Multi-task Spatiotemporal Deep Learning-based Arctic Sea Ice Prediction

Jamal Bourne[1], Michael Hu[2], Eliot Kim[3], Peter Kruse[4], Skylar Lama[5], RA: Sahara Ali[6], Mentor: Dr. Jianwu Wang[6], Collaborator: Dr. Yiyi Huang[6, 7]

[1]Department of Mathematics and Computer Science, McDaniel College
[2]Department of Computer Science, Georgia Institute of Technology
[3]Nelson Institute of Environmental Studies and Department of Statistics, University of Wisconsin-Madison
[4]Department of Accounting, Business, and Economics, Juniata College
[5]Department of Atmospheric and Oceanic Science, University of Maryland, College Park
[6]Department of Information Systems, University of Maryland, Baltimore County
[7]Science Systems and Applications, Inc.

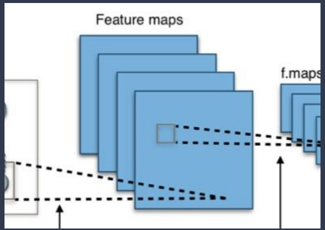
Acknowledgements: NSF, UMBC, NASA, HPCF

Outline

Background

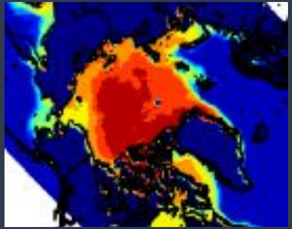
- Topic Intro
- Problem Statement
- Dataset

Models and Results



- VAR
- CNN
- ConvLSTM
- Multi-Task

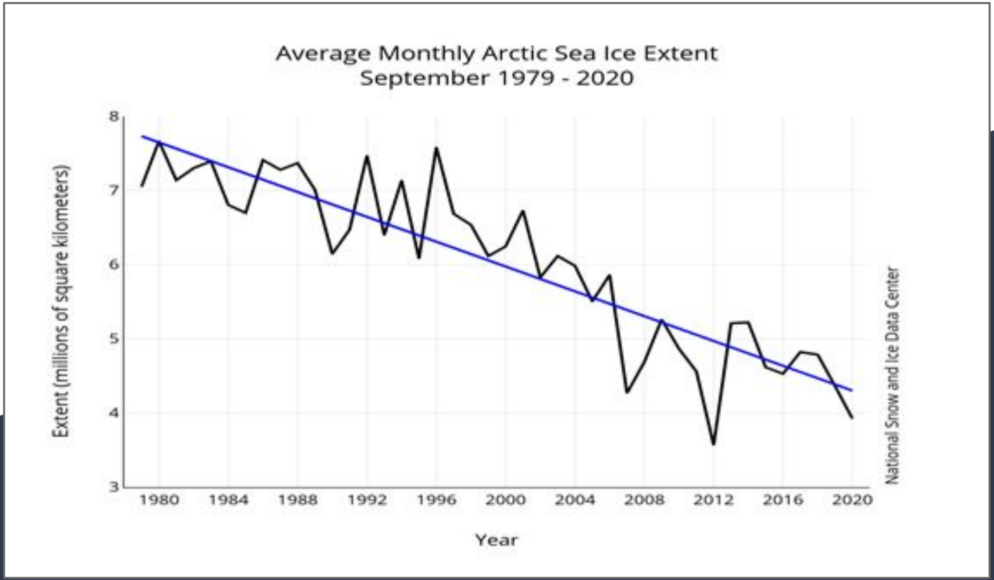
Conclusions



- Overall Results
- Comparisons to Literature

Arctic Sea Ice

- Over past decades Arctic summer sea ice has decreased by about 50%
  - Declining at rate of about 13.1% per decade
- Decline acceleration in early 21 century
  - Large effect on communities of stakeholders



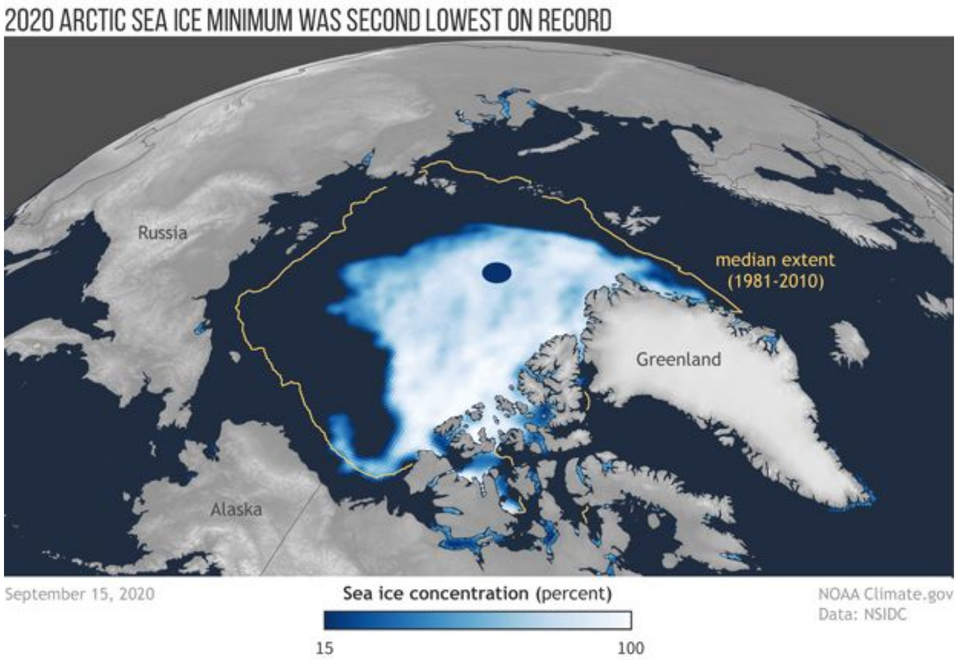
What are we predicting?

Arctic Sea Ice Concentration (SIC)

- Total area of ice covered ocean in relation to a total given area of the ocean
- Given as a fraction or percentage of (sea ice area) / (total area)
- We are focusing on SIC **per pixel** (in %)

Arctic Sea Ice Extent (SIE)

- Total area of ice covered ocean
- Include areas with 15% or greater SIC
- We are focusing on **total** SIE (in km^2)



https://nsidc.org/cryosphere/seaice/data/terminology.html

https://www.climate.gov/news-features/understanding-climate/2020-arctic-report-card-climategov-visual-highlights

Data

Feature	Source	Units	Range
Sea Ice Concentration	NSIDC	% per pixel	0-100
Surface Pressure	ERA-5 (ECMWF)	Pa	40000-110000
10m Wind Speed	ERA-5 (ECMWF)	m/s	0-40
Near-Surface Humidity	ERA-5 (ECMWF)	kg/kg	0-0.1
2m Air Temperature	ERA-5 (ECMWF)	K	200-350
Shortwave Radiation	ERA-5 (ECMWF)	W/m^2	0-1500
Longwave Radiation	ERA-5 (ECMWF)	W/m^2	0-300
Rain Rate	ERA-5 (ECMWF)	mm/day	0-800
Snow Rate	ERA-5 (ECMWF)	mm/day	0-200
Sea Surface Temperature	ERA-5 (ECMWF)	K	200-350

Physical Variables

Problem Statement

Given $\boldsymbol{n}$ months of historical data $\boldsymbol{X}$ comprising of the 10 atmospheric and ocean variable measurements in Arctic region for each pixel, learn a function to forecast pixel-wise sea-ice concentration $\boldsymbol{Y_s}$ and total sea-ice extent $\boldsymbol{Y_e}$ for the next month

$$Y_{s_{t+1}} = f(X_{t-n}, X_{t-n+1}, \ldots, X_t)$$

$$Y_{e_{t+1}} = f(X_{t-n}, X_{t-n+1}, \ldots, X_t)$$

Challenges

- How to predict both sea ice concentration and extent through models
- Noise caused by land and open ocean values caused an increase in RMSE
  - How to train model to only focus on sea ice and not land and open ocean

Our Work

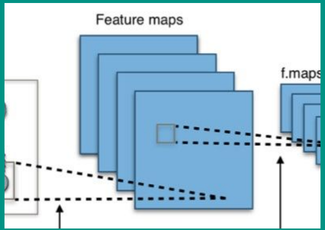
- Forecast Arctic sea ice concentration and extent
- Eliminate noise through post-processing
- Use deep learning models with novelty approaches
  - Custom loss
  - Multi-task
- Create accurate SIC and SIE predictions that are comparable or better than previous works

Outline

Background

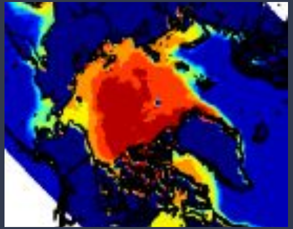
- Topic Intro
- Problem Statement
- Dataset

Models and Experiments



- VAR
- CNN
- ConvLSTM
- Multi-Task

Conclusions



- Overall Results
- Comparisons to Literature

Models and Experiments

Statistical:

- VAR

Deep Learning:

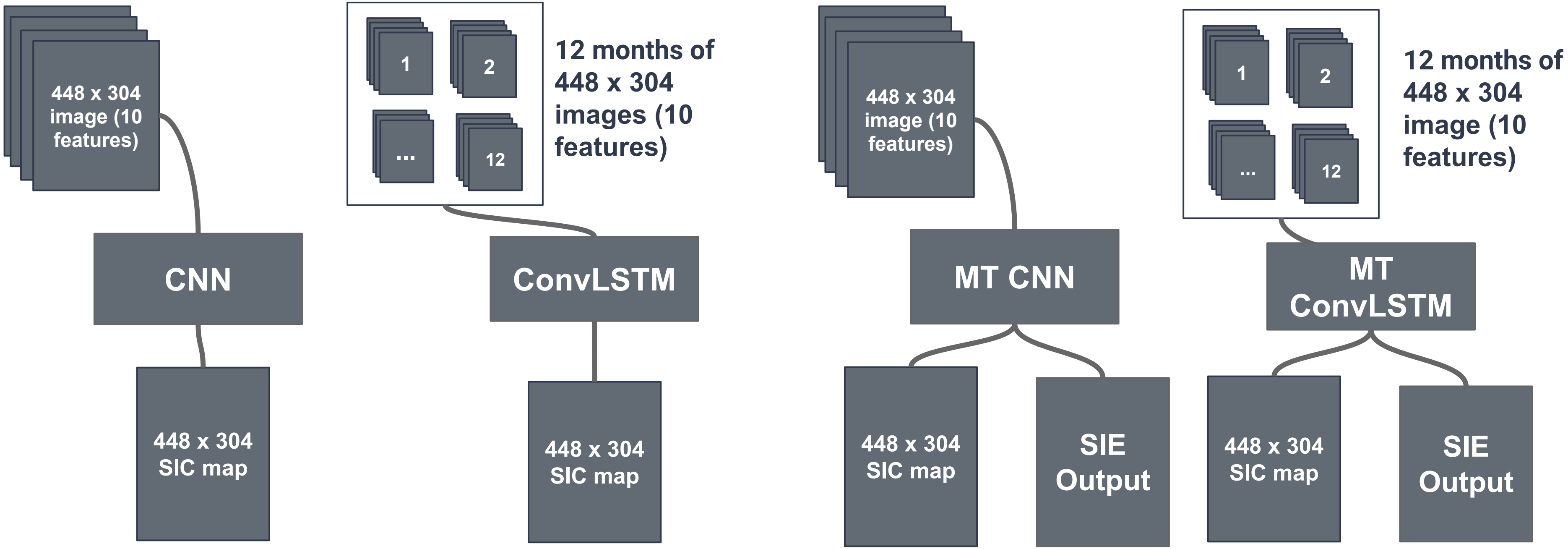
- ConvLSTM
- CNN
- Multi-task ConvLSTM
- Multi-task CNN

https://www.hakaimagazine.com/videos-visuals/frozen-ocean/

Data Split and Usage

Data split between training and testing:

- Train: Jan. 1979 - Dec. 2012
- Test: Jan. 2014 - Dec. 2020



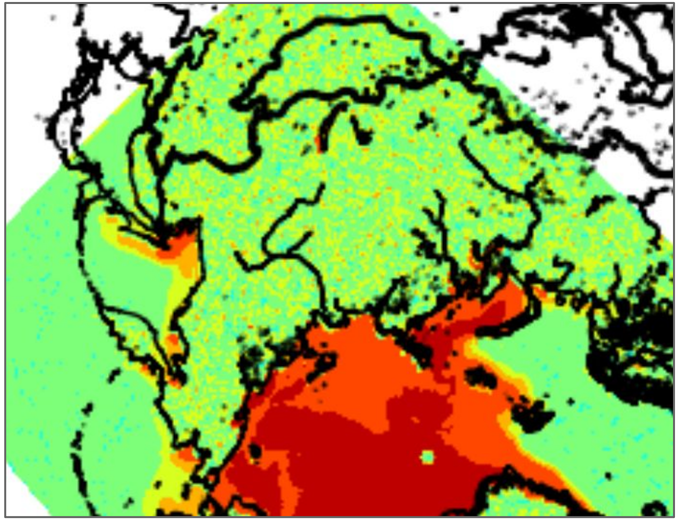
Data Processing

Derived sea ice extent:

- Calculated using sea ice concentration and per-pixel area
- Sum of areas of pixels with >15% SIC

Post-processing:

- North Pole Hole pixels are ignored due to lack of observations
- Land pixels are ignored
- Values below 0 are converted to 0; values above 100 are converted to 100



Example of predicted SIC prior to post-processing

Evaluation Metrics

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

NRMSE = RMSE / $\bar{y}$

Baseline Model: VAR (Vector Autoregression)

How it works:

- VAR models learn the relationship between multiple variables as they change through time
- From this learning, forecasts can then be made to predict future values
- **Lag:** Number of prior time-steps used to predict values for the current time-step

Task: Create a spatially averaged prediction for sea ice concentration

VAR: $y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t$

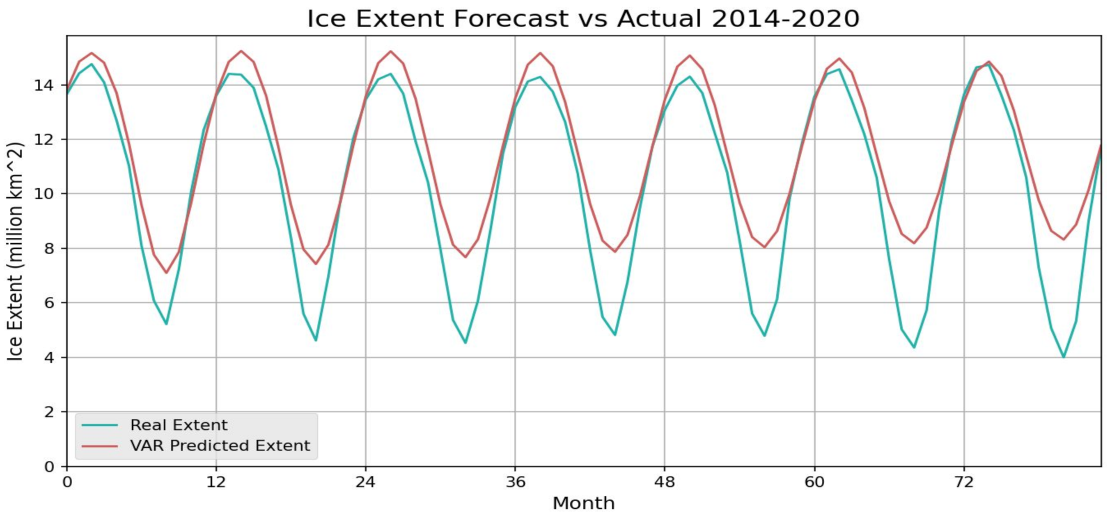
Model Configurations:

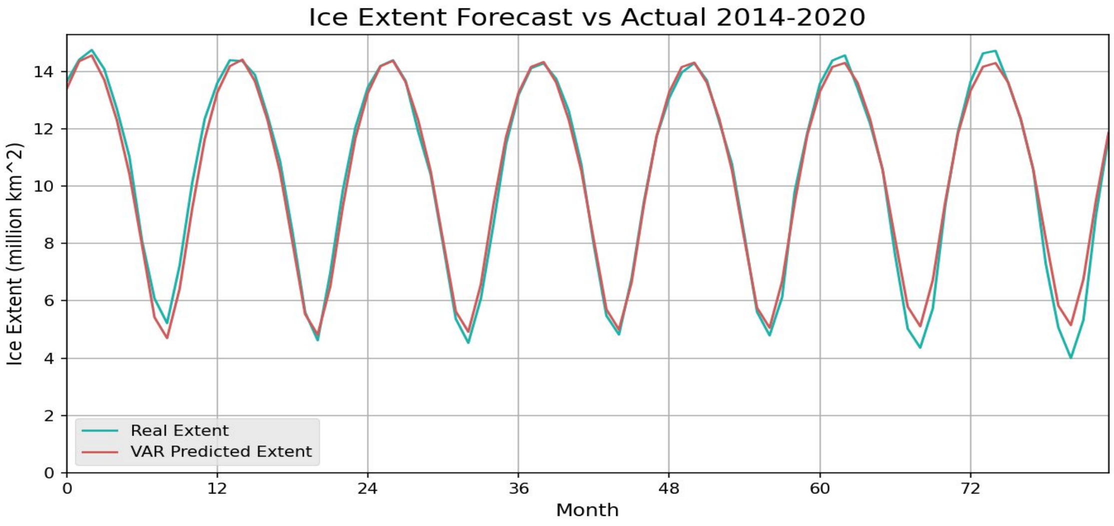
1. VAR with lag two, namely VAR(2), chosen based on BIC (Bayesian information criterion)
2. VAR with lag ten, VAR(10), chosen based on AIC (Akaike information criterion)

VAR Results

- **Top: VAR with lag 2 based on BIC**
- **RMSE: 1.536 million km²**
- Predicts March maxima better than September minima

- **Bottom: VAR with lag 10 based on AIC**
- **RMSE: 0.424 million km²**
- Predicts both March maxima and September minima accurately





CNN

How it works:

- The model takes each image input and passes it through a series of **convolutional**, **max-pooling**, and **fully connected layers**
- Features are extracted from images to help the model learn and produce an image output forecasting future predictions
- Includes **Custom Loss Function**

Task: For time *t* and lead time *1*, use samples from month *t* to predict SIC per-pixel at time *t+1*

Data:

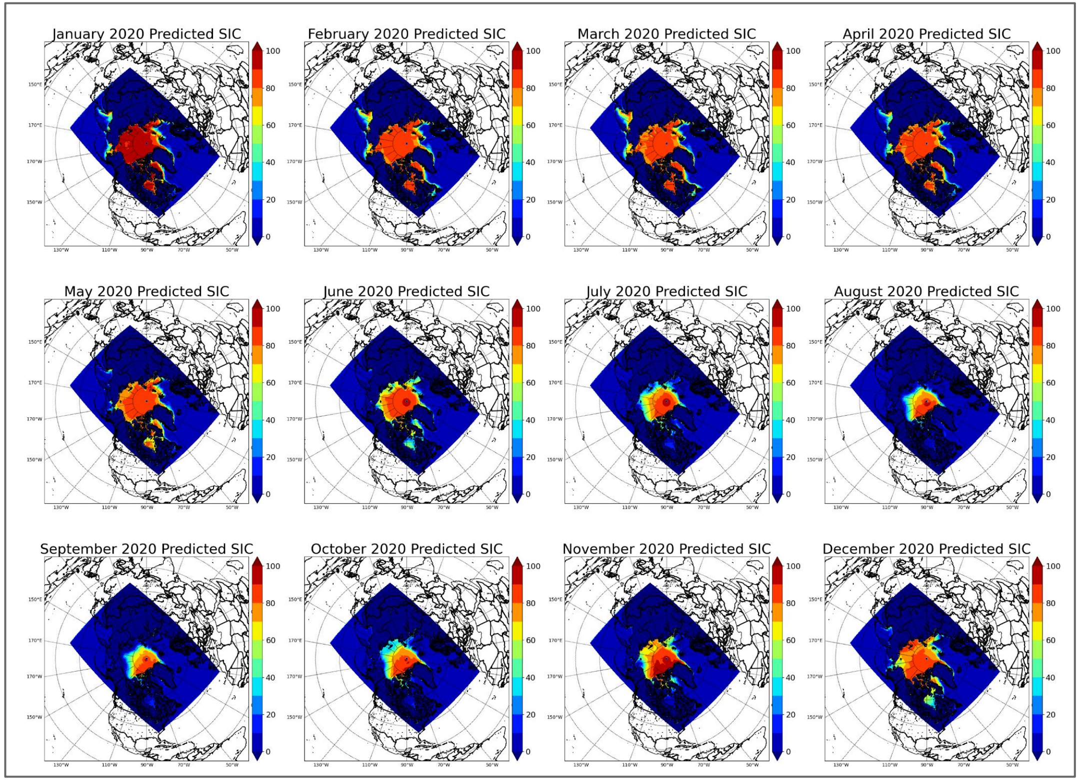
- North Pole Hole filled for training, removed during post-processing

Network Structure:

- Convolutional Layer (128 filters, 5x5 kernel, input shape of (448, 304, 10))
- Max Pooling (2x2)
- Convolutional Layer (32 filters, 5x5 kernel, relu activation)
- Max Pooling (2x2)
- Convolutional Layer (8 filters, 5x5 kernel, relu activation)
- Fully Connected Layer (256 nodes, relu activation)
- Output Layer (448*304 nodes, linear activation)

Base CNN SIC Results

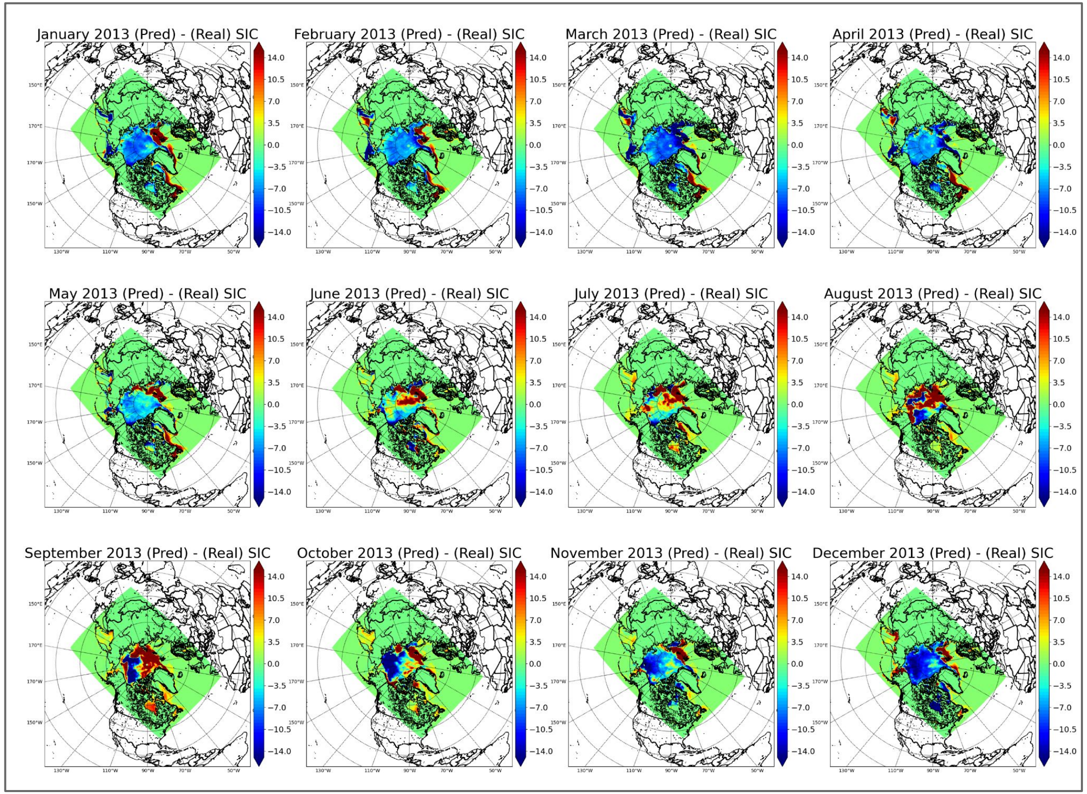
- Predicts distinct spatial distribution of sea ice for each month
- Reasonable RMSE of 7.231%



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc SIC RMSE (%)
31	4	11.738%	12.005%	0.862 million km^2	7.231%

Base CNN SIC Difference Plot

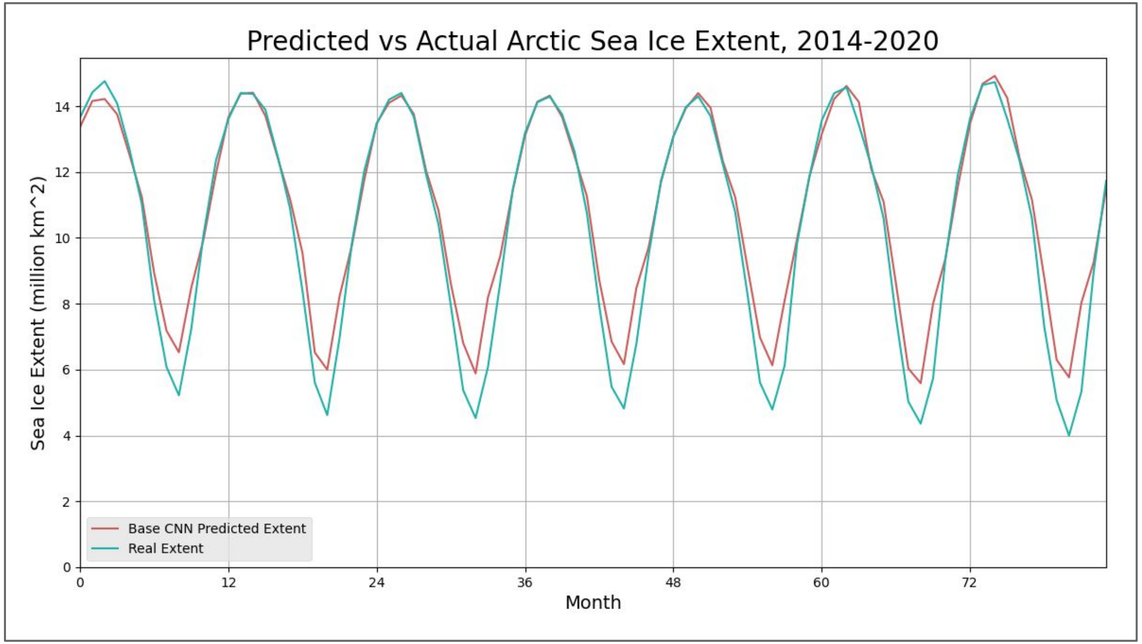
- Consistent underestimates of SIC during winter and spring
- Greater differences during August through October



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc SIC RMSE (%)
31	4	11.738%	12.005%	0.862 million km^2	7.231%

Base CNN SIE Results

- Predicted SIE derived from SIC values
- Accurate March maxima predictions
- Significant overestimate of September minima
- Improved over VAR(2), worse performance than VAR(10)



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc SIC RMSE (%)
31	4	11.738%	12.005%	0.862 mil. km^2	7.231%

Extent Loss CNN

How it works:

- Same model architecture as Base CNN

Custom Loss Function: Incorporates SIE error in the custom loss function

  - Model optimizes for both SIC and SIE predictions

Task: For time *t* and lead time *1*, use samples from month *t* to predict SIE per-pixel at time *t+1*?

Data:

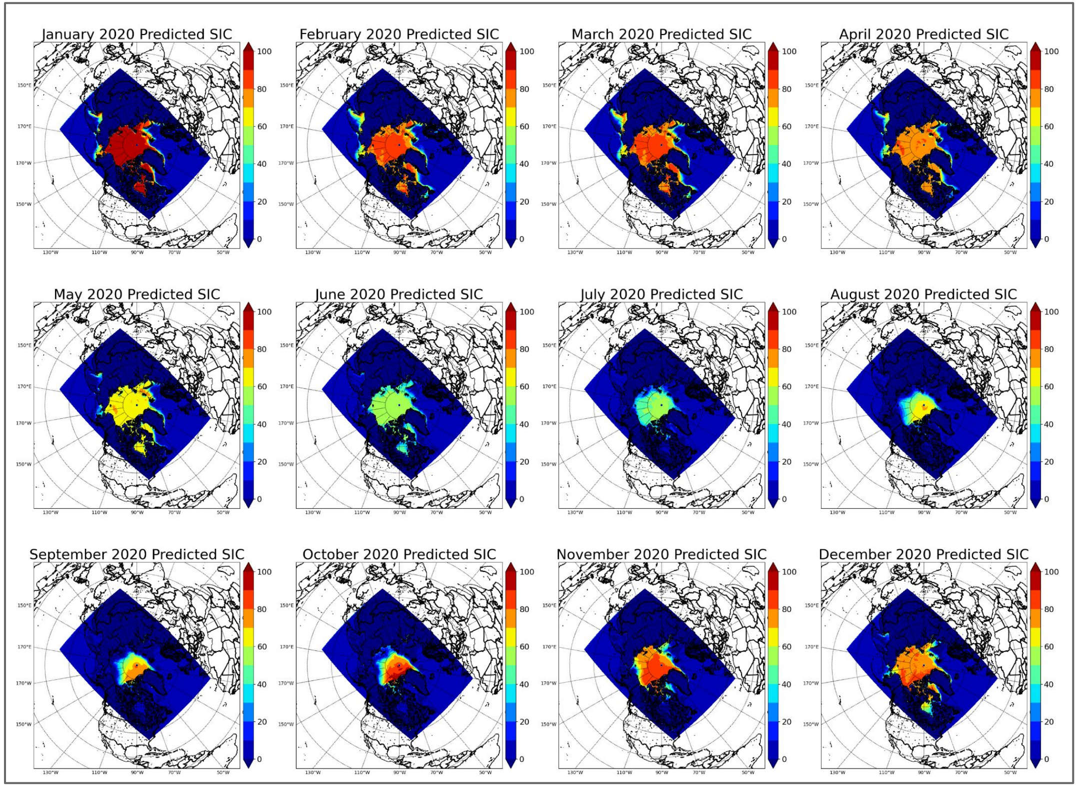
- North Pole Hole filled for training, removed during post-processing

Network Structure:

- Convolutional Layer (128 filters, 5x5 kernel, input shape of (448, 304, 10))
- Max Pooling (2x2)
- Convolutional Layer (32 filters, 5x5 kernel, relu activation)
- Max Pooling (2x2)
- Convolutional Layer (8 filters, 5x5 kernel, relu activation)
- Fully Connected Layer (256 nodes, relu activation)
- Output Layer (448*304 nodes, linear activation)

Extent Loss CNN SIC Results

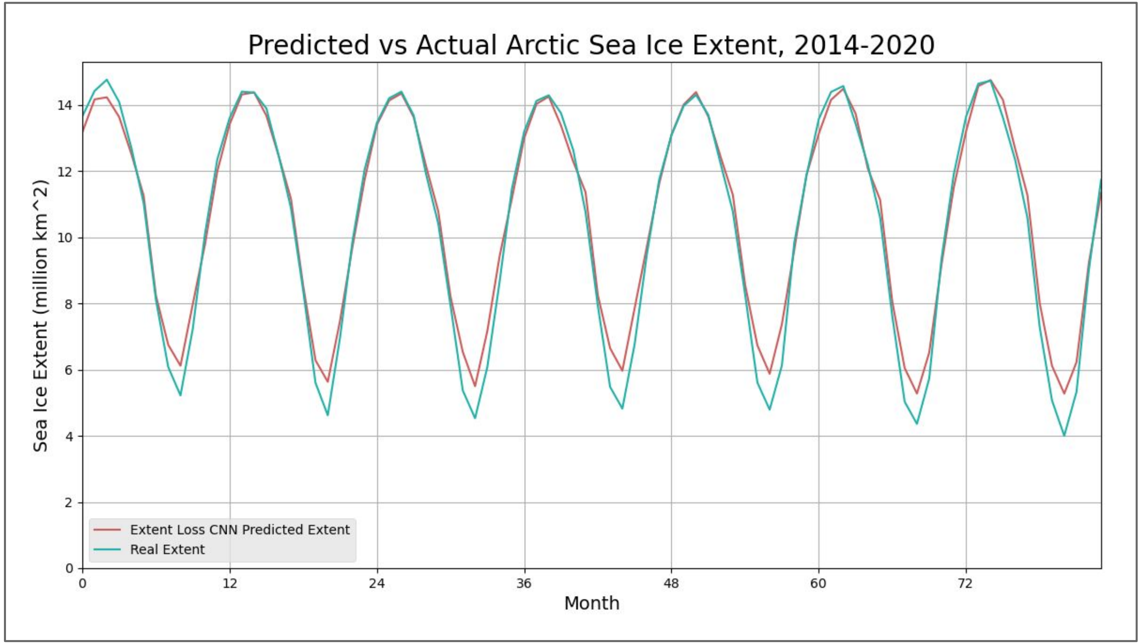
- Improved SIC prediction performance
- Very low SIC values for April through July



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km²)	Post-Proc SIC RMSE (mil. km²)
57	4	11.911%	12.228%	0.571 mil. km²	7.150%

Extent Loss CNN SIE Results

- Predicted SIE derived from SIC values
- Improved prediction of September minima compared to Base CNN
- Significantly lower Extent RMSE



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc SIC RMSE (%)
57	4	11.911%	12.228%	0.571 mil. km^2	7.150%

ConvLSTM

How it works:

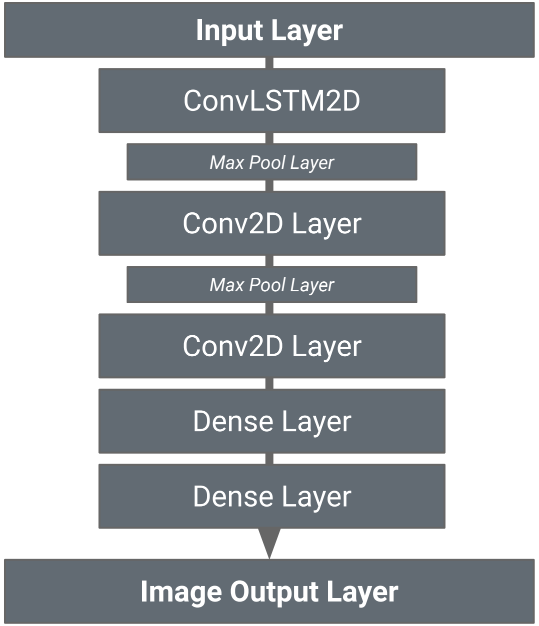
- Combines image processing capabilities of CNN modeling with the temporal processing capabilities of LSTM modeling
- Allows the model to more easily understand patterns over a spatial and temporal domain

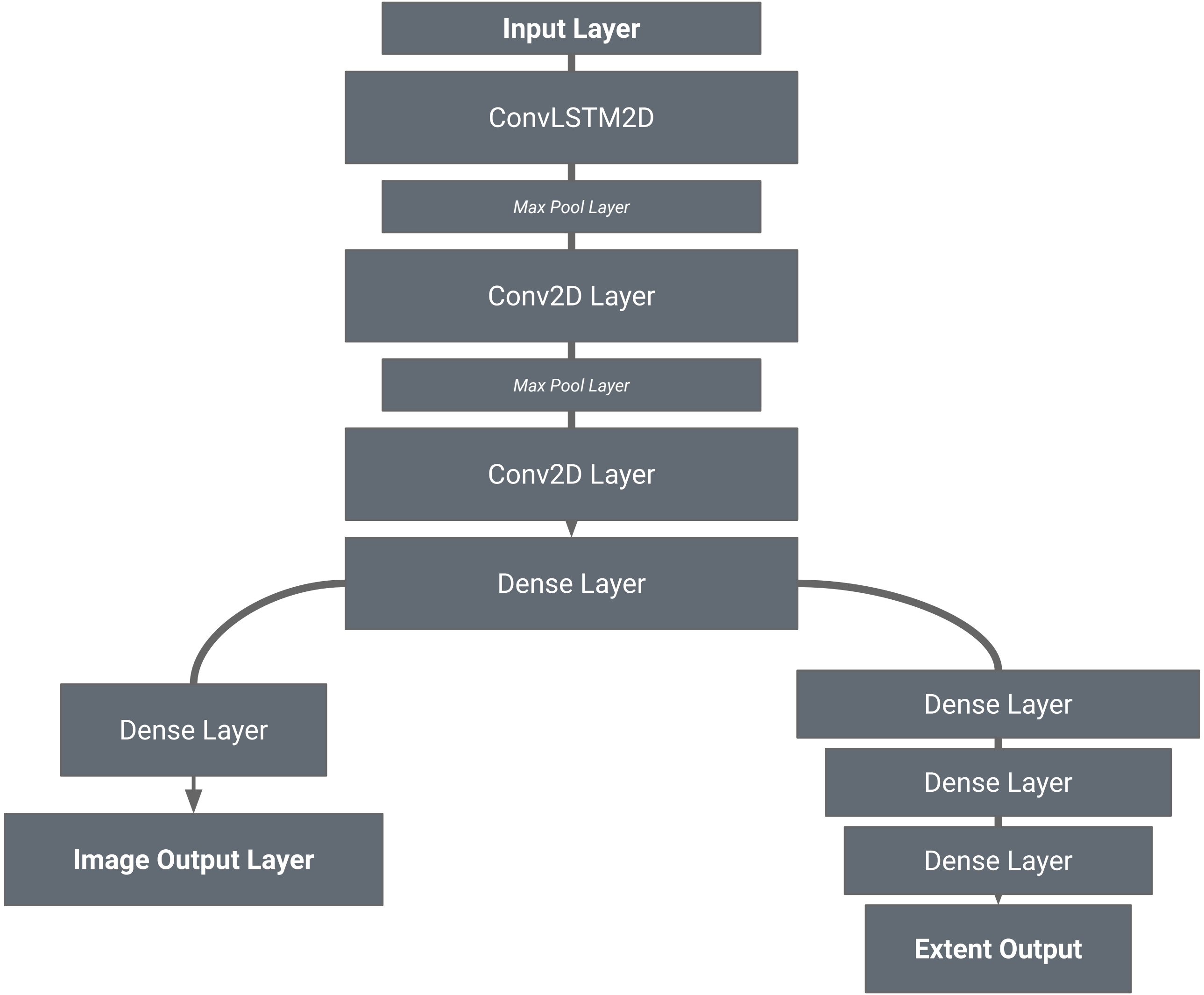
Task: For time *t* and lead time *l*, use samples from months *t-12, t-11, ..., t* to predict SIC per-pixel at time *t+l*

Data:

- 1 month lead time, unstandardized
- Rolling window
- Inputs: samples of shape (12 months, 448 x 304 spatial map, 10 features)
- Outputs: SIC image maps of size 448 x 304

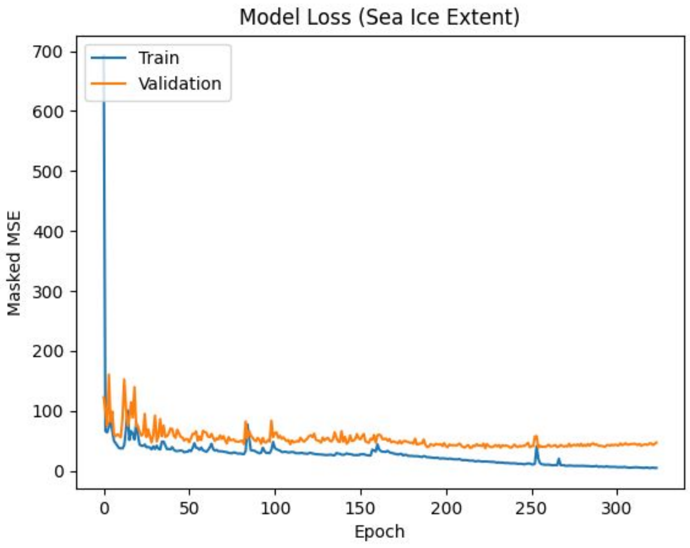
Architecture:





Input Layer
ConvLSTM2D
Max Pool Layer
Conv2D Layer
Max Pool Layer
Conv2D Layer
Dense Layer
Dense Layer
Image Output Layer
Dense Layer
Dense Layer
Dense Layer
Extent Output

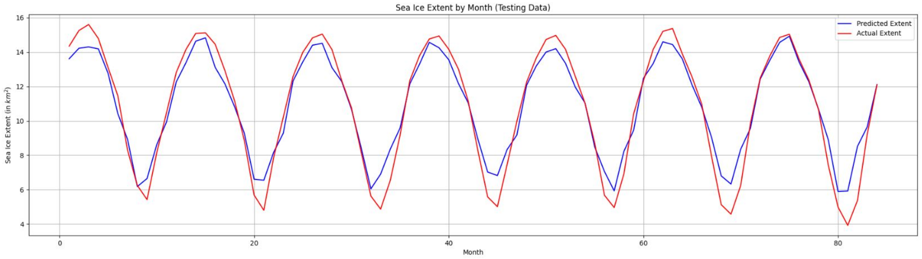
ConvLSTM Results



SIC Image Loss

Results

Batch Size	Epochs	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Train RMSE (mil. km^2)	Derived Extent Test RMSE (mil. km^2)	Post-Process RMSE (%)
4	324	10.054%	11.478%	0.908 mil. km^2	0.938 mil. km^2	8.162%



Predicted Sea Ice Extent (Blue) vs. Actual Sea Ice Extent (Red) on test data

Multi-Task CNN

How it works:

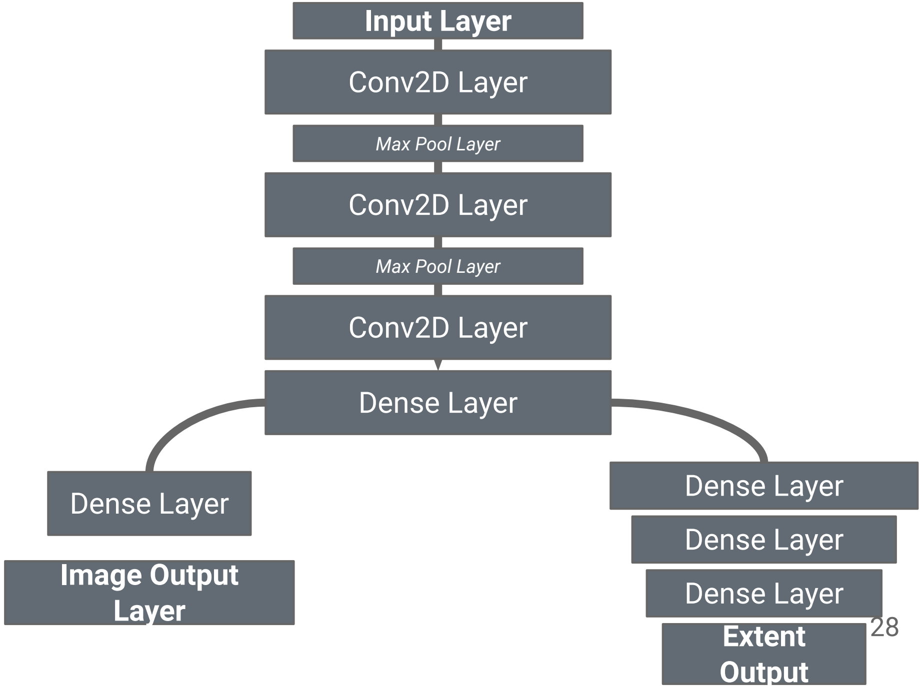
- Similar to a normal CNN and ConvLSTM, but uses branch architecture to learn two tasks at once
- Comprised of a shared "root" and two "branches"
  - One branch predicts SIC images, while the other predicts sea ice extent

Task: For time *t* and lead time *l*, use samples from month *t* to predict SIC per-pixel at time *t+l* **and** total SIE at time *t+l*

Data:

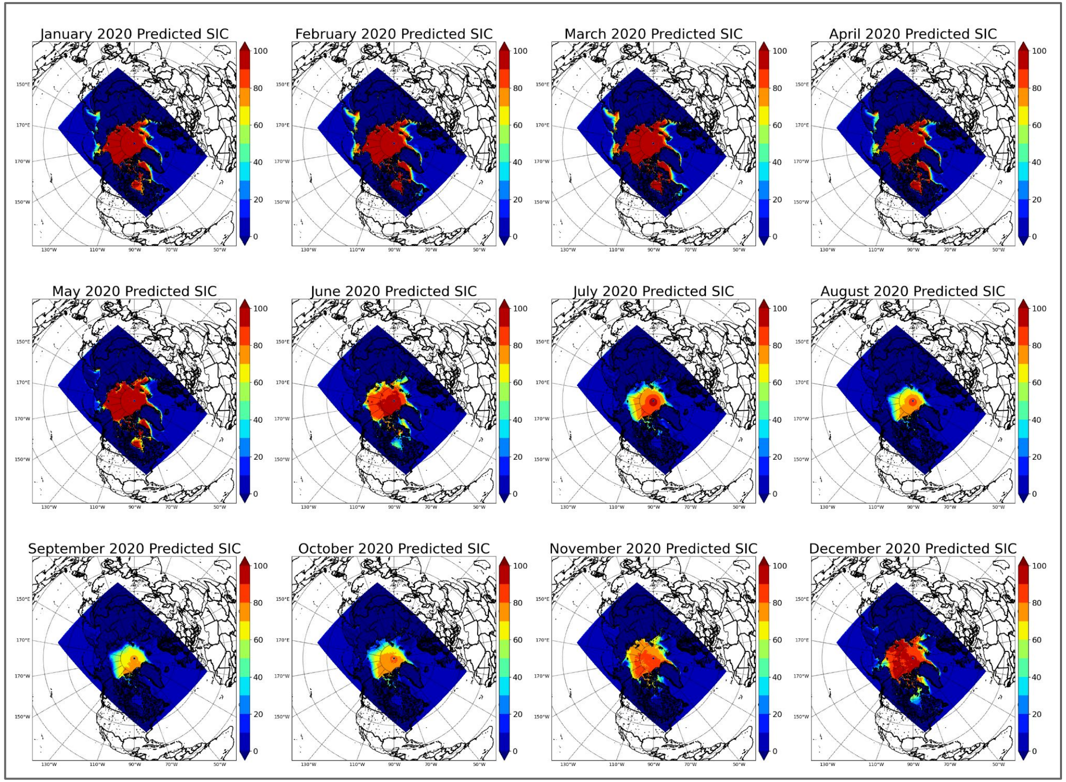
- 1 month lead time, unstandardized
- Inputs: samples of shape (448 x 304 spatial map, 10 features)
- Outputs: SIC image maps of size 448 x 304

Architecture:



Multi-Task CNN SIC Results

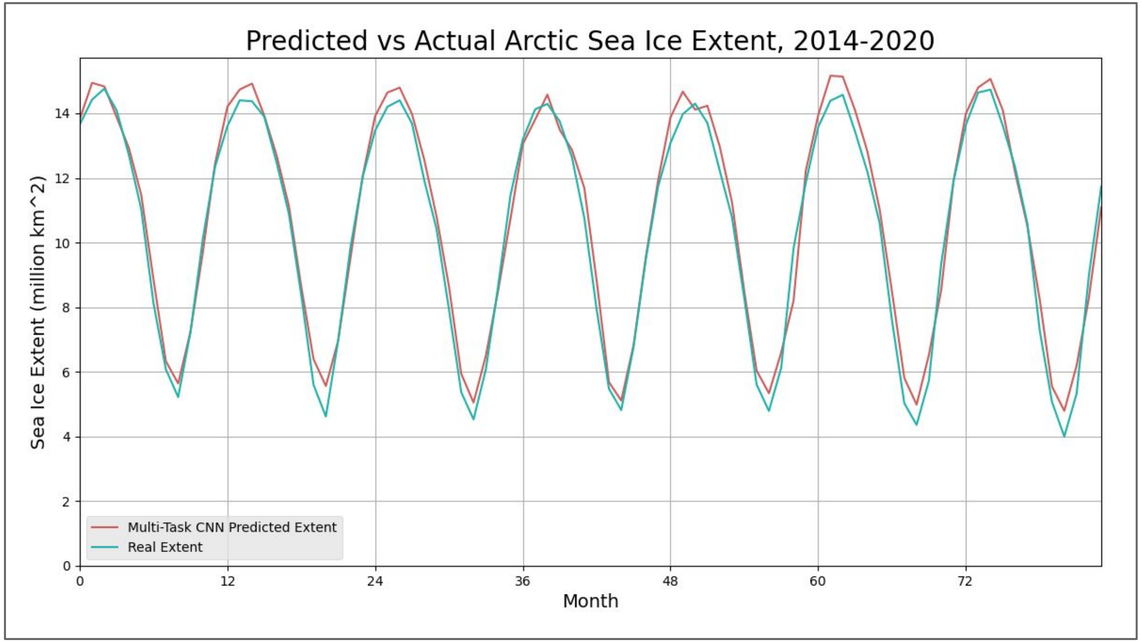
- Slightly decreased SIC prediction performance from Base CNN



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc RMSE (%)
143	32	13.108%	13.348%	0.515 mil. km^2	7.527%

Multi-Task CNN SIE Results

- Improved performance over single output CNN models
- More accurate September minima predictions
- Slightly worse March maxima predictions



Epochs	Batch Size	SIC Train RMSE (%)	SIC Test RMSE (%)	Derived Extent Test RMSE (mil. km^2)	Post-Proc RMSE (%)
143	32	13.108%	13.348%	0.536 mil. km^2	7.527%

Multi-Task ConvLSTM

How it works:

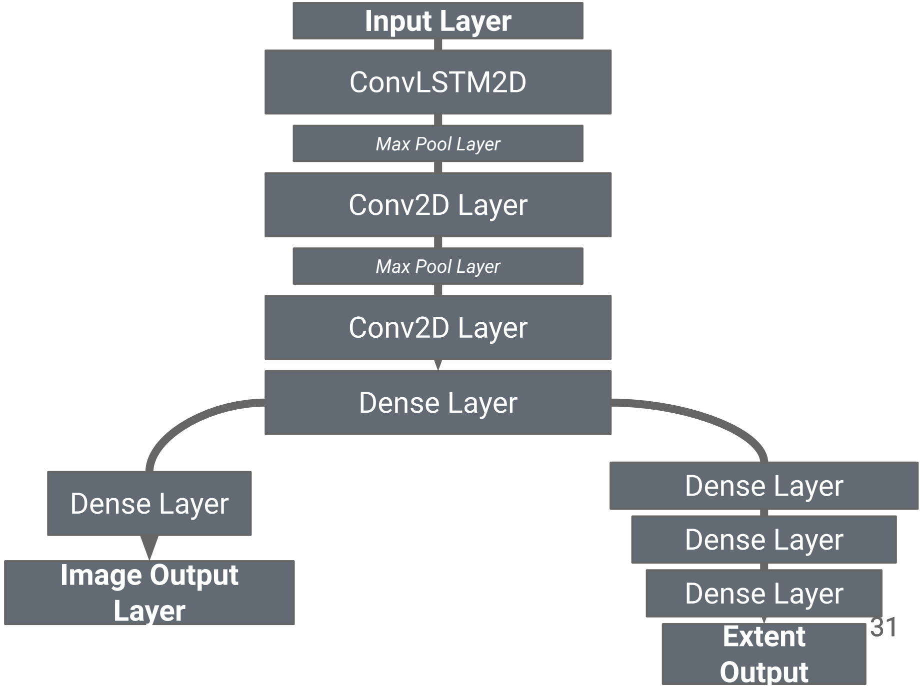
- Similar to a normal ConvLSTM, but uses branch architecture to learn two tasks at once
- Comprised of a shared “root” and two “branches”
  - One branch predicts SIC images, while the other predicts sea ice extent

Task: For time *t* and lead time *l*, use samples from months *t-12*, *t-11*, …, *t* to predict SIC per-pixel at time *t+l* **and** total SIE at time *t+l*

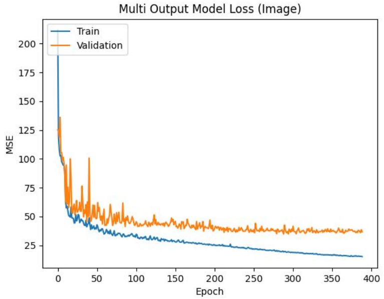
Data:

- 1 month lead time, unstandardized
- Rolling window
- Inputs: samples of shape (12 months, 448 x 304 spatial map, 10 features)
- Outputs: SIC image maps of size 448 x 304

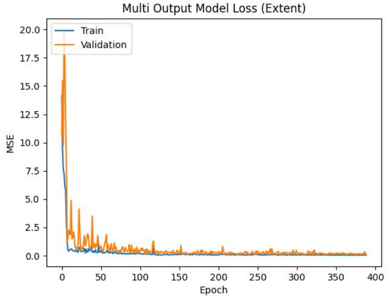
Architecture:



Multi-Task ConvLSTM Results



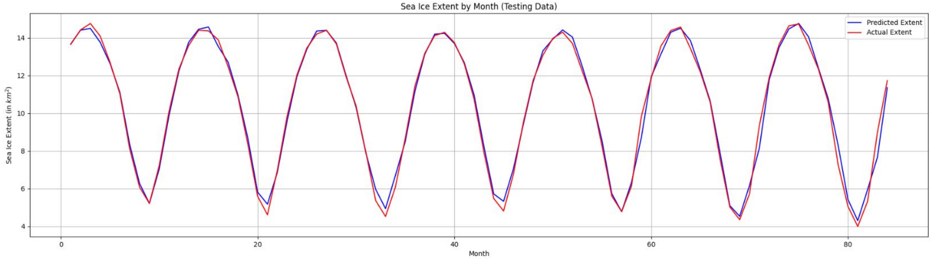
SIC Image Loss



SIE Loss

Results

Batch Size	Epochs	SIC Train RMSE (%)	SIC Test RMSE (%)	Extent Train RMSE (mil. km^2)	Extent Test RMSE (mil. km^2)	Post-Process RMSE (%)
4	251	9.846%	10.785%	0.2678 mil. km^2	0.441 mil. km^2	7.192%



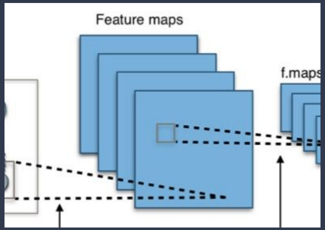
Predicted Sea Ice Extent (Blue) vs. Actual Sea Ice Extent (Red) on test data

Outline

Background

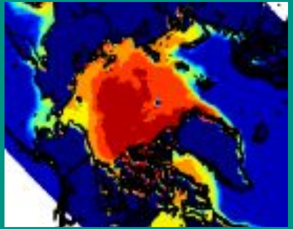
- Topic Intro
- Problem Statement
- Dataset

Models and Results



- VAR
- CNN
- ConvLSTM
- Multi-Task

Conclusions



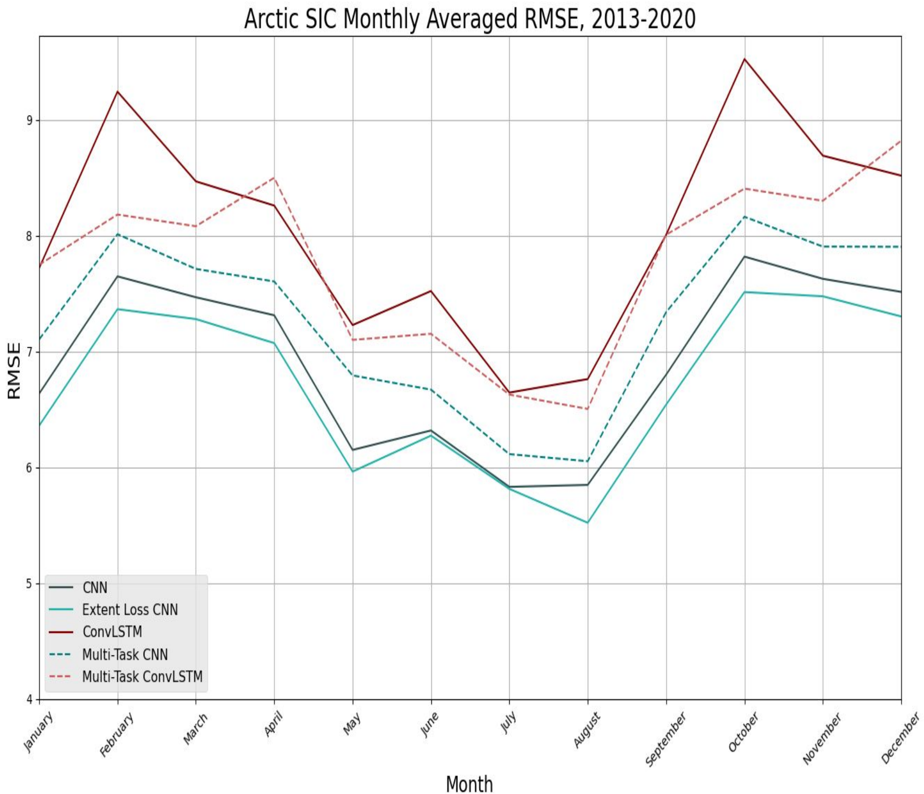
- Overall Results
- Comparisons to Literature

Overall Results

- Similar SIC prediction errors
- Extent Loss CNN does best by small margin
- LSTM model has best SIE predictions
- MultiTask models are comparable to VAR, LSTM
- Time Series-only models have better SIE performance

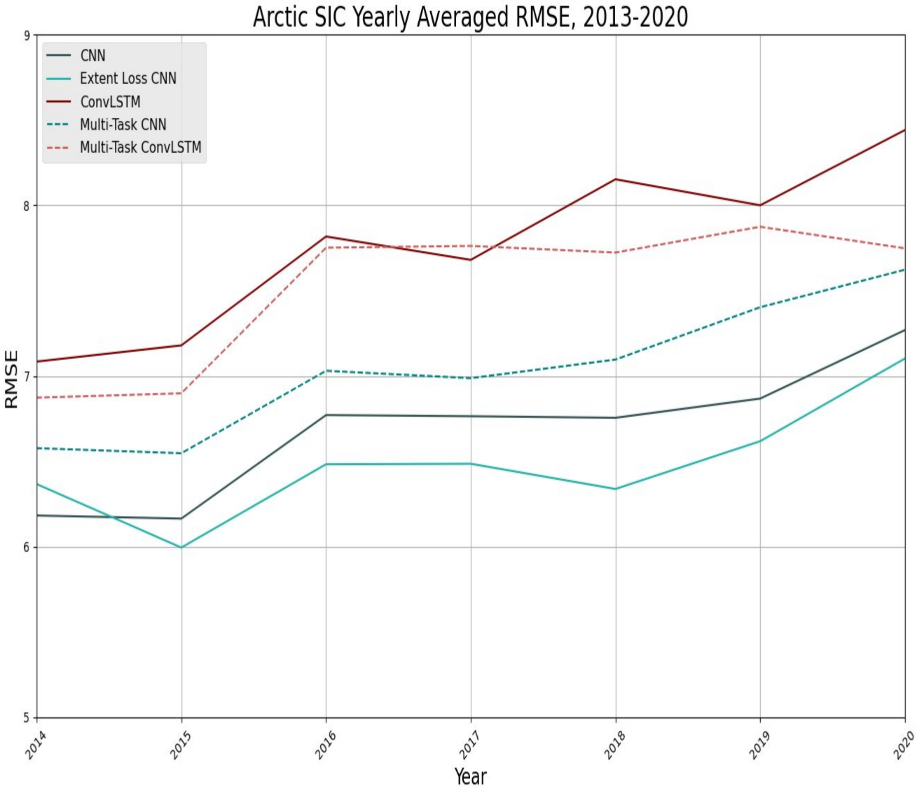
Method	Concentration Training RMSE (unit: SIC %)	Concentration Testing RMSE (unit: SIC %)	Post-Processed Concentration RMSE (unit: SIC %)	Extent Training RMSE (unit: million km^2)	Extent Testing RMSE (unit: million km^2)
VAR	N/A	N/A	N/A	N/A	0.424
LSTM	N/A	N/A	N/A	**0.179**	**0.314**
CNN	11.734	12.005	7.231	N/A	0.862*
Extent Loss CNN	11.911	12.228	**7.150**	N/A	0.670*
ConvLSTM	10.054	11.478	8.162	0.908*	0.938*
Multi-Task CNN	13.108	13.348	7.527	0.375	0.536
Multi-Task ConvLSTM	**9.846**	**10.785**	7.192	0.268	0.441

Results: Sea Ice Concentration (SIC)



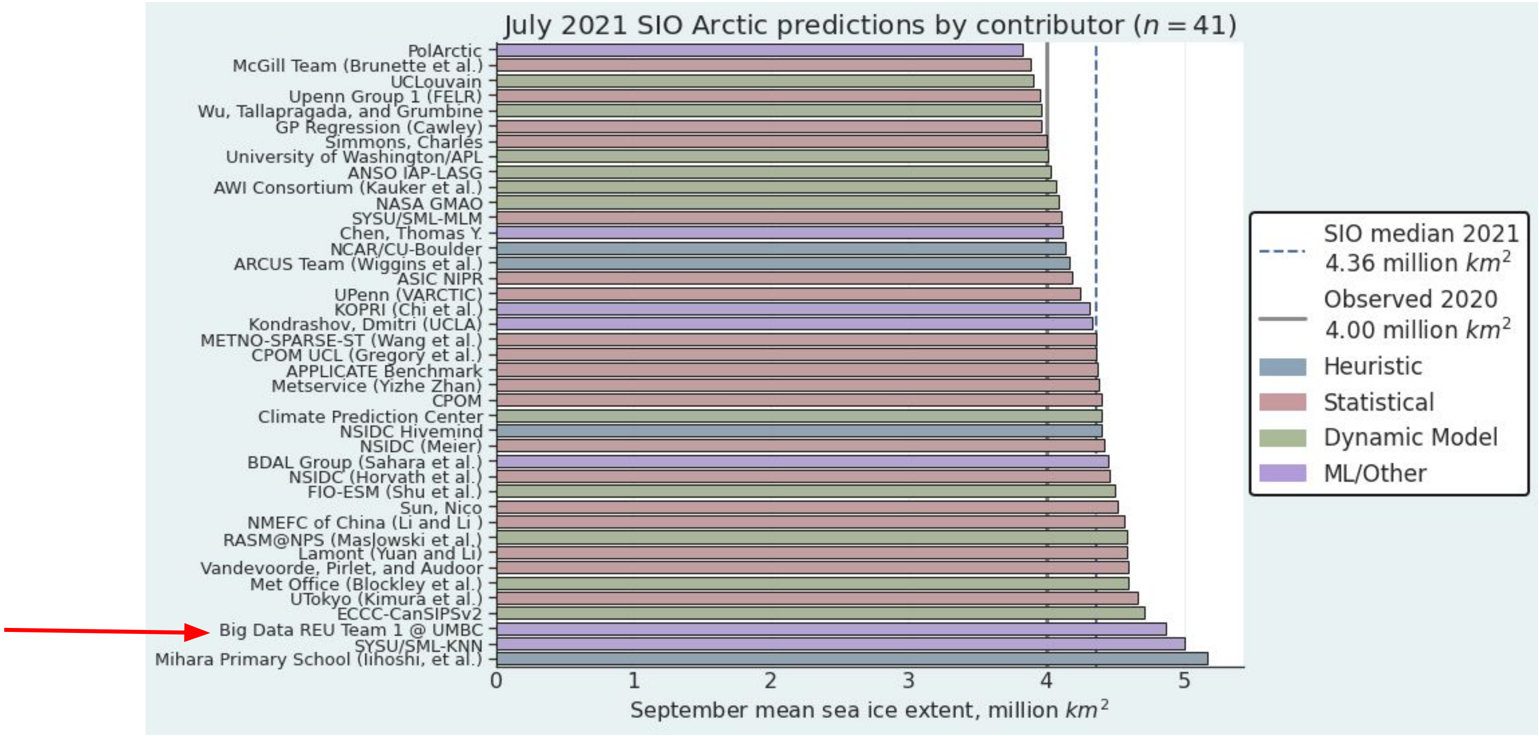
- CNN models perform better than ConvLSTM
- Extent Loss CNN has overall best performance
- Lower RMSE for May-August
- Higher RMSE for Jan-April, September-December
- Greater RMSE during periods with greater temporal variability in sea ice

Results: Sea Ice Concentration (SIC)



- CNN models again had lower RMSEs than ConvLSTM models
- Extent Loss CNN has lowest RMSEs
- Error increases over time
  - Testing data becomes more dissimilar from training data

Sea Ice Prediction Network Competition



Comparison of Related Work Predicting Sea Ice Concentration

Team	Model	Data	Physical Variables	Temporal Resolution	Lead Time	SIC % RMSE	SIC % NRMSE (RMSE / $\bar{y}$)
Liu	ConvLSTM	25 x 25 km	✓	Daily	1 day	11.2%	N/A
Liu	CNN	25 x 25 km	✓	Daily	1 day	13.7%	N/A
RS Kim	BMA/DNN	25 x 25 km	✓	Monthly	1 month	NA	0.8%
EGU Kim	CNN	25 x 25 km	✓	Monthly	1 month	5.76%	N/A
RS Chi	LSTM	Daily averaged monthly inputs	✕	Monthly	1 month	8.89%	N/A
Team1	ConvLSTM	25x25km monthly avg.	✓	Monthly	1 month	8.162%	0.860%
Team1	CNN	25x25km monthly avg.	✓	Monthly	1 month	5.635%	N/A
Team1	Multi-Task ConvLSTM	25x25km monthly avg.	✓	Monthly	1 month	7.197%	0.759%
Team1	Multi-Task CNN	25x25km monthly avg.	✓	Monthly	1 month	7.394%	N/A

Conclusions

- CNN and ConvLSTM models provide similar performance for SIC prediction
    - After adding a temporal dimension, our ConvLSTM model does not appear to greatly improve SIC model performance
    - Results are comparable to similar studies in the literature
- Deep learning models perform comparably to VAR when predicting SIE
- Multi-task learning allows us to effectively predict both monthly SIC and SIE with error rates comparable to or better than other ML/DL methods

Next Steps

- Better approach to discover and utilize temporal patterns
  - Daily data, different window size, etc.
- Scaled loss function
  - Removes the need for post-processing
- Add previous month SIE values to input data
- Test models on varying lead times
- Hyperparameter tuning and large-scale studies
  - Reduce overfitting
- Work towards conference/journal paper submission