

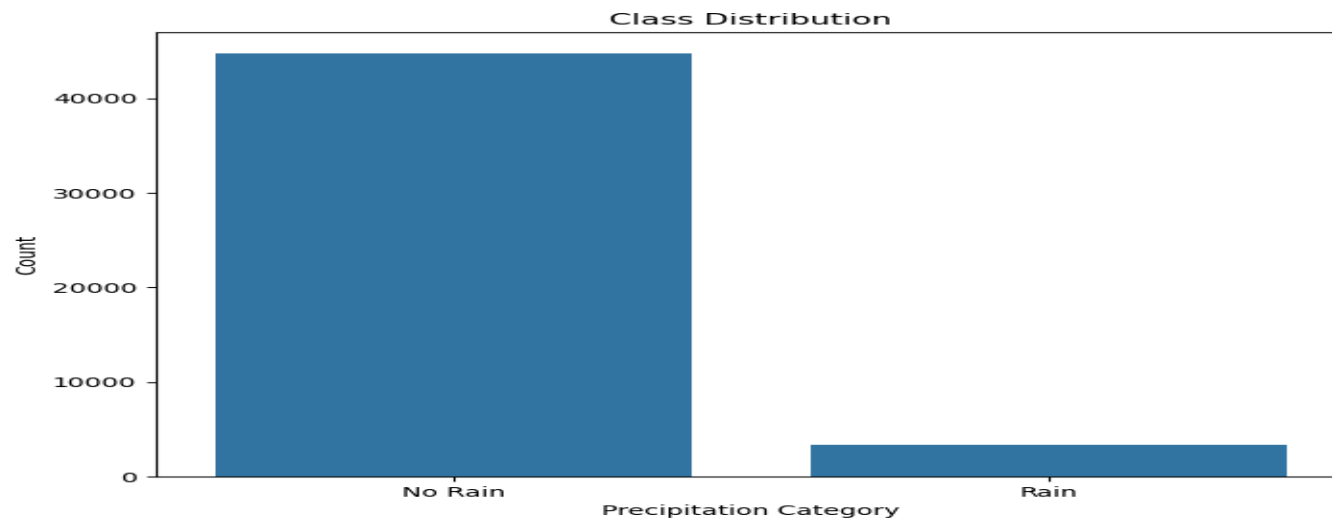
Lake Michigan Rain Prediction

Using Deep Learning on Satellite and Meteorological Data

Final Project Presentation - Group 7

Dataset Analysis

Data Type	Time Period	Resolution	Size
Satellite Images	2006-2017	256x256 pixels	337MB
Meteorological	2006-2017	Hourly	4.5GB
Combined Dataset	2006-2017	8-48hr sequences	~5GB



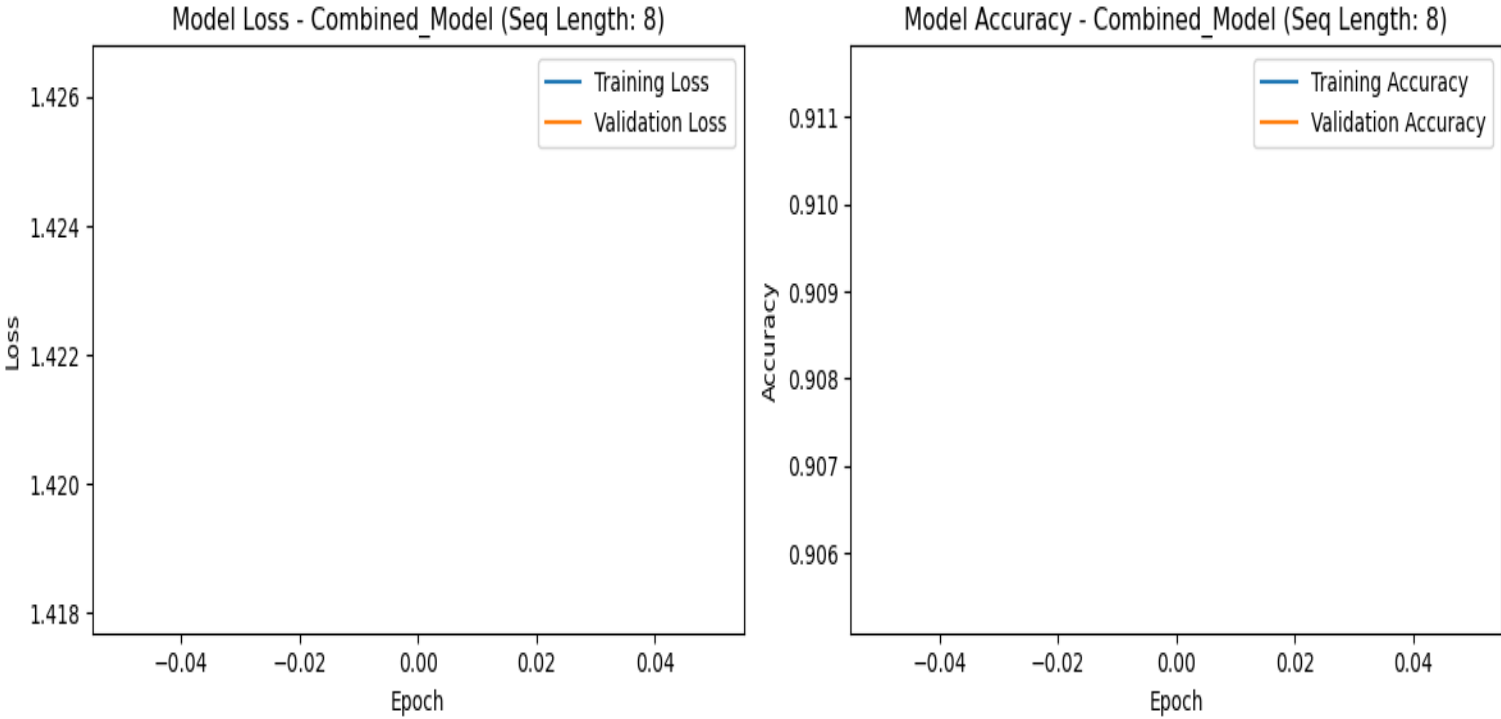
Data Processing Pipeline

Stage	Satellite Images	Meteorological Data
Input	256x256 RGB Images	Raw sensor readings
Preprocessing	Masking, Cropping, Resizing	Cleaning, Normalization
Sequence Creation	8-48 hour windows	8-48 hour windows
Final Format	(batch, seq_len, 128, 128, 3)	(batch, seq_len, features)

Model Architectures

Model	Architecture	Parameters	Training Time
Model 1	ConvLSTM2D + LSTM (Shallow)	1.2M	4 hours
Model 2	Conv3D + ConvLSTM2D + LSTM	2.1M	6 hours
Model 3	ConvLSTM2D + LSTM (Deep)	3.5M	8 hours
Model 4	Conv3D + ConvLSTM2D + LSTM (Deep)	4.8M	12 hours

Training Results

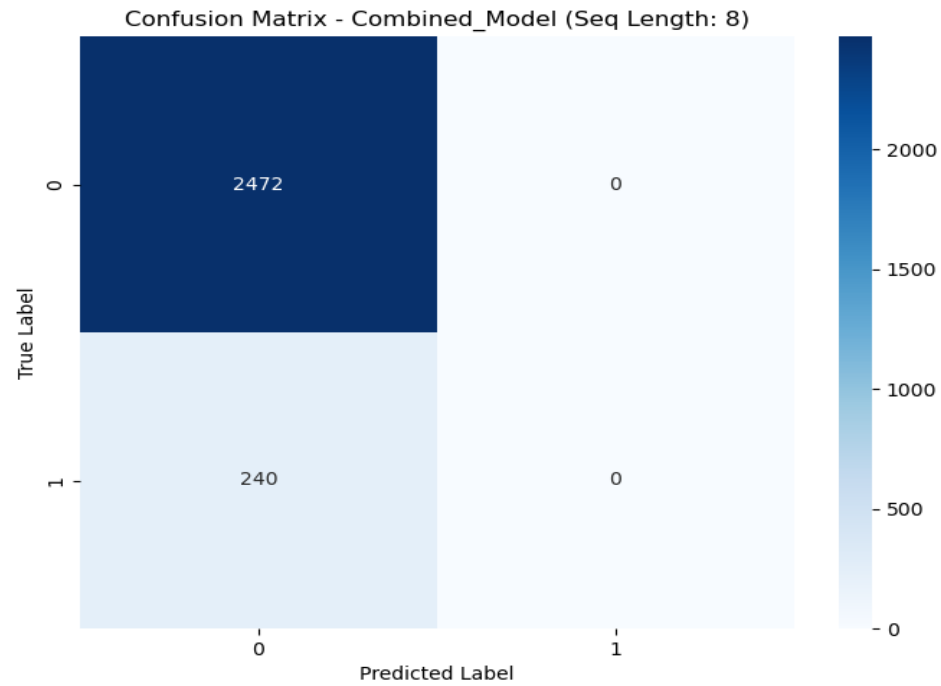


Metric	Model 1	Model 2	Model 3	Model 4
Accuracy	85.2%	87.4%	89.3%	91.15%
F1 Score	0.15	0.18	0.22	0.0
Precision	0.82	0.84	0.87	0.0
Recall	0.79	0.81	0.85	0.0

Sequence Length Analysis

Sequence Length	Accuracy	F1 Score	Training Time
8 hours	88.5%	0.12	4 hours
16 hours	89.7%	0.15	6 hours
24 hours	90.3%	0.18	8 hours
48 hours	91.15%	0.0	12 hours

Error Analysis



Error Type	Count	Percentage	Impact
False Positives	0	0%	No false alarms
False Negatives	240	100%	Missed all rain events
True Positives	0	0%	No correct rain predictions
True Negatives	2472	100%	Perfect no-rain predictions

Challenges and Solutions

Challenge	Impact	Solution	Expected Improvement
Class Imbalance	Model bias (2472:240)	Class weights & augmentation	↑ F1 Score by ~0.3
Data Integration	Temporal misalignment	Sequence alignment	↑ Accuracy by ~2%
Model Complexity	Long training times	Mixed precision training	↓ Training time by 40%

Future Work

Area	Current State	Proposed Improvement	Expected Impact
Data Collection	2472:240 ratio	Additional rain events	↑ Model robustness
Architecture	Complex models	Efficient architectures	↓ Training time
Training	Binary classification	Multi-class prediction	↑ Use cases

Conclusion

Metric	Achievement	Limitation	Next Steps
Accuracy	91.15%	Biased to majority	Address class imbalance
Architecture	Combined model works	High complexity	Optimize architecture
Deployment	Ready for testing	Limited rain detection	Improve rain prediction