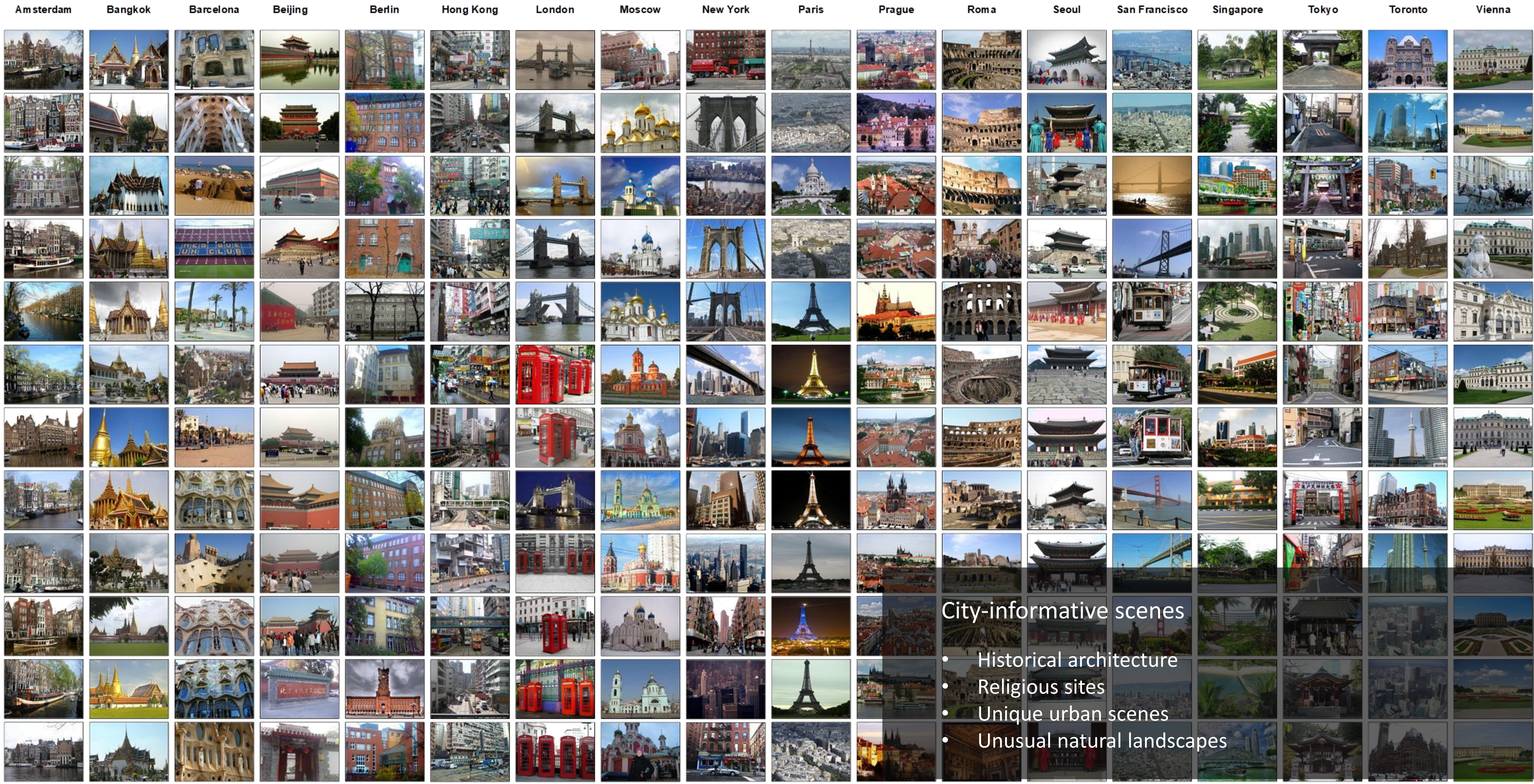


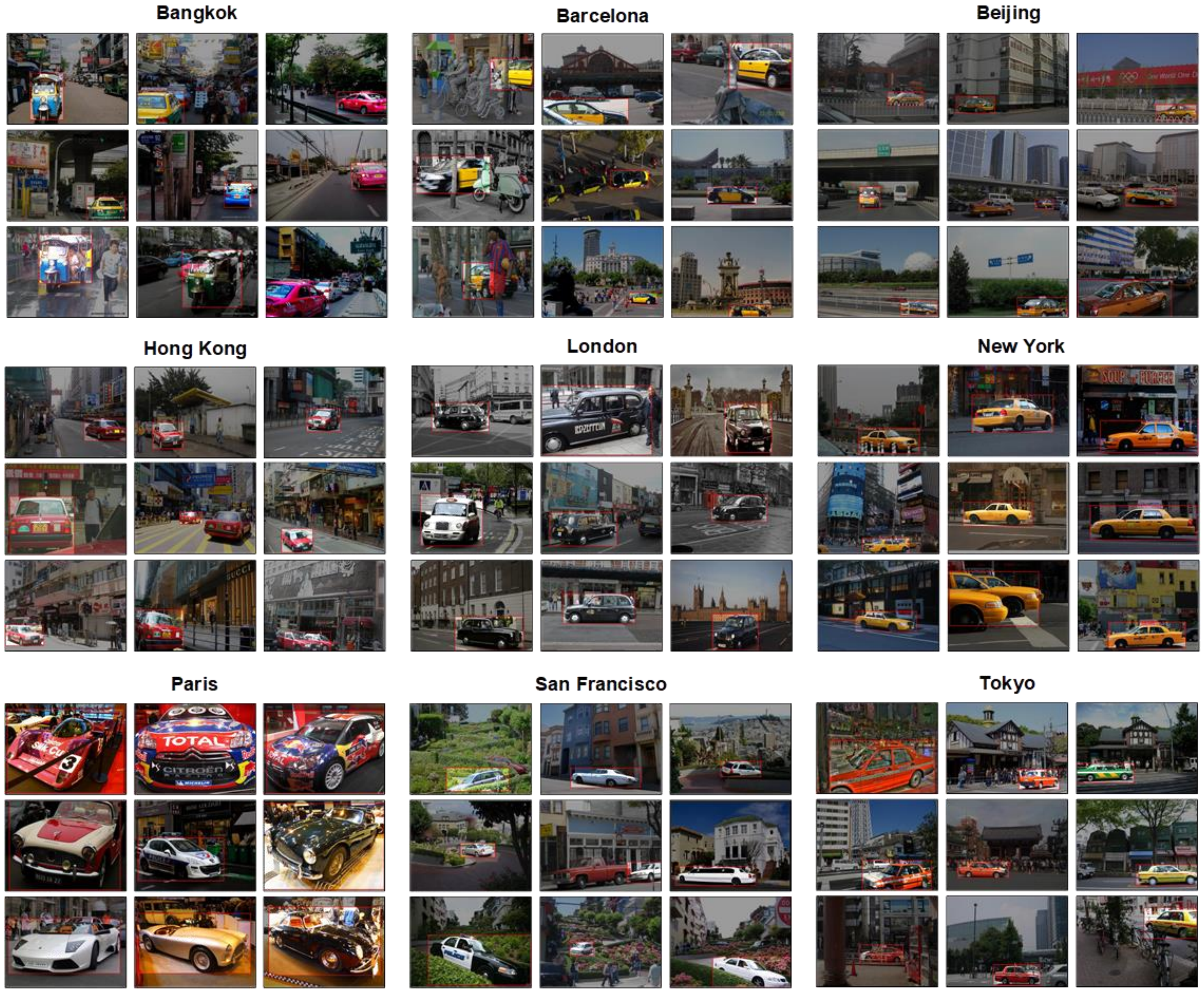
- Learning deep representations of urban scenes using deep learning
- Metric of similarity among cities; mining most representative image samples



- 36% accuracy in the 18-city image recognition task
- (Given any street image, there is a 36% probability of identifying the city it comes from)

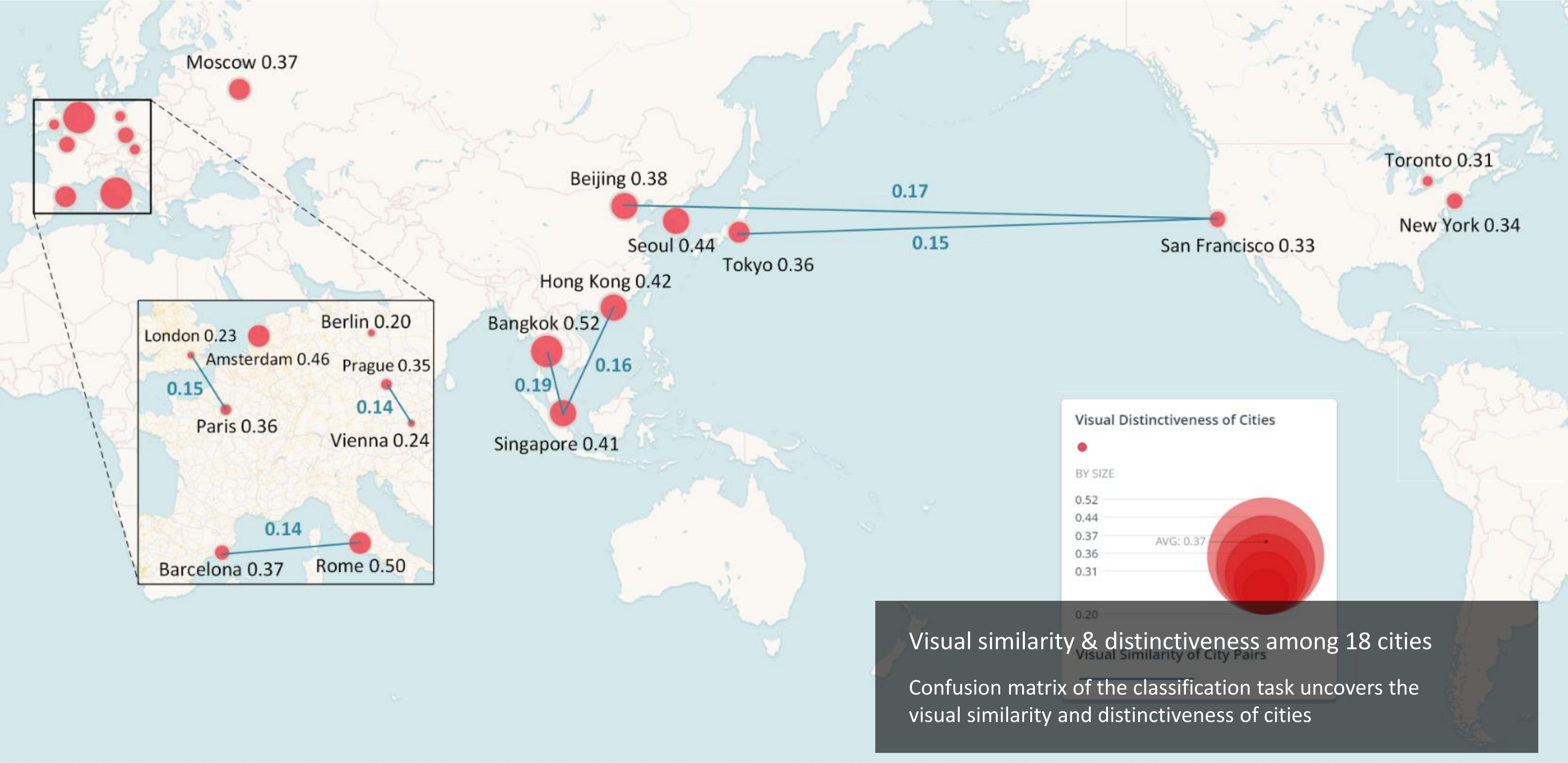
Zhang, Fan, Bolei Zhou, Carlo Ratti, and Yu Liu. "**Discovering place-informative scenes and objects using social media photos**." *Royal Society open science* 6, no. 3 (2019): 181375.


Amsterdam
Bangkok
Barcelona
Beijing
Berlin
Hong Kong
London
Moscow
New York
Paris
Prague
Roma
Seoul
San Francisco
Singapore
Tokyo
Toronto
Vienna
City-informative scenes
• Historical architecture
• Religious sites
• Unique urban scenes
• Unusual natural landscapes



❑ City-informative Objects (vehicles)

- Globalization has led to cities around the world using the same vehicle brands and similar shapes; however, interestingly, the model still uncovers the vehicles that representative of each city, such as cabs, buses, police cars, etc.
- Other objects? Building façade, trees, store signs, and dressing styles, etc.
- Understand cities from a different perspective



Visual similarity & distinctiveness among 18 cities

Confusion matrix of the classification task uncovers the visual similarity and distinctiveness of cities

http://senseable.mit.edu/indistinct_cities/

senseable
city lab.

Instagram
@Senseable_City_Lab

Facebook
@SenseableCity

Twitter
@SenseableCity

Fan Zhang
zhangfan@mit.edu

Fábio Duarte
fduarte@mit.edu