A Machine Learning Approach to Automated Road-Surface Condition Predictions in Collaboration with the New York State Department of Transportation

Presented by: Carly Sutter
Dr. Kara Sulia, Dr. Nick Bassill, Dr. Christopher Wirz, Dr. Christopher Thorncroft

HMT Seminar 02/18/2025

Outline

<u>Part 1:</u>

- Data
- ML algorithms
- Model design
- Model performance

<u>Part 2:</u>

- Application and end-user perspective (NYSDOT)

All camera images in this presentation are from 511ny.org

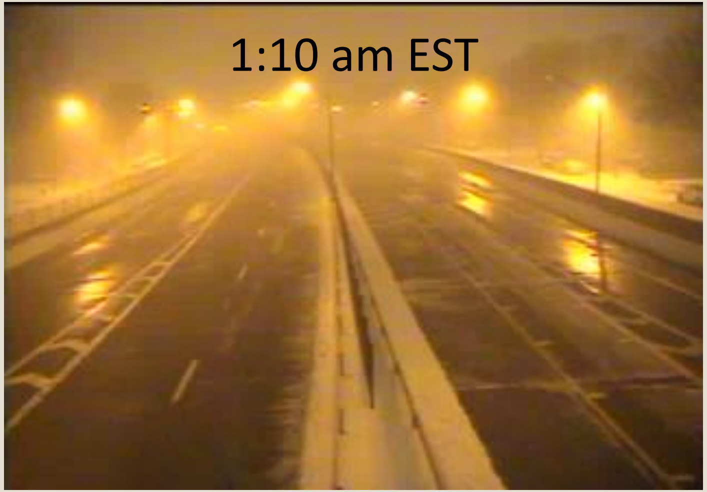
Traffic, Travel and Transit Info
www.511ny.org

Introduction

High impact weather affects NYSDOT resource allocation, maintenance of road conditions, and traveler safety.

Goal: automatically detect weather-related road surface conditions using camera images and weather data

Example from Buffalo NY on November 19, 2022, *5 minutes apart*



Model overview

Objective: road surface condition (“RSC”) classification

Image data *(6 road surface conditions)*

Severe snow (n=1100)

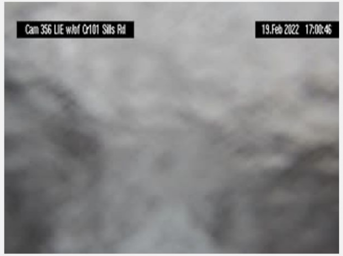
Snow (n=2000)

Wet (n=6800)

Dry (n=6200)

Poor Visibility (n=600)

Obstructed (n=1000)



Hand-labeled dataset
- 6 classes
- 30 locations
- 17.7k observations

For **model training**, the **ground truth labels** for a given observation (time, lat, and lon) is the **hand-labeled classification**

Image source: 511ny.org
Archive: stored at UAlbany’s xCITE lab

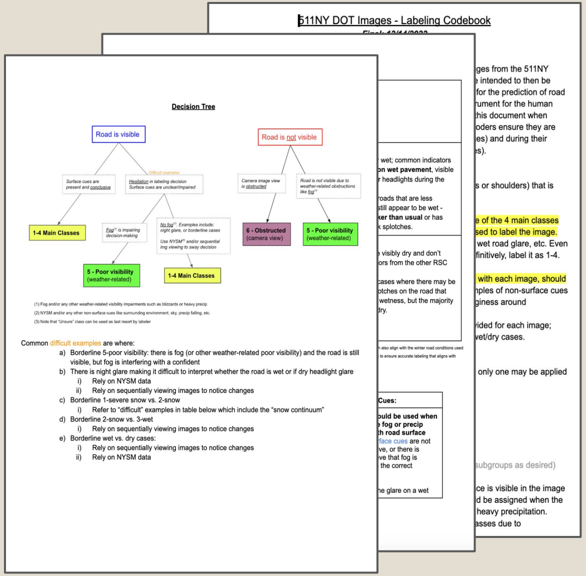
Hand-labeled dataset

Published paper on ***Quantitative Content Analysis (QCA)*** using this project as case study:

Wirz, C. D., Sutter, C., Demuth, J. L., Mayer, K. J., Chapman, W. E., Cains, M. G., et al. (2024). Increasing the reproducibility and replicability of supervised AI/ML in the Earth systems science by leveraging social science methods. *Earth and Space Science*, 11, e2023EA003364. https://doi.org/10.1029/2023EA003364

Labeling **"codebook"** following QCA framework

- Carefully **define & set rules** for labeling ⇒ **consistency** and **trust** in the data going into the ML model
- Codebook & reliability trial data and results published on Zenodo: DOI: 10.5281/zenodo.8370665



Model data (input data)

Data used for model training: both image data <u>and</u> forecast data

- Image data

- HRRR forecast data: High-Resolution Rapid Refresh, valid at time of image, collocated for each image observation

2m air temp	2m relative humidity	10m average wind	accumulated snow	total precipitation	total cloud cover

Model data summary

	Model ***input***		Model ***output***
Observation (lat, lon, time)	**Image Data**	**HRRR Data**	**Road surface condition** Hand-labeled
#1		T = 271.8K RH = 93% Accumulated snow = 2”	Snow
…	…	…	…
# 17,717		T = 273.6K RH = 86% Accumulated snow = 0”	Wet

Machine learning algorithms

Convolutional neural network (CNN)

Convolutional neural network (CNN) deep learning algorithms for tasks that use ***visual information***, or any data where ***position matters*** (including spatial and/or temporal). These are commonly used for ***image recognition tasks***.

Support vector machine (SVM)

An SVM classifies data by finding the optimal ***decision boundary*** that separates the data into groups and maximizes the margin between the groups.

Random forest (RF)

A ***combination of decision tree*** classifiers that are commonly used ML algorithms that capture nonlinear relationships

Model selection:
Multiple ML algorithms
Hyperparameter tuning

Model selection

Consider end-user priorities ⇒ ***consider 8 metrics*** (not just one metric, e.g. accuracy)

High importance:

1. overall accuracy
2. severe snow

High-medium importance:

3. snow

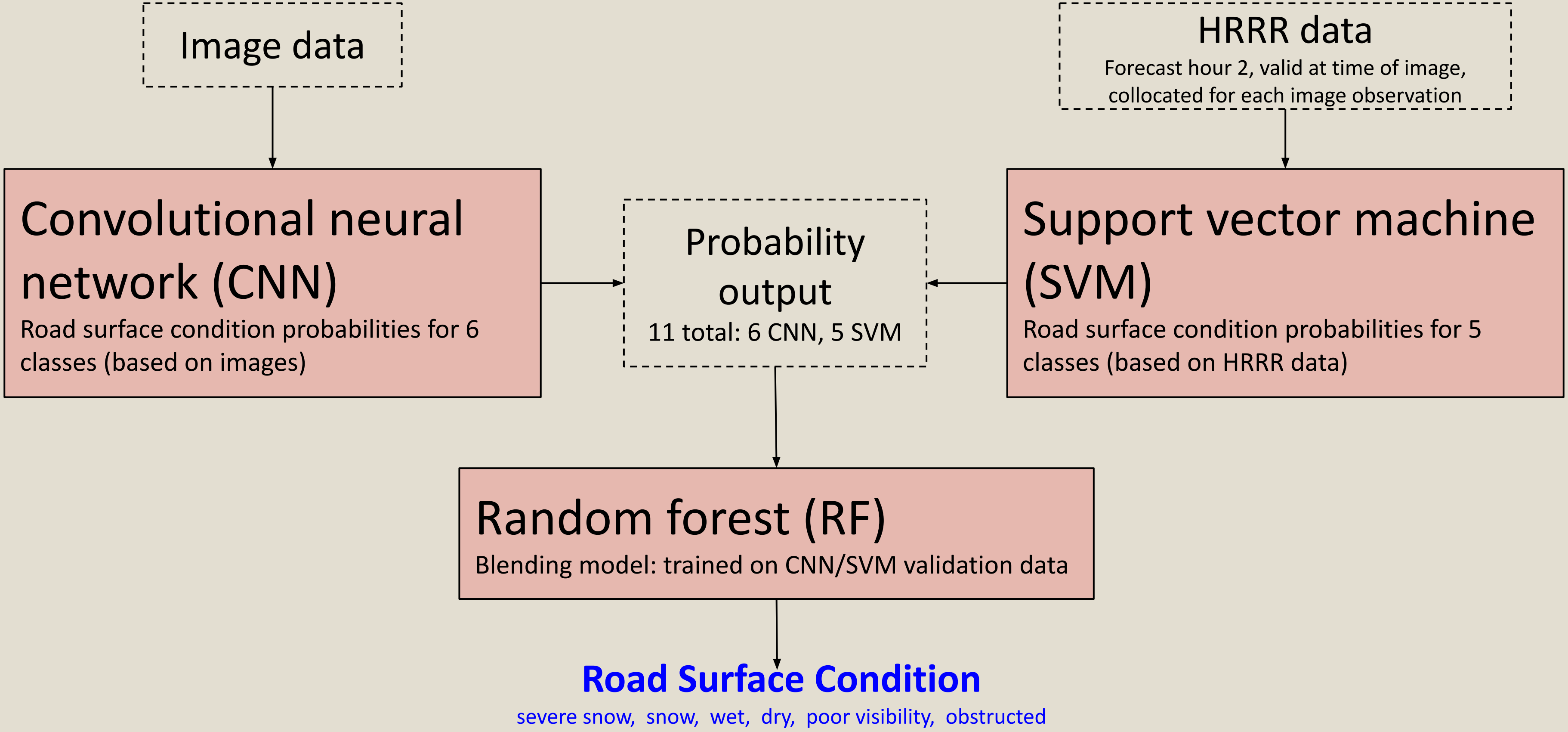
Medium importance:

4. adjusted accuracy
5. wet
6. dry
7. poor visibility

Low importance:

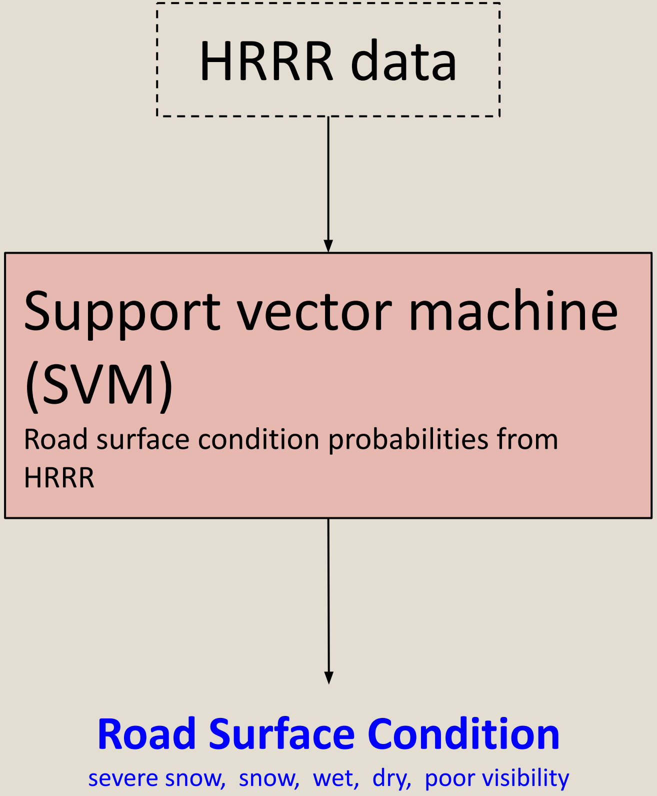
8. Obstructed

Model process - **nowcast** road surface conditions



Model process - <u>**forecast**</u> road surface conditions

- Only use HRRR data (no image data for future times)
- Valid at time of labeled observations



Model process - **forecast** road surface conditions

- Only use HRRR data (no image data for future times)
- Valid at time of labeled observations
- Forecast hours
  Short: 2, 3, 4, 5, 6
  Medium: 9, 12, 15
  Long: 18, 24, 30, 36, 42, 48

Highlighted is the forecast hour that the selected model was trained on for that forecast group (short/medium/long)

HRRR data

↓

Support vector machine (SVM)

Road surface condition probabilities from HRRR

↓

Road Surface Condition

severe snow, snow, wet, dry, poor visibility

Data splitting

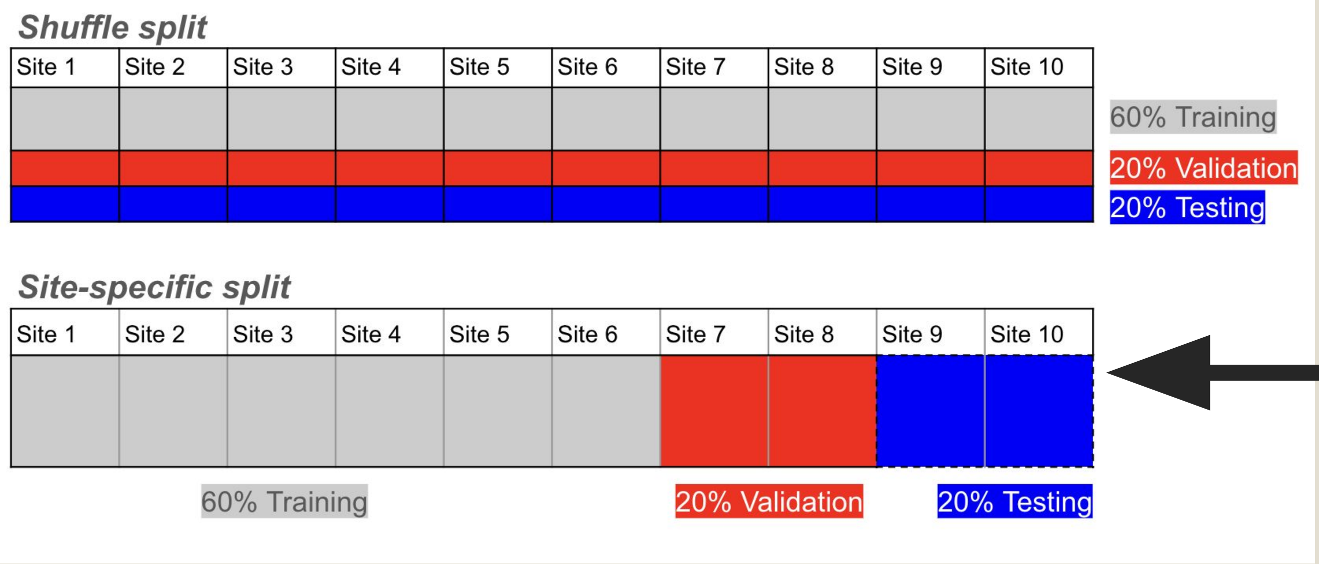
Using 5-fold cross validation

- Every observation is represented in validation and test datasets
- Training data is the data that the model sees and learns from
- Validation is used for model tuning/selecting the best models
- Testing is used for assessing performance

Data splitting

Emphasis on generalizability

Evaluation on *unseen* camera sites → operational application



Use this method!

Results - Nowcast

Road surface condition nowcasting skill:

Testing accuracy: 76.1%

Note: shuffle split (each site represented in train and test) is 88.9%.

Adjusted accuracy: 82.8%

Adjacent classes counted half correct:

Dry ↔ Wet ↔ Snow ↔ Severe snow

Results - Nowcast

Road surface condition nowcasting skill:

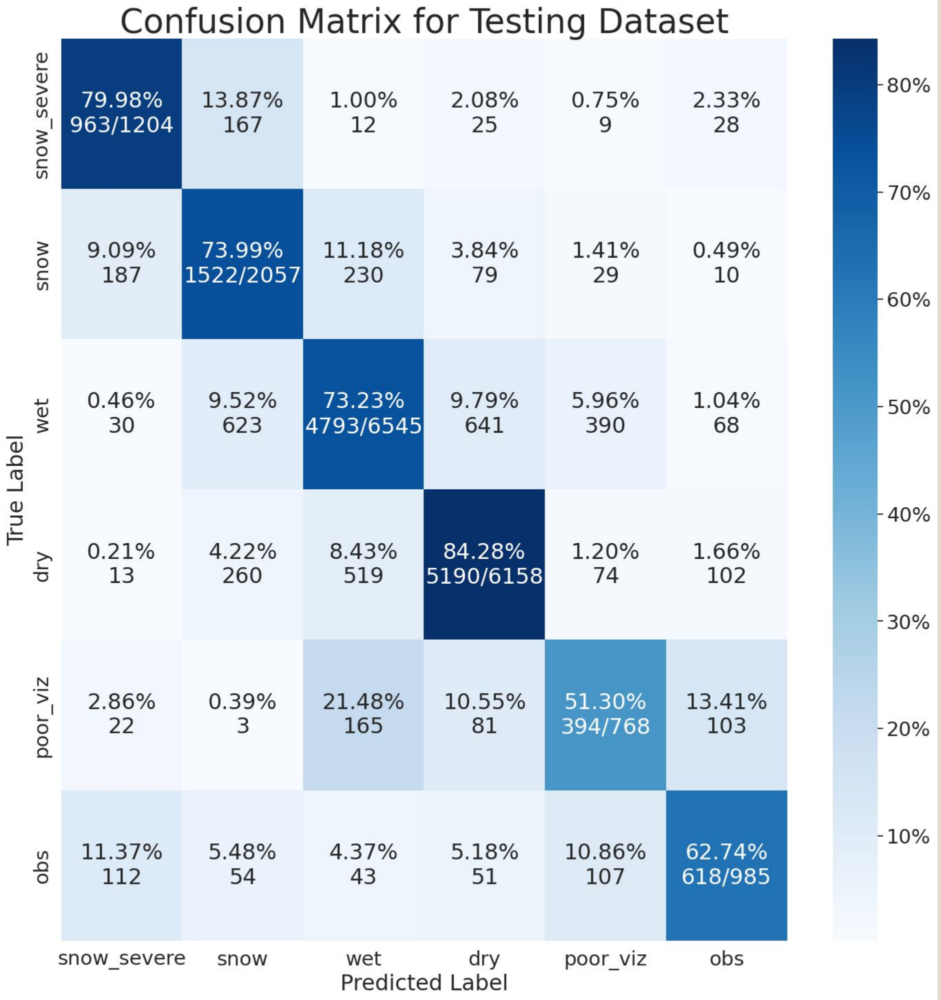
Testing accuracy: 76.1%

Note: shuffle split (each site represented in train and test) is 88.9%.

Adjusted accuracy: 82.8%

Adjacent classes counted half correct:

Dry ↔ Wet ↔ Snow ↔ Severe snow



Metric shown is recall (out of total labeled in that class). Calculation is recall = probability of detection (POD) = True Positive / (True Positive + False Negative)

Compared to CNN, adding weather data for the nowcast adds…

Severe snow: +5%

Snow: +15%

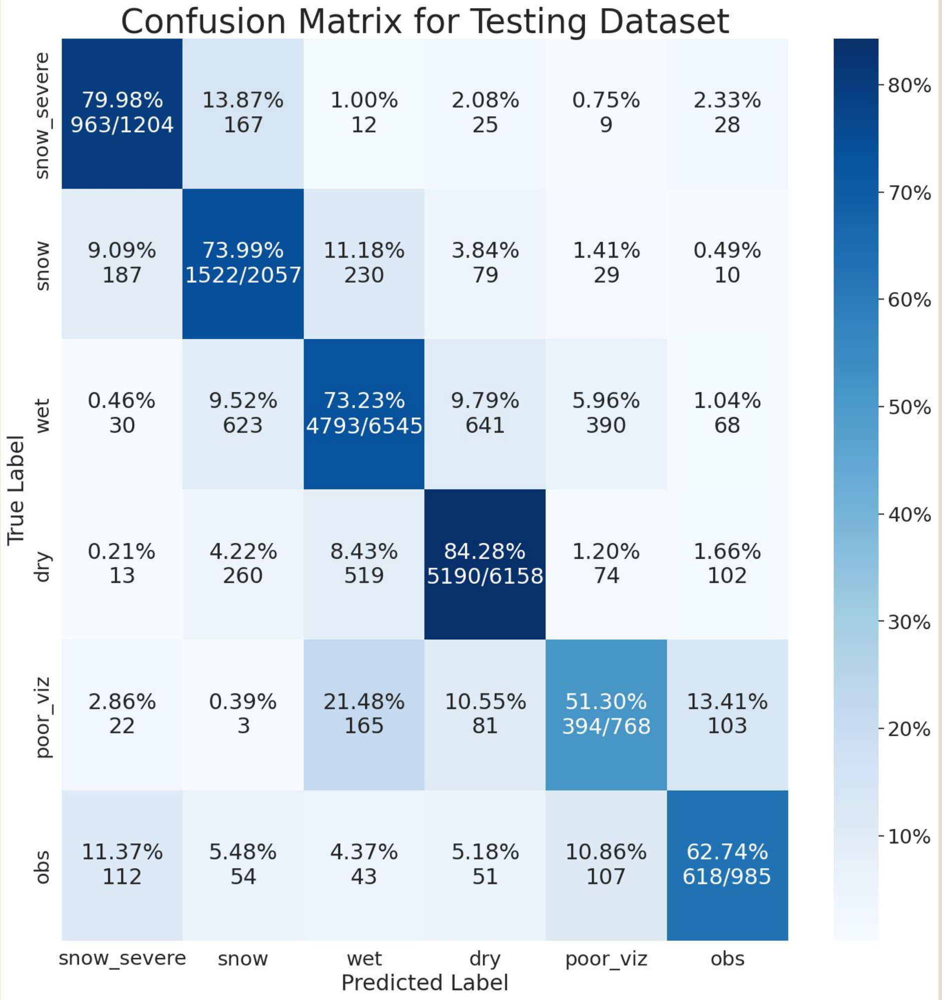
Wet: +8%

Dry: +18%

Poor visibility: +10%

N/A

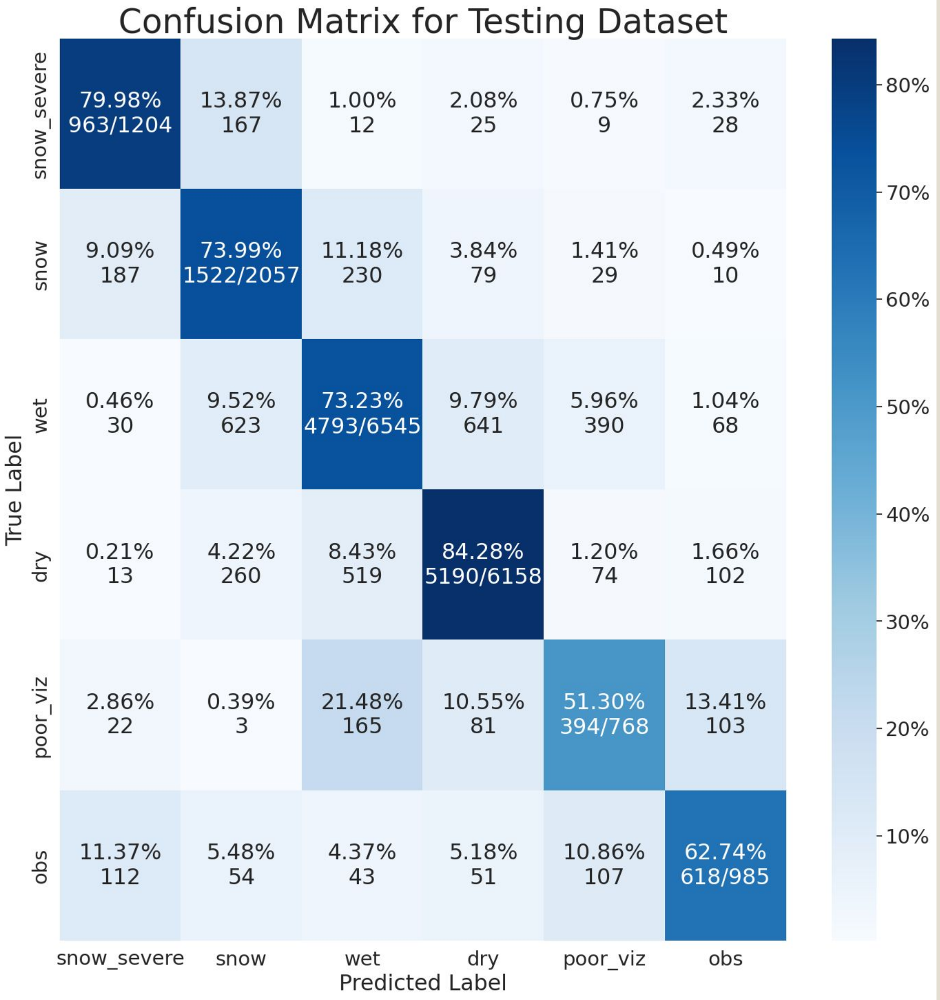
Metric shown is recall (out of total labeled in that class). Calculation is recall = probability of detection (POD) = True Positive / (True Positive + False Negative)



Adding weather data fixes some camera predictions

Example:

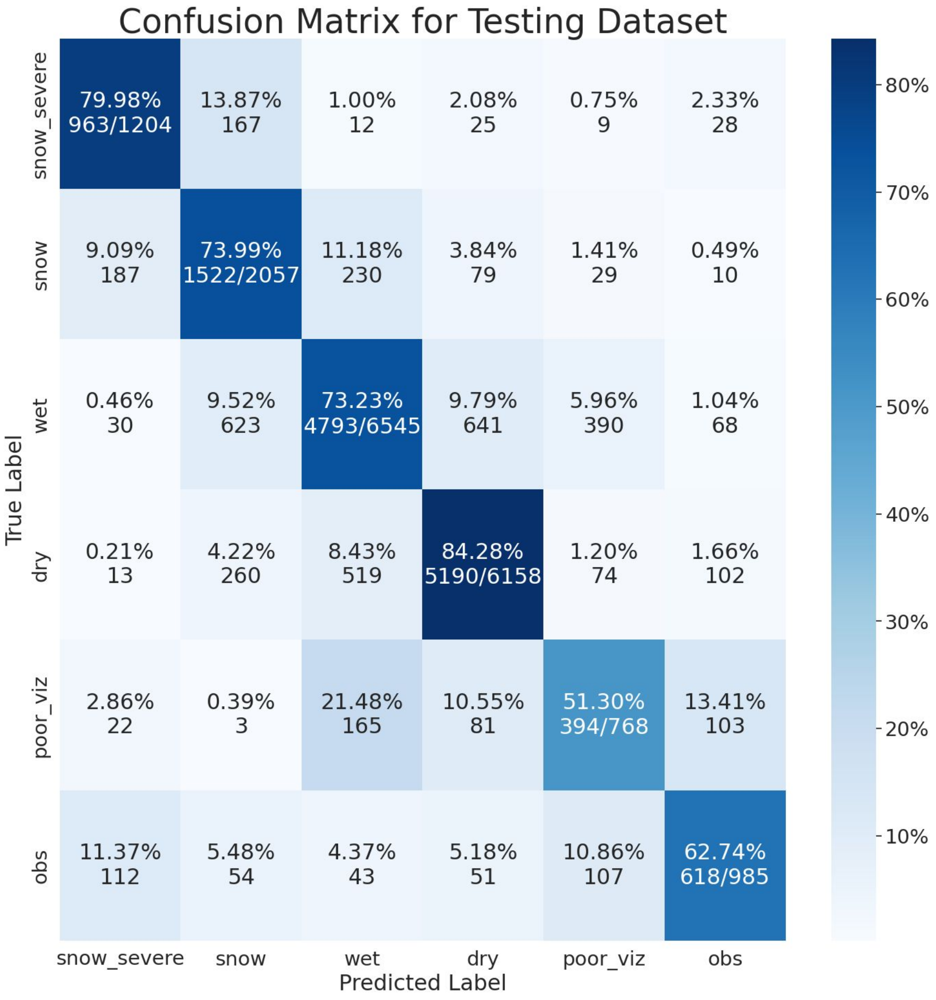
- Ground truth hand-label: Snow
- CNN-predicted: Severe snow
- ***Final model: Snow***



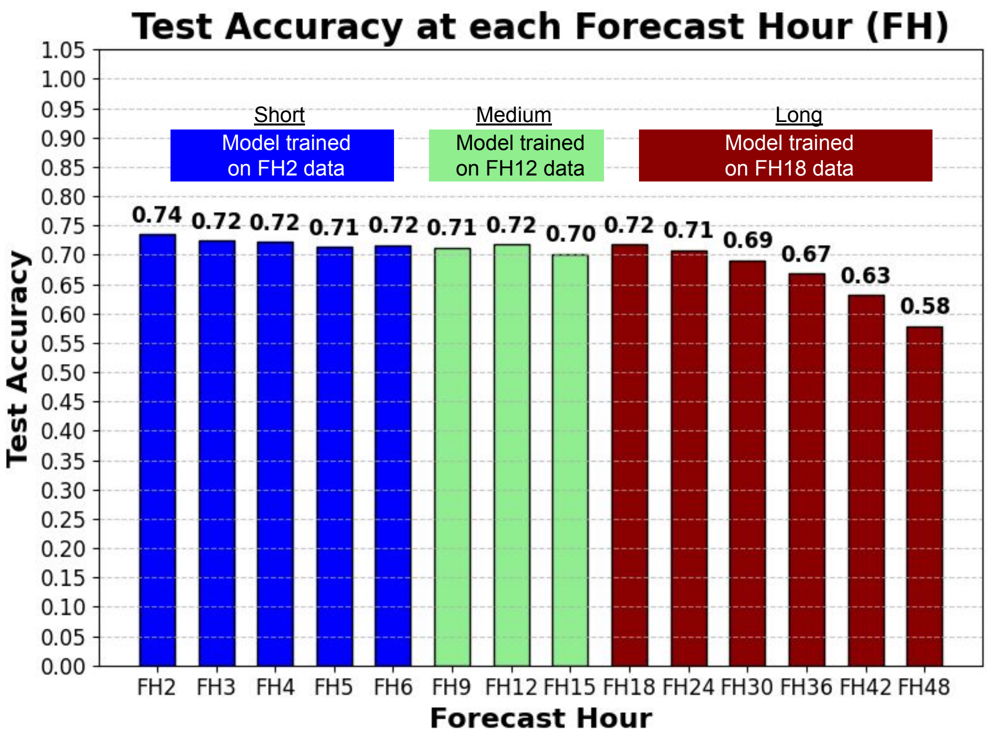
But some are still “borderline” cases that are still difficult to predict correctly

Example:

- Ground truth hand-label: Snow
- CNN-predicted: Wet
- ***Final model: Wet (wrong)***



Results - Forecast

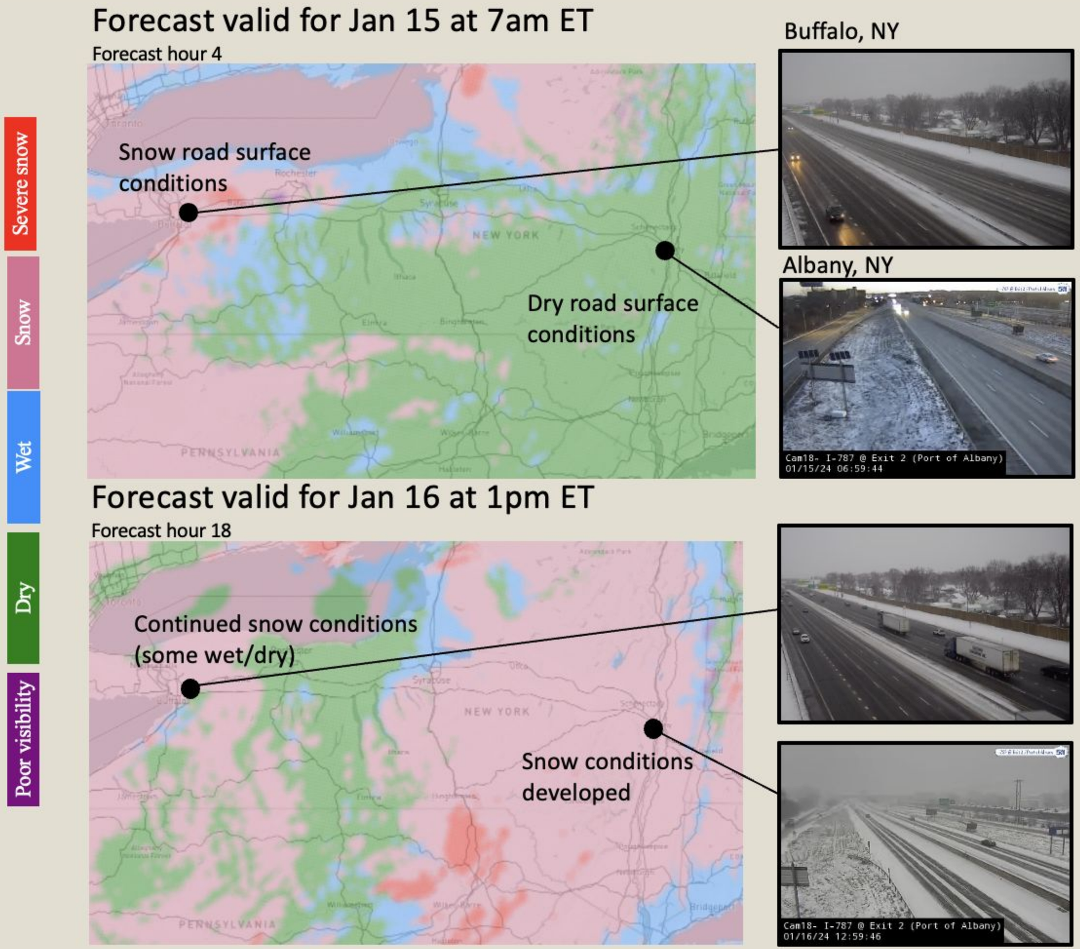


Recall by Class & Forecast Hour (FH)

Forecast hour group	Road surface condition class				
	Severe Snow	Snow	Wet	Dry	Poor Visibility
Short (FH 2-6 average)	77.9%	71.2%	61.8%	84.0%	61.4%
Medium (FH 9-15 average)	80.3%	65.4%	58.8%	85.6%	60.5%
*FH 18-48 average**	*72.8%*	*67.0%*	*56.8%*	*82.2%*	*62.6%*

** Limited labeled data available for FH24-48*

Forecast examples: New York State Jan 15-16, 2024

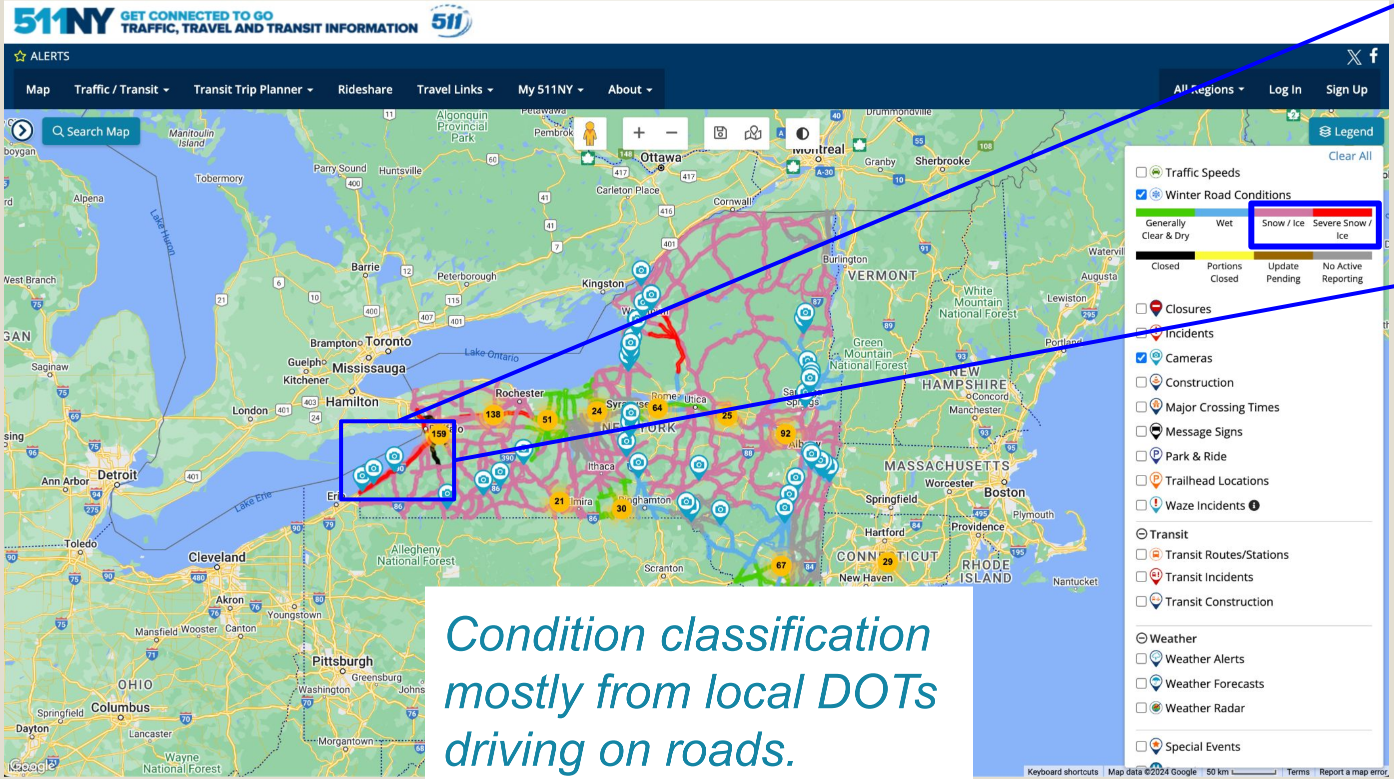


Part 2

Application and end-user perspective

NEW YORK STATE OF OPPORTUNITY | Department of Transportation

Current tool: NYSDOT - 511ny.org



Condition classification mostly from local DOTs driving on roads.

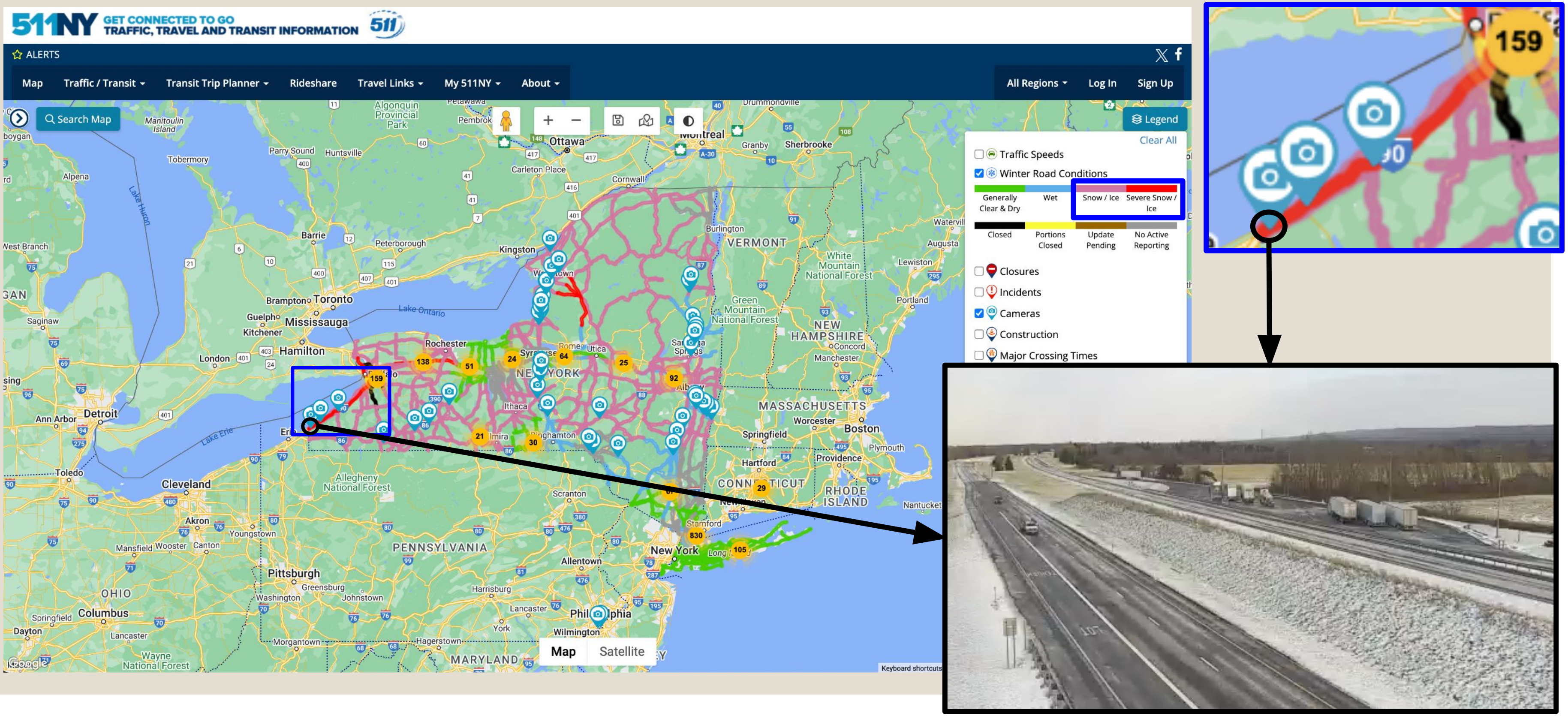
Severe Snow

Snow

January 4, 2024
3:20pm

NEW YORK STATE OF OPPORTUNITY | Department of Transportation

Current tool: NYSDOT - 511ny.org



Addressing societal needs

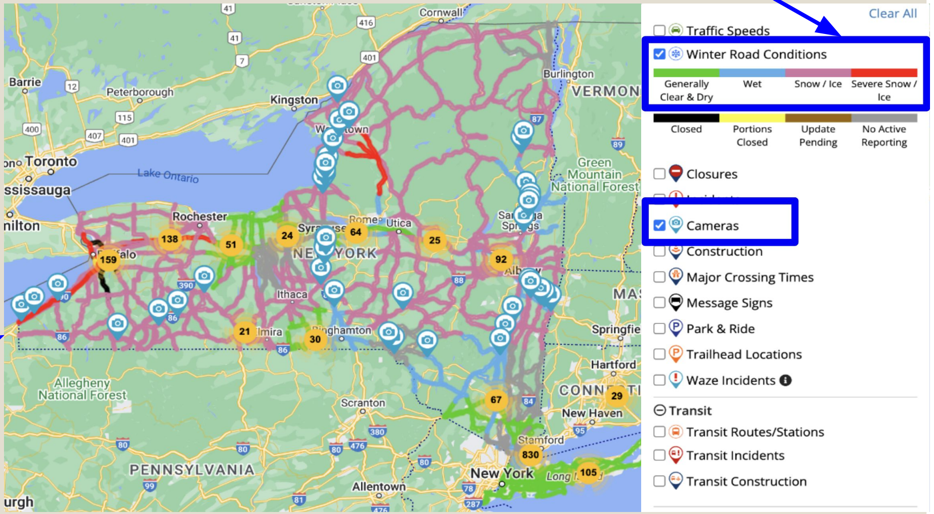
NYSDOT operational interests

Benefits/opportunity for road surface condition model:

- ***Frequent*** updates of road surface conditions
- Covering ***large geographic area (NYS)***
- Specific ***camera-level predictions*** (local/granular)
- Provides ***another perspective*** of road surface conditions (current methods <u>and</u> model)
- Save time for NYSDOT employees!
    - Manually logging conditions
    - Public tool – every 4 hours

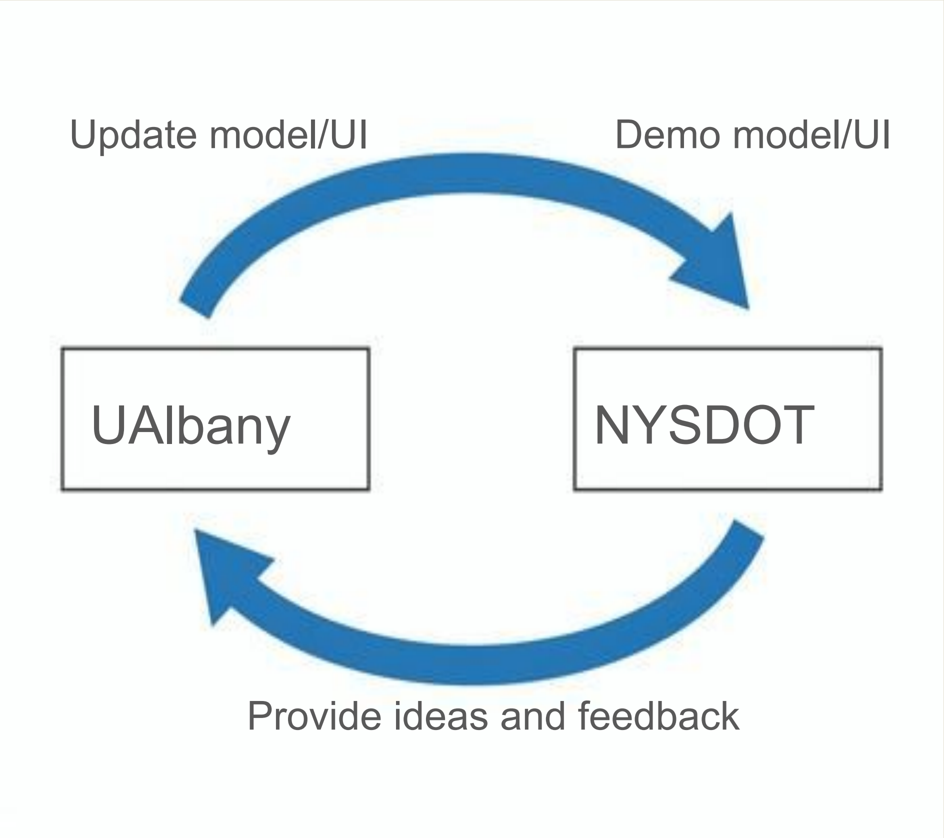
Public interests:

- Avoid 511ny.org stale conditions which impact public trust/reliance



January 4, 2024 at 3:20pm

Co-design with NYSDOT



Ongoing collaboration with the end-user is crucial create a useful tool

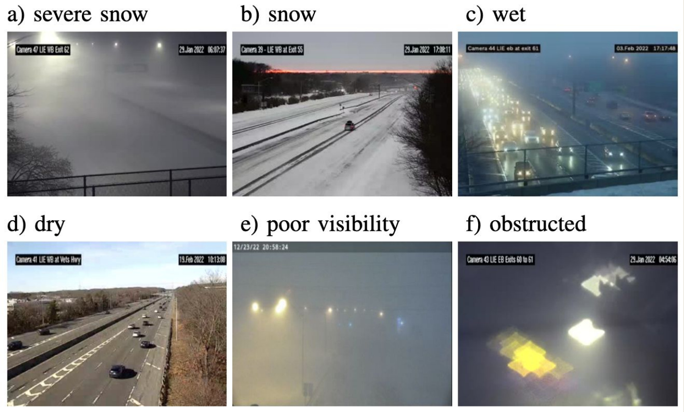
Collect feedback &

tailor features iteratively

Co-design with NYSDOT

Working iteratively and regularly with NYSDOT ⇒ their needs guide the development of the model

Labeled-dataset curation for supervised ML



Deciding which classes to include in the model

Deciding what ***visually*** (imgs) constitutes each class

Designing model splitting by site ⟶ operational use

Model selection

What considerations are most important for model selection?

High priority
1. Validation accuracy
2. Adjusted validation accuracy(2)
3. **Severe snow** class recall

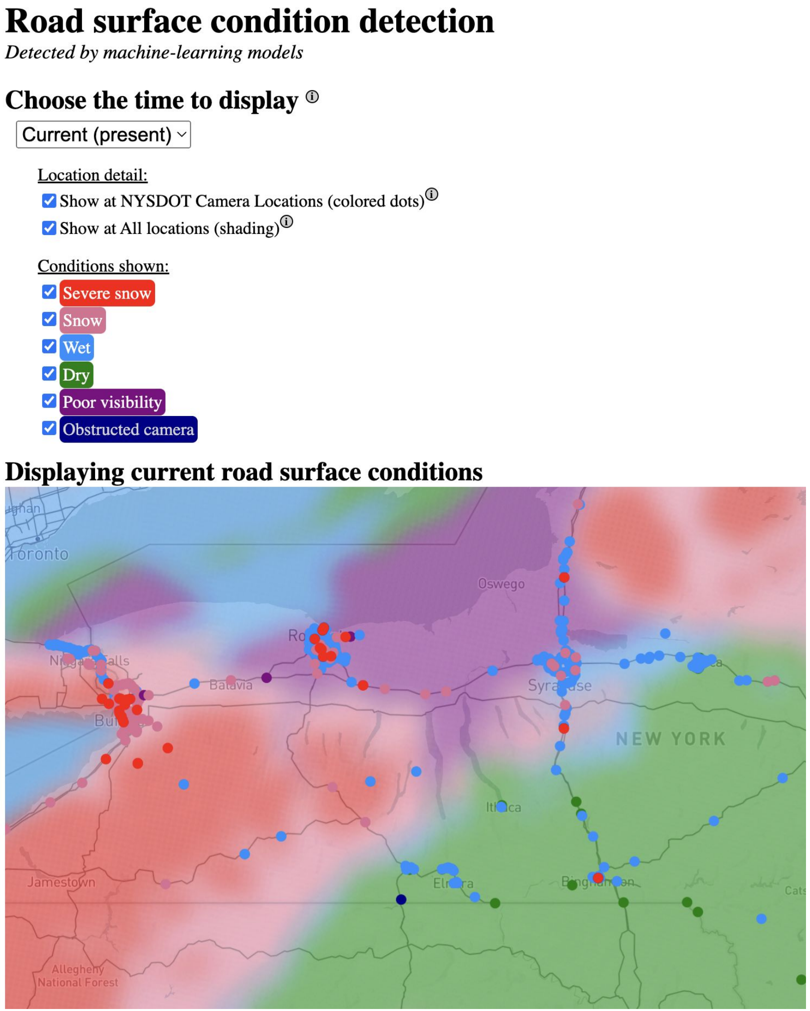
Medium priority
4. **Snow** class recall
5. **Wet** class recall
6. **Dry** recall

Low priority
7. **Poor visibility** recall(3)
8. **Obs** recall(3)

Co-design with NYSDOT

Dashboard: A user interface will allow the NYSDOT to ***visualize*** and easily ***use*** the model predictions

Recently: Social science interviews



End-user engagement and co-design

Social science interviews (IRB) for understanding decision making and co-development of a dashboard

Completed Jan-Feb 2025:

- 15 employees total
- 13 interviews
  - 11 were individual (one-on-one)
  - 2 were done in pairs
- Interview length was ~45 minutes each
- Two parts:
  - Part 1: understanding decision making
  - Part 2: demo ML model and dashboard

End-user engagement and co-design

<u>*Preliminary* takeaways from interviews:</u>

Wind is a major concern for NYSDOT. Blowing snow: can't maintain clear roads, poor visibility, trees/power lines

Ice prediction model would be useful

Interest in dashboard displaying the **confidence** of the predictions

Forecasting *probably* more useful than the nowcasting tool, but both have uses – a successful example of co-design ✅

Tool provides a holistic (statewide) picture **where the storm is and isn't** – for reallocation of resources

Trust in our tool was widely varied! Most say **"need to use it first"**

Conclusion

NYSDOT's current manual methods & Automated road-surface detection → **Understanding of weather-related road surface conditions**

Future directions:
- Social science data analysis from interviews & subsequent model/UI updates from those
- Modeling: LSTM forecasting models, ice model, assessing model drift of selected models
- Operationalizing dashboard

This material is based upon work supported by the National Science Foundation under Grant No. RISE-2019758

Contact information: Carly Sutter csutter@albany.edu