Forecasting heavy rainfall in the Southeastern US using deep learning

Suhaib Mansour[1], Zijin Ye[2], Maan Adam[3], Yan Pan[4]

[1] Duke University, Durham, NC; [2]Carnegie Mellon University, Pittsburg, PA;

[3] Kaust University, Thuwal, Saudi Arabia; [4] Nicholas School of the Environment, Duke University, Durham, NC

Project Mentor: Yan Pan

Abstract

Heavy rainfall and flash flooding pose persistent risks across the Southeastern United States, driven by frequent convective storms and strong spatial–temporal variability in precipitation. We develop a deep learning–based nowcasting system that leverages multi-sensor radar imagery to forecast short-term heavy rainfall. Using FURENet—an improved U-Net–style architecture that captures spatial and temporal patterns—we generate 40-minute reflectivity forecasts conditioned on MRMS radar inputs. Initial experiments demonstrate promising skill, with competitive hit rates and Heidke Skill Scores at operationally relevant reflectivity thresholds, while highlighting degradation beyond one-hour lead times. We summarize our objectives, data and modeling pipeline, preliminary results, key challenges (compute/storage constraints and longer-lead accuracy), and a roadmap for next steps (expanded data, additional polarimetric inputs such as KDP, and model/learning improvements).

Introduction

The Southeast experiences high precipitation variability and frequent convective storms that are difficult to forecast precisely at short lead times. Traditional methods often struggle to capture rapid convective development and evolution. Recent progress in deep learning offers a complementary path by learning complex spatial–temporal patterns directly from radar observations. Our project explores a cost-efficient and generalizable approach to near-term precipitation nowcasting (≈40 minutes) using MRMS radar data and a specialized neural architecture, FURENet, designed to fuse spatial features with temporal evolution for heavy-rainfall events.

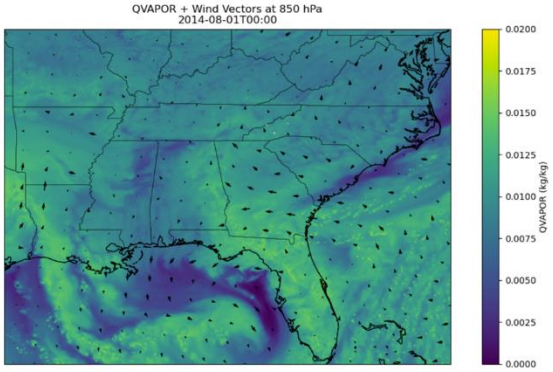
Project Objectives and Goals

1. **Improve short-term heavy rainfall nowcasting** (≤ 1 hour) over the Southeastern U.S. using deep learning on radar imagery.
2. **Enhance flash flood risk prediction** by accurately capturing the spatial distribution and intensity evolution of convective systems.
3. **Quantify skill at operational thresholds** via metrics such as Probability of Detection (POD), False Alarm Rate (FAR), Heidke Skill Score (HSS), and an overall evaluation score.

Methodology and Approach

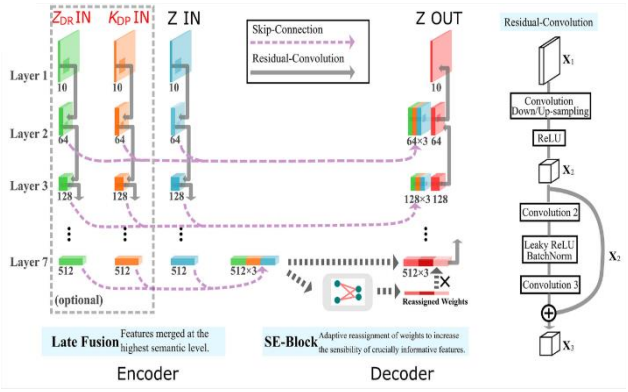
Dataset and Preprocessing

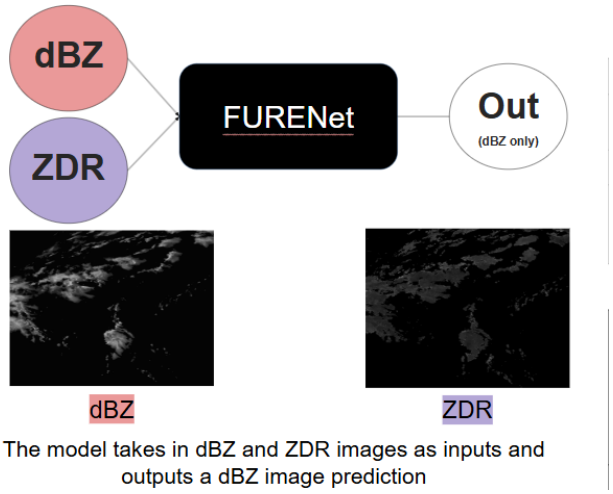
- **Source:** Multi-Radar Multi-Sensor (MRMS) product suite.
- **Inputs:** Radar reflectivity **dBZ** and polarimetric **ZDR** (drop-size distribution proxy).
- **Selection:** Heavy rainfall days identified using NOAA CPC resources.
- **Temporal resolution:** dBZ at ~2-minute cadence; ZDR at ~5-minute cadence.
- **Preprocessing:** Radar frames converted into normalized PNG images; curated event-based subsets for heavy-rain training vs. broader daily sets for comparison.



Model Architecture (FURENet)

- U-Net-style encoder–decoder augmented to better capture **spatial + temporal** patterns of intense precipitation.
- Designed for **nowcasting** by learning from high-impact rainfall days; outputs future reflectivity fields (dBZ) at a specified lead time (here, 40 minutes).

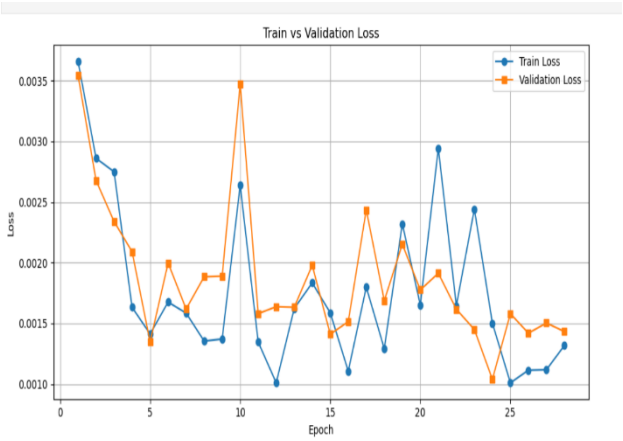




The model takes in dBZ and ZDR images as inputs and outputs a dBZ image prediction

Training Setup

- Supervised learning on sequences of radar frames (dBZ, with/without ZDR) to forecast dBZ at +40 minutes.
- Baseline loss: mean squared error on predicted vs. target reflectivity fields; monitored validation loss for early stopping.



Training (blue) and **validation** (orange) loss across 28 epochs. Both curves drop sharply through epoch 5, then plateau with minor fluctuations. Early stopping at epoch 25, where validation loss bottoms out, helps prevent overfitting, yielding a final validation loss of approximately 0.0010 and demonstrating strong generalization.

Evaluation Protocol

- Threshold-based verification at operational reflectivity cutoffs (e.g., 15 and 35 dBZ) to compute:
    - **POD (hit rate)** — fraction of observed events correctly predicted.
    - **FAR** — fraction of predicted events that did not occur.
    - **HSS** — skill over random chance, balancing hits, misses, and false alarms.
- Aggregate **evaluation score** to compare input/selection strategies (e.g., heavy-rain-only vs. all-days; with/without ZDR).

Current Progress and Milestones Achieved

- Assembled a **multi-sensor radar dataset** (dBZ, ZDR) for the Southeast with curated heavy-rain days.
- Implemented the **FURENet** architecture and trained multiple ablations:
    - Heavy-rain days with 3 months of dBZ only.
    - All days (3 months) with dBZ only.
    - Heavy-rain days (3 months) with **dBZ + ZDR** inputs.
- Established a **40-minute nowcasting** target; trained for **~28 epochs** with early stopping near epoch 25 at the validation loss minimum.
- Produced **skillful forecasts** that capture the spatial structure and intensity evolution of convection, with a clear benefit from including ZDR.

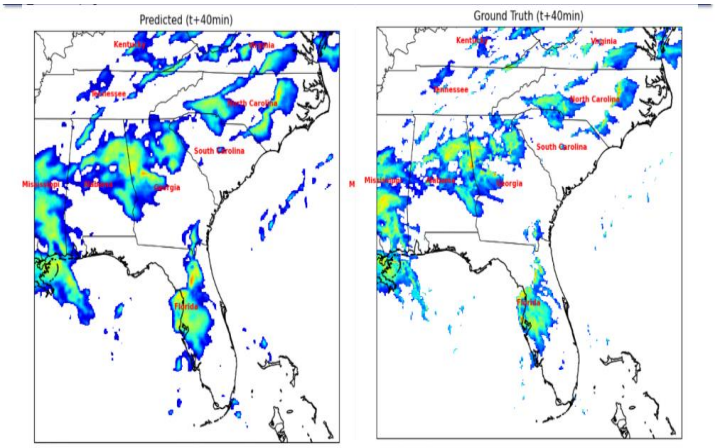
Results:

- **Evaluation score** (higher is better):

Statistics for 40 minute forecast

	dBZ threshold = 15	dBZ threshold = 35
Probability of detection (Hit rate)	.843	.664
False alarm rate	.138	.344
Heidke Skill Score (comparison to random forecast)	.846	.657

Data available for training	Train with only heavy rain days with 3 month dBZ	Train with all days from 3 month dBZ data	Train with only heavy rain days data with 3 month dBZ and ZDR
Evaluation Score	53%	50%	76%



Predicted reflectivity in 40 minutes

Actual reflectivity in 40 minutes

- **Qualitative behavior:** Predicted reflectivity fields at +40 minutes reproduce storm placement and broad intensity gradients, with some smoothing at finer scales—typical for data-driven nowcasting.

Interpretation

- **ZDR augments dBZ** by injecting polarimetric information related to drop size and hydrometeor characteristics, improving detection and intensity prediction of convective cores.
- **Event-focused training** (heavy-rain subsets) helps the model prioritize patterns relevant to flash-flooding scenarios.
- Skill **degrades beyond ~1 hour**, motivating architecture and data-fusion strategies that better capture mesoscale organization and propagation.

Challenges and Obstacles Encountered

- **Storage & compute constraints:** Radar archives and multi-sensor inputs are large, which prevented us from adding more inputs such as KDP
- **Longer-lead accuracy:** Forecast skill **drops beyond ~60 minutes**, this is expected but it could be improved by increasing the data.

Next Steps

1. **Expand data coverage** across more years and seasons to diversify convective regimes and improve generalization.
2. **Add polarimetric and environmental predictors:** KDP, wind vectors, water vapor/thermodynamic indices, and potentially satellite channels to enhance physical context.

Resources and Tools Used

- Data: MRMS (dBZ, ZDR) for the Southeastern U.S.; heavy-rain day identification via NOAA CPC.
- Modeling: FURENet (U-Net–style, spatiotemporal), PyTorch; training with early stopping.
- Software stack: Python, PyTorch, CUDA, OpenCV, NumPy, xarray, plotting/analysis utilities; job scripts for HPC execution.
- Infrastructure: Duke Compute Cluster (DCC) with NVIDIA A100 GPUs; storage optimized via image conversion and curated event selection.

Timeline & Schedule

Date / Window	Key Work & Deliverables
May 19–23, 2025	Kickoff, scope, literature review; DCC A100 setup & access.
May 24–31, 2025	MRMS **dBZ** pipeline (download→convert→organize); heavy-rain event selection.
Jun 1–14, 2025	Preprocessing & normalization; dataset loader; **FURENet** baseline; eval tools; first **+40 min** baseline.
Jun 15–23, 2025	Heavy-rain vs. all-days comparison; LR/batch tuning; regularization; **First presentation (Jun 23)**.
Jun 24–Jul 12, 2025	Added **ZDR**; multi-input ablations; threshold sweeps; observed skill gains.
Jul 13–25, 2025	Final runs & best checkpoint (~epoch 25); **Second presentation (Jul 21)**; poster prep; **Poster session (Jul 25)**.

Group Members' Contributions

Member	Contributions
Suhaib Mansour (Duke)	Data curation and preprocessing; dataset organization; assistance with evaluation & visuals.
Zijin Ye (CMU)	Evaluation pipeline and visualization; metric analysis; figures for poster.
Maan Adam (KAUST)	Heavy-rain day identification using CPC (NOAA Climate Prediction Center); training the model on DCC (A100).
Yan Pan (Duke)	Experiment planning and coordination; literature review; presentation prep.
All	Model development.

References

- Pan et al. (2021). Improving nowcasting of convective development by incorporating polarimetric radar variables into a deep-learning model. *Geophysical Research Letters*, 48(21), e2021GL095302.
- Zhang et al. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *BAMS*, 97(4), 621–638.

Acknowledgments

- We are grateful for guidance from Climate+ mentors and AI faculty, and for support from our partner institutions (Duke University and KAUST Academy). Their mentorship and infrastructure made this work possible.