Deep Learning for Automatically Detecting Sidewalk Accessibility Problems Using Streetscape Imagery

Galen Weld, Esther Jang, Anthony Li, Aileen Zeng, Kurtis Heimerl, Jon Froehlich

MAKEABILITY LAB

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

UNIVERSITY of WASHINGTON


30.6
million U.S. adults
have a mobility impairment
Source: 2010 U.S. Census


15.2
million use an assistive aid
Source: 2010 U.S. Census


Olive Wy
© 2012 Google


CURB RAMPS
Olive Wy


MISSING CURB RAMPS
CHANEL
ONE WAY
DIESEL


OBSTACLES


SURFACE PROBLEMS

OBSTACLE
MISSING CURB RAMP
SURFACE PROBLEM

Problem: there is no comprehensive information on "the degree to which sidewalks are accessible" in cities.

- National Council on Disability, 2007

Traditional methods for gathering this information are time-consuming, laborious, and expensive.

https://sdotblog.seattle.gov/2017/08/28/seattle-sidewalk-survey-update/

Some automated methods
have been attempted…

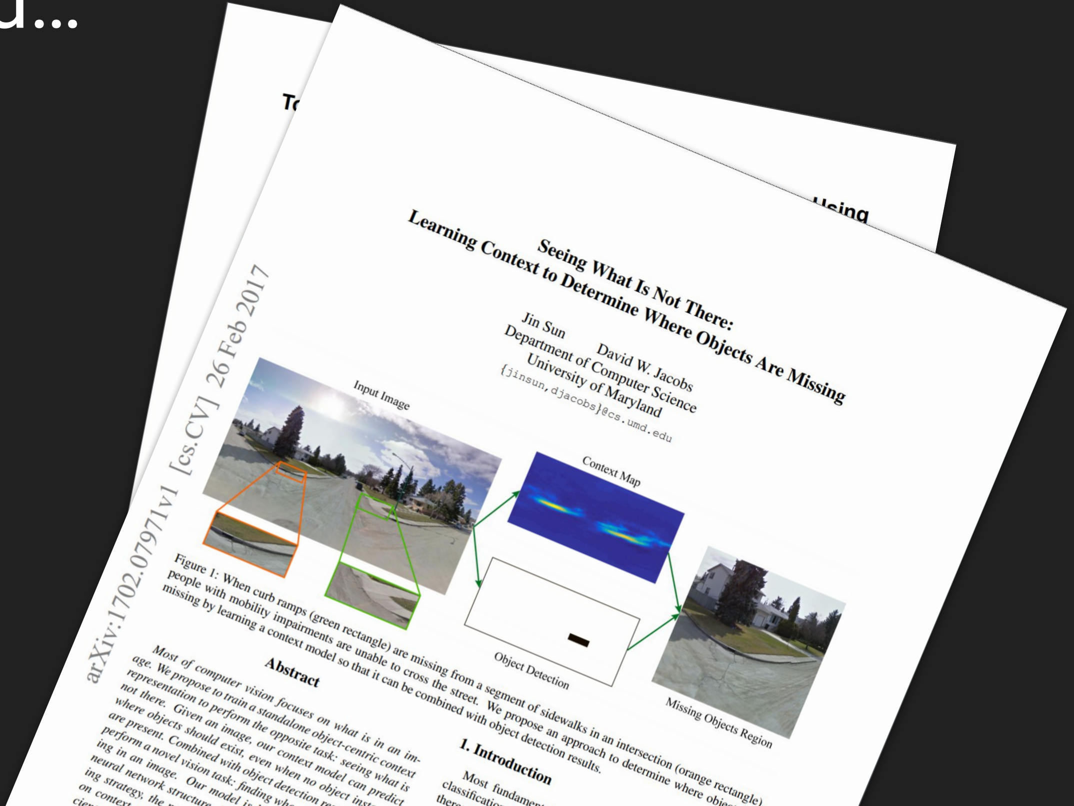
Some automated methods have been attempted…

26% precision
67% recall
for curb ramps

Kotaro Hara, Jin Sun, Robert Moore, David Jacobs, and Jon Froehlich. 2014. Tohme. In Proceedings of the 27th annual ACM Symposium on User interface software and technology - UIST '14.

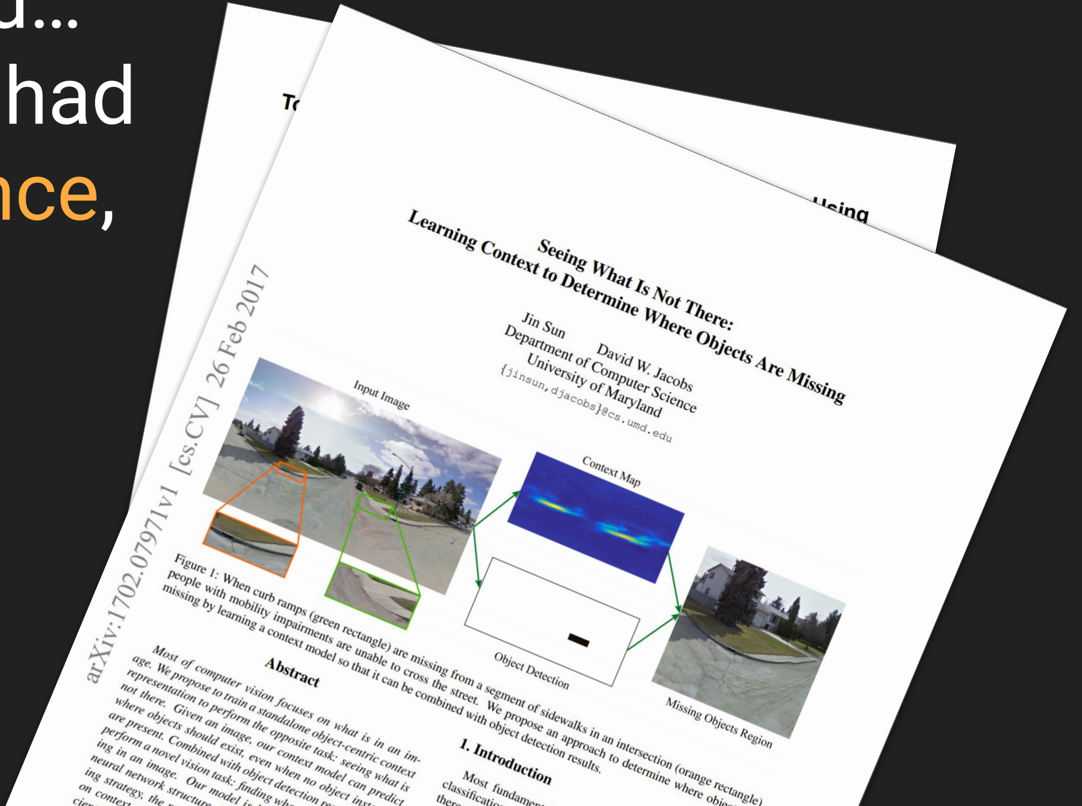
Some automated methods have been attempted…

27% recall
for missing curb ramps

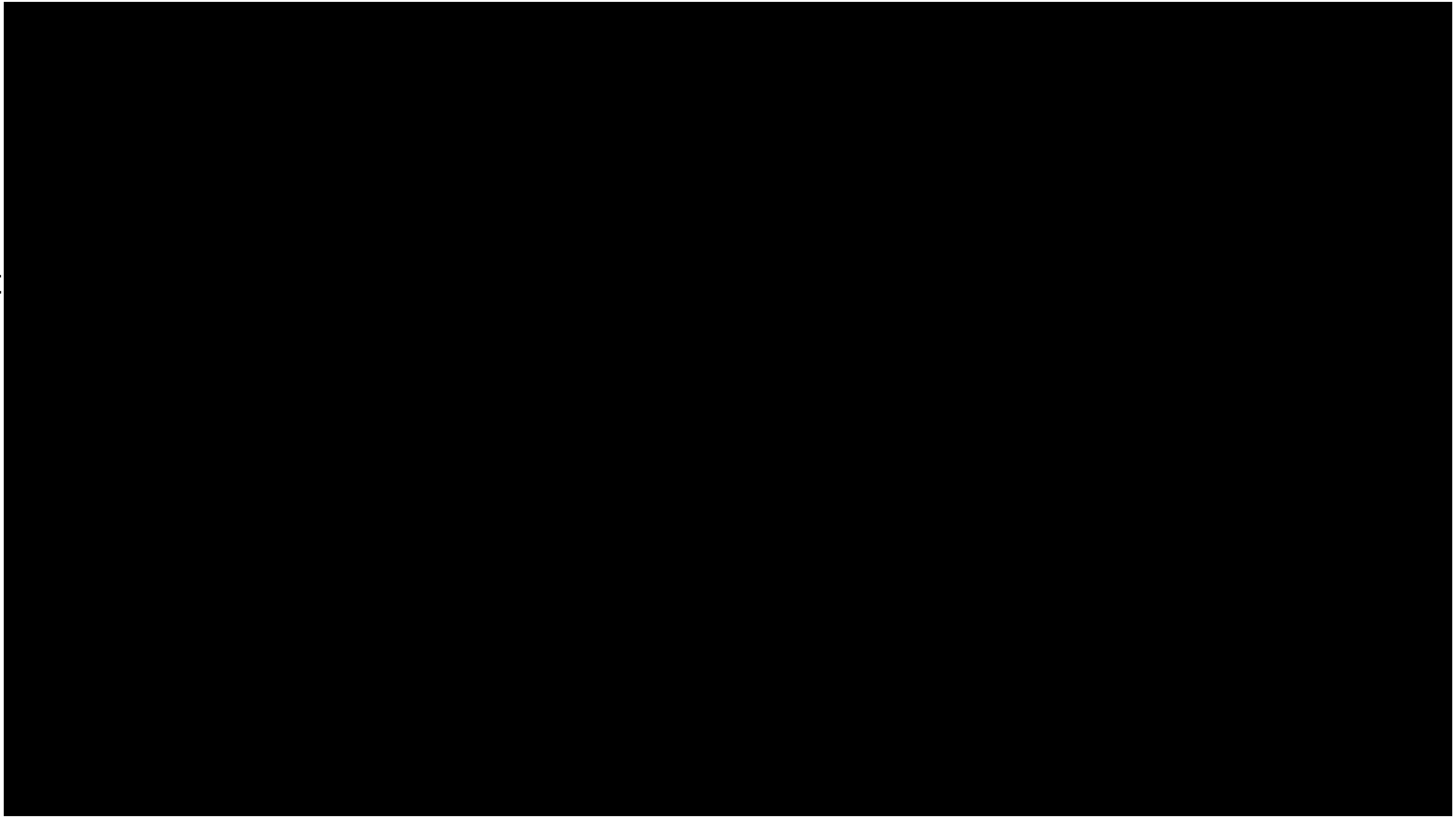


Jin Sun and David W. Jacobs. 2017. Seeing What is Not There: Learning Context to Determine Where Objects are Missing. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 1234–1242.

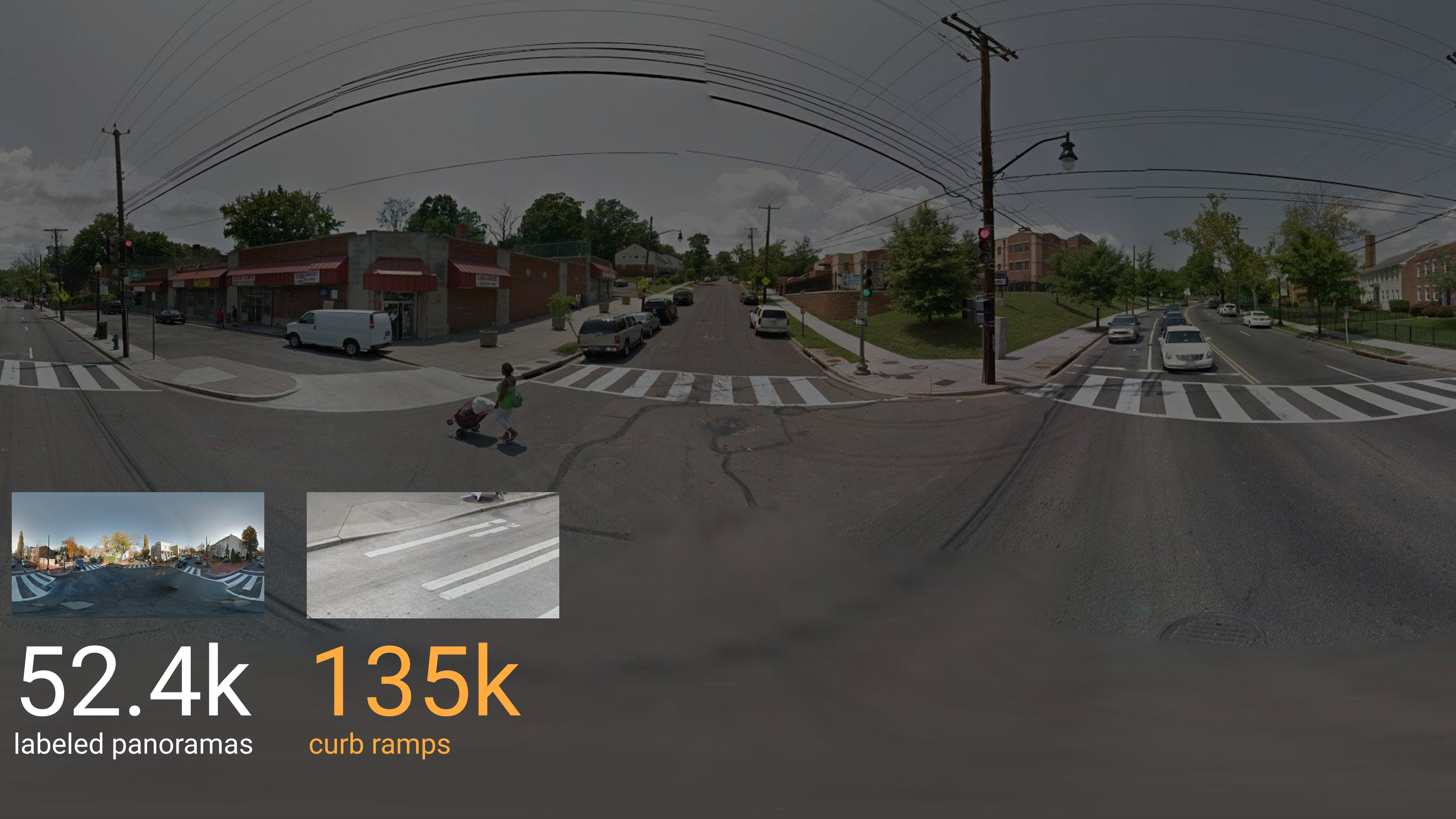
Some automated methods have been attempted… however these have had moderate performance, and narrow focus.



Crowdsourcing tools offer better performance, but are still slow and expensive.




52.4k
labeled panoramas



52.4k
labeled panoramas

135k
curb ramps

52.4k
labeled panoramas

135k
curb ramps

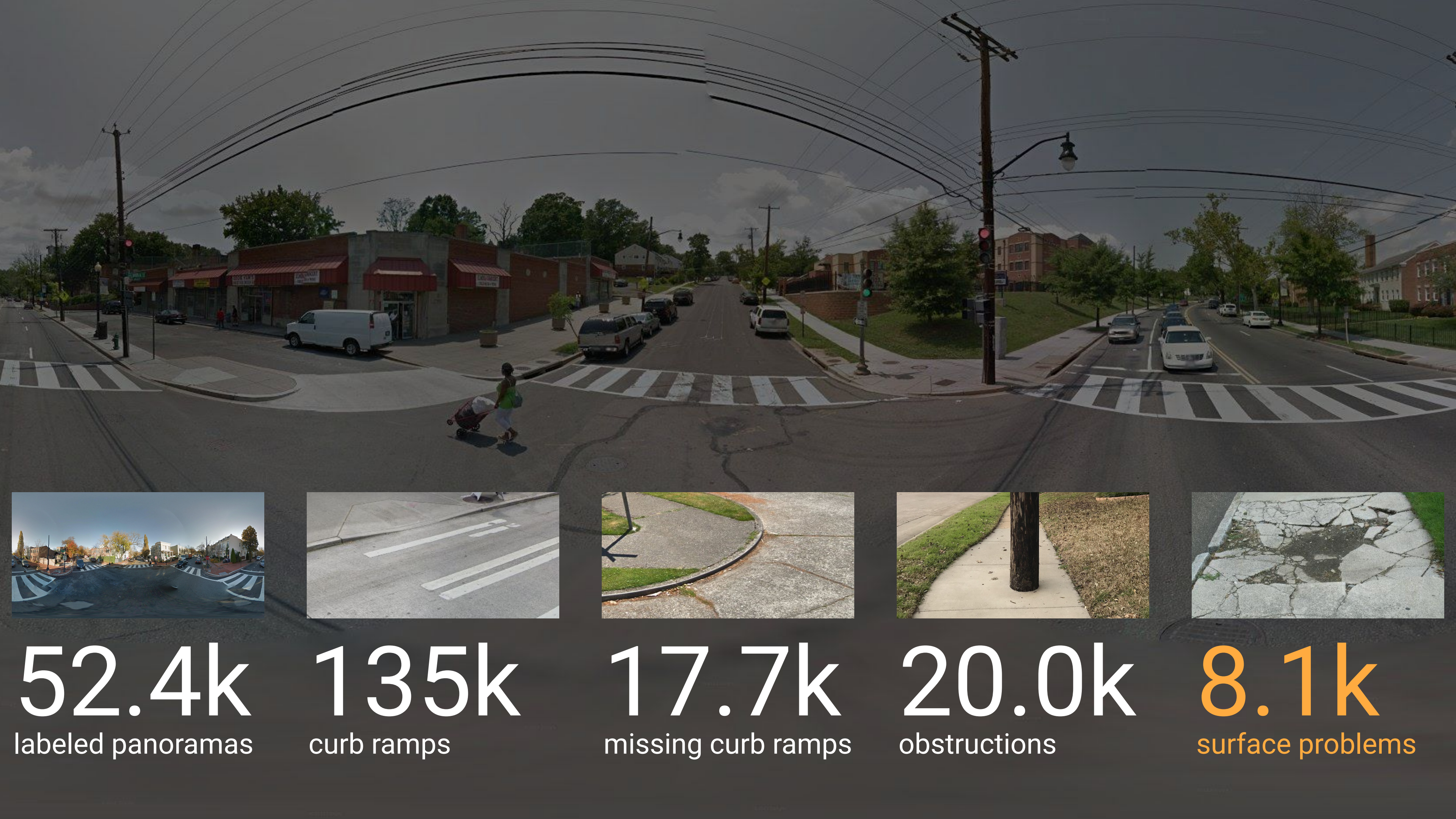
17.7k
missing curb ramps

52.4k
labeled panoramas

135k
curb ramps

17.7k
missing curb ramps

20.0k
obstructions


52.4k
labeled panoramas
135k
curb ramps
17.7k
missing curb ramps
20.0k
obstructions
8.1k
surface problems

Our Goal

Develop a system to automatically detect different types of sidewalk problems using streetscape imagery.

This system should be accurate, and generalizable to any city.

Two Automated Tasks

Validation

Is this an **obstruction**?

Labeling

What problems are in this pano?

Two Automated Tasks

Validation

Is this an
obstruction?

Two Automated Tasks

Validation

Is this an
obstruction?

Is this a
missing curb ramp?

Two Automated Tasks

Validation

Is this an
obstruction?

Is this a
missing curb ramp?

Is this a
curb ramp?

Two Automated Tasks

Validation

Is this an
obstruction?

Is this a
missing curb ramp?

Is this a
curb ramp?

Is this an
obstruction?

Two Automated Tasks

Validation

Is this an **obstacle**?

Labeling

What problems are in this pano?

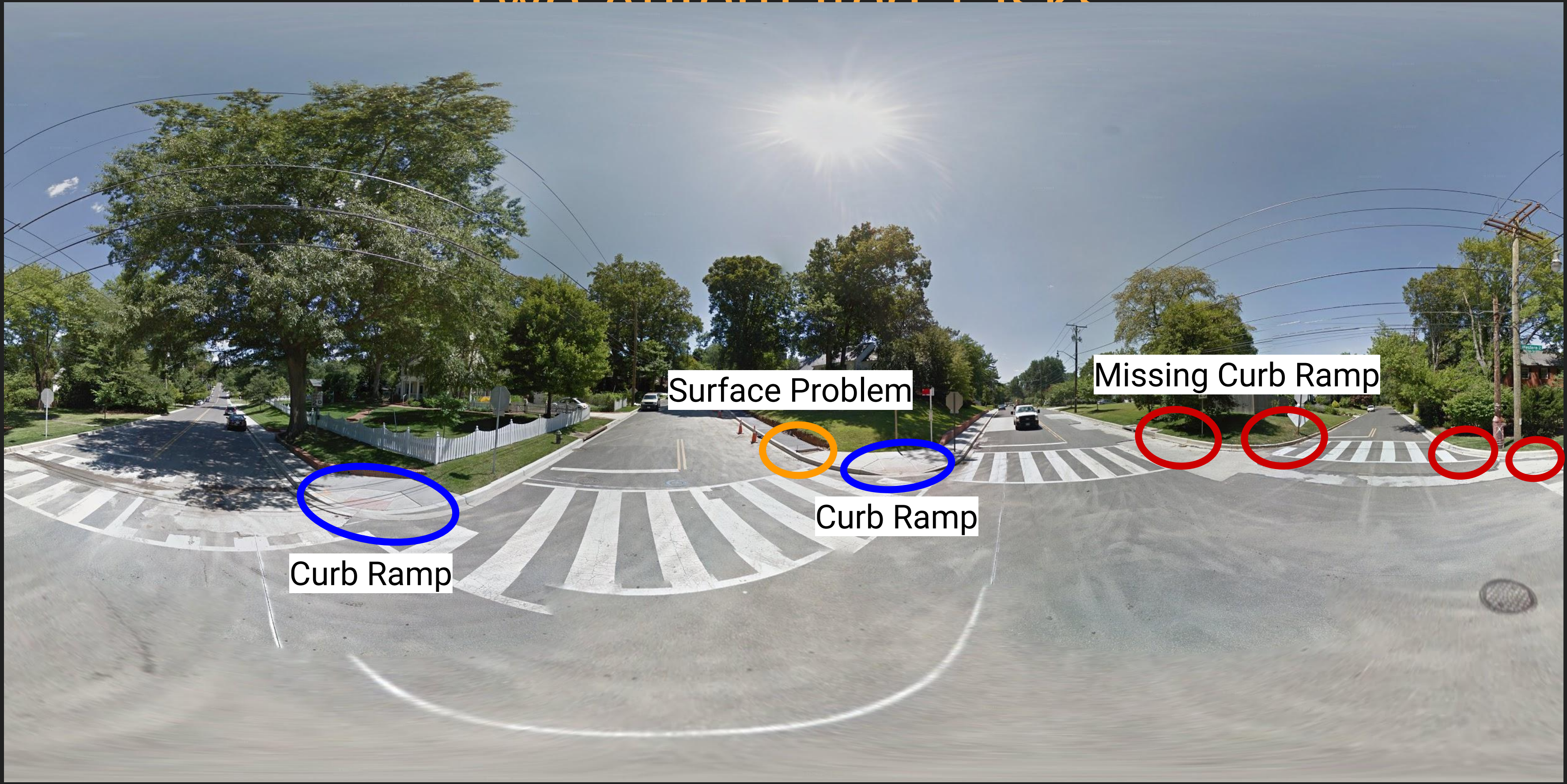
Two Automated Tasks

Two Automated Tasks

Two Automated Tasks

Two Automated Tasks

Two Automated Tasks



How do we automate these tasks?

1. Start with 181k labeled problems from Project Sidewalk dataset.
2. Compute 3 types of features for each human-placed label.
3. Train two different neural networks, one for validation, one for labeling.
4. Use a sliding window to label panoramas.
5. Evaluate on a researcher-created ground-truth test dataset.

3 types of features

image features

positional features

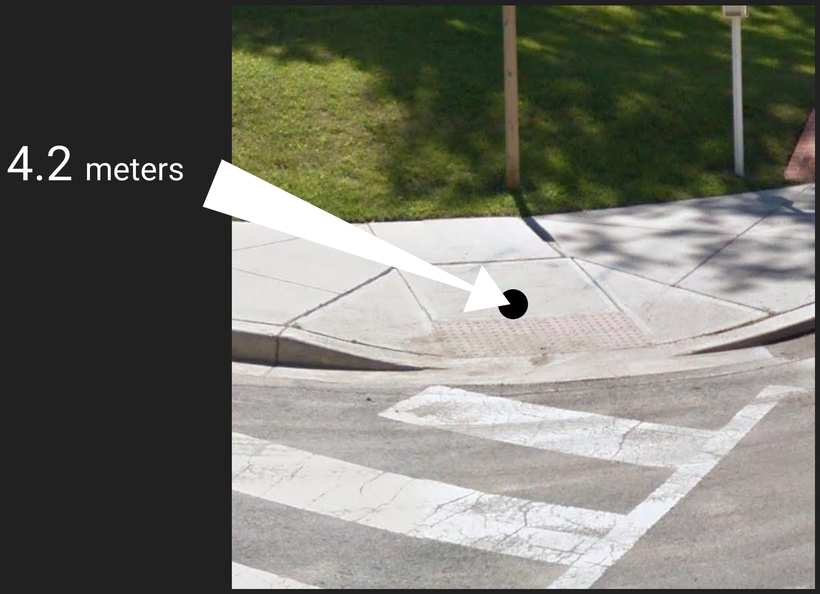
geographic features

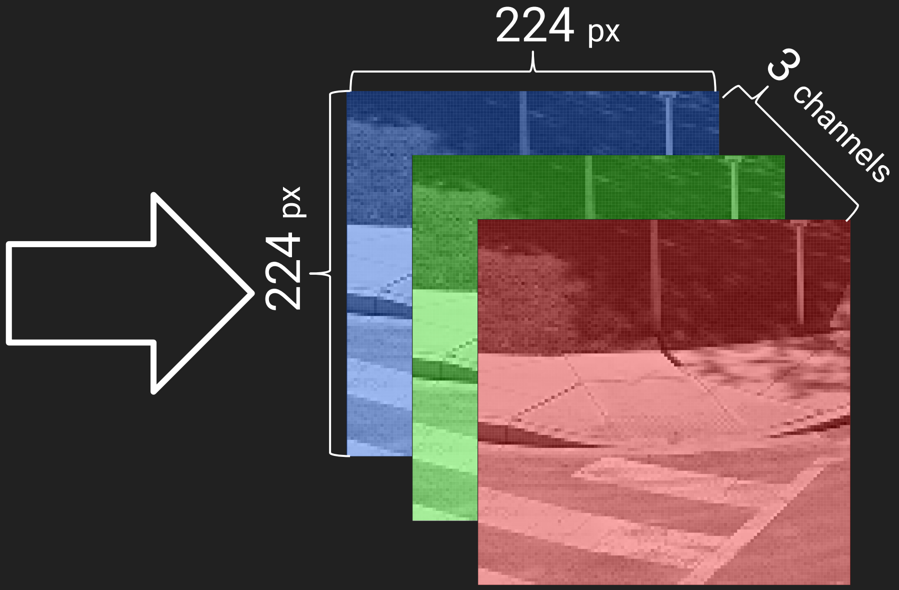
3 types of features

image features

positional features

geographic features





3 types of features

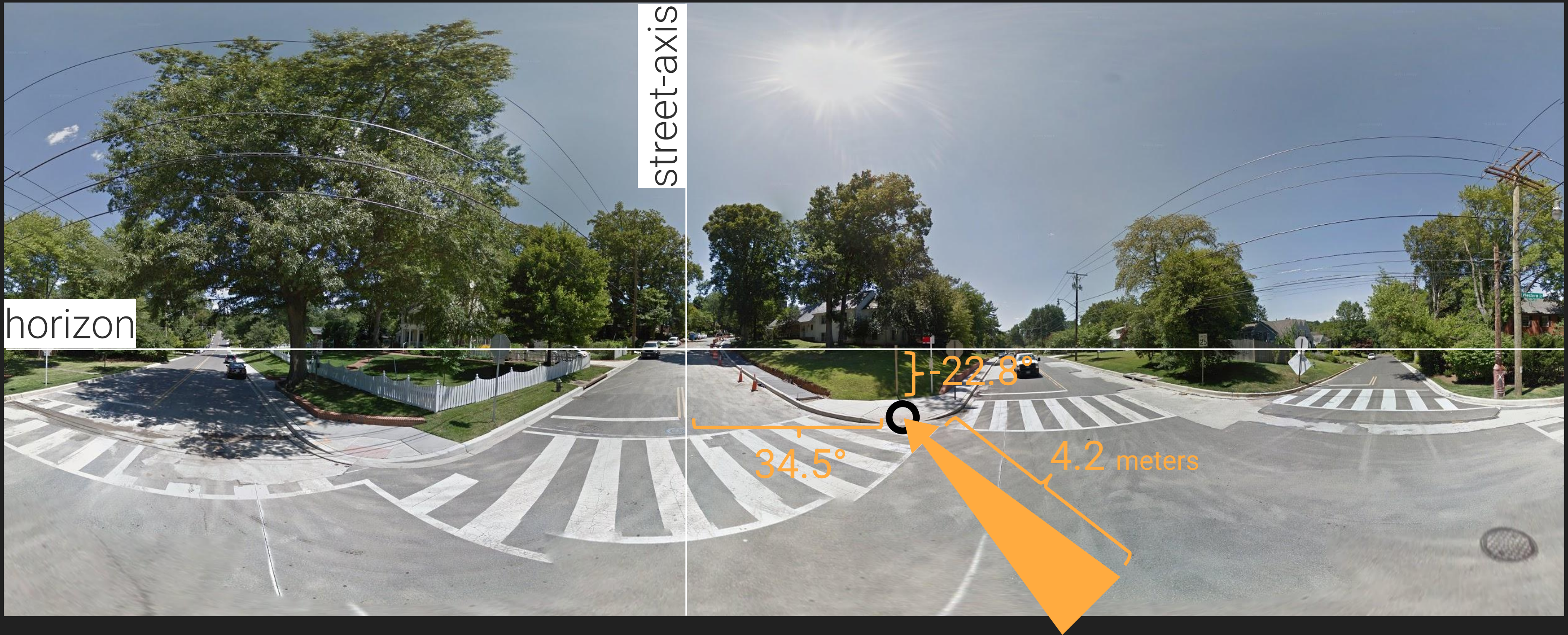
image features   positional features   geographic features

3 types of features

image features

positional features

geographic features

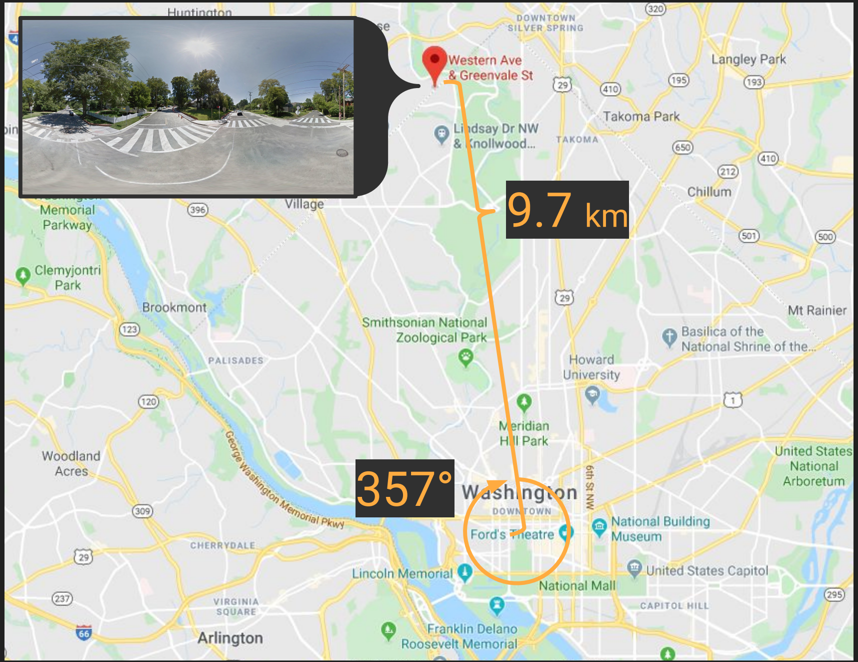


3 types of features

image features

positional features

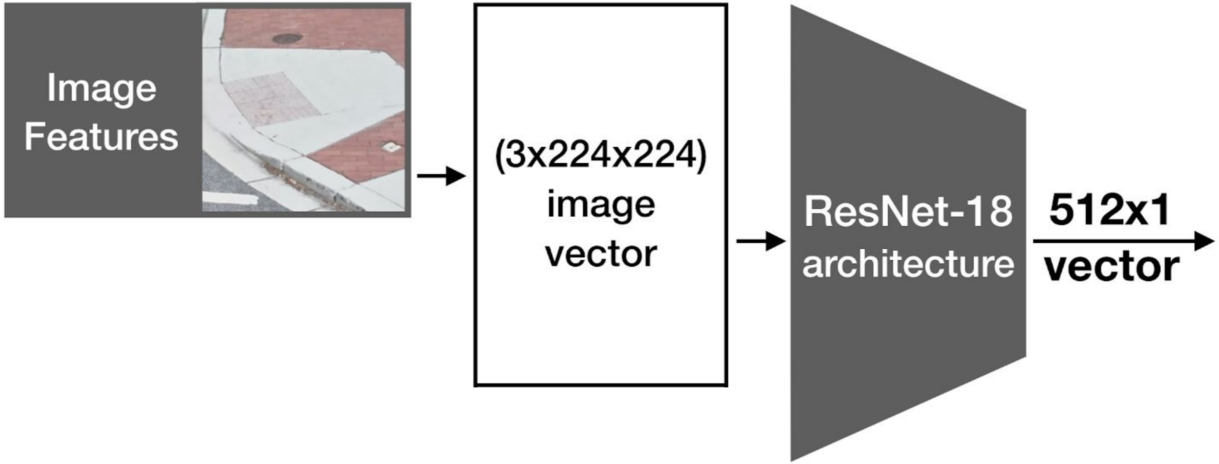
geographic features



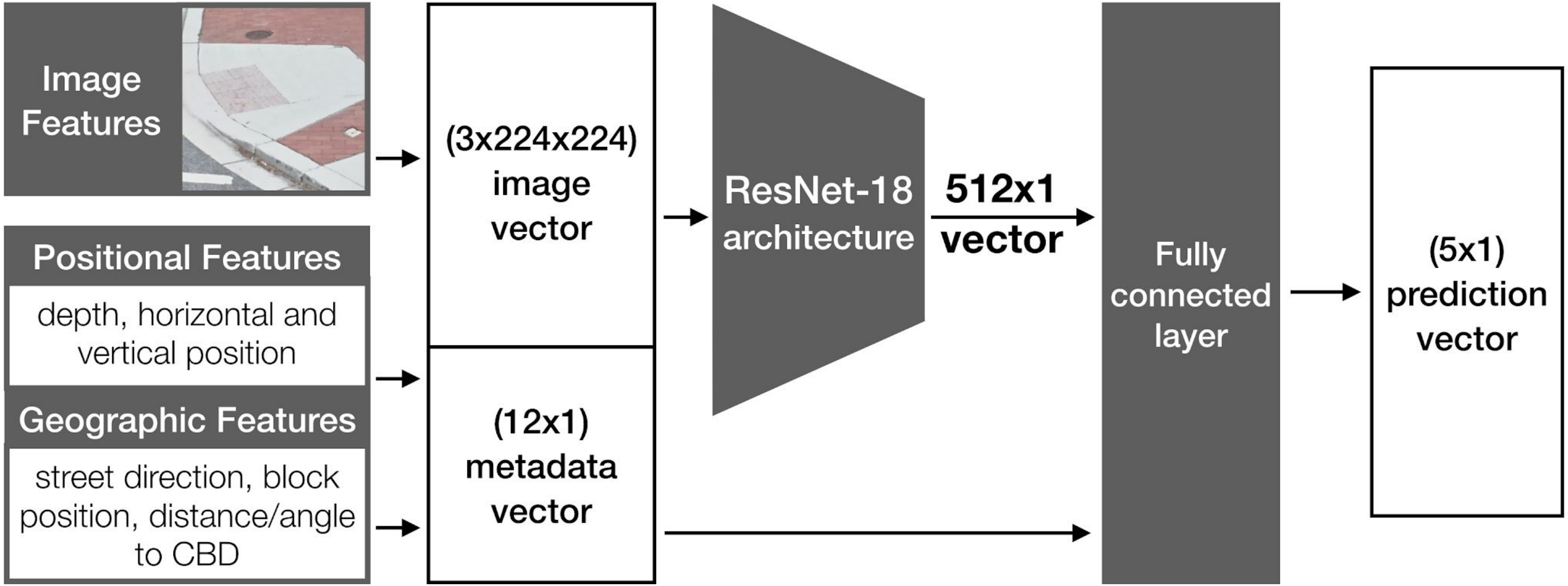
How do we automate these tasks?

1. Start with 181k labeled problems from Project Sidewalk dataset.
2. Compute 3 types of features for each human-placed label.
3. Train two different neural networks, one for validation, one for labeling.
4. Use a sliding window to label panoramas.
5. Evaluate on a researcher-created ground-truth test dataset.

Neural Network Architecture



Neural Network Architecture



How do we automate these tasks?

1. Start with 181k labeled problems from Project Sidewalk dataset.
2. Compute 3 types of features for each human-placed label.
3. Train two different neural networks, one for validation, one for labeling.
4. Use a sliding window to label panoramas.
5. Evaluate on a researcher-created ground-truth test dataset.

Sliding Window

Sliding Window

Sliding Window

Sliding Window

4

curb ramp

Sliding Window

4

curb ramp

Sliding Window

4

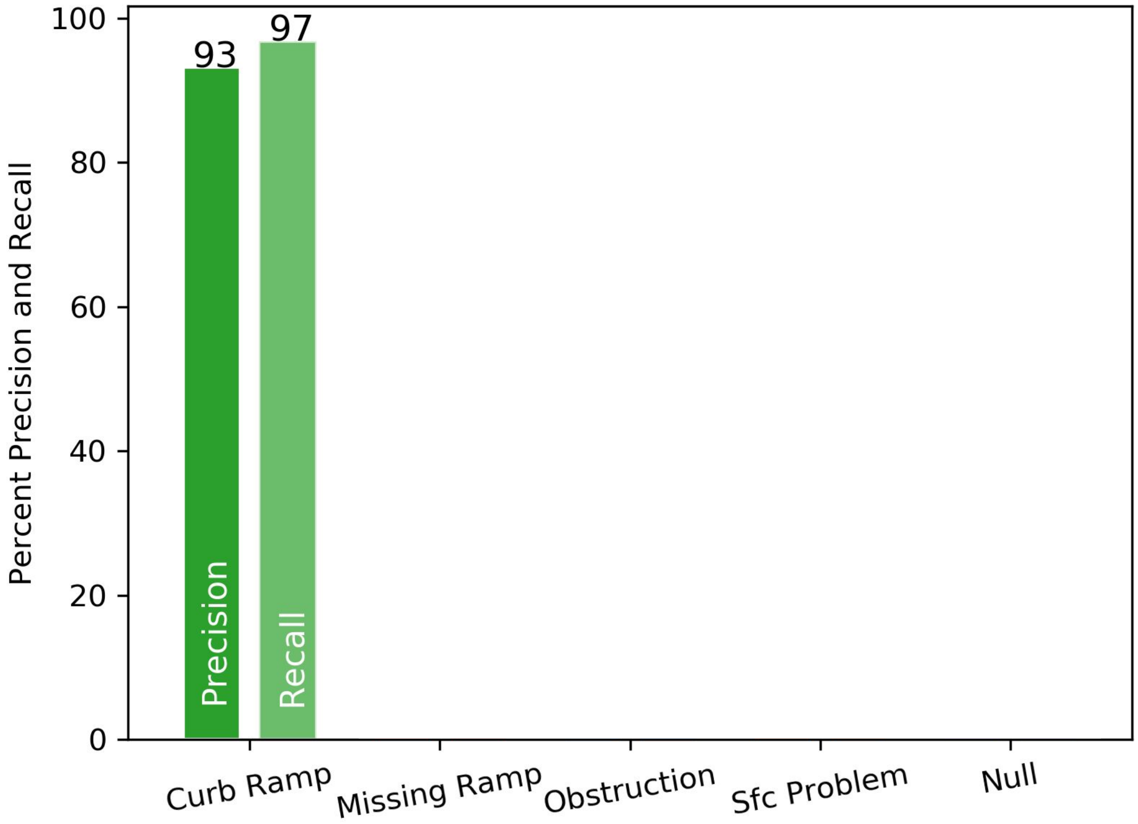
curb ramp


Sliding Window
4
curb ramp
3
curb ramp

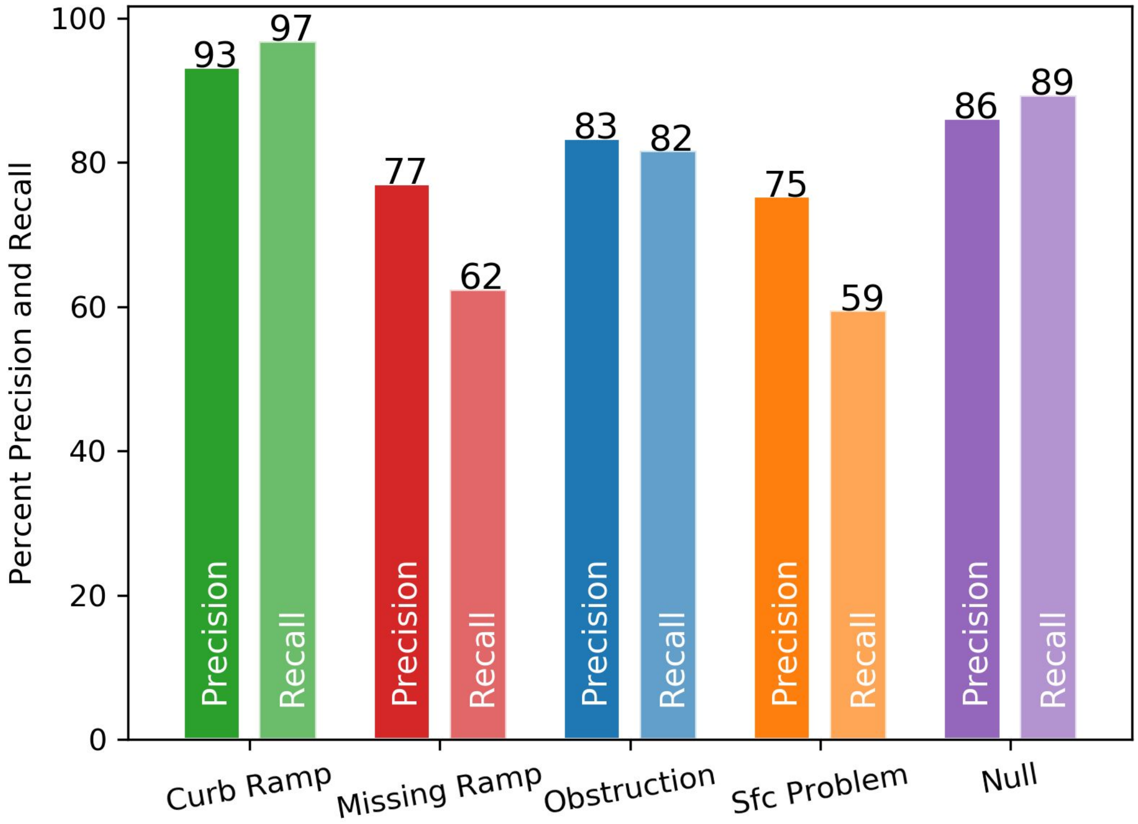
How do we automate these tasks?

1. Start with 181k labeled problems from Project Sidewalk dataset.
2. Compute 3 types of features for each human-placed label.
3. Train two different neural networks, one for validation, one for labeling.
4. Use a sliding window to label panoramas.
5. Evaluate on a researcher-created ground-truth test dataset.

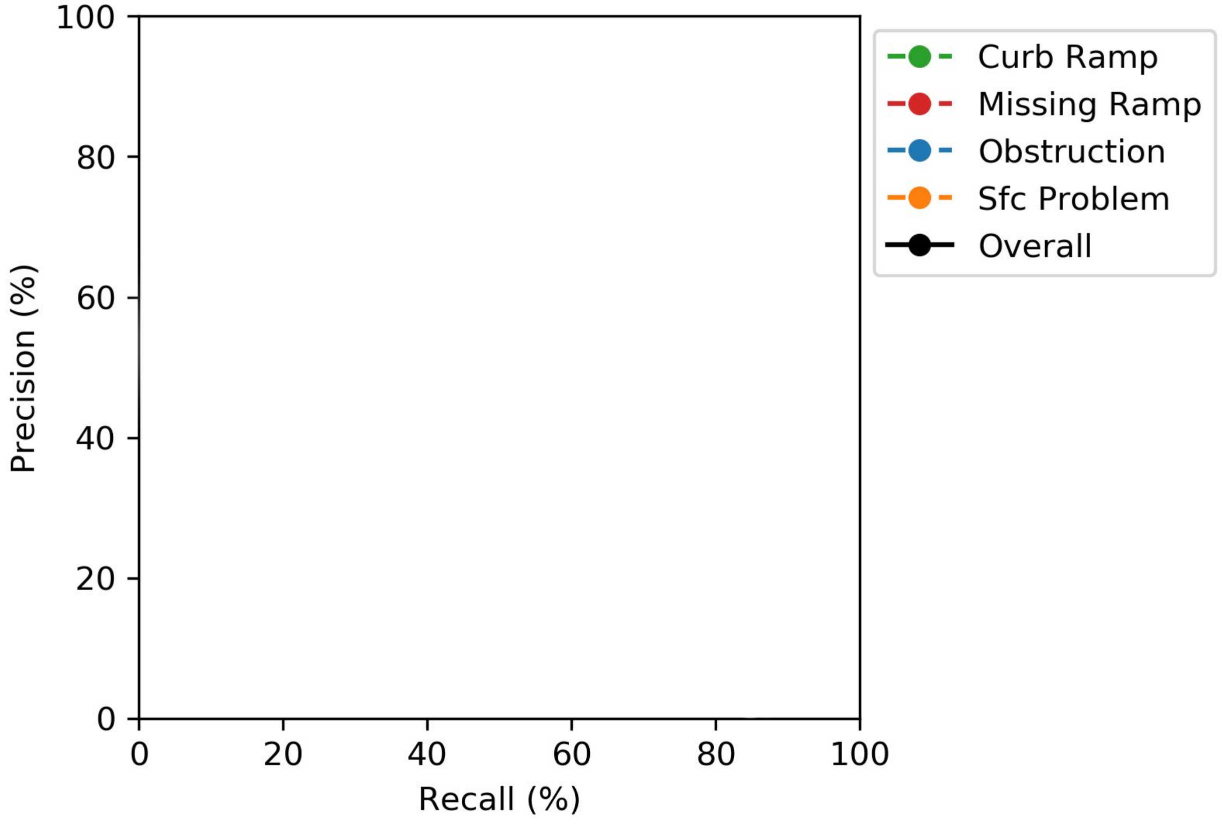
Validation Performance



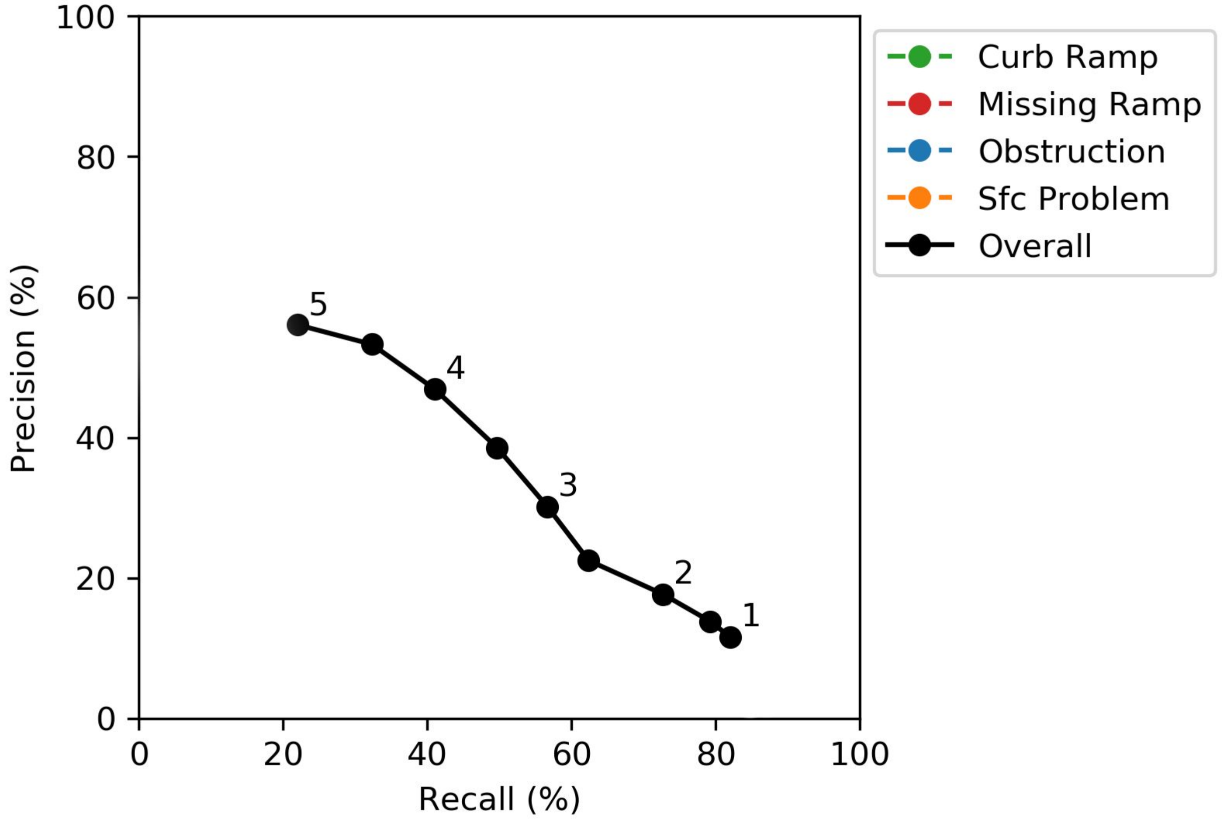
Validation Performance



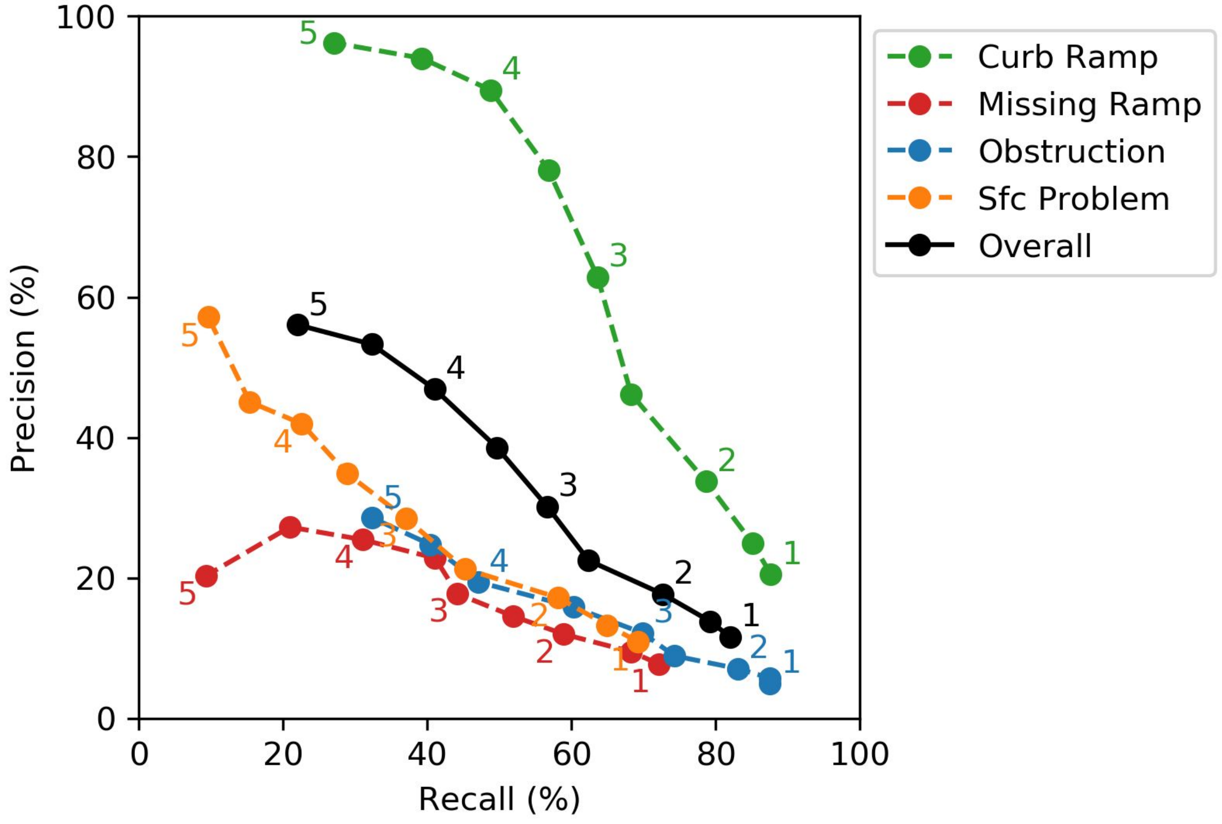
Labeling Performance



Labeling Performance

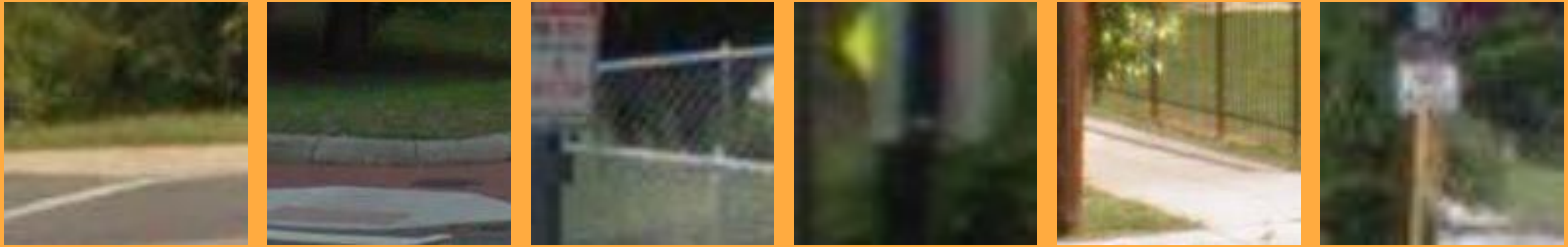


Labeling Performance



Validation Errors - Common Factors

low resolution imagery



difficult lighting

Validation Errors - False Positives

curb ramp

27%

crosswalk

Validation Errors - False Positives

curb ramp

27%

crosswalk

missing curb ramp

86%

curb

Validation Errors - False Positives

curb ramp

27%

crosswalk

missing curb ramp

86%

curb

obstruction

58%

not on path

Comparison with Automated Systems

Fully Automated Systems

		Tohme [1]	Our Model	Change
Curb Ramp	*precision*	26%	33.7%	+30%
	recall	67%	78.7%	+17%

		Sun *et al.* [2]	Our Model	Change
Missing Ramp	*precision*	not reported	12.0%	N/A
	recall	27%	58.6%	+117%

[1] Kotaro Hara, Jin Sun, Robert Moore, David Jacobs, and Jon Froehlich. 2014. *Tohme*. In Proceedings of the 27th annual ACM Symposium on User interface software and technology - UIST '14.

[2] Jin Sun and David W. Jacobs. 2017. *Seeing What is Not There: Learning Context to Determine Where Objects are Missing*. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 1234–1242.

Comparison with Human Systems

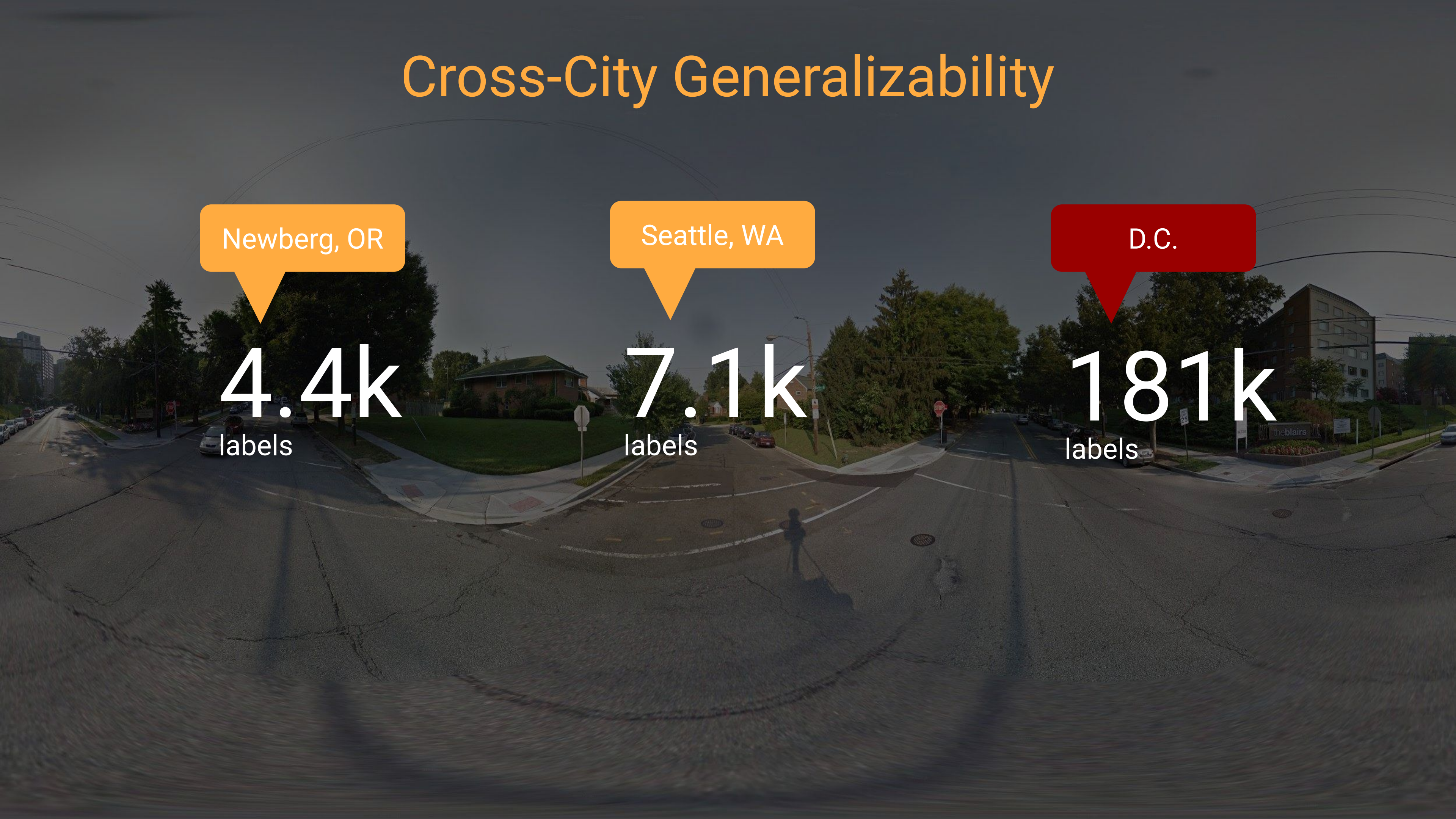
Majority Vote of
5 Crowdworkers

		Hara *et al.* [3]	Our Model	Change
Overall	*precision*	37%	39%	+5%
	recall	46%	50%	+9%

[3] Kotaro Hara, Vicki Le, and Jon Froehlich. 2013. *Combining crowdsourcing and google street view to identify street-level accessibility problems*. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13.

Cross-City Generalizability

Cross-City Generalizability



Cross-City Generalizability

baseline	D.C. model
three experiments	D.C. + new city
	new city only
	new city, *pretrained* with D.C.

Cross-City Generalizability

baseline	D.C. model
three experiments	D.C. + new city
	new city only
	new city, *pretrained* with D.C.

Cross-City Generalizability

Newberg, OR

90.2%

overall recall

Seattle, WA

82.8%

overall recall

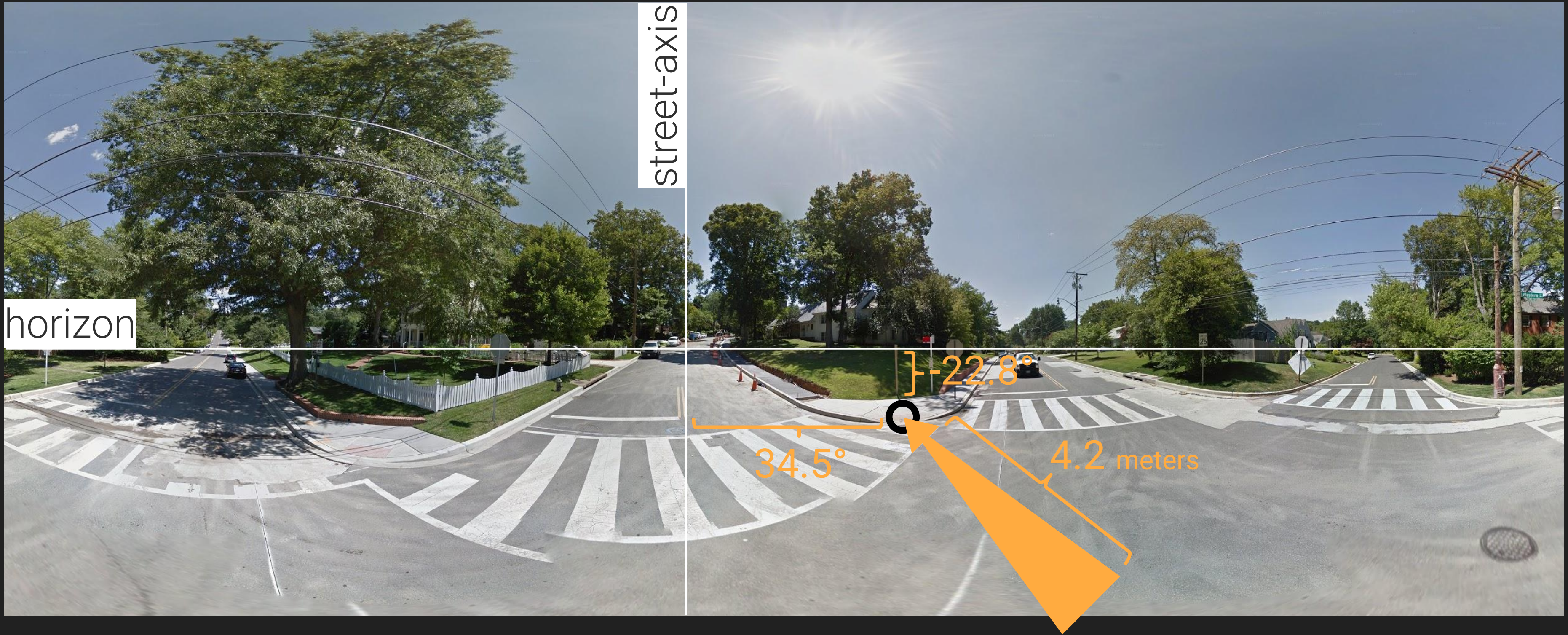
D.C.

89.9%

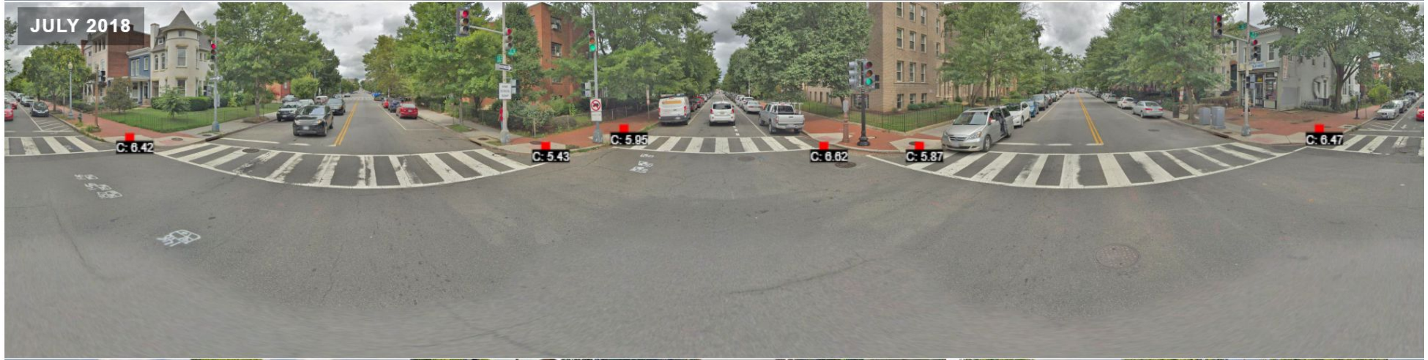
overall recall


what’s next?

richer contextual features



What new work can we enable?




our vision is to
map the accessibility of
all sidewalks in the world

acknowledgements

NSF

ALFRED P. SLOAN FOUNDATION
1934

Alfred P. Sloan
FOUNDATION

acknowledgements

Alfred P. Sloan FOUNDATION

and thanks...

Esther Jang, Anthony Li, Aileen Zeng, Kurtis Heimerl, and Jon Froehlich

Thank You. Questions?

Possibilities to include (that I haven’t already)

How do we generate null-crops?

Differences between the sliding-window training set and the centered-crop training set.

Validation Errors - False Negatives

curb ramp

41%

bad delineation

Validation Errors - False Negatives

curb ramp

41%

bad delineation

missing curb ramp

30%

no crosswalk

Validation Errors - False Negatives

curb ramp

41%

bad delineation

missing curb ramp

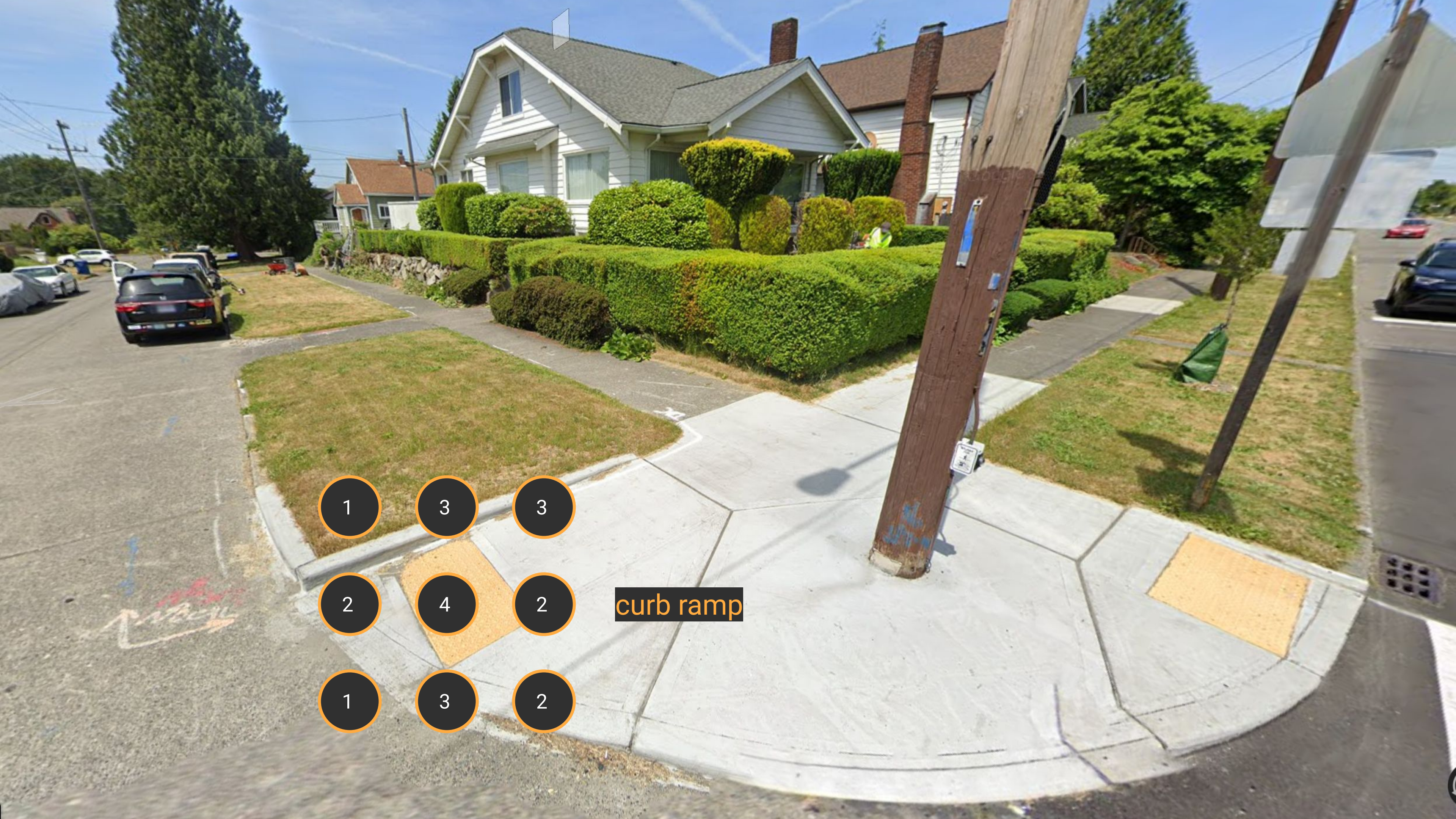
30%

no crosswalk

surface problem

48%

grass


1
3
3
2
4
2
curb ramp
1
3
2


4
curb ramp

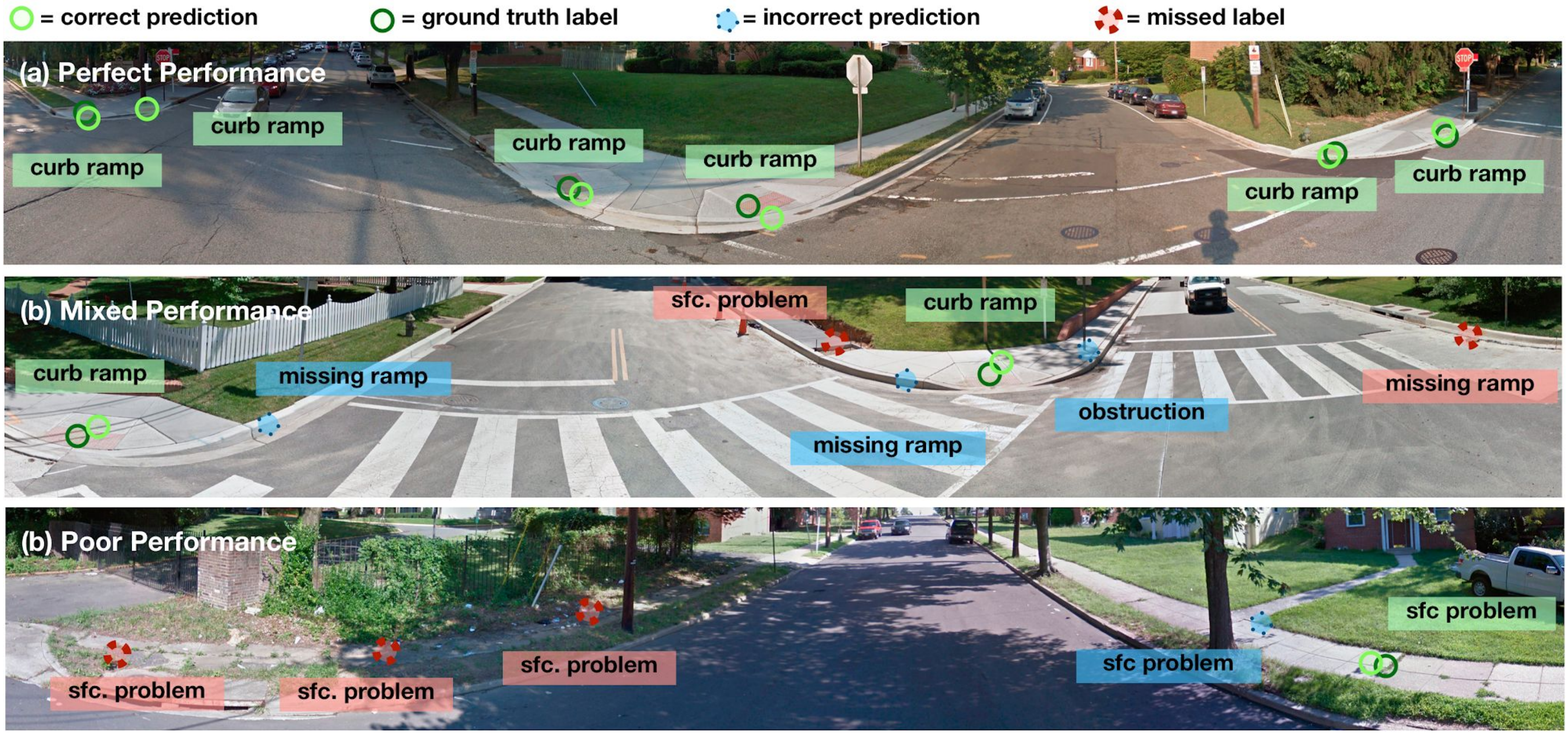

1
curb ramp
4
curb ramp
3
curb ramp


4
curb ramp
3
curb ramp

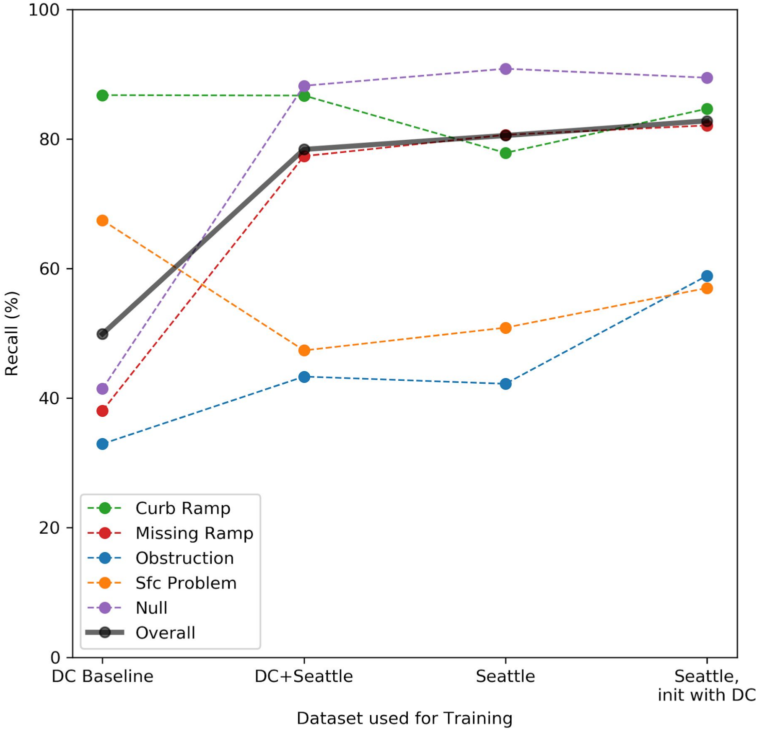
Effect of Extra Input Features

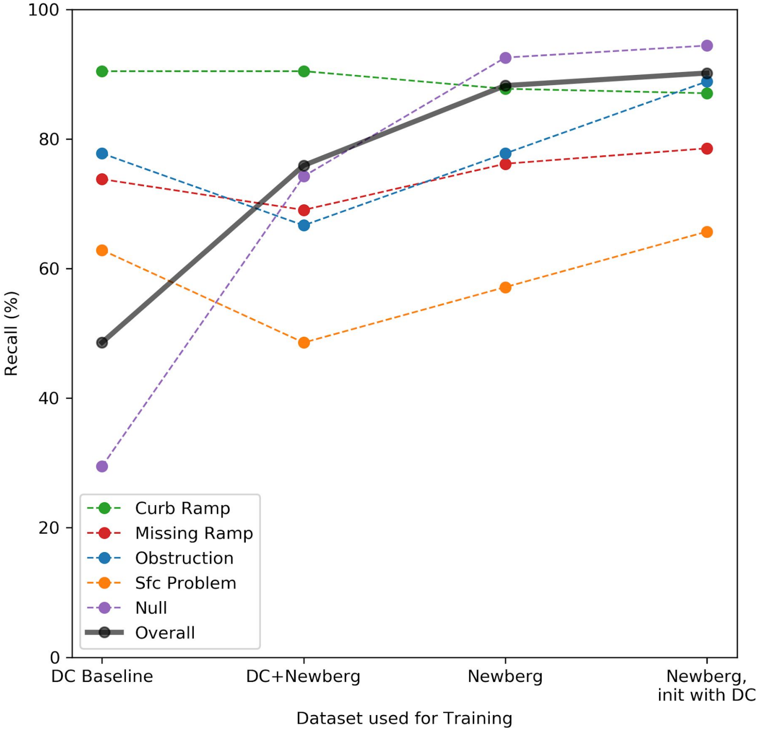
	Precision			Recall		
	Image	Img. + Position	All	Image	Img. + Position	All
Overall	80.3	79.5	79.7	79.6	80.0	80.1
Curb Ramp	81.5	80.1	79.7	90.7	93.2	93.6
Missing Ramp	80.2		80.6	50.7		51.8
Obstruction	84.9	84.9	85.4	73.0	71.9	69.8
Sfc Problem	79.3		73.5	48.5	50.8	56.7
Null	75.6		79.3	89.4		

Labeling Performance



Cross-City Generalizability





3 types of features

image features - describe *appearance* of object

geographic features

positional features

3 types of features

image features - describe *appearance* of object

geographic features - where is the object *within a panorama?*

positional features

3 types of features

image features - describe *appearance* of object

geographic features - where is the object *within a panorama?*

positional features - where is the panorama *within the city?*

Cross-City Generalizability