Classifying Rainfall Regions in Weather Radar Mosaics

ALEX HABERLIE

LOUISIANA STATE UNIVERSITY

AHABERLIE1@LSU.EDU

PROJECT COLLABORATORS: WALKER ASHLEY (NIU), VICTOR GENSINI (NIU), RUSS SCHUMACHER (CSU)

Useful Information

https://github.com/ahaberlie/unidata-workshop-2018

Haberlie and Ashley (2018)

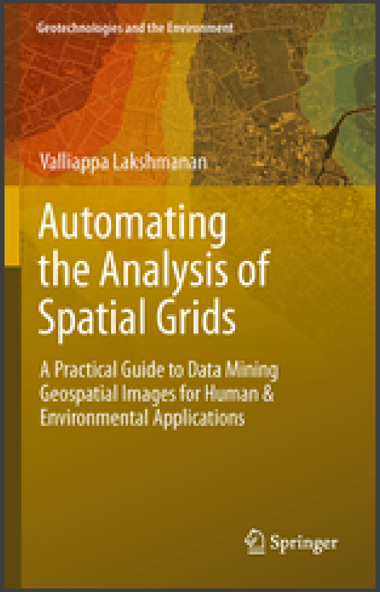
McGovern et al. (2017) – Great overview of current machine learning trends in Meteorology

Storm Identification and Feature Extraction

- WDSS-II (Lakshmanan et al. 2007)
- TITAN (Han et al. 2009)
- THoR (Houston 2015)
- Hagelslag (Gagne 2018)

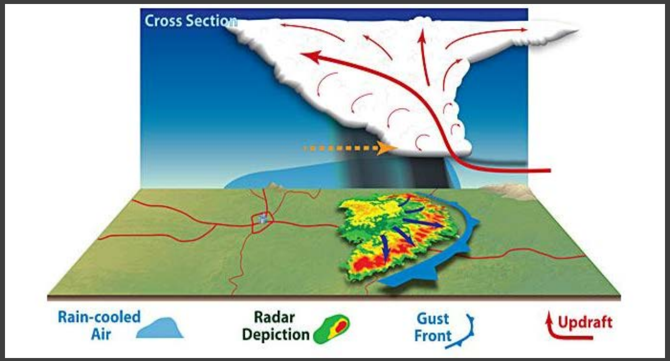
Machine Learning and Forecasting:

- MCS initiation (Ahijevych et al. 2016)
- Damaging Straight-Line Wind Prediction (Lagerquist et al. 2017)
- Heavy rain forecasting (Herman and Schumacher 2018)



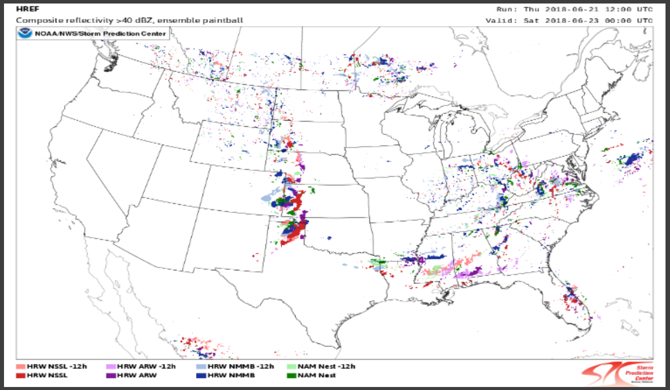
Mesoscale Convective Systems

Source: *NOAA Storm Prediction Center*

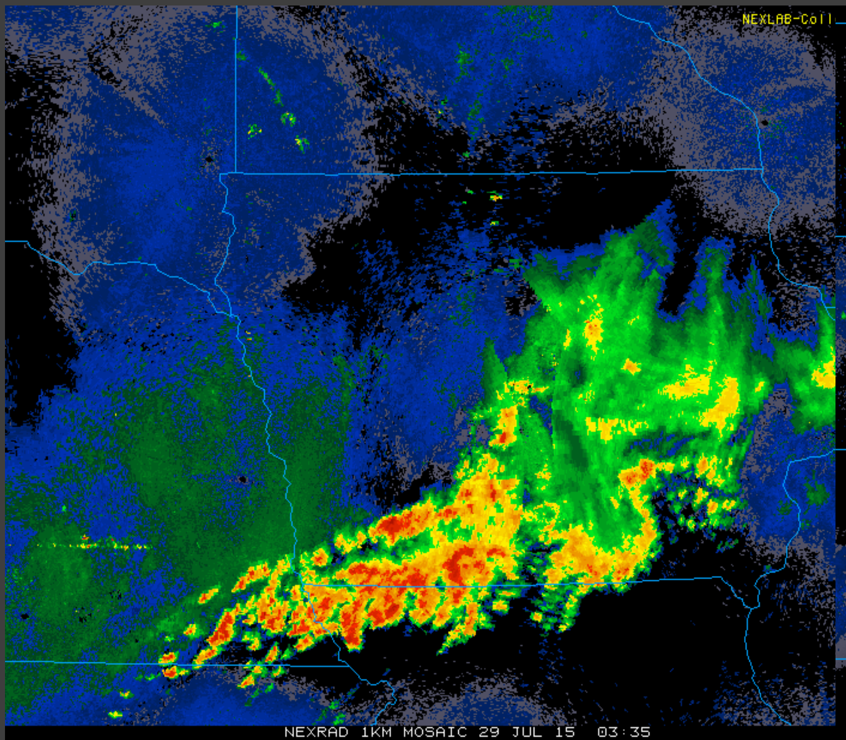


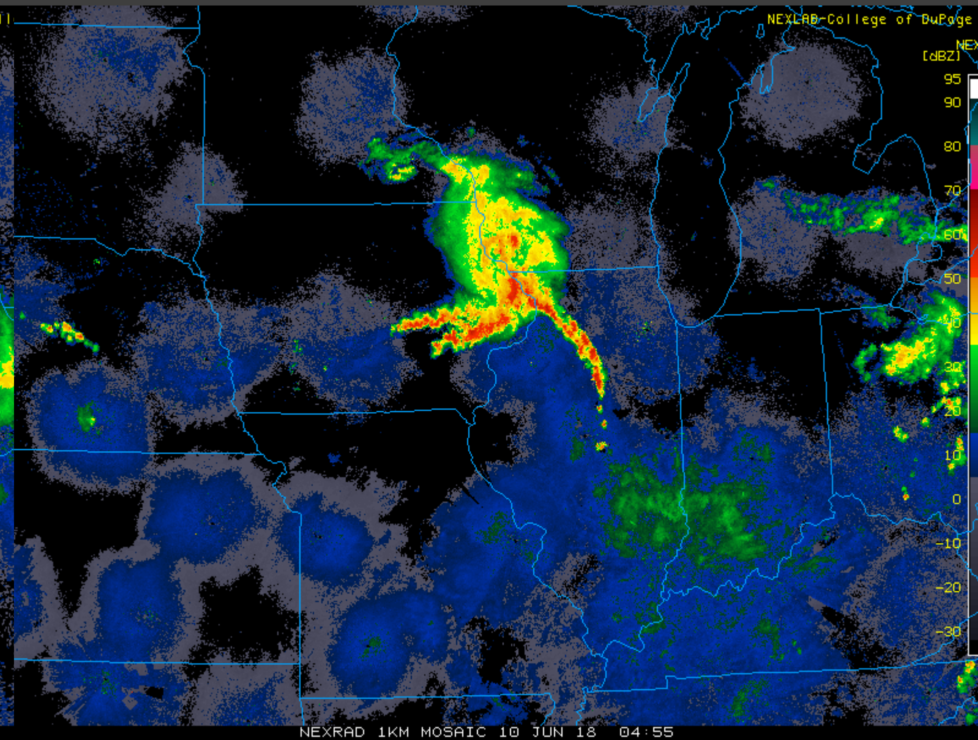
Source: Walker Ashley

Source:Arturo Fernandez/Rockford Register Star via AP.



Source: *NOAA Storm Prediction Center*

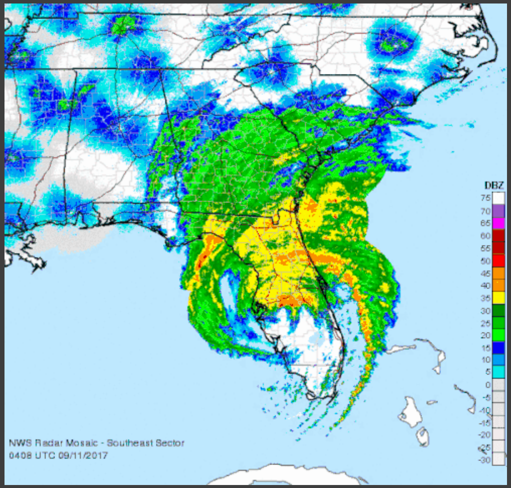


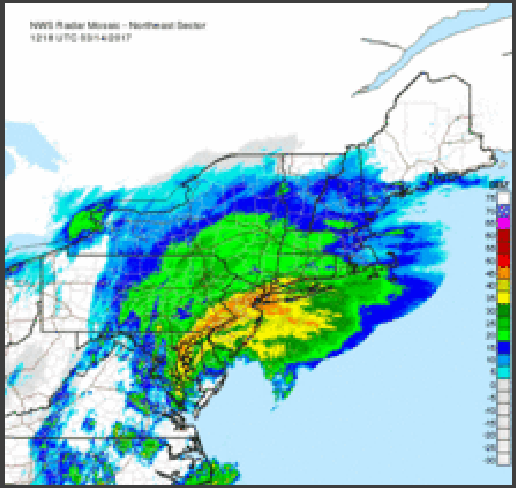


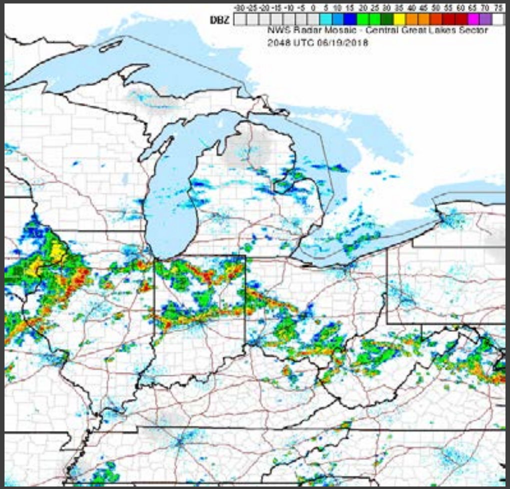
Parker and Johnson (2000)

Gallus et al. (2008)

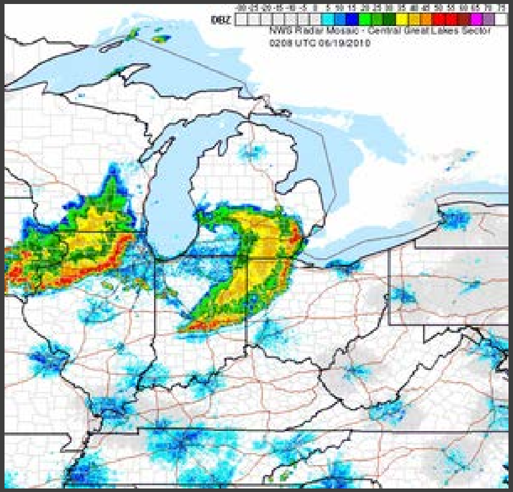
Are These MCSs?

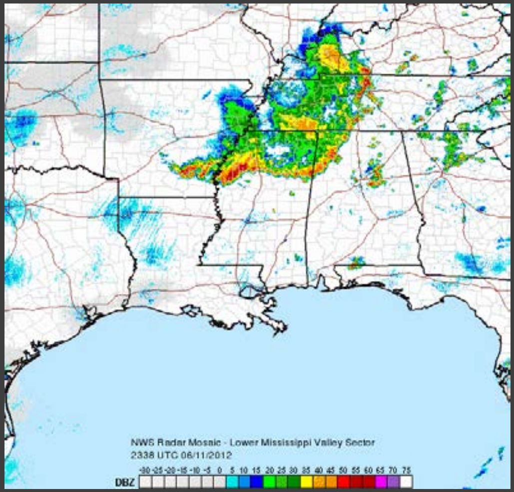






What about these?





Problem Description

How do you generate a climatology of mesoscale convective systems (MCSs) with a huge dataset of composite radar mosaics?

- NOWrad (~2 km)
- Over 95% of 15 minute periods from 1996-2017
- ~10^6 images
- Many well-known issues, but the analyses can be useful (Fabry et al. 2017)

Parker and Johnson (2000) **objective definition**:

- Convective cells organized on a horizontal scale of at least 100 km
- Must last for at least 3 hours

Computing Resources:

- Ryzen 1700 (8 C, 16 T), nVidia GTX 1070, 32 gb RAM

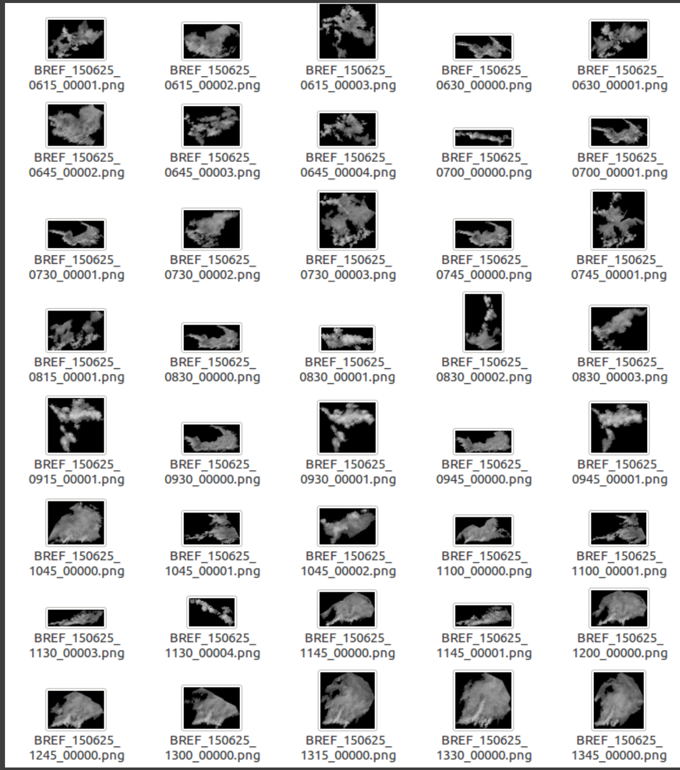
Why Machine Learning?

"Reducing time to science"

~5.5 million "MCS Snapshots"

Automate classification of MCSs and four common false positives after segmentation

- Tropical Systems
- Synoptic Systems
- Unorganized clusters
- Ground Clutter / Noise / Etc.



Related Work

Baldwin et al. 2005

- Linear, cellular, stratiform

Gagne et al. 2009

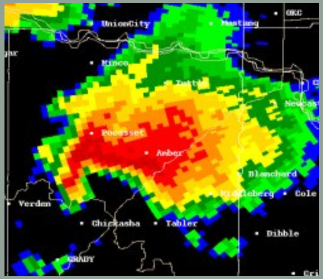
- Pulse, multicellular, MCSs

Lack and Fox 2012

- Supercell, QLCS, rotating storms, pulse, etc.

Hobson et al. 2012

- Supercell, pulse, multicellular, linear

Visual Depiction	Features
 User Label: Supercell	Area: 9,000 sq. km Mean Intensity: 35 dBZ Eccentricity: 0.3 …
 User Label: Linear	Area: 70,000 sq. km Mean Intensity: 31 dBZ Eccentricity: 0.8 …

Sample Training Workflow

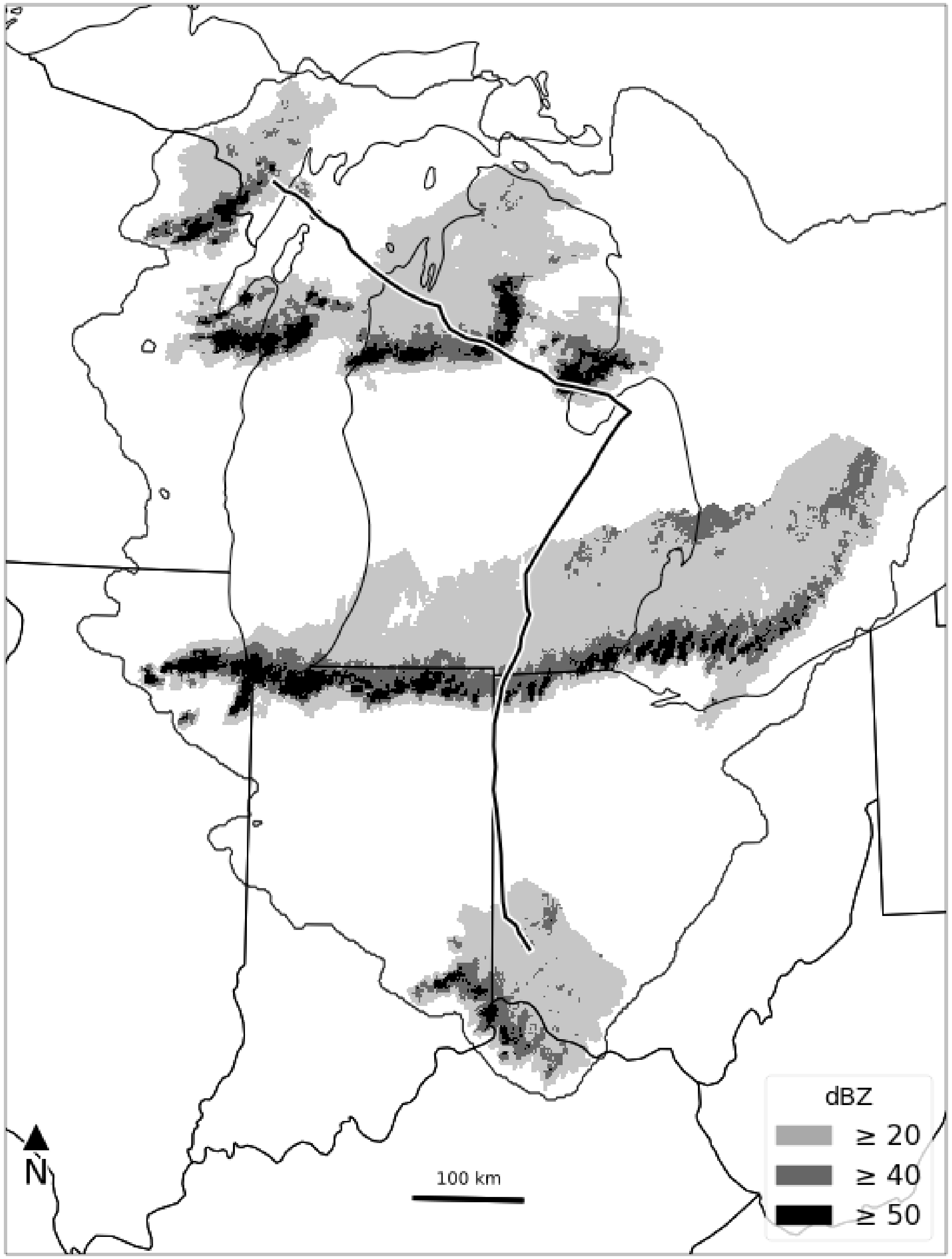
1) Ask yourself a few questions:
   - What are the classes you want to identify?
   - What are distinguishing features of each class?
   - What data do you need to gather samples?
   - What algorithm should I use?

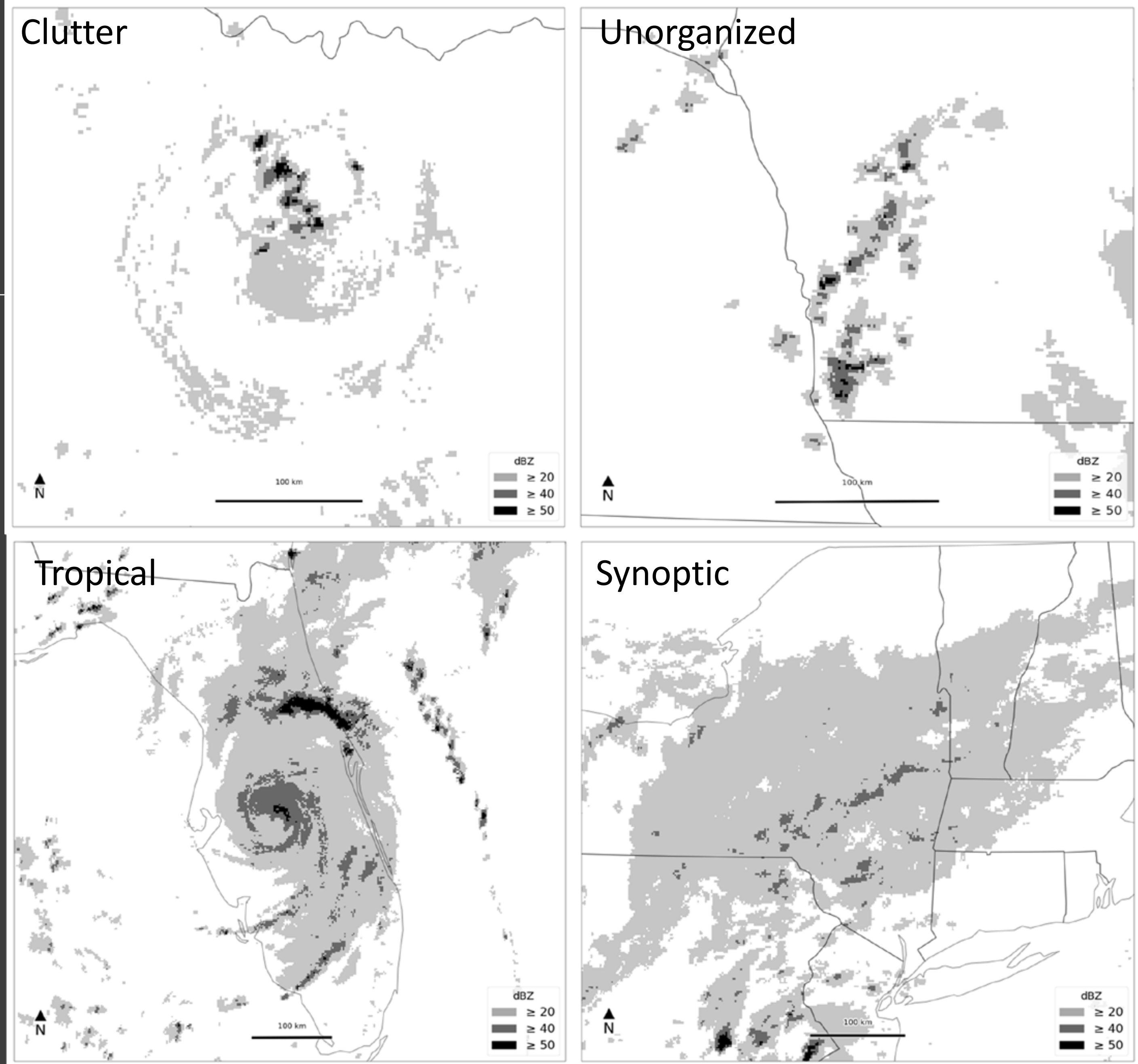
2) Identify class examples

3) Extract features
   - Area, Shape, Intensity, etc.

4) Generate training and testing data

5) Train machine learning model
   - Always test model performance on data **not** used to generate model


dBZ
≥ 20
≥ 40
≥ 50
100 km
N



Clutter
Unorganized
Tropical
Synoptic
dBZ
≥ 20
≥ 40
≥ 50
N
100 km

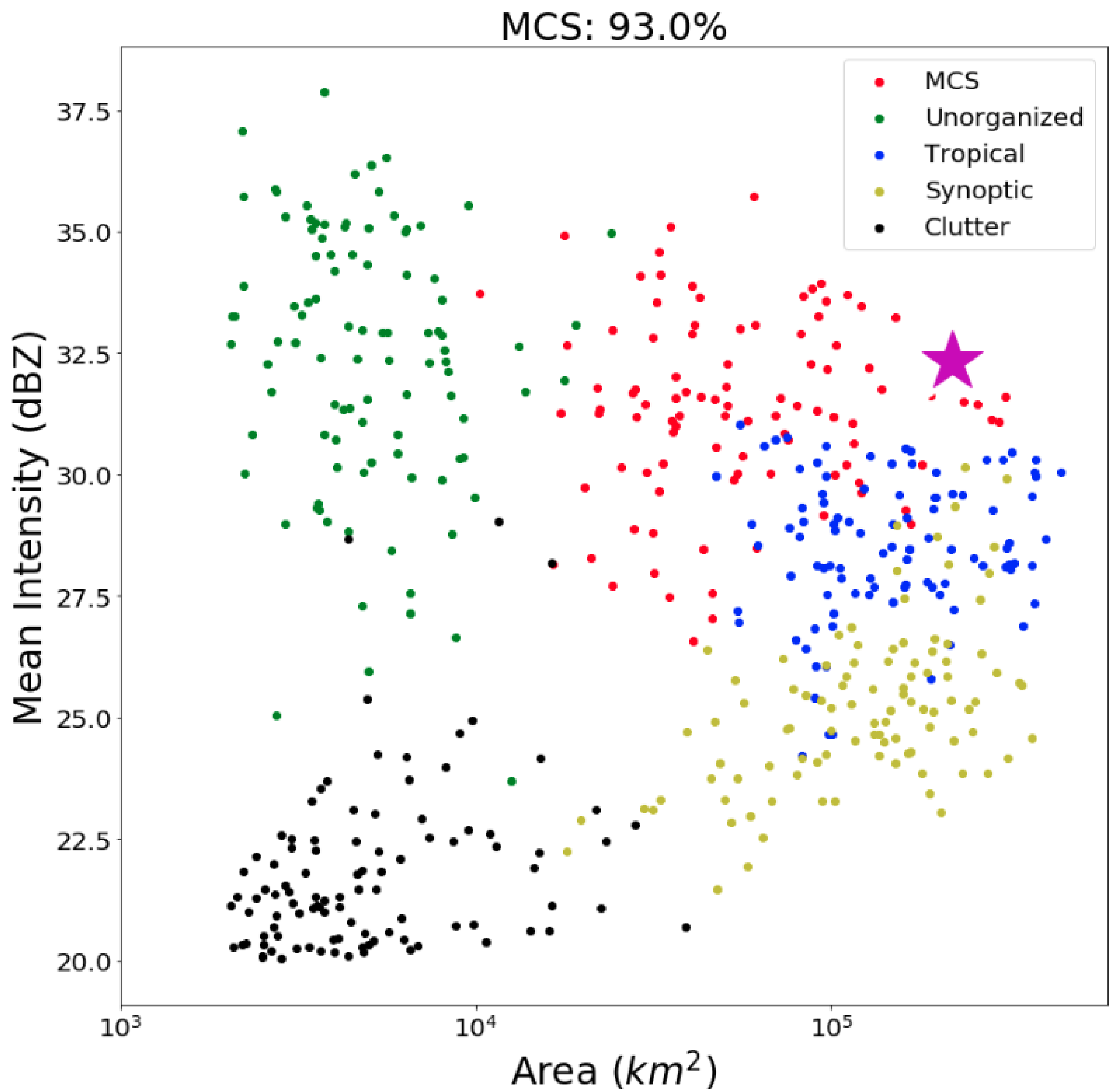
Notebook Examples

Training Process

Extraction Process

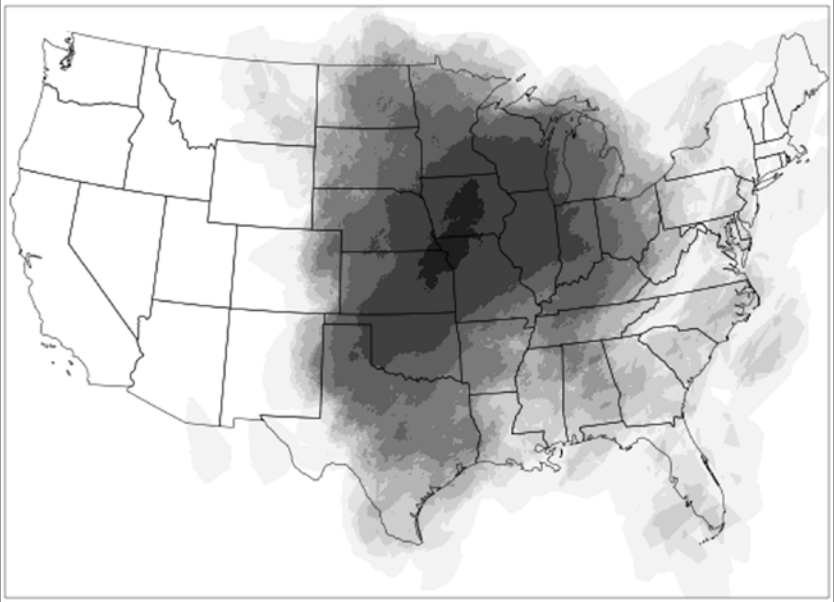
Area 218872

Mean Inten... 32.3364340802

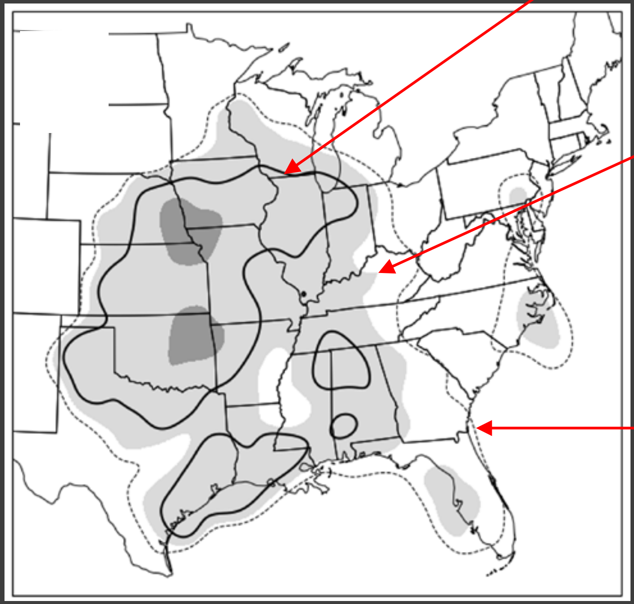


Machine Learning vs. Manual

Manual MCS slice positions (2003-2013)



Automated Approach (2015)



0.95 threshold 40-hr isopleth

0.5 threshold 40-hr fill

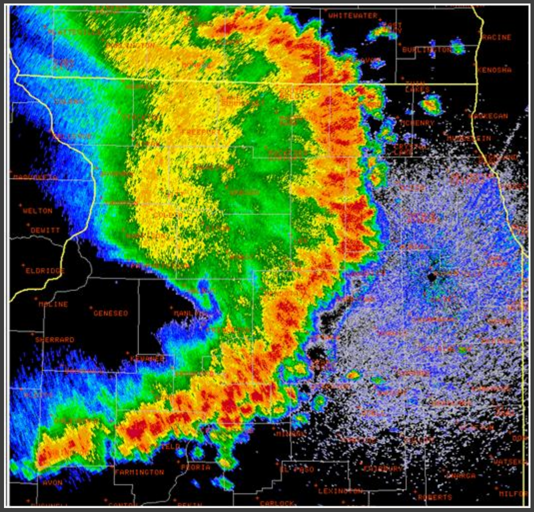
No threshold 40-hr isopleth

Quasi-linear convective systems

Quasi-linear convective systems (QLCSs) can produce severe weather (Trapp et al. 2005)

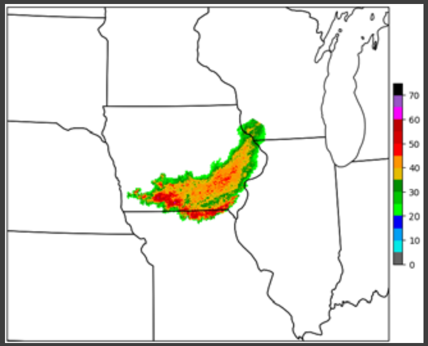
Convection-permitting models have trouble with QLCSs (Lawson and Gallus 2016)

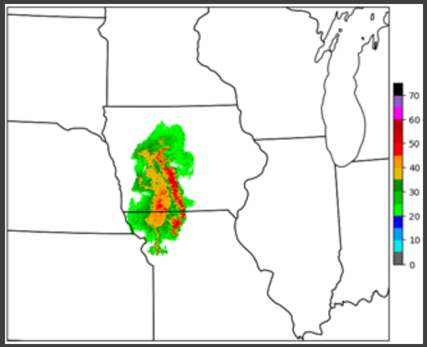
Implications for people, weather forecasting, and high resolution climate simulations

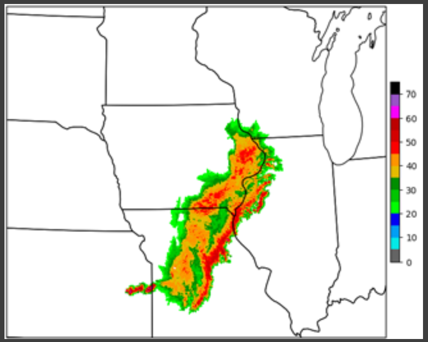


Visual differences

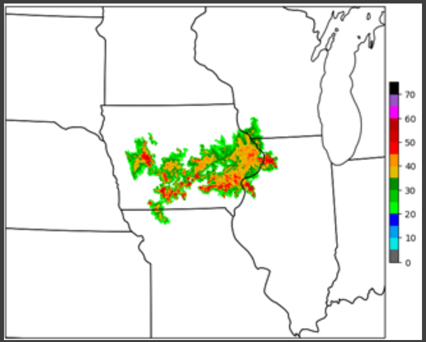
QLCS

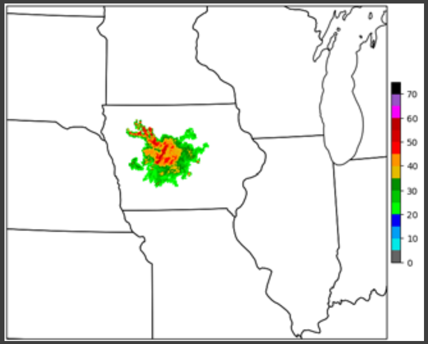


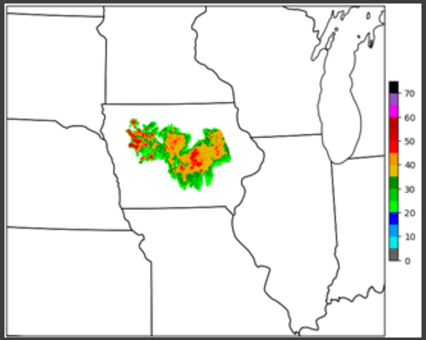




Non-QLCS







First Try

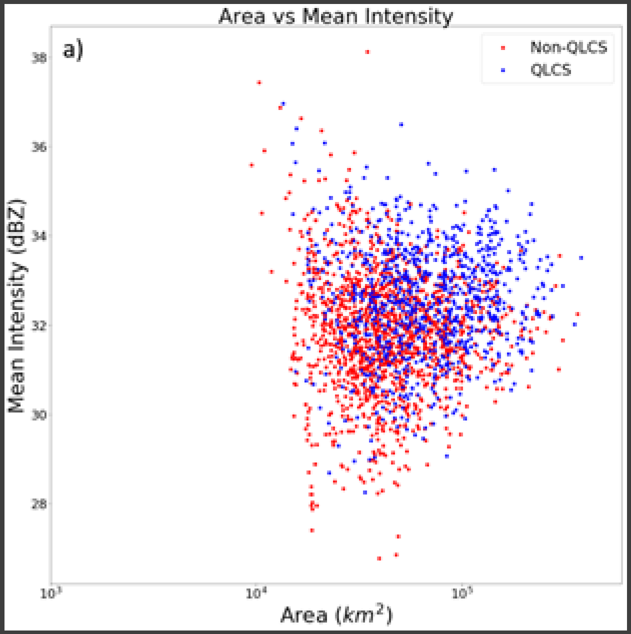
Select 3000 random high-probability MCS "snapshots"

Label as QLCS or Non-QLCS based only on their visual features

- Subjective
- Looking for common visual traits

Use features to train tree-based ensemble

  - ~70% accuracy

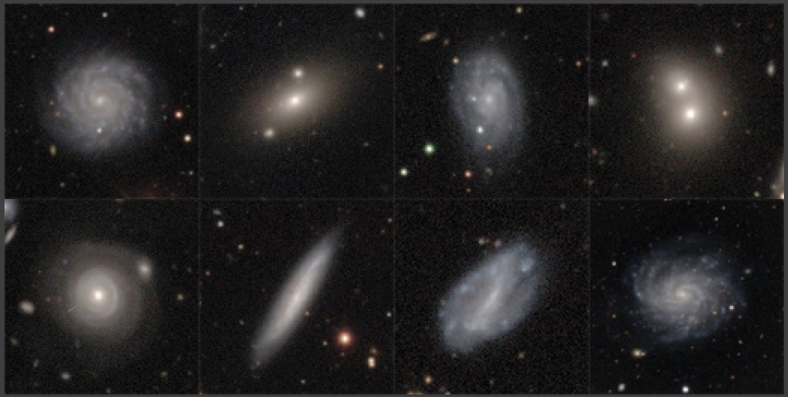


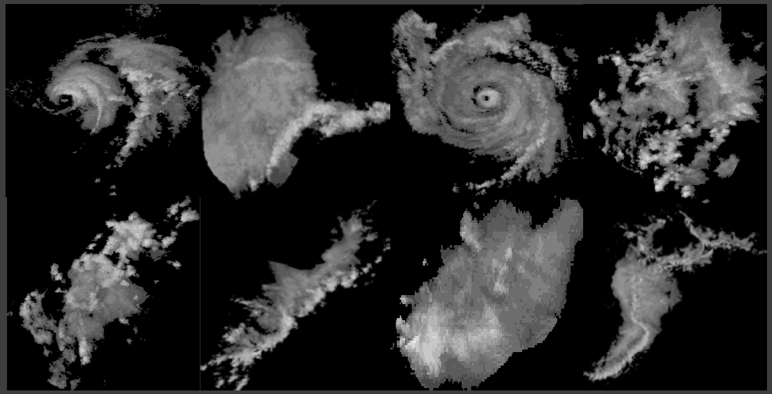
Second Try

Employ a convolutional neural network (Krizhevsky et al. 2012)

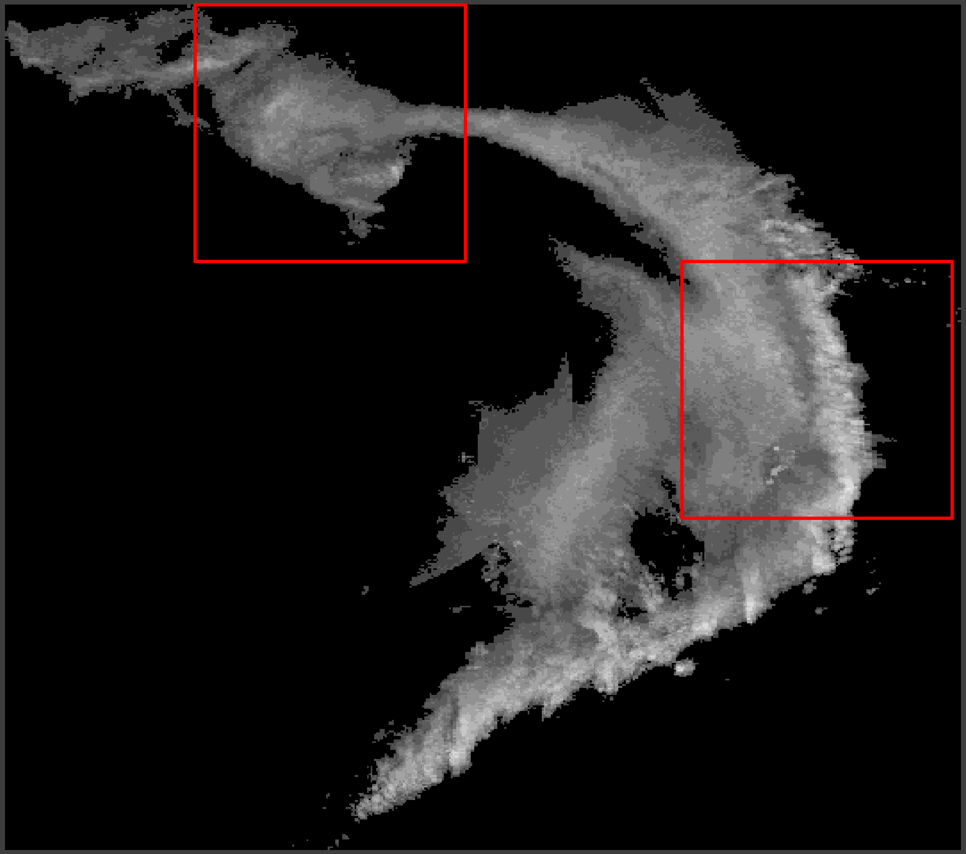
Inspiration / model configuration came from astronomy (Dieleman et al. 2015)

Much harder to generate training / testing data for CNNs



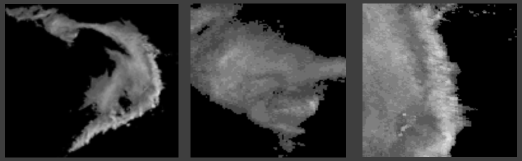


Training data approach



What is visually important for classifying this as a QLCS?

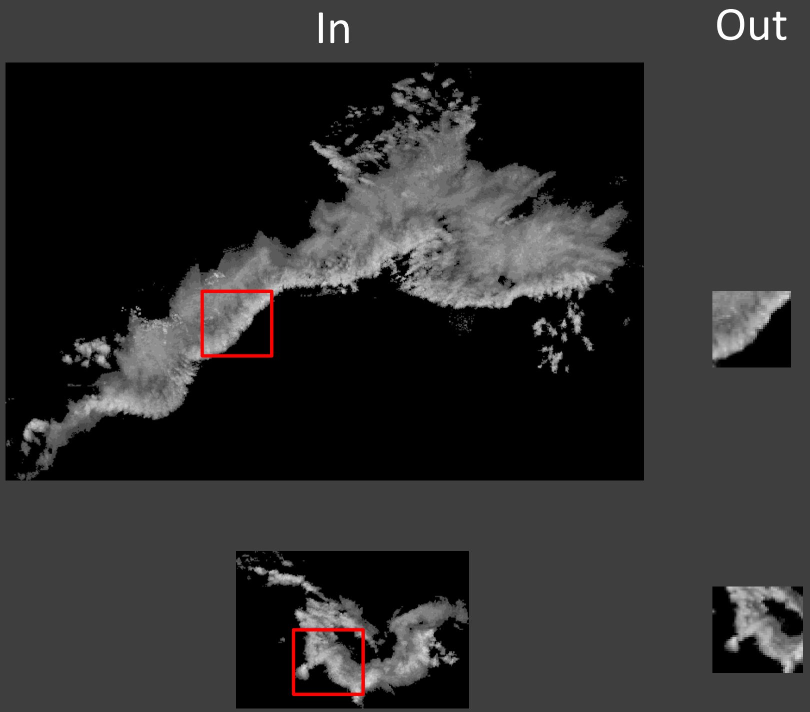
1) The entire structure?
2) Stratiform features?
3) Convection features?



"I'm looking for an intense line with a strong reflectivity gradient."

Training / testing data creation

- All images must be the same size
- Find largest contiguous region of 50+ dBZ
- Center a box on the intensity-weighted centroid
- Extract intensity information within box



Data Augmentation

Addressing overfitting

- Only ~3000 samples

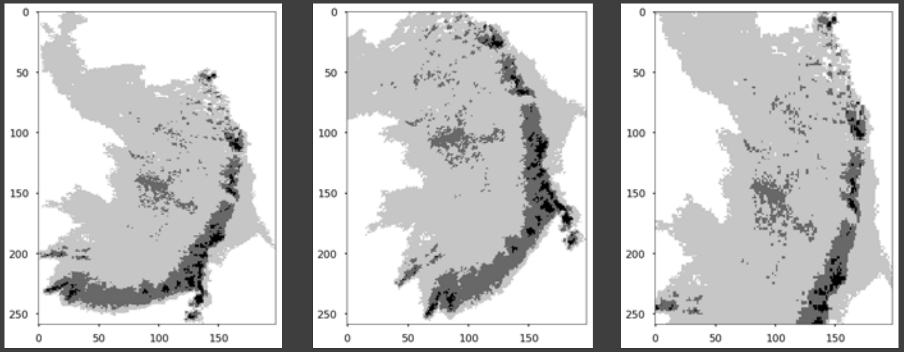
Keras ImageDataGenerator

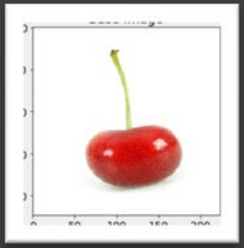
- Randomly apply slight modifications to images during training

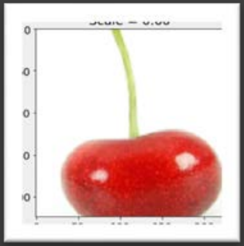
Physically Reasonable?

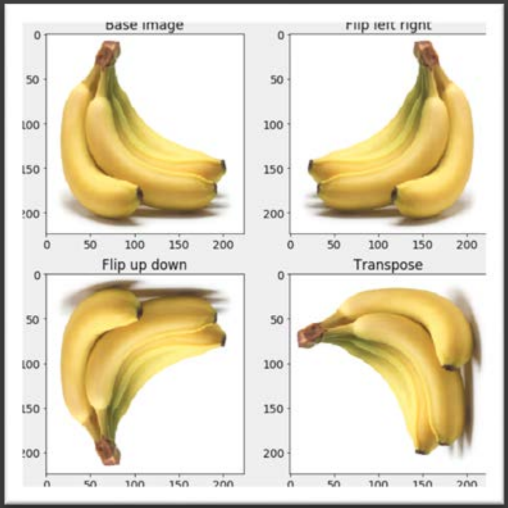
- Scale is important
- Orientation might be important

Original Rotation Zoom

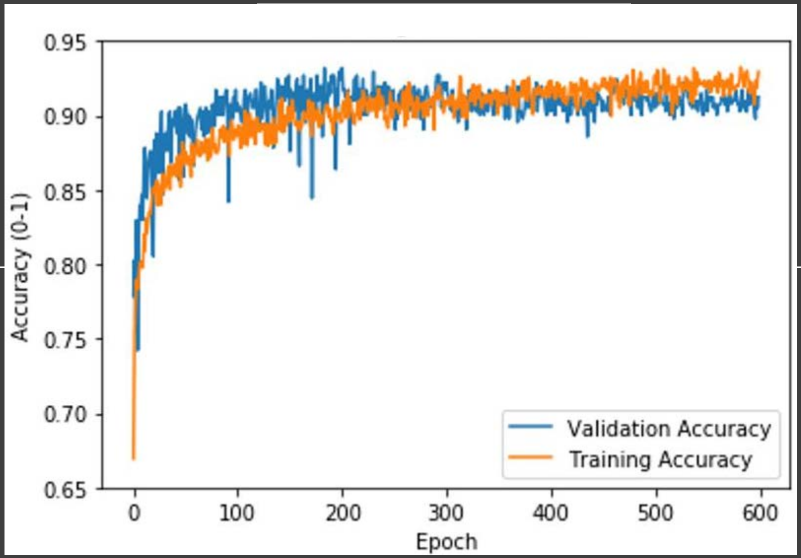




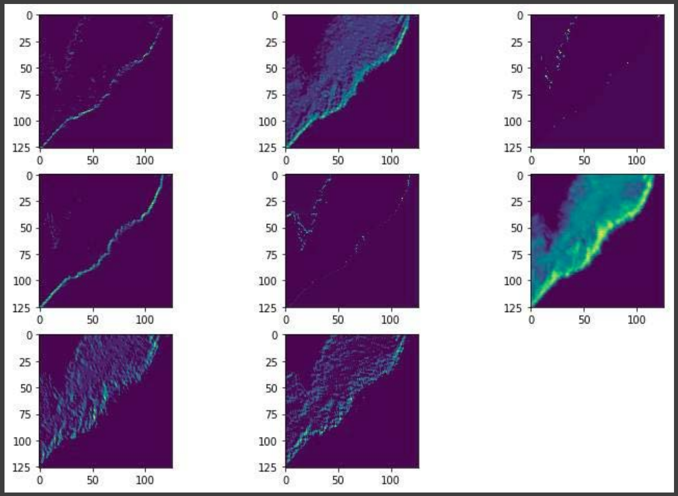




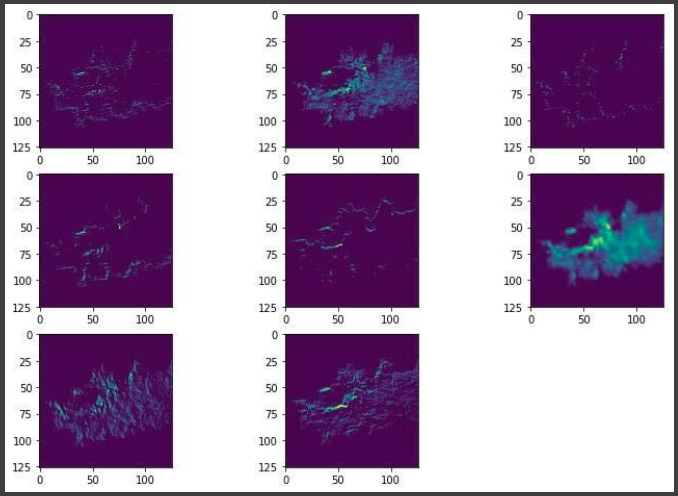
Fruit examples from : https://medium.com/ymedialabs-innovation/data-augmentation-techniques-in-cnn-using-tensorflow-371ae43d5be9

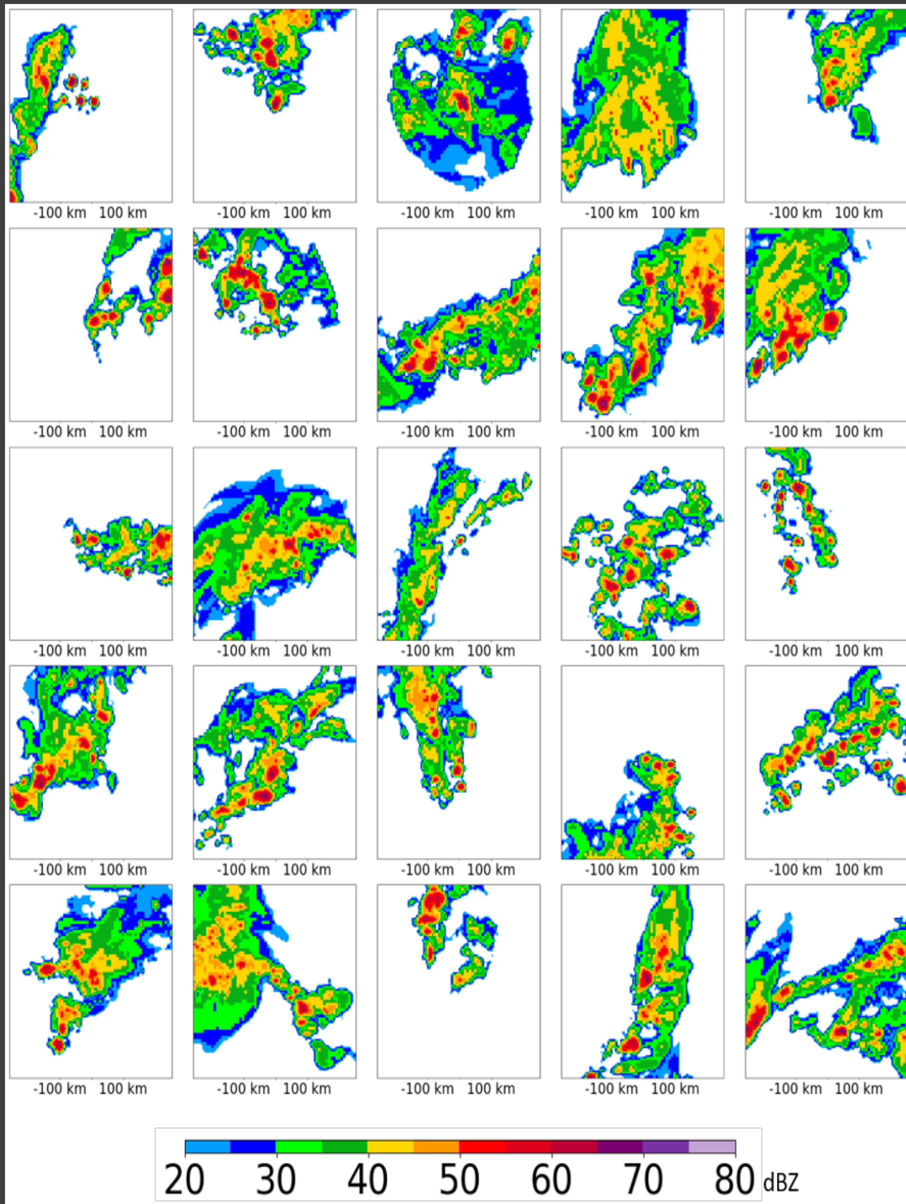


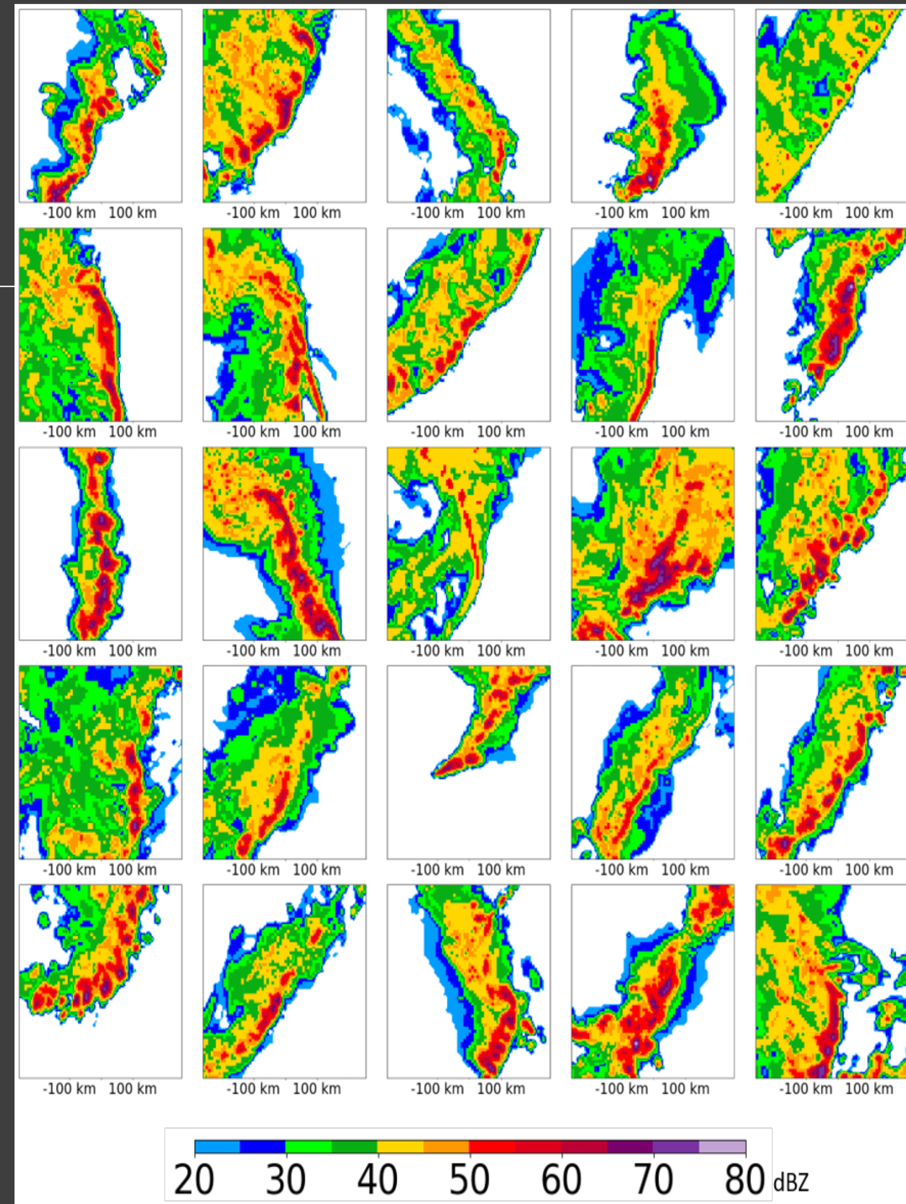
QLCS Sample



Non-QLCS Sample



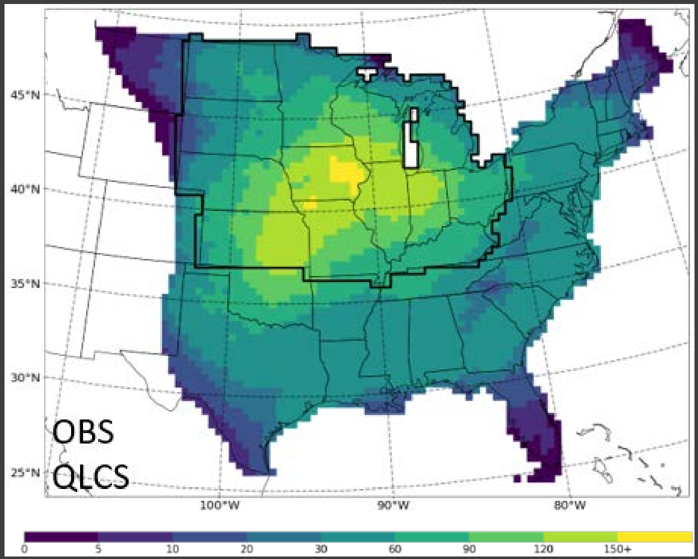

-100 km 100 km
20 30 40 50 60 70 80 dBZ


-100 km 100 km
20 30 40 50 60 70 80 dBZ

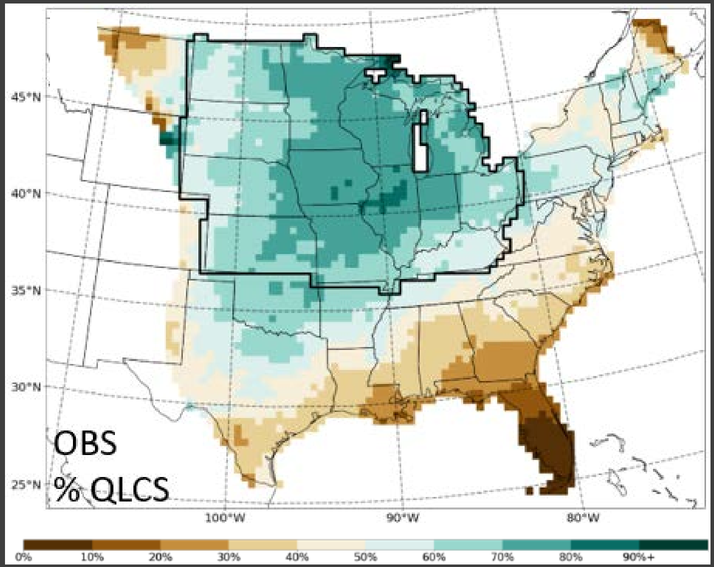
Notebook Examples

QLCS Detection

Application



QLCS Occurrence
June – August (2001-2013)



Percent of MCS events that were QLCSs
June – August (2001-2013)

References

HABERLIE, ALEX M., and WALKER S. ASHLEY. "A Method for Identifying Midlatitude Mesoscale Convective Systems in Radar Mosaics. Part I: Segmentation and Classification." *Journal of Applied Meteorology and Climatology* 2018 (2018).

McGovern, Amy, Kimberly L. Elmore, David John Gagne, Sue Ellen Haupt, Christopher D. Karstens, Ryan Lagerquist, Travis Smith, and John K. Williams. "Using Artificial Intelligence to Improve Real-Time Decision-Making for High-Impact Weather." *Bulletin of the American Meteorological Society* 98, no. 10 (2017): 2073-2090.

Lakshmanan, Valliappa, Travis Smith, Gregory Stumpf, and Kurt Hondl. "The warning decision support system–integrated information." *Weather and Forecasting* 22, no. 3 (2007): 596-612.

Han, Lei, Shengxue Fu, Lifeng Zhao, Yongguang Zheng, Hongqing Wang, and Yinjing Lin. "3D convective storm identification, tracking, and forecasting—An enhanced TITAN algorithm." *Journal of Atmospheric and Oceanic Technology* 26, no. 4 (2009): 719-732.

Houston, Adam L., Noah A. Lock, Jamie Lahowetz, Brian L. Barjenbruch, George Limpert, and Cody Oppermann. "Thunderstorm observation by radar (ThOR): An algorithm to develop a climatology of thunderstorms." *Journal of Atmospheric and Oceanic Technology* 32, no. 5 (2015): 961-981.

Gagne II, David John. "hagelslag Documentation." (2018).

Ahijevych, David, James O. Pinto, John K. Williams, and Matthias Steiner. "Probabilistic forecasts of mesoscale convective system initiation using the random forest data mining technique." *Weather and Forecasting* 31, no. 2 (2016): 581-599.

Herman, Gregory R., and Russ S. Schumacher. "Money Doesn't Grow on Trees, but Forecasts Do: Forecasting Extreme Precipitation with Random Forests." *Monthly Weather Review* 146, no. 5 (2018): 1571-1600.

Lagerquist, Ryan, Amy McGovern, and Travis Smith. "Machine Learning for Real-Time Prediction of Damaging Straight-Line Convective Wind." *Weather and Forecasting* 32, no. 6 (2017): 2175-2193.

Baldwin, Michael E., John S. Kain, and S. Lakshmivarahan. "Development of an automated classification procedure for rainfall systems." *Monthly weather review* 133, no. 4 (2005): 844-862.

Gagne, David John, Amy McGovern, and Jerry Brotzge. "Classification of convective areas using decision trees." *Journal of Atmospheric and Oceanic Technology* 26, no. 7 (2009): 1341-1353.

Lack, Steven A., and Neil I. Fox. "Development of an automated approach for identifying convective storm type using reflectivity-derived and near-storm environment data." *Atmospheric research* 116 (2012): 67-81.

Hobson, Angelyn G. Kolodziej, Valliappa Lakshmanan, Travis M. Smith, and Michael Richman. "An automated technique to categorize storm type from radar and near-storm environment data." *Atmospheric research* 111 (2012): 104-113.

Lawson, John, and William A. Gallus Jr. "On contrasting ensemble simulations of two Great Plains bow echoes." *Weather and Forecasting* 31, no. 3 (2016): 787-810.

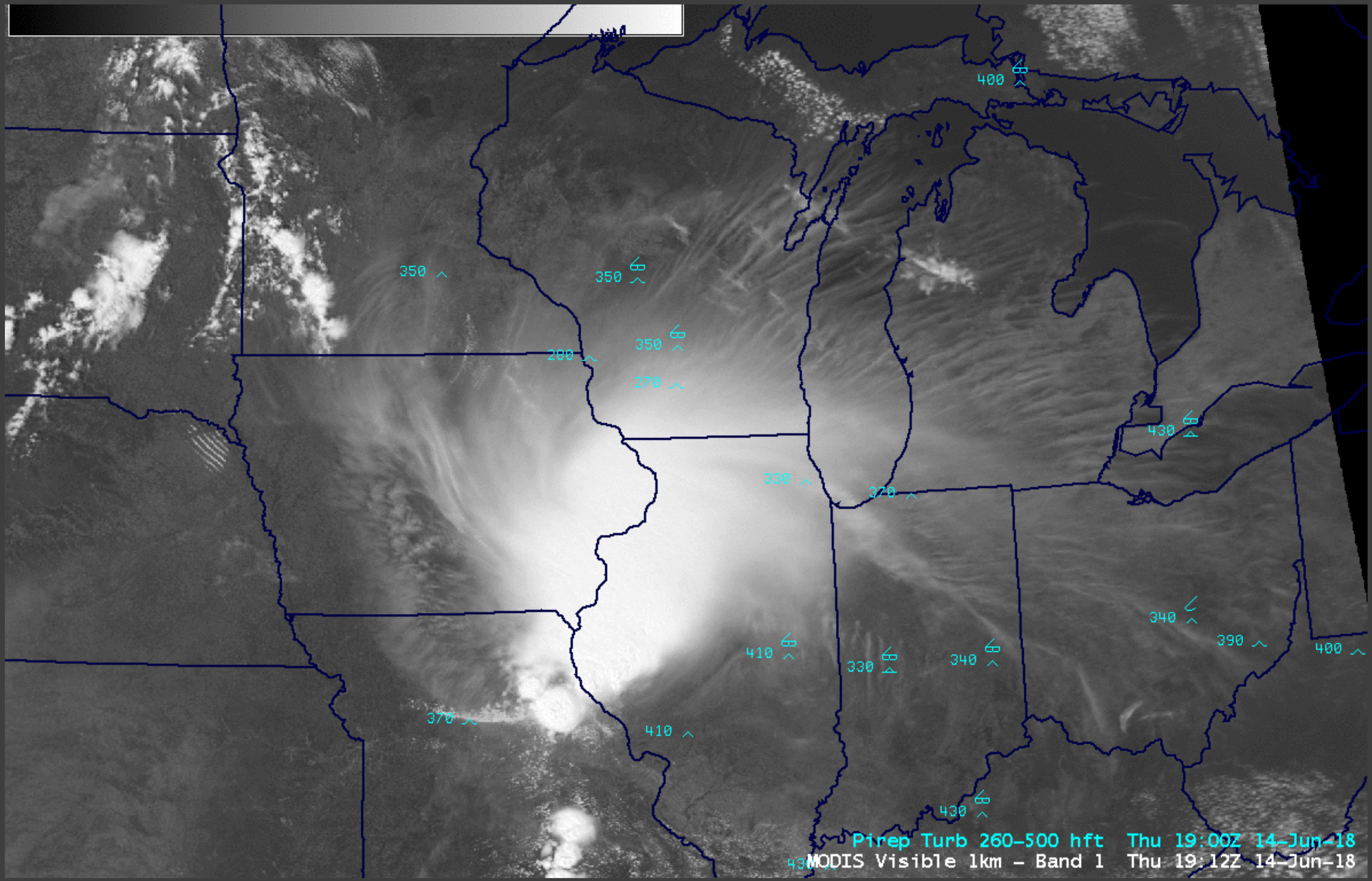
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105. 2012.

Dieleman, Sander, Kyle W. Willett, and Joni Dambre. "Rotation-invariant convolutional neural networks for galaxy morphology prediction." *Monthly notices of the royal astronomical society* 450, no. 2 (2015): 1441-1459.

Gallus Jr, William A., Nathan A. Snook, and Elise V. Johnson. "Spring and summer severe weather reports over the Midwest as a function of convective mode: A preliminary study." *Weather and Forecasting* 23, no. 1 (2008): 101-113.

Parker, Matthew D., and Richard H. Johnson. "Organizational modes of midlatitude mesoscale convective systems." *Monthly weather review* 128, no. 10 (2000): 3413-3436.

Thank You



Pirep Turb 260-500 hft Thu 19:00Z 14-Jun-18
MODIS Visible 1km - Band 1 Thu 19:12Z 14-Jun-18