MACHINE-LEARNING CLASSIFICATION OF BLOWING SNOW FROM WEBCAM IMAGERY

Sarah McCorkle[1], Z. Lebo[1], B. Geerts[1], R. Capella[1], E. Collins[1], R. Cox[2], A. Lyons[2]

1 University of Wyoming Department of Atmospheric Sciences

2 NOAA/NWS

June 6th, 2020

19th Conference on Mountain Meteorology

smccorkl@uwyo.edu

OVERVIEW

- CSTAR overview
- Convolutional neural networks overview
- My research goals
- First round of training – binary model
- Second round of training – classification model
- Meteorological characteristics of each category
- Future work

CSTAR INTRO

- NOAA/NWS Collaborative Science Technology, and Applied Research
- Aims to improve hazardous winter weather forecasting
- Wyoming impacted by several types of hazardous winter wx
  - High winds
  - Snow squalls
  - **Blowing snow (BLSN)**

- Focus on HRRR (High-Resolution Rapid Refresh) model
  - Validation of HRRR forecasts
  - HRRR-based predictions
  - BLSN product

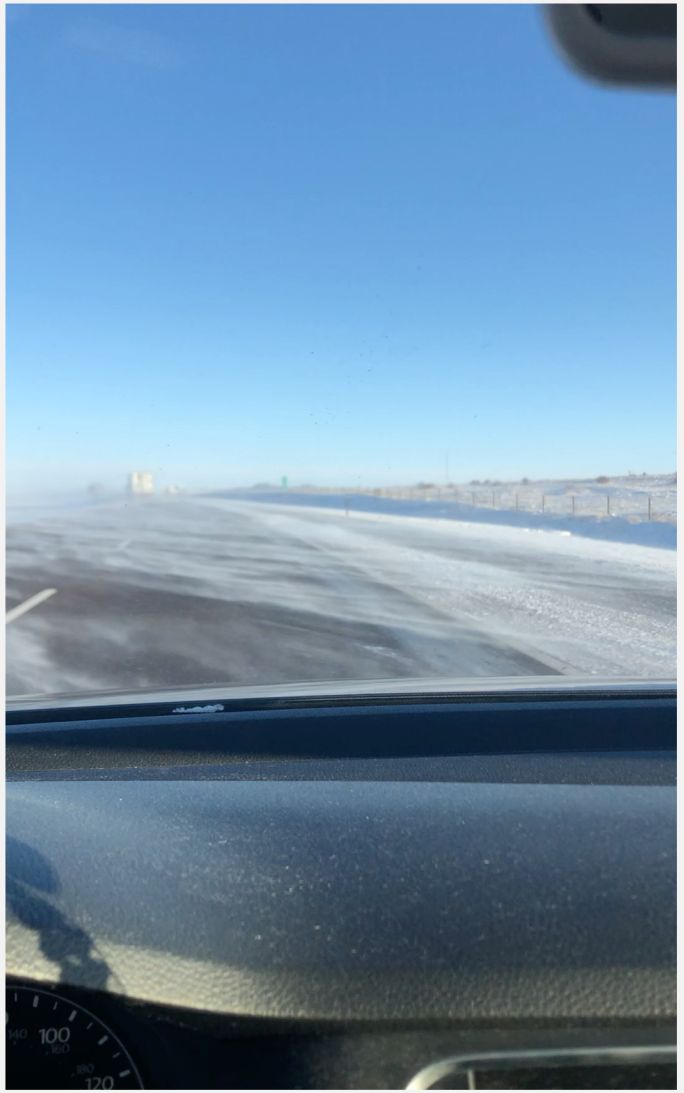
BLOWING SNOW RISKS

Blowing snow affects travel

- Low visibility
- Hard to predict

Depends on:

- Snow cover
- Snow age
- Other meteorological variables



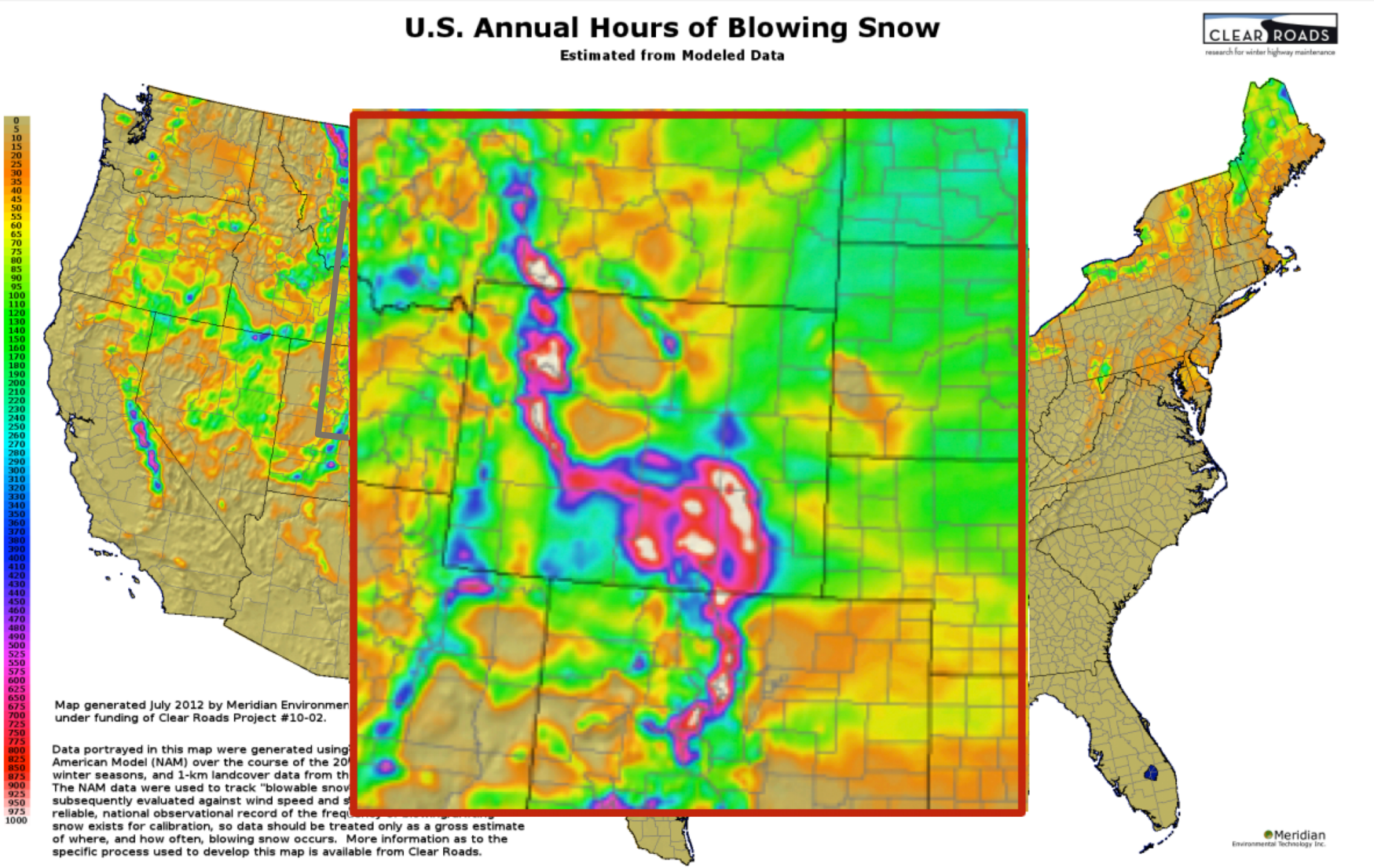


Figure 4: Map of Average Annual Duration of Blowing/Drifting Snow (in hours) as developed under this project.

Source: Clear Roads 2012: Mapping Weather Severity Zones

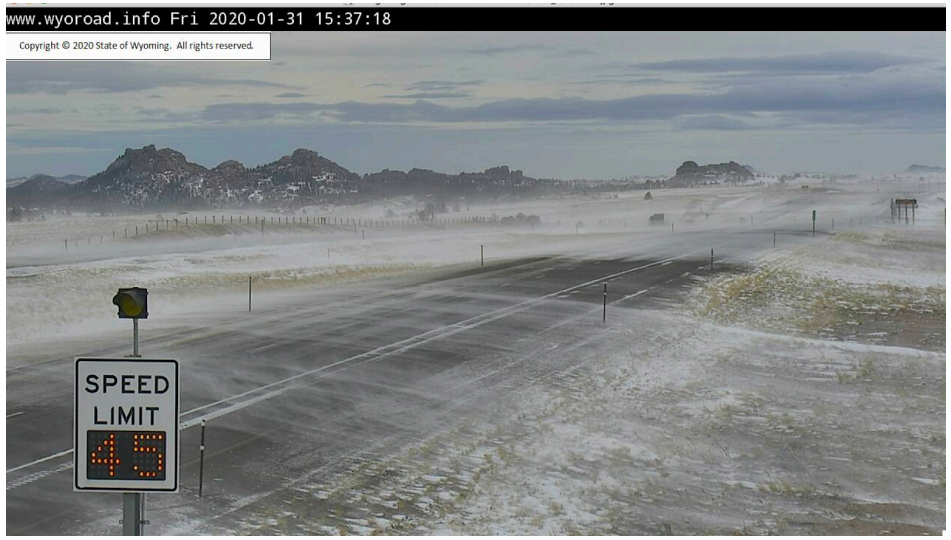
CONVOLUTIONAL NEURAL NETWORKS

- Image classification uses deep convolutional neural networks (CNNs)
- Defined as "a family of **parametric**, **non-linear** and **hierarchical representation learning functions**, which are massively optimized…to **encode domain knowledge**, i.e. domain invariances, stationarity." -- Efstratios Gavves

CNNS USEFUL FOR IMAGE CLASSIFICATION

- CNNs used for images of blowing snow from webcam images
  - Identifies patterns and features from training dataset
  - Classify based on characteristics only computer sees

What I see



What a computer sees

Adopted from https://medium.com/@ksusorokina/image-classification-with-convolutional-neural-networks-496815db12a8

CNNS CONVOLVE OVER THEIR INPUT

Input – image RGB values

3	3	2	1	0
0	0	1	3	1
3	1	2	2	3
2	0	0	2	2
2	0	0	0	1

Shaded: Sliding convolution kernel with size 3x3

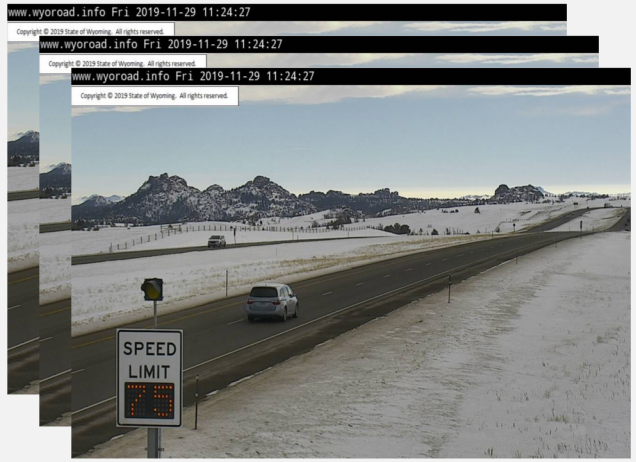
3.0	3.0	3.0
3.0	3.0	3.0
3.0	2.0	3.0

Max numbers being retained – most important features are carried on through the model (pooling)

https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

GOALS OF THE BLOWING SNOW ML MODEL: DEEP LEARNING ALGORITHM OVERVIEW

Input:
real-time webcam images



Output:
Is there blowing snow in these images? No

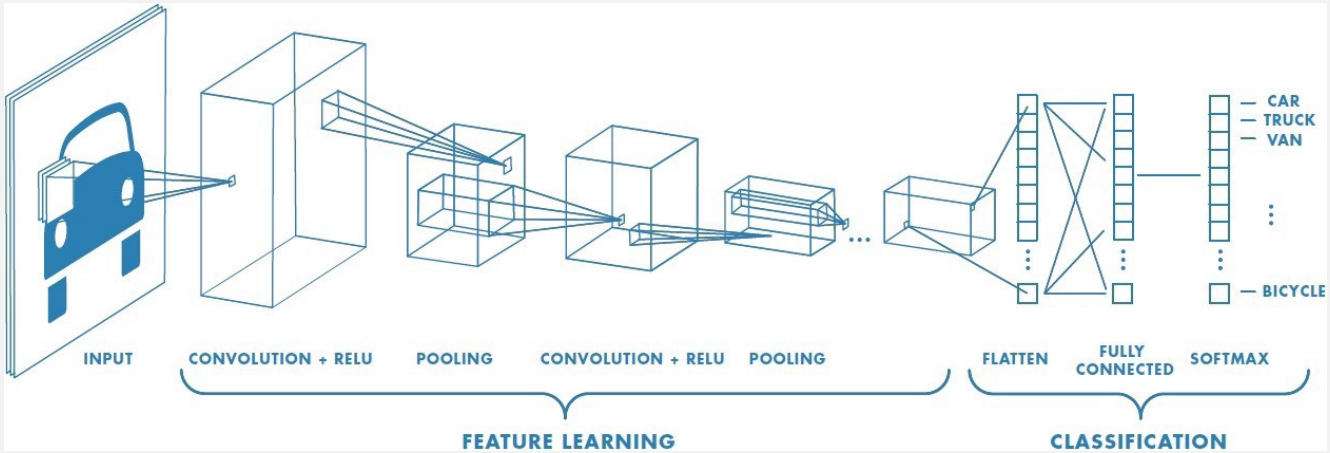
Adopted from https://www.xenonstack.com/blog/log-analytics-deep-machine-learning/

GOALS OF THE BLOWING SNOW ML MODEL

- Model accurately classifies blowing snow in real-time upon seeing new images
- Operational use→ model notifies forecaster when a webcam detects blowing snow
  - Aid forecasters in blowing snow events with automated webcam detection
- Develop an observational database of BLSN intensity for HRRR-based BLSN product
  - Outputs a probability of a certain intensity of BLSN

MODEL ARCHITECTURE OVERVIEW

- 3 convolutional layers
  - First layer (input layer) – 32 filters
    - Takes in pixel height, width, **features (RGB values)**
    - (720,1280,3)
    - Kernel size = (3,3)
    - Max pooling = (2,2)
  - Second layer – 32 filters
  - Third layer – 64 filters
- Flattened – output layer
- Weights of nodes saved
- 10 epochs – one complete presentation of the data to the model



https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-w
3bd2b1164a53

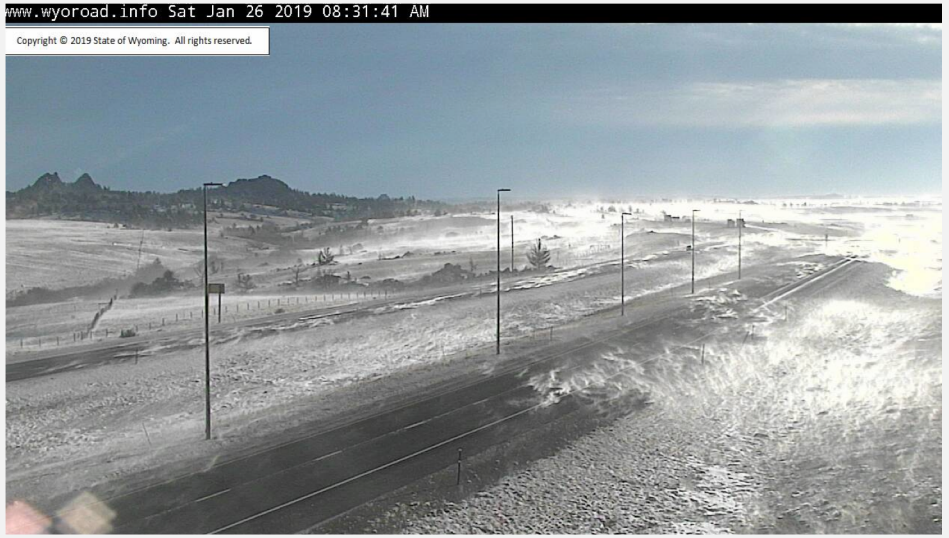
TRAINING DATASET - BROAD

- Archived WYDOT images
  - October 1, 2018 - January 31, 2019
  - November 1, 2019 - Feb 28, 2020
- Focused along Interstate 80 - 54 cameras
- Angles facing east, west, road sfc

BINARY MODEL

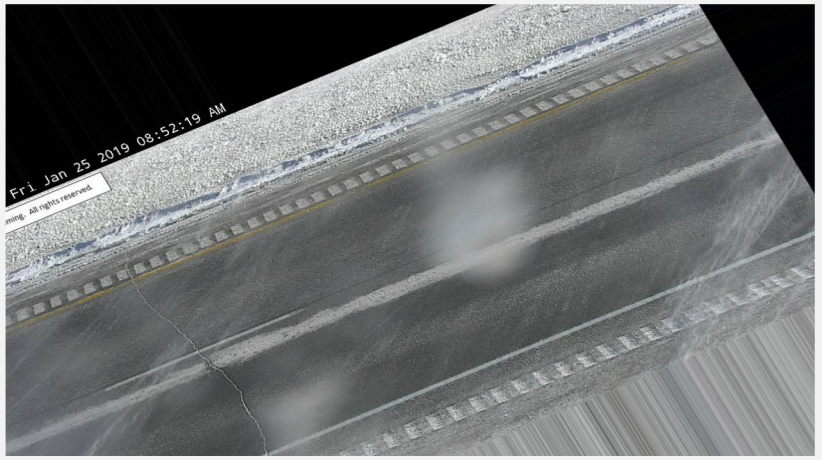
TRAINING DATASET

- 2 categories – binary model
  - Blowing snow vs not blowing snow (clear road)
  - Multiple camera angles and locations
  - 645 images



TRAINING DATASET

- Data Augmentation
  - Good for smaller datasets
  - Some road surface images at different angles

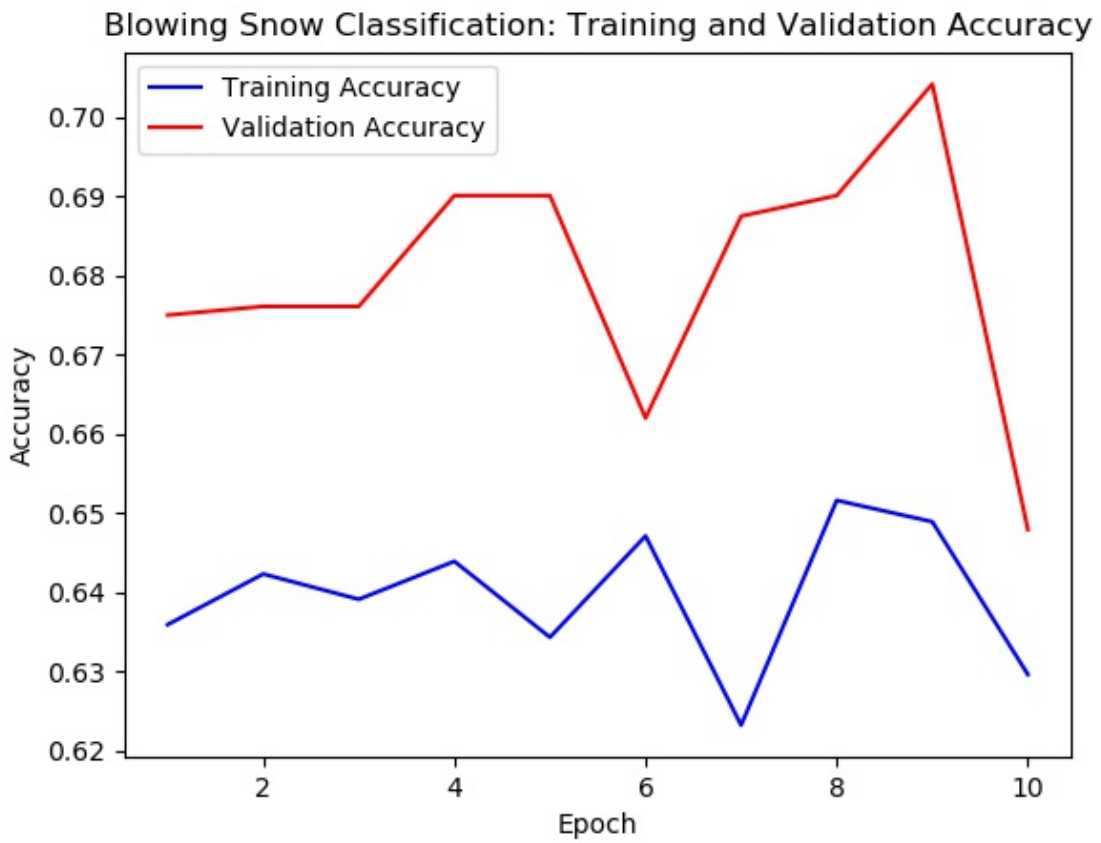


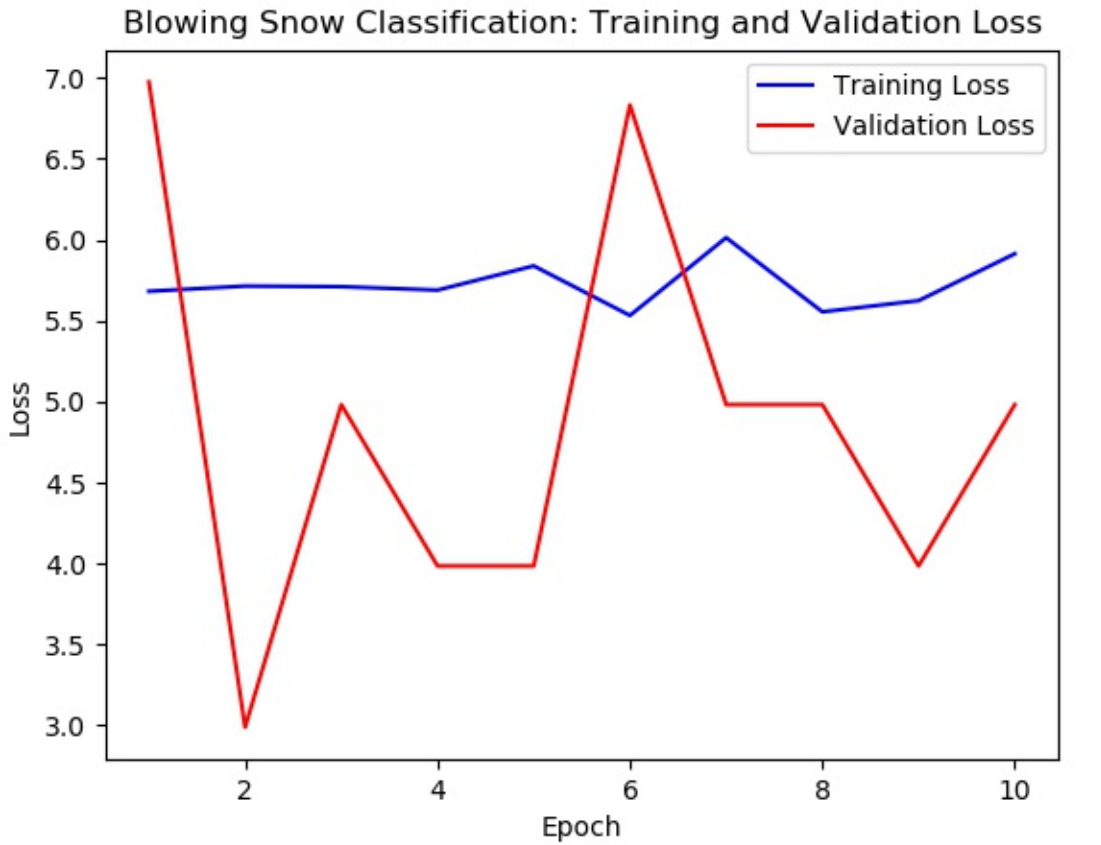
TESTING/VALIDATION DATASET

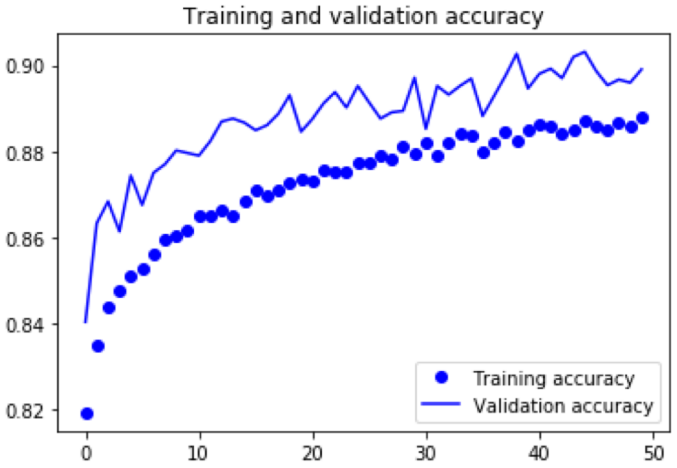
- Provides unbiased evaluation of the model
- Separate images from the training dataset
- 2 categories as well
  - Blowing snow vs not blowing snow
  - 87 images

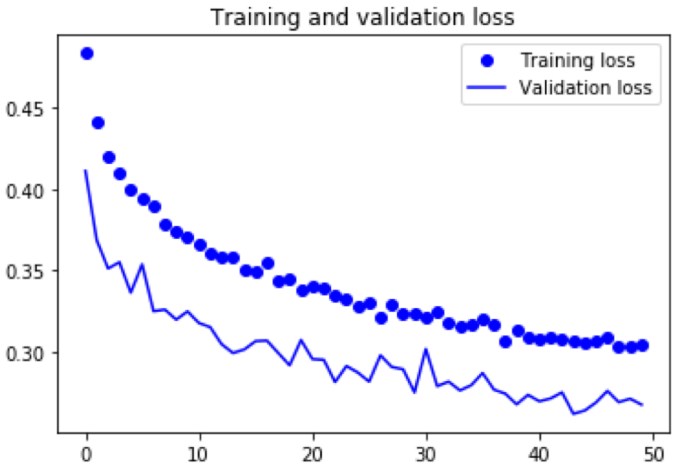
Binary Model Training: Initial Results

- 645 in training (2 classifications), 87 in validation (2 classifications)
- Accuracies not improving







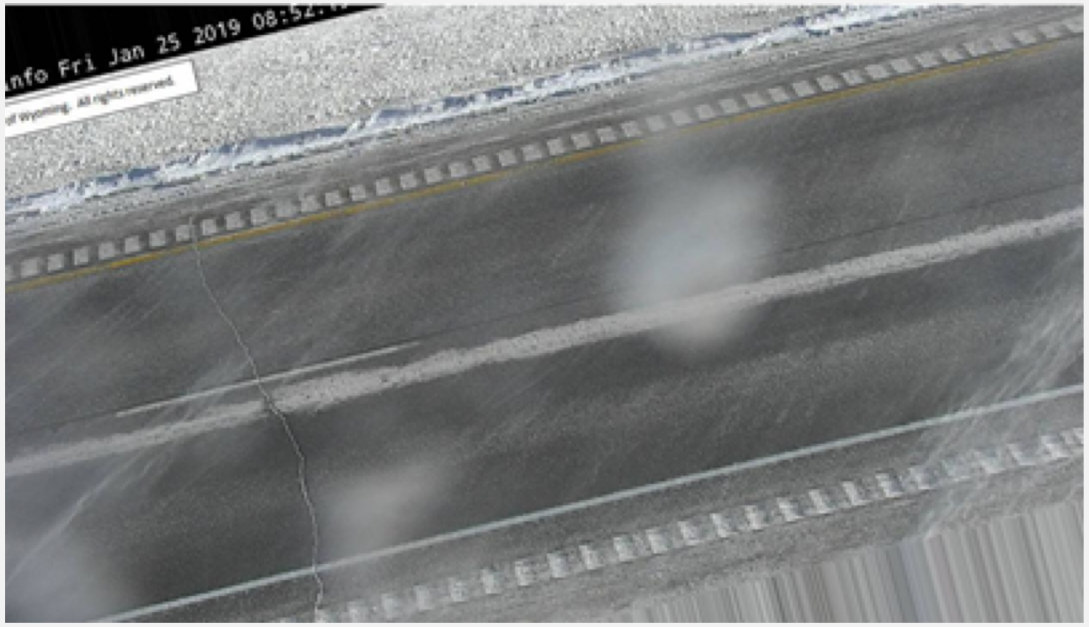


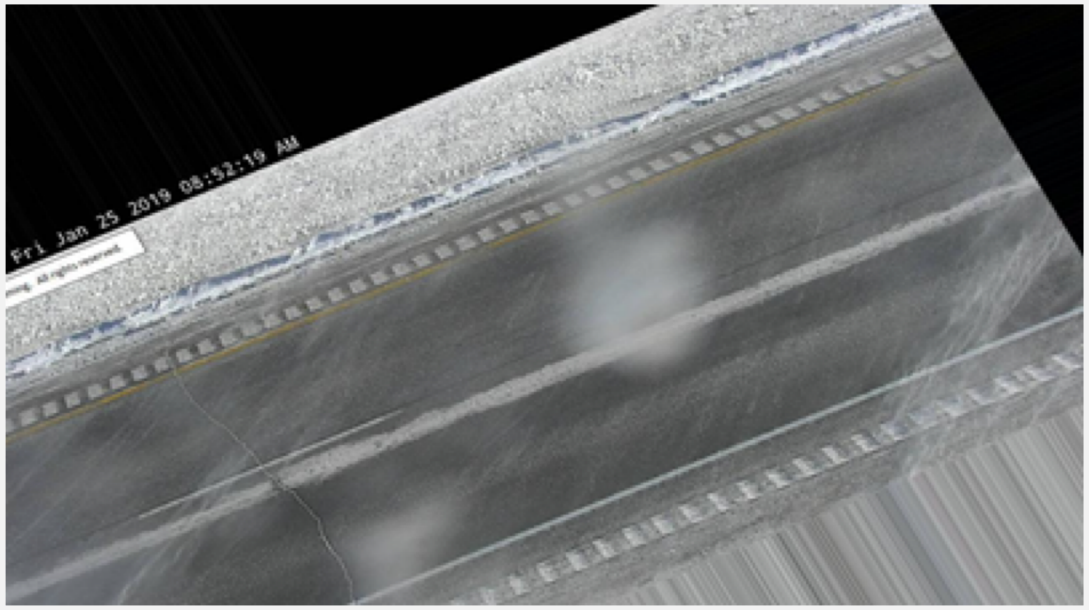
WHAT TO EXPECT

https://towardsdatascience.com/the-4-convolutional-neural-network-models-that-can-classify-your-fashion-images-9fe7f3e5399d

IMPROVING THE TRAINING DATASET

- Got rid rid these images in training dataset:





- Cropped images with cameras facing East and West

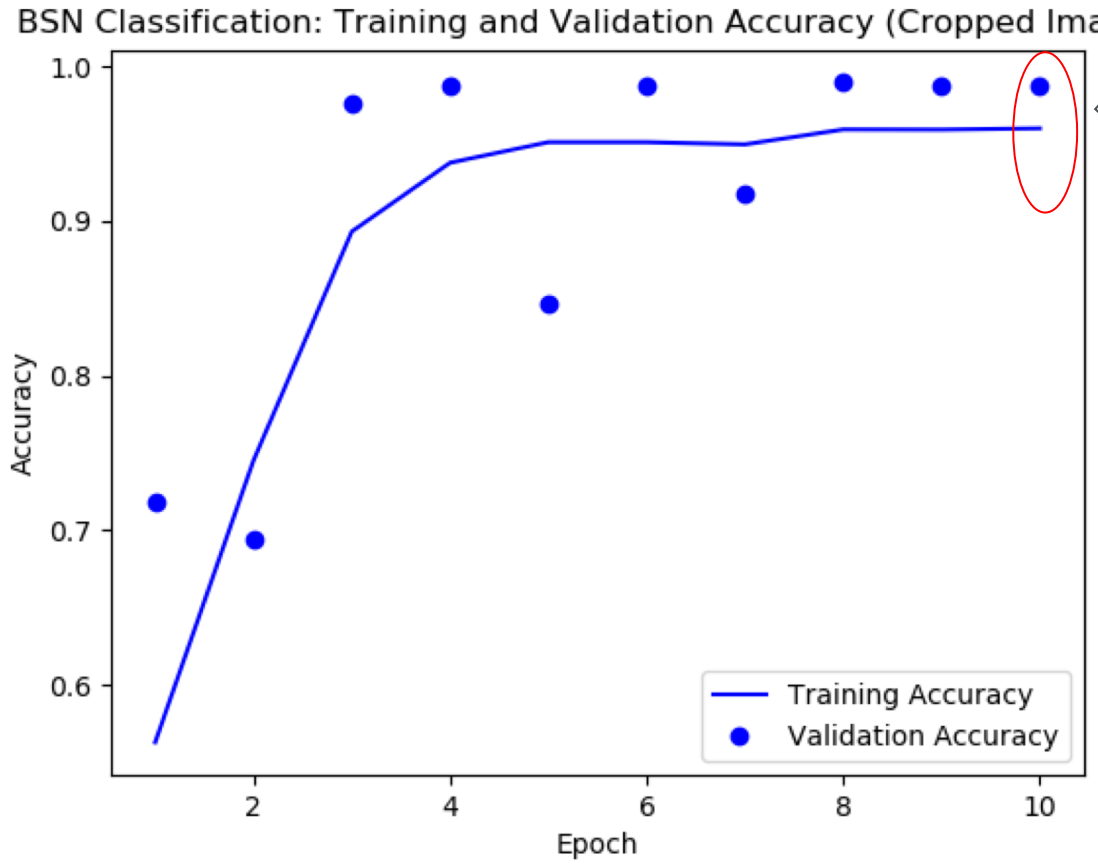
CROP IMAGES

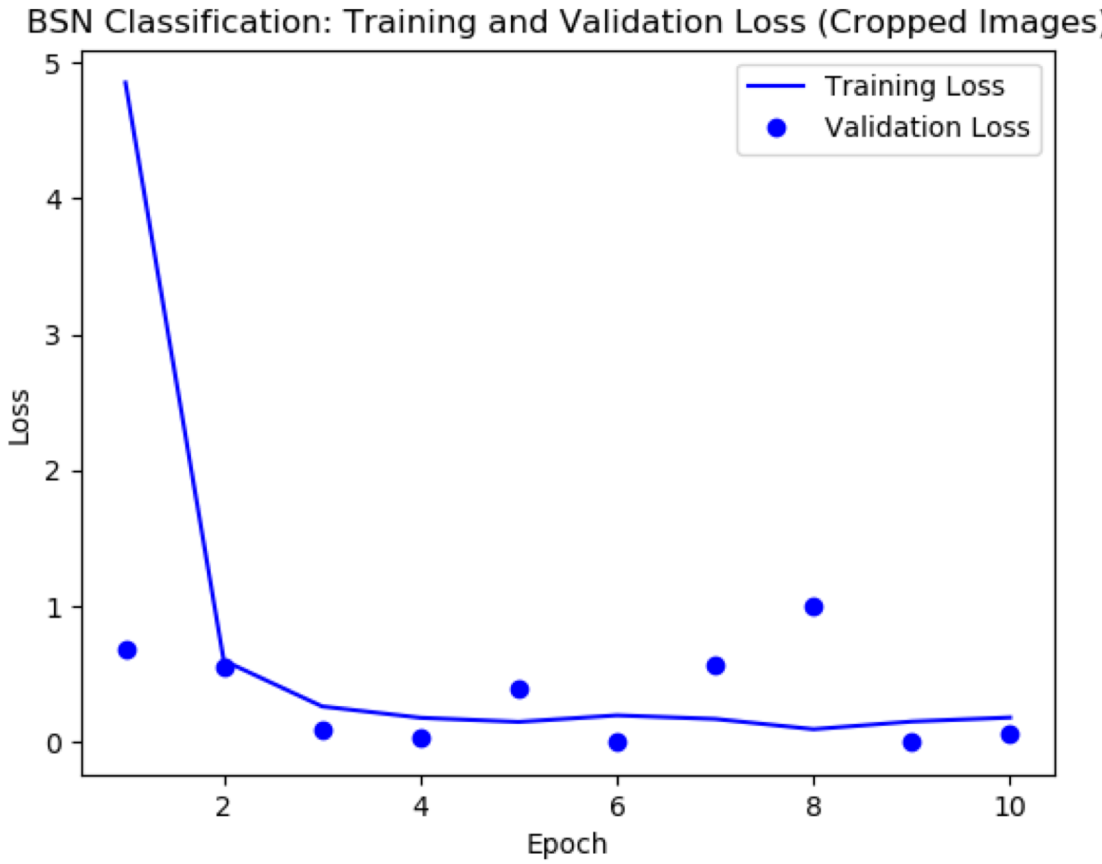
- Example: Vedauwoo East

- (720,1280) to (360, 1210)
- Hope that training the model on just the road will improve accuracy

Improved Accuracy

Training acc = 96.0%
Validation acc = 98.8%





CLASSIFICATION MODEL

CLASSIFICATION MODEL: BLOWING SNOW INTENSITY CATEGORIES

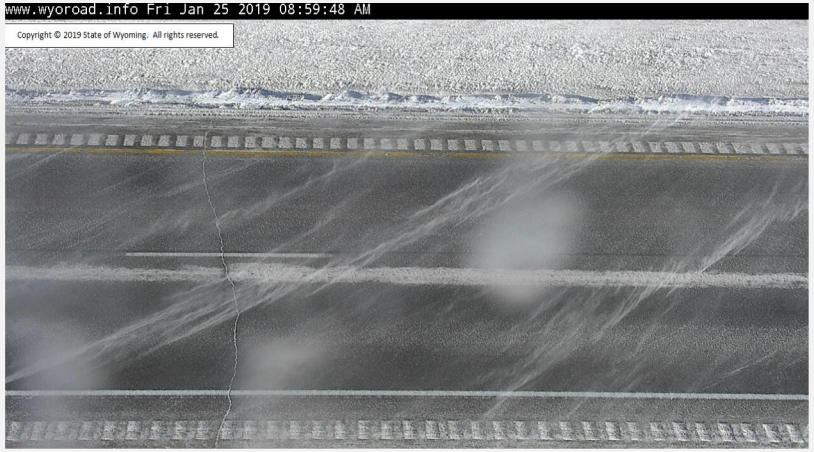
- Adopted from the Cheyenne NWS office's categories for blowing snow intensity
  - Patchy = Drifting Snow Only
  - Areas = Blowing Snow
  - Def = Advisory Conditions (Near Blizzard) - Sporadic to Occasional whiteout conditions
  - Def 1/4sm = Blizzard Conditions - Persistent whiteout conditions

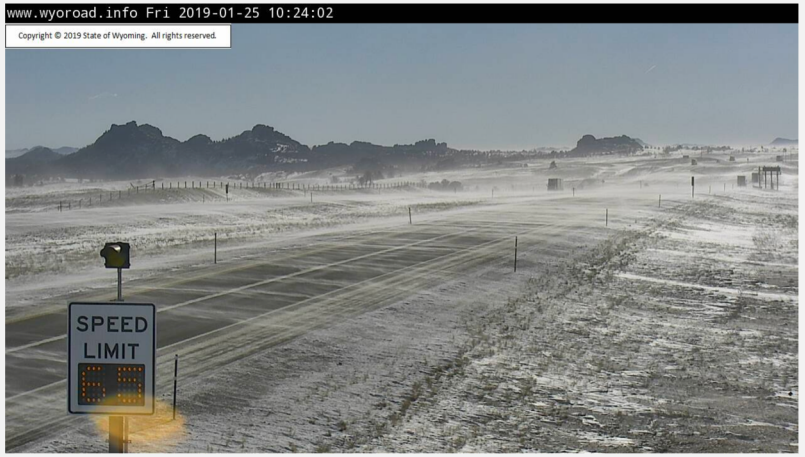
*Courtesy of Matthew Brothers, NOAA/NWS Weather Forecast Office, Cheyenne, WY

CATEGORIAL MODEL TRAINING AND TESTING DATASET

- Training dataset contains 717 images falling into 4 categories:
  - Strong, medium, drifting, no BLSN
- Testing dataset: 178 images
  - Falling into the same categories

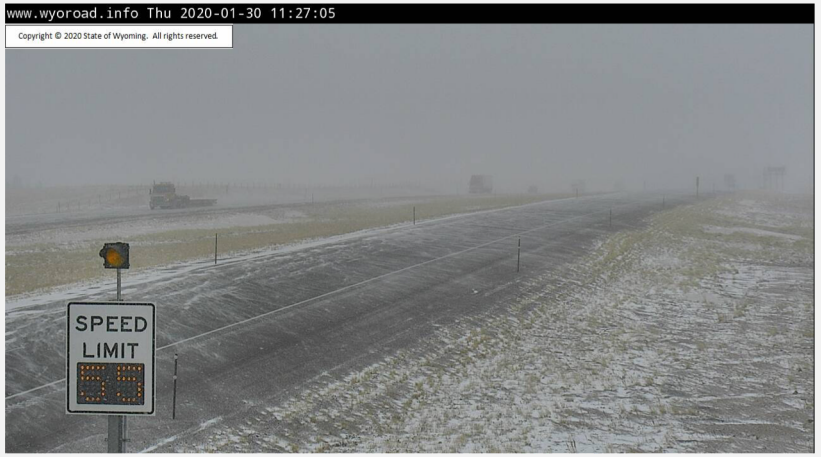
Drifting





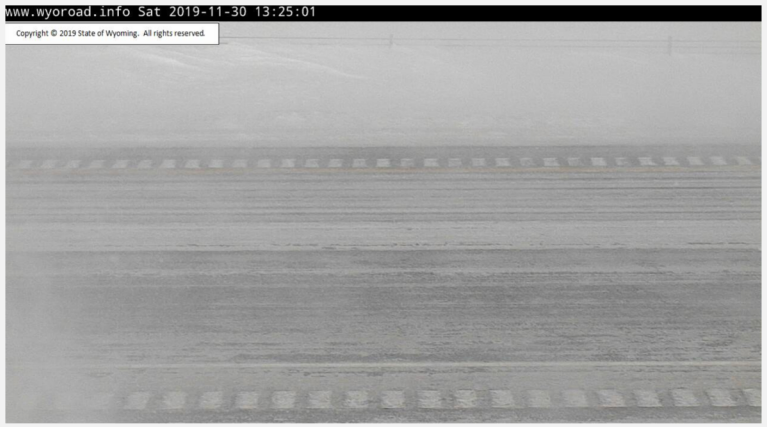
Jan 25, 2019

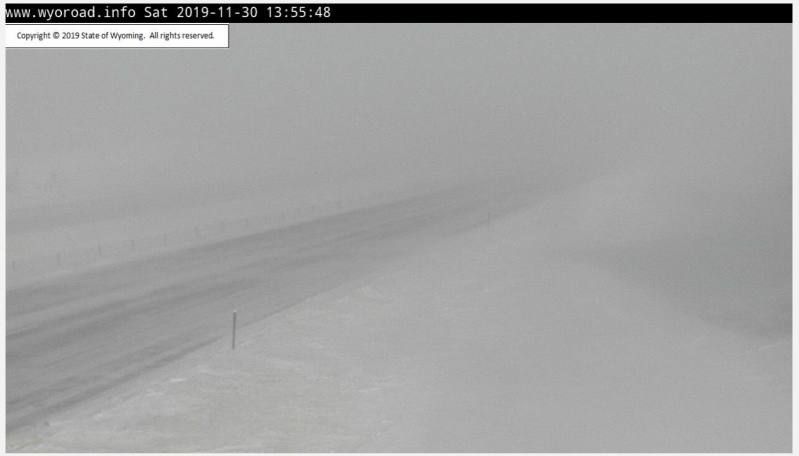
Medium



Jan 30, 2020

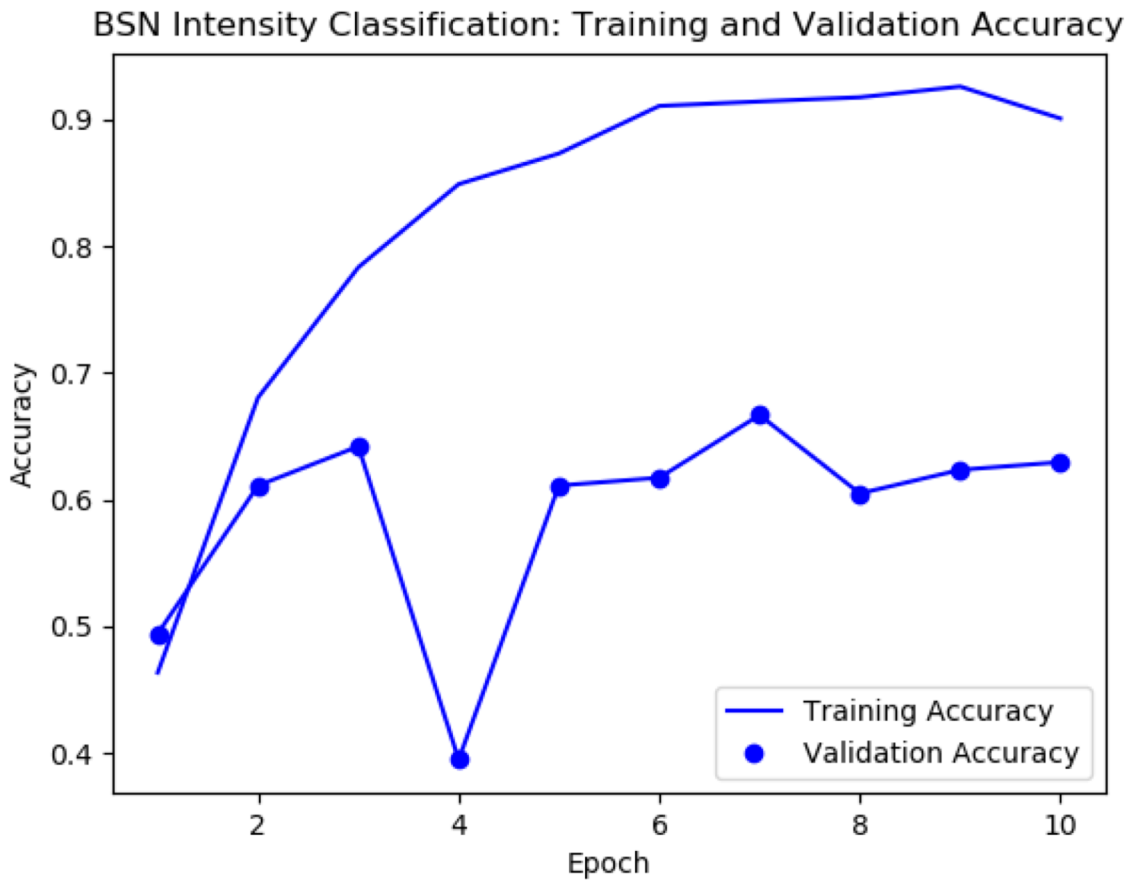
Strong

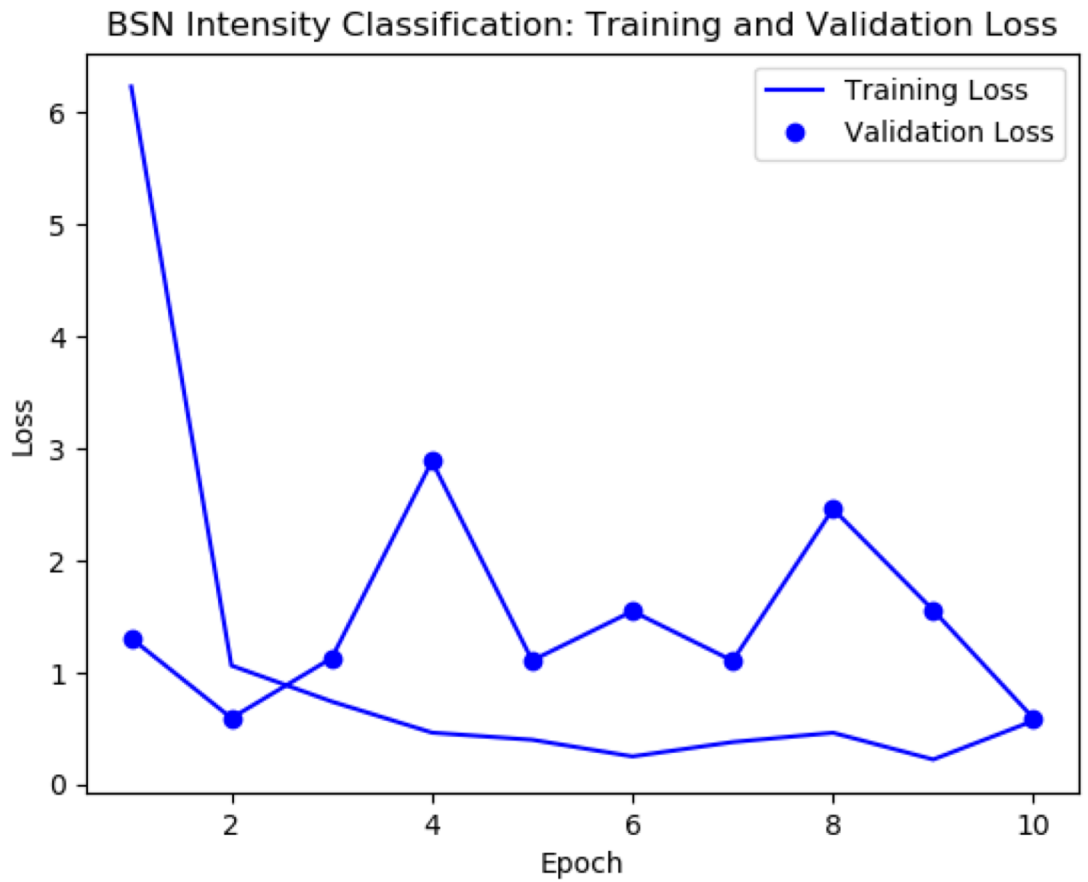




Nov 30, 2019

Classification Model: Training Results





-Training accuracy improves with each epoch

-Validation accuracy not the best → evaluates model performance upon seeing new images

-Training loss drops off nicely, but want it to be close to validation loss

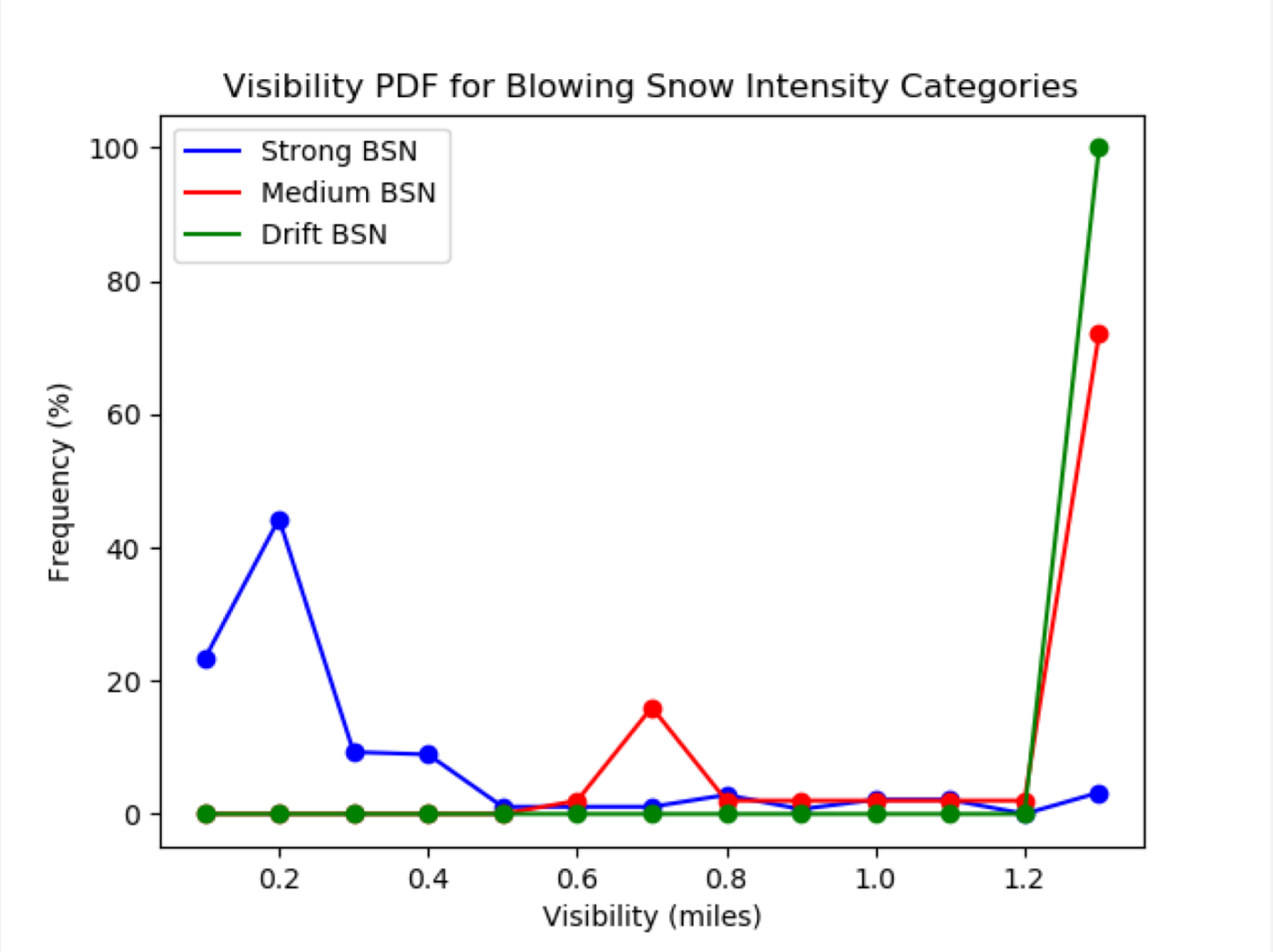
-See possible overfitting here since validation loss > training loss

METEOROLOGICAL CHARACTERISTICS OF EACH CATEGORY

- Gather observations within each category → strong, medium, drifting
  - Air temperature, road temperature, visibility, wind speed, wind gust, RH (liquid), RH (ice)
- Look at probability distributions to see which meteorological variables are relevant for causing BLSN

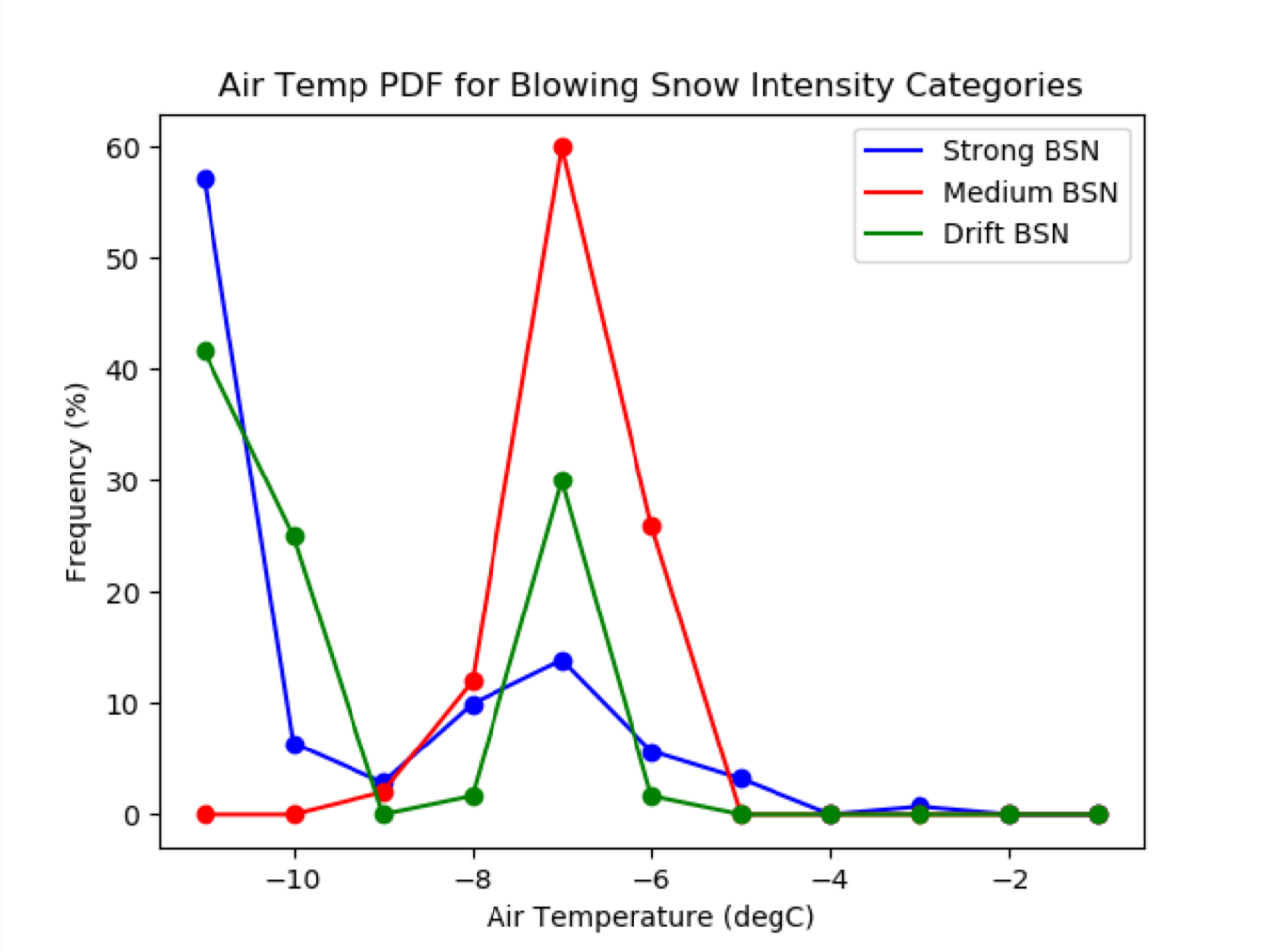
- Strong: 280 samples
- Medium: 60 samples
- Drift: 80 samples

VISIBILITY PDF

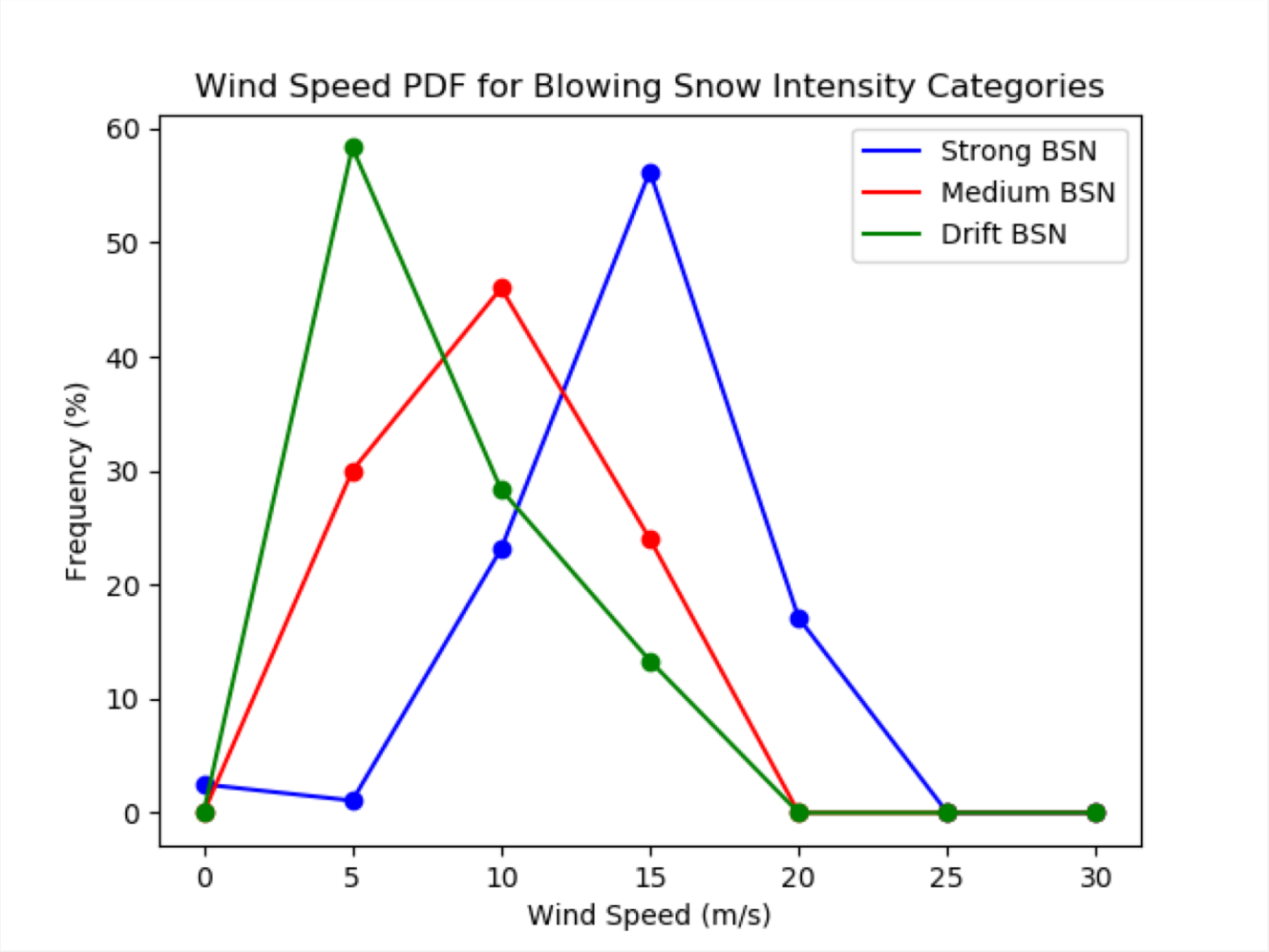


*1.24 miles is max visibility measured

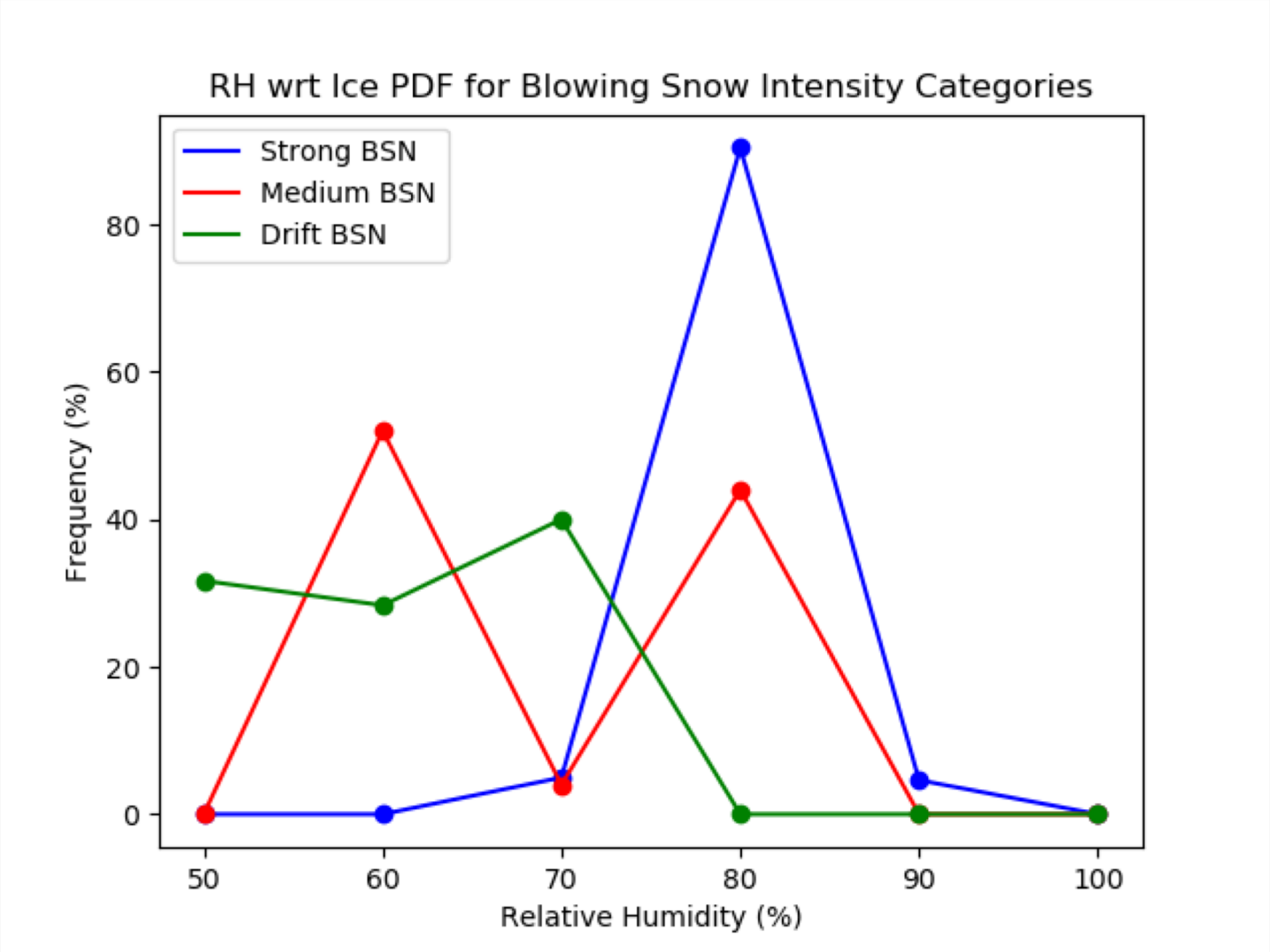
TEMPERATURE PDF



WIND PDF

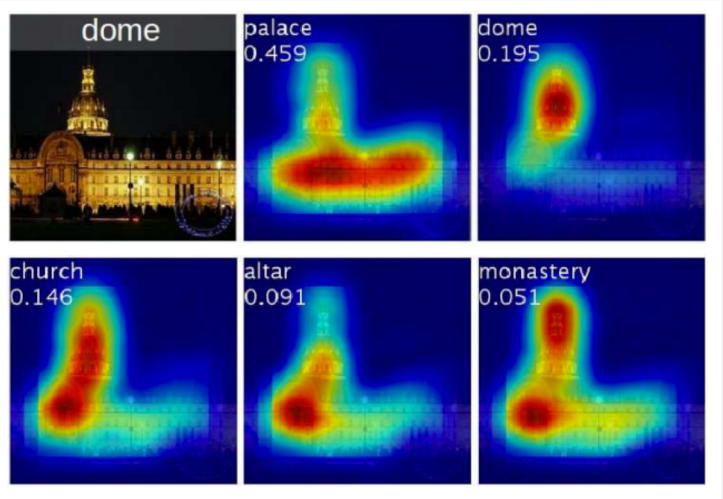

Wind Speed PDF for Blowing Snow Intensity Categories
Strong BSN
Medium BSN
Drift BSN
Frequency (%)
0
10
20
30
40
50
60
0
5
10
15
20
25
30
Wind Speed (m/s)

RELATIVE HUMIDITY PDF



NEXT STEPS

- Improve model validation accuracy/loss
  - Add more images to dataset
  - Optimize hyperparameters in model
- Create new model with different input types – images and meteorological data to improve accuracies
- Increase model interpretability
  - ML algorithms more transparent, less of a blackbox
  - Use visualizations such as saliency maps, class-activation maps, etc.



https://rajpurkar.github.io/mlx/visualizing-cnns/

Thank you to:

- Zach Lebo
- CSTAR lab group
- WYDOT for the archived images
- Cheyenne NWS
- NOAA Grant NA19NWS4680005

Questions?

Q&A session Monday, July 13th at 3:15 MT

Contact:
Sarah McCorkle
smccorkl@uwyo.edu