

# ARMING THE EDGE

Using AI and Edge computing to improve cloud particle data from Lidar

Robert Jackson<sup>1</sup>, Scott Collis<sup>1</sup>, Pete Beckman<sup>1</sup>, Nicola Ferrier<sup>1</sup>, Raj Sankaran<sup>1</sup>  
 1. Argonne National Laboratory, Argonne, IL, USA

## WHAT IS EDGE COMPUTING?

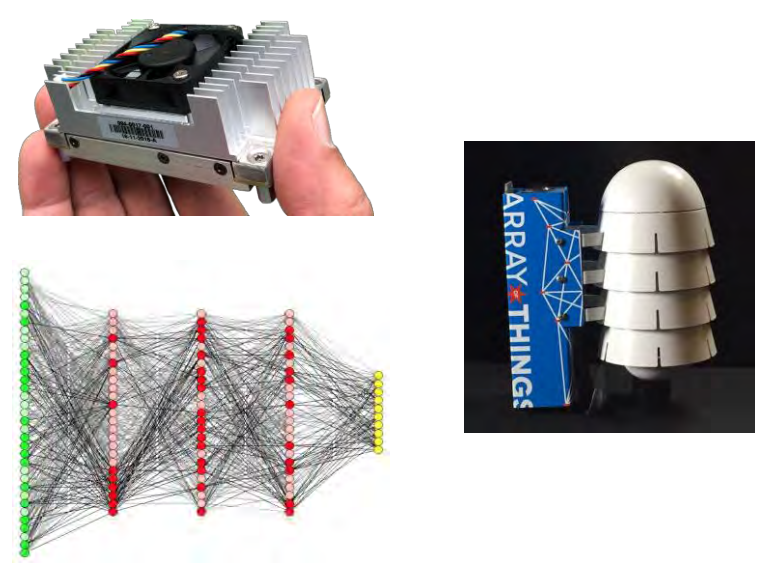
"A supercomputer is a device for turning compute-bound problems into I/O-bound problems."  
 (Ken Batcher)

"An edge computer is a device for turning I/O-bound problems onto compute-bound problems."



Larry Smarr at NCSA, 1986

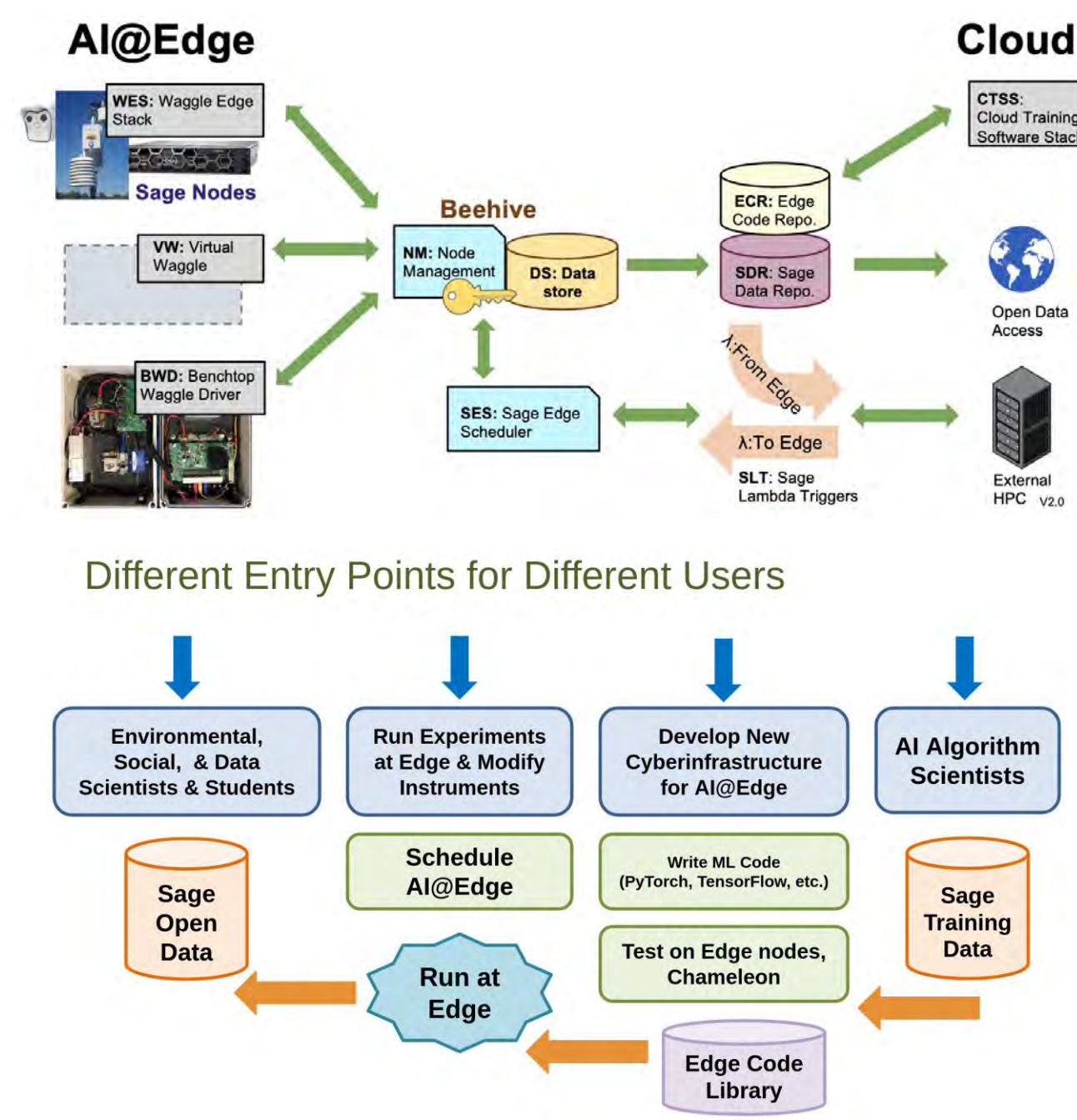
AI@Edge  
 Parallel Computing



- Artificial Intelligence
- Deep Learning Inference
- Lightweight Edge Learning

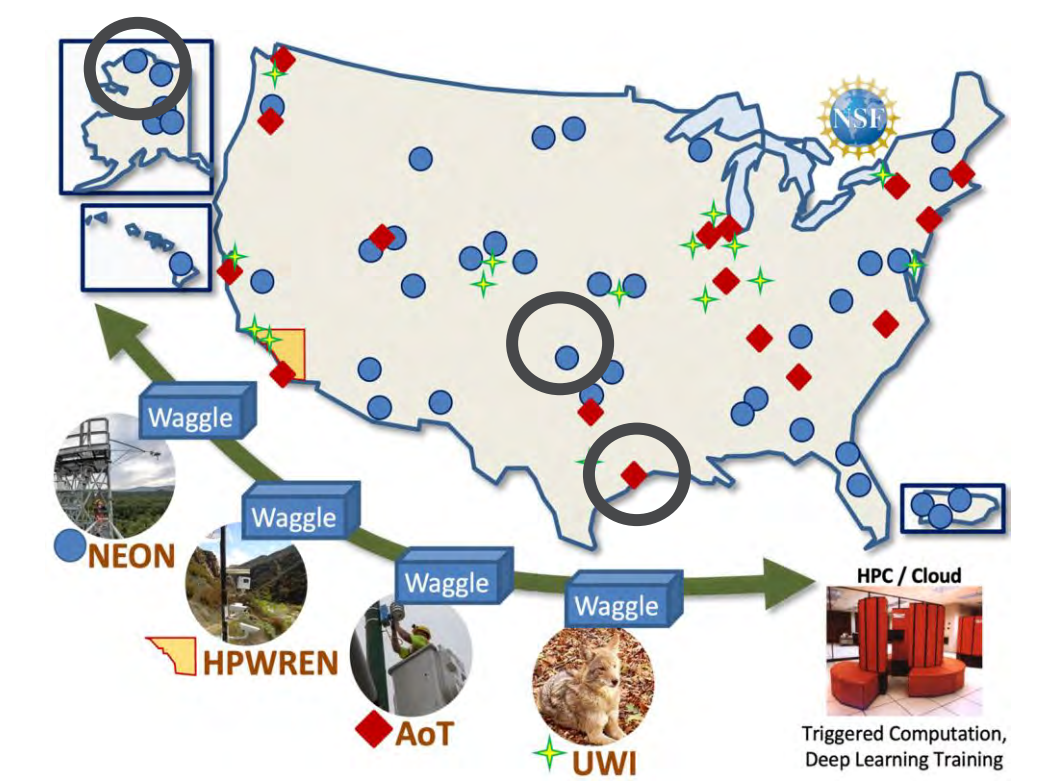
## WAGGLE PLATFORM AND SAGE

An Open Platform for AI@Edge Computing and Intelligent Sensors

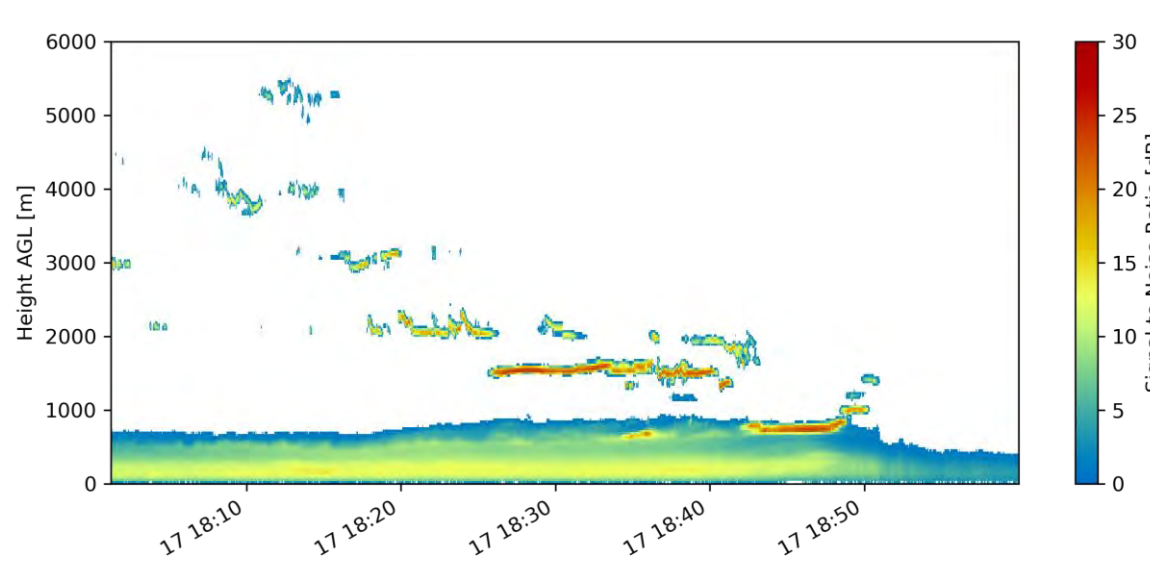


**SAGE** [sagecontinuum.org](http://sagecontinuum.org)  
 Cyberinfrastructure for AI at the Edge

ARMing The Edge → Implement Sage @ 3 U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) sites

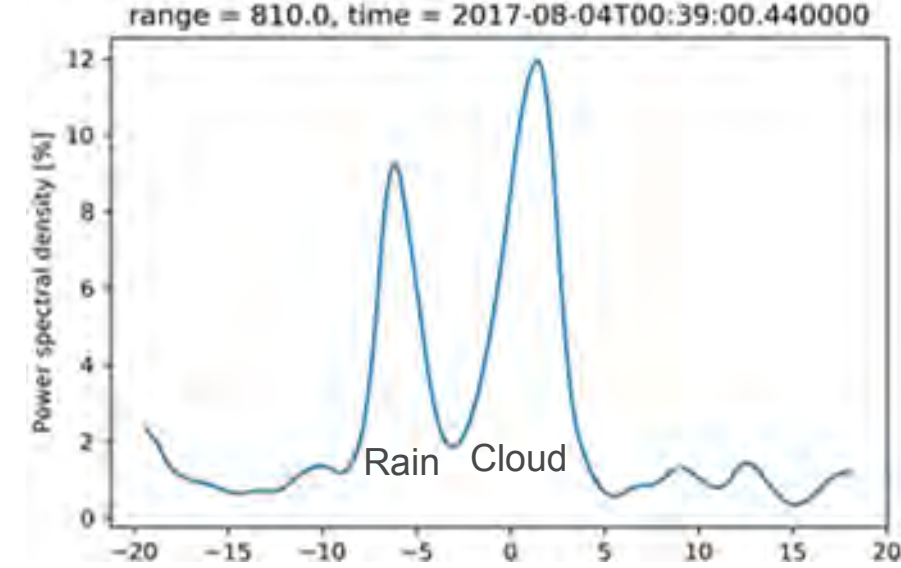


## USE EXAMPLE: DOPPLER LIDAR



U.S. DOE ARM Southern Great Plains facility in Lamont, OK.

- Decades of data from numerous meteorological/aerosol sensors
- Years of doppler lidar images of clouds, rain, and clear air

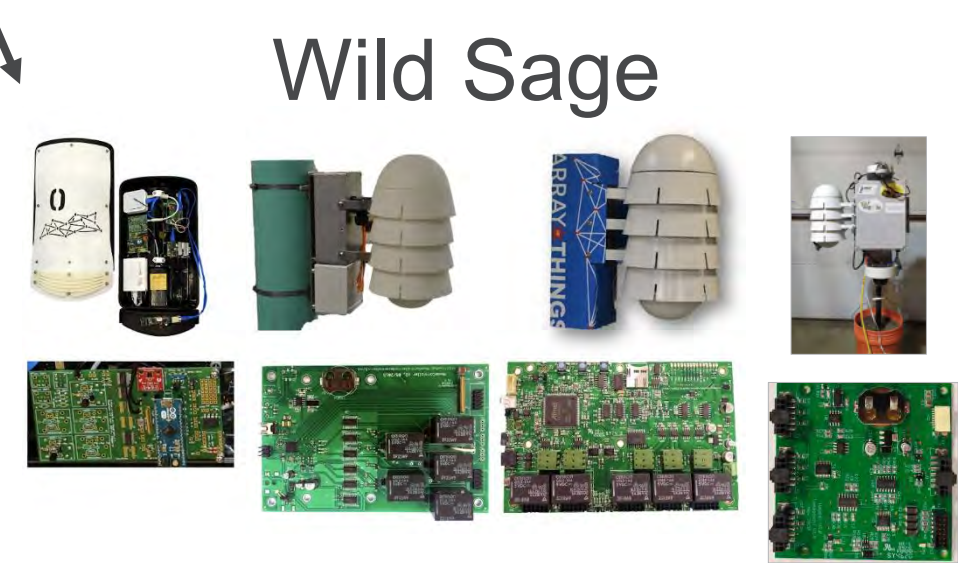


Doppler lidar records terabytes/month of cloud particle data:

- Particle fall speed spectra (above)
- Information about rain + cloud drops

Due to bandwidth constraints, vital information about cloud and rain drops currently discarded.

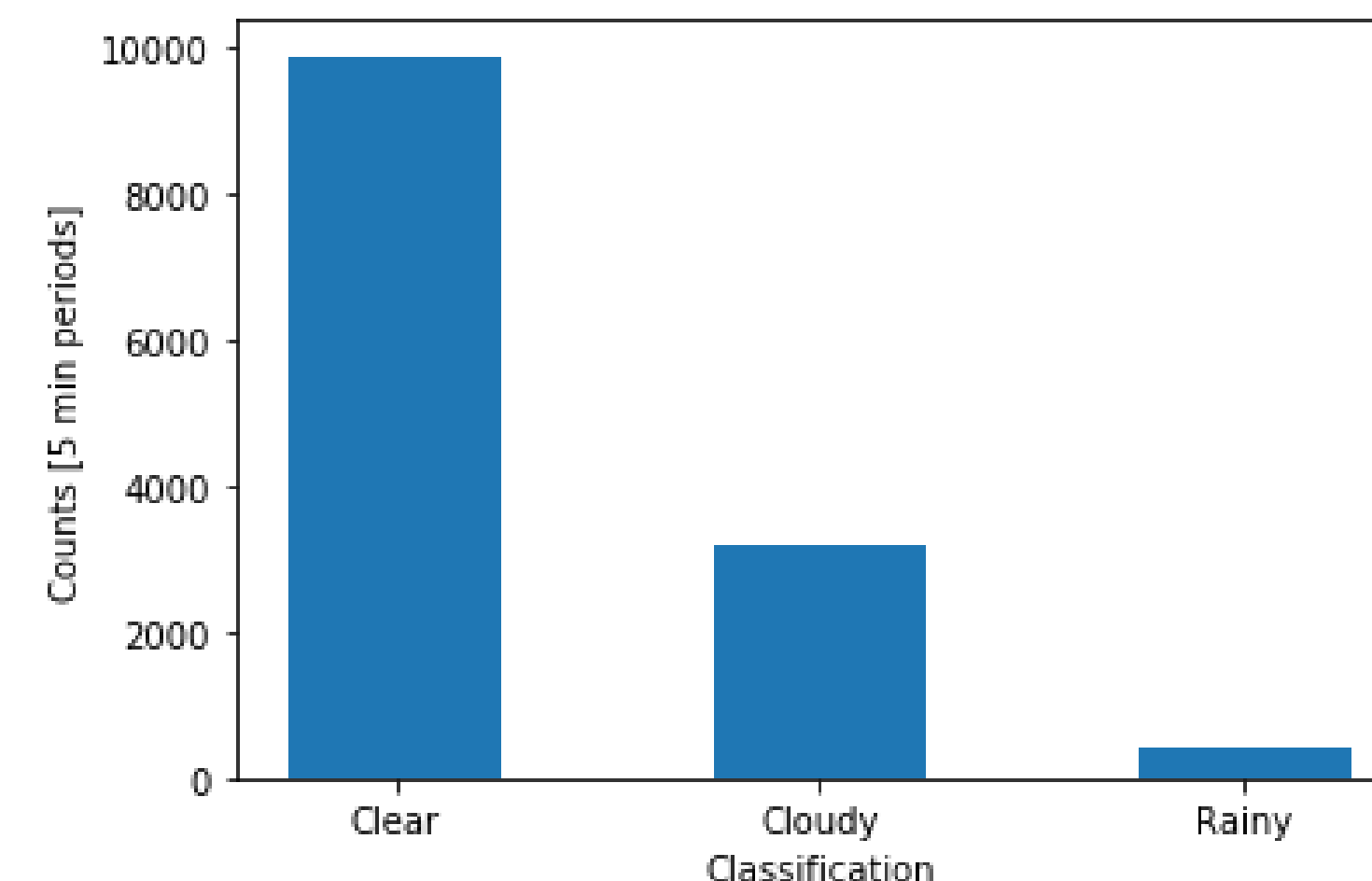
GPU-based software HighIQ developed/tested for Doppler Lidar processing @ Edge



Can we utilize Array of Things + AI to identify time periods when detailed cloud and rain drop information needed?

## ML-BASED IDENTIFICATION FOR EDGE

- Algorithm 1: XGBoost. Early stopping + 5-fold cross-validation: 1300 trees,
- Train on % of 5 min period w/ SNR > 3, 5 dB @ 200 m height intervals.
- Classify into clear/cloudy/rainy.

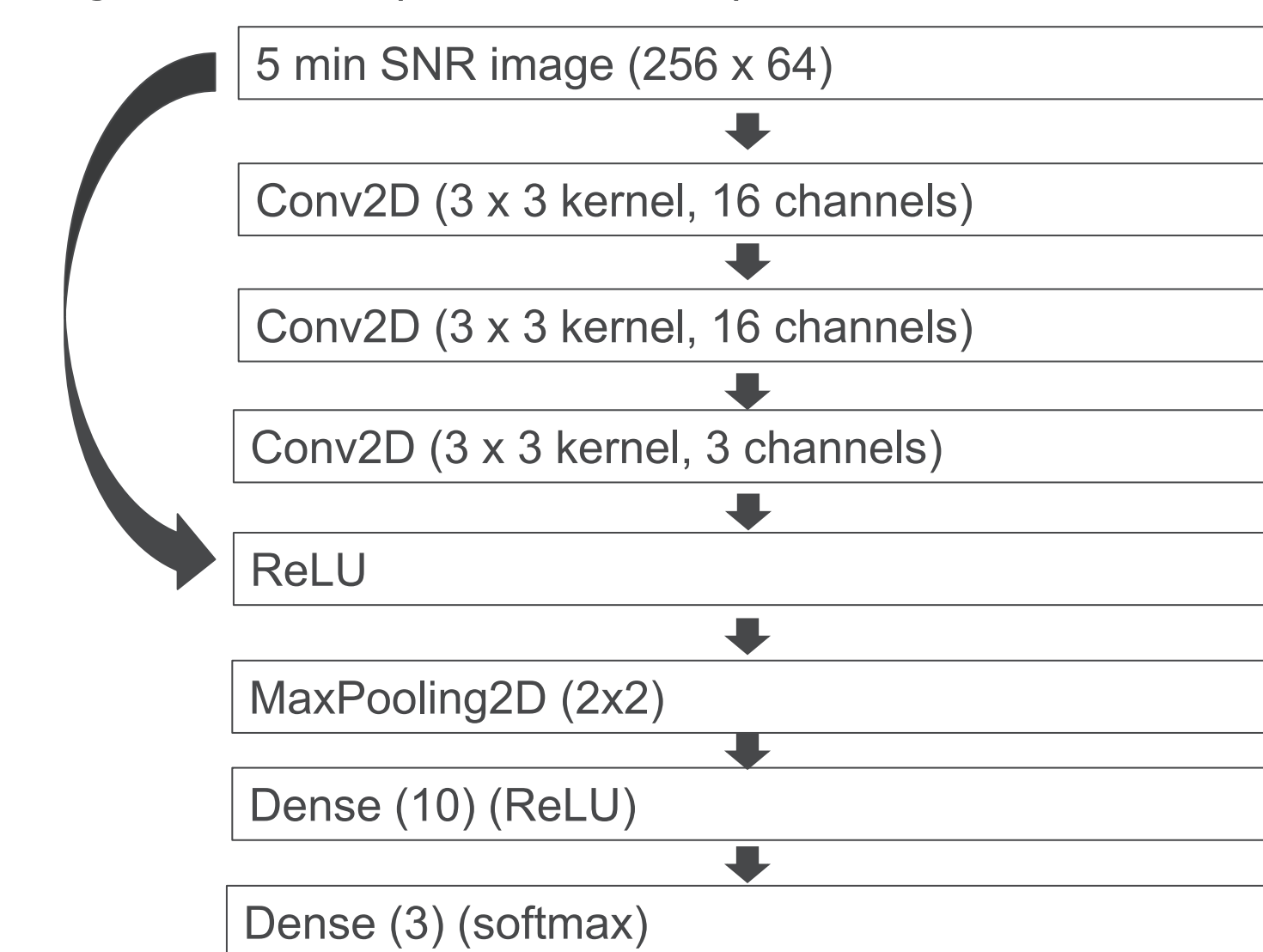


20% clouds/rain

	Clear	Cloudy	Rain
Clear	9745	310	166
Cloudy	130	2898	46
Rain	14	4	226

94% accuracy

Algorithm 2: (Keras CNN)



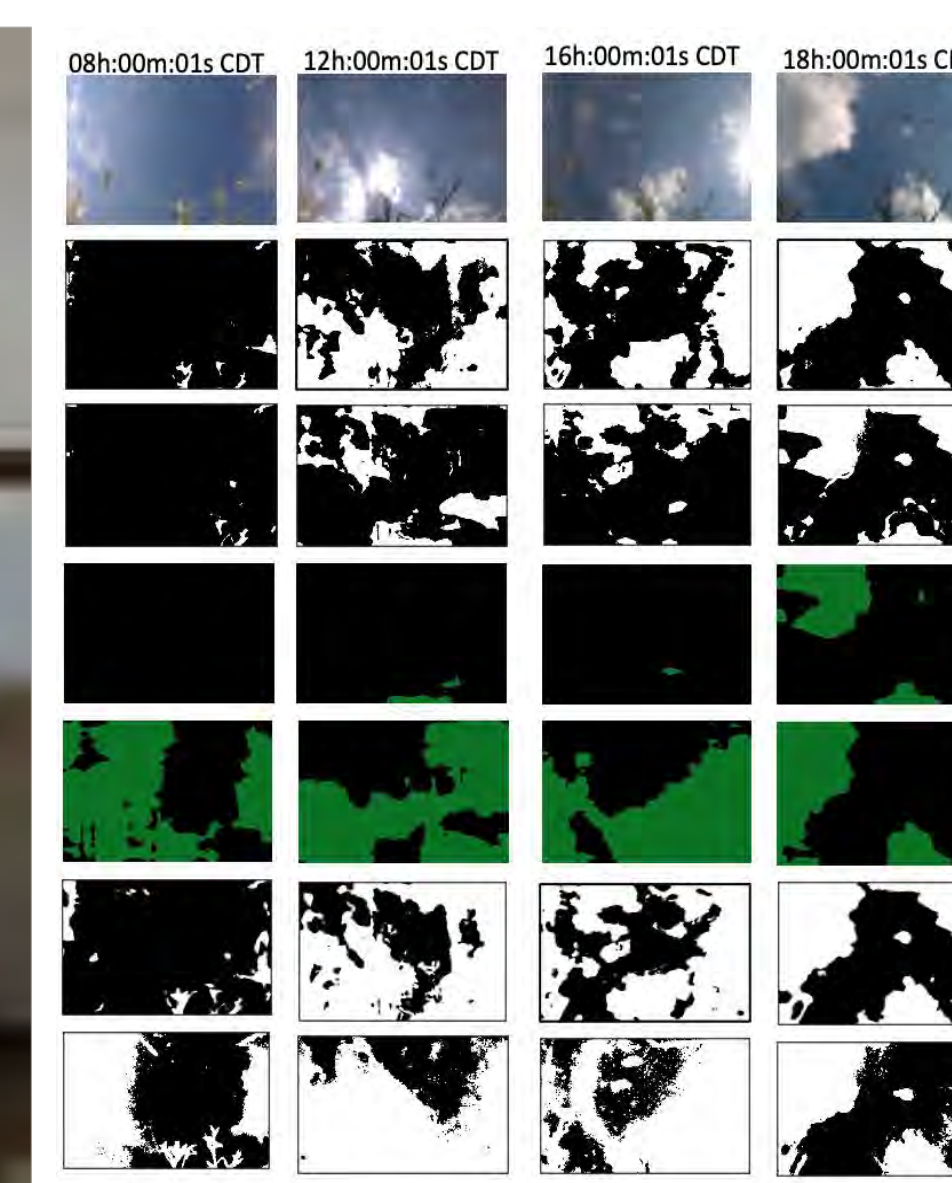
	Clear	Cloudy	Rain
Clear	7207	2497	162
Cloudy	2345	818	55
Rain	709	244	17

88% accuracy

## CONCLUSIONS

- Current ARM facility infrastructure does not permit transfer of detailed cloud microphysical data from Doppler Lidar. Much information about cloud/drizzle properties lost in simple transfer → hinders cloud process studies
- Software developed/tested for GPU edge computing processing of cloud microphysics
- XGBoost applied to signal-to-noise ratio guides when more detailed processing needed (~20% of time) and outperforms CNNs
- Machine learning algorithms available at: [https://github.com/rcjackson/arming\\_the\\_edge](https://github.com/rcjackson/arming_the_edge)
- HighIQ software available at: <https://github.com/openradar/HighIQ>

## FUTURE STEPS



- WAGGLE node with pan/tilt camera ready for SGP deployment

- Can we determine cloud fraction from total sky imager to develop cloud climatology?