A Comparison of Machine Learning Techniques for Convective Morphology Classification from Radar Imagery

Jonathan E. Thielen[1], William A. Gallus, Jr.[1], and Alex M. Haberlie[2]

[1]Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA
[2]Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA

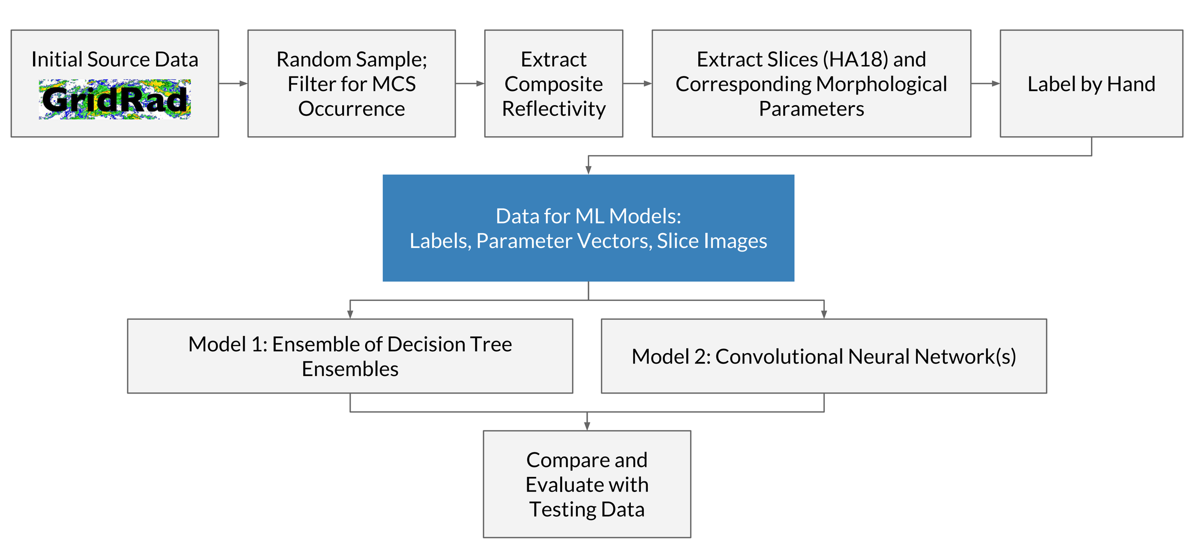
Iowa State University of Science and Technology

Background

Convective morphology (shape, structure, and organization) is an important characteristic in the analysis and forecasting of convective systems.

However, most past studies of convective morphology have relied upon manual methods, which are tedious and potentially inconsistent. Recent work by Haberlie and Ashley (2018) found success in applying machine learning techniques in classifying mesoscale convective systems (MCSs) vs. non-MCSs. This study extends the methodology of Haberlie and Ashley (2018) to the detailed, nine-category scheme of Gallus et al. (2008).

Methodology

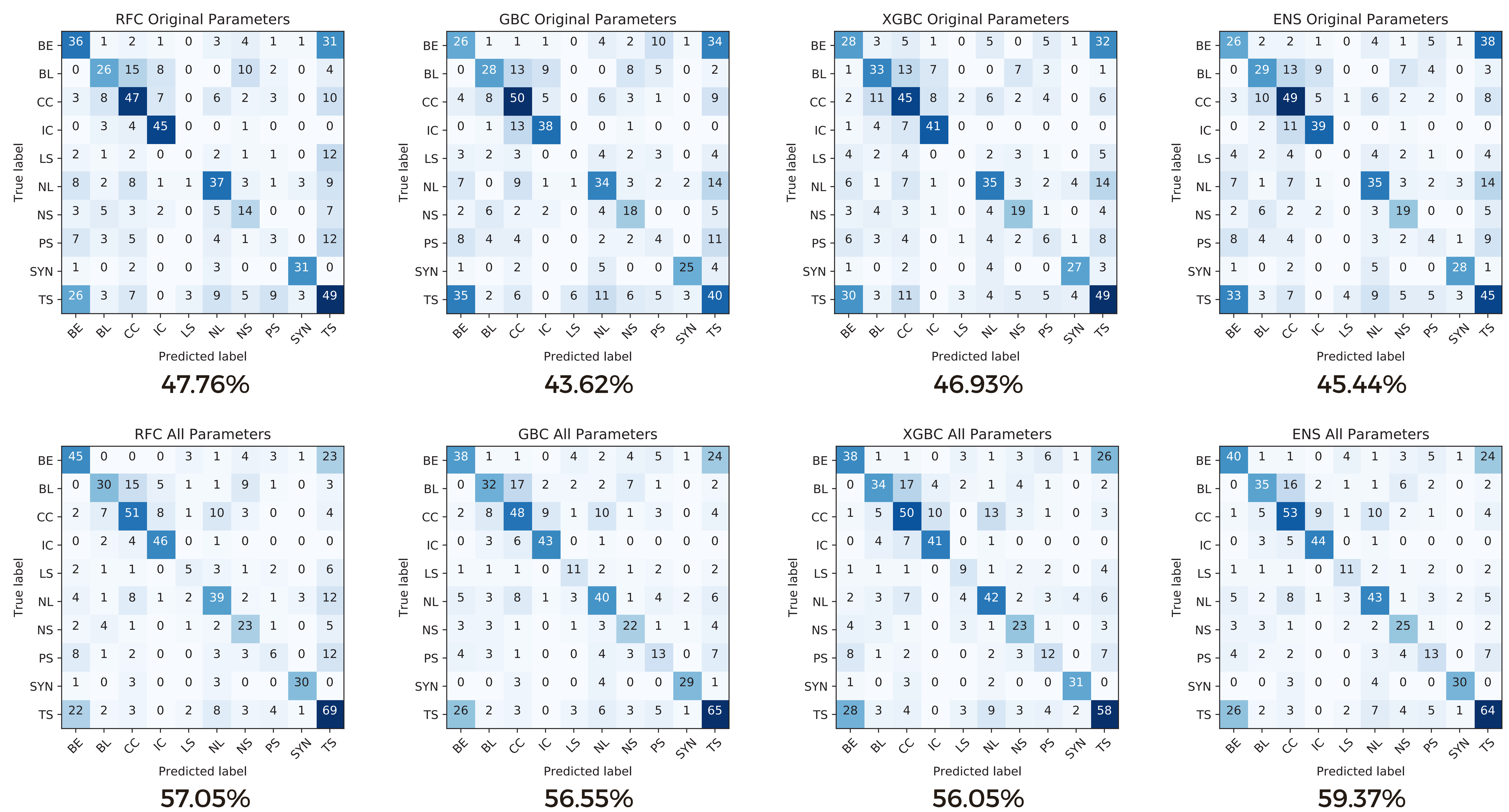


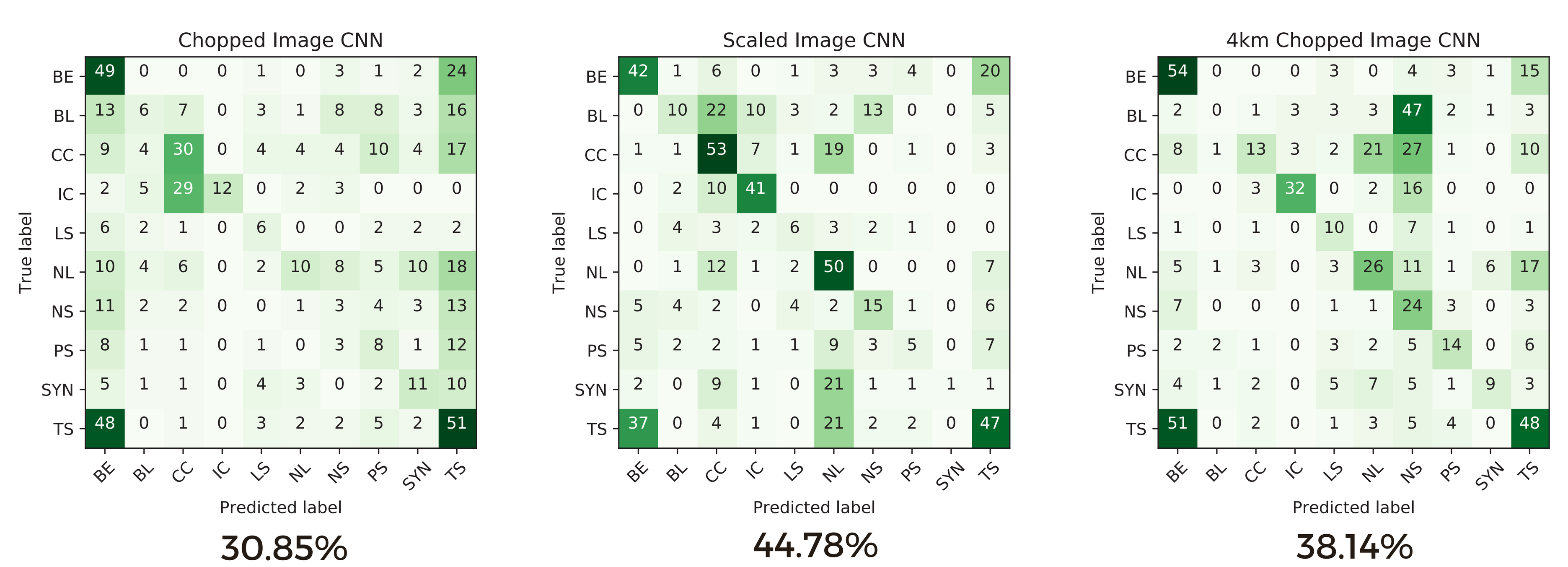
Two general machine learning approaches:

1) Ensembles of decision trees
  - As in Haberlie and Ashley (2018)
  - Algorithms:
    - Random forest (RFC)
    - Gradient boosted trees (GBC)
    - XGBoost (XGBC)
    - Ensemble of above (ENS)
  - Tested using original set of 14 areal and intensity features and expanded set with 24 added scheme-specific features

2) Convolutional Neural Networks (CNNs)
  - 128 by 128 pixel, single channel input
  - Three image extraction processes: scaling, cropping, and cropping after upscaling to 4 km

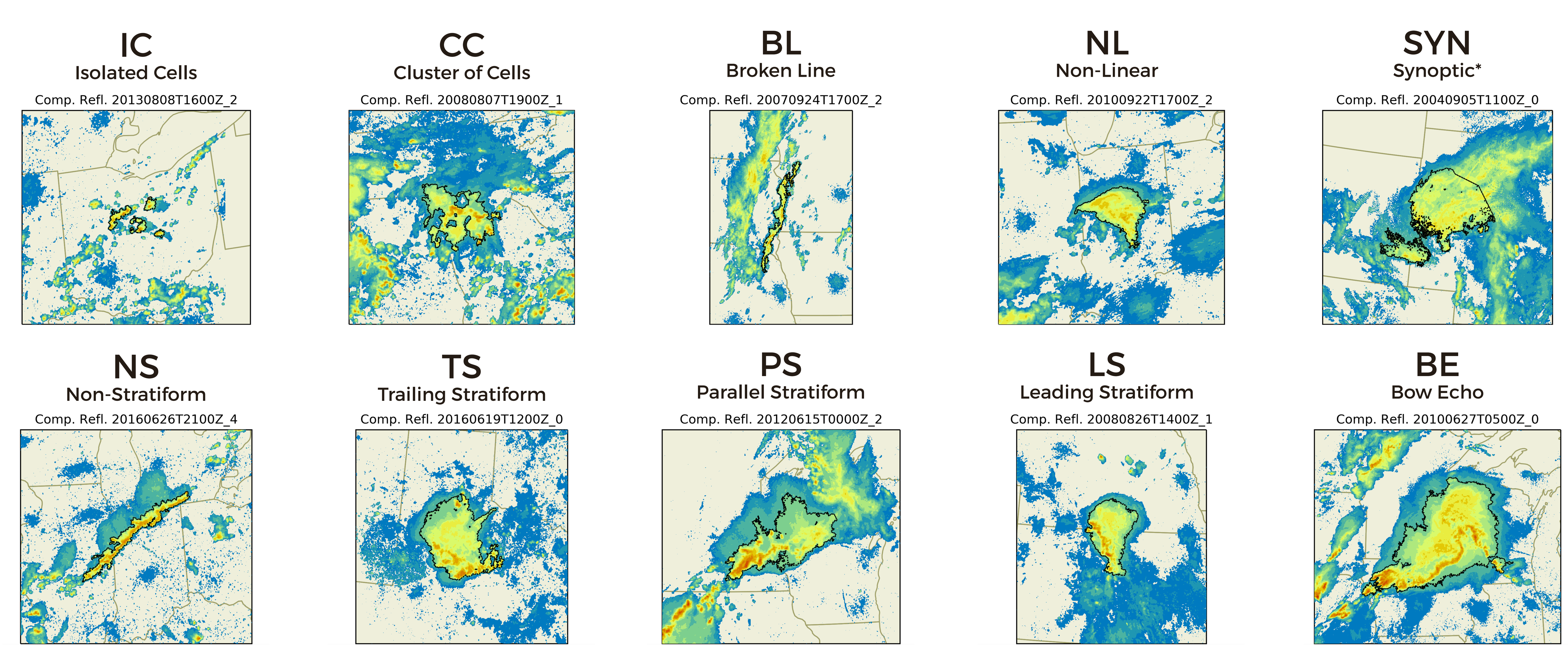
3,000 hand-labeled cases with composite reflectivity data from GridRad (regridded to 2 km Lambert Conformal)
Training data from 2007-2016 (2397 cases)
Testing data from 2004-2006 (603 cases)





Figures (Above): Confusion matrices and accuracy scores for the 11 primary classifiers tested in this study (8 decision tree ensemble classifiers, in blue, and 3 CNNs, in green)

Figures (Below): System examples of highest probability of classification from best performing classifier (ENS All Parameters)



Conclusions

Ensemble of decision tree ensemble classifiers (with full parameter/feature set) performed best
- Still lacking in discriminating:
  - Amount of bowing
  - CC vs. NL
  - Between cellular modes
- Shows promise, but insufficient for reliable use in future studies

CNNs relatively poor performing
- Model complexity and robustness constrained by computational limits
- Sample size likely too small (even with data augmentation procedures)
- Small image sizes may be preventing depiction of details

Future work will
- Utilize larger samples
- Explore more robust/complex CNNs
- Investigate improvements in morphology feature extraction
- Examine probabilistic classifications
- Investigate applications in studying severe weather hazards

References:

Gallus, W. A., N. A. Snook, and E. V. Johnson, 2008: Spring and Summer Severe Weather Reports over the Midwest as a Function of Convective Mode: A Preliminary Study. Wea. Forecasting, 23, 101-113, https://doi.org/10.1175/2007WAF2006120.1.

Haberlie, A. M. and W. S. Ashley, 2018: A Method for Identifying Midlatitude Mesoscale Convective Systems in Radar Mosaics. Part I: Segmentation and Classification. J. Appl. Meteor. Climatol., 57, 1575-1598, https://doi.org/10.1175/JAMC-D-17-0293.1.

Interested in learning more about this ongoing project? Find it on GitHub:

github.com/jthielen/ml_convective_morphology