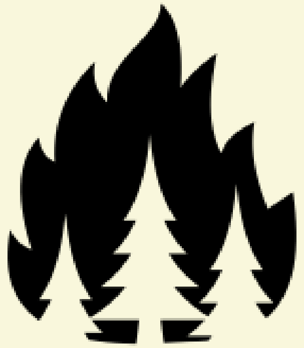
Wildfire Forecasting

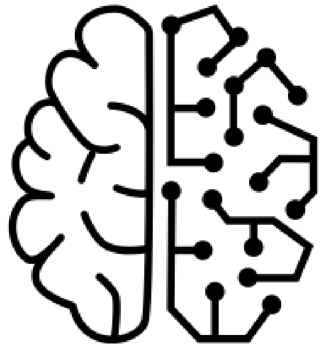
Shrreya Behll, Yibing Chen, Hongxuan Jiang, Peiyi Leng, Yixin Li, Sarah Nesti, Aoife Robertson, Robert Smith

Motivation



Rising Wildfire Threats

Limitations of
Traditional Models



ML approach
efficient and flexible

Objectives

1

Build **RNN** surrogate wildfire predictive model

2

Build a **Generative AI** surrogate wildfire generative model

3

Perform **data assimilation** on objectives 1 & 2

Data Analysis

Training and Test Data:

- 175 wildfire cycles
  (125 train, 50 test)
- 100 images per cycle
- time step: 1 unit

Observation Data

- 5 Satellite images
- time step: 10 units

s

Background Data

- 5 images
- time step: 10 units

Weakpoints/Challenges

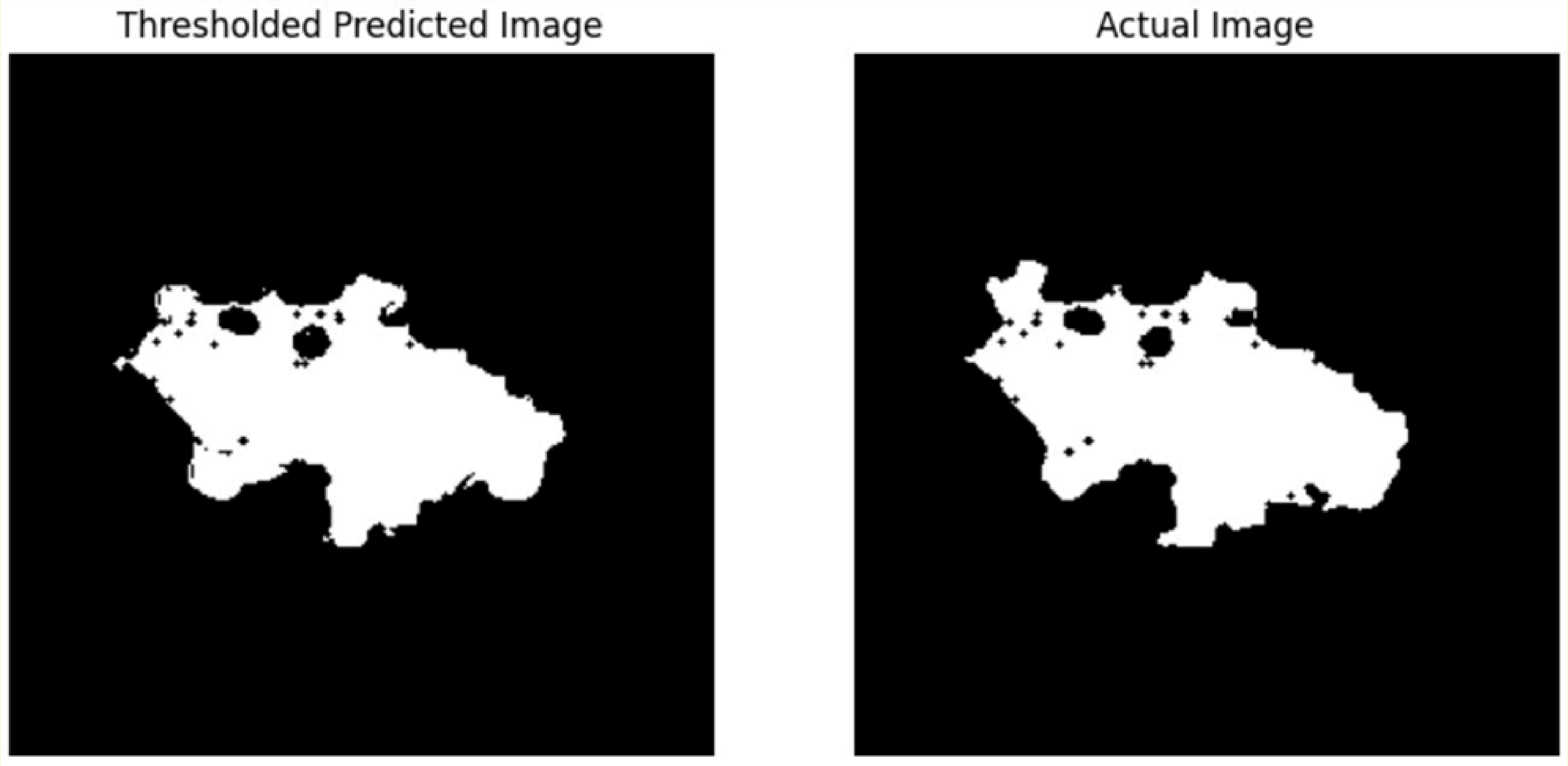
- Lack of “ground truth” for assimilation
- Sparse Observation and Background Data
- Size of Images in Training Data Set
- Variation of wildfire cycles (only ignition)

Mitigation Strategy

- less complex models to manage computational load
- exploration of undersampling ignition-cycles
- balanced DA

Objective 1

Build **RNN** surrogate wildfire predictive model

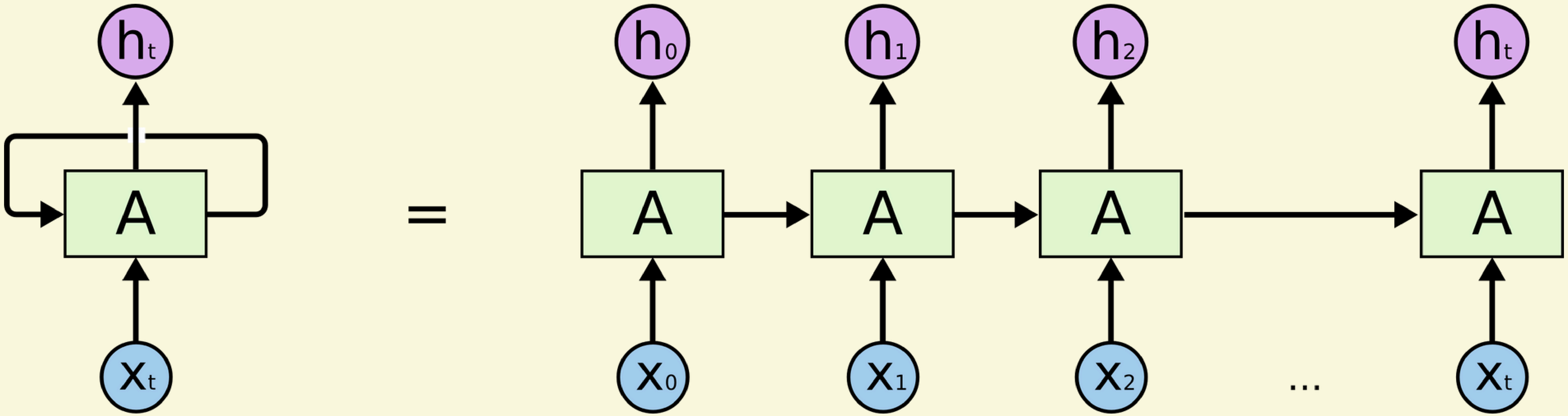


Our plan

Build a surrogate model using RNN

Forecast with Ferguson fire background data

Compare forecast results with satellite data



https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Trialled methods

Flatten LSTM (Linear)

- Basic model for exploration
- Weak in capturing non-linear information

Convolutional LSTM

- Capture both spatial and temporal correlations
- Provide some robustness to variations
- Computational complexity

Methods

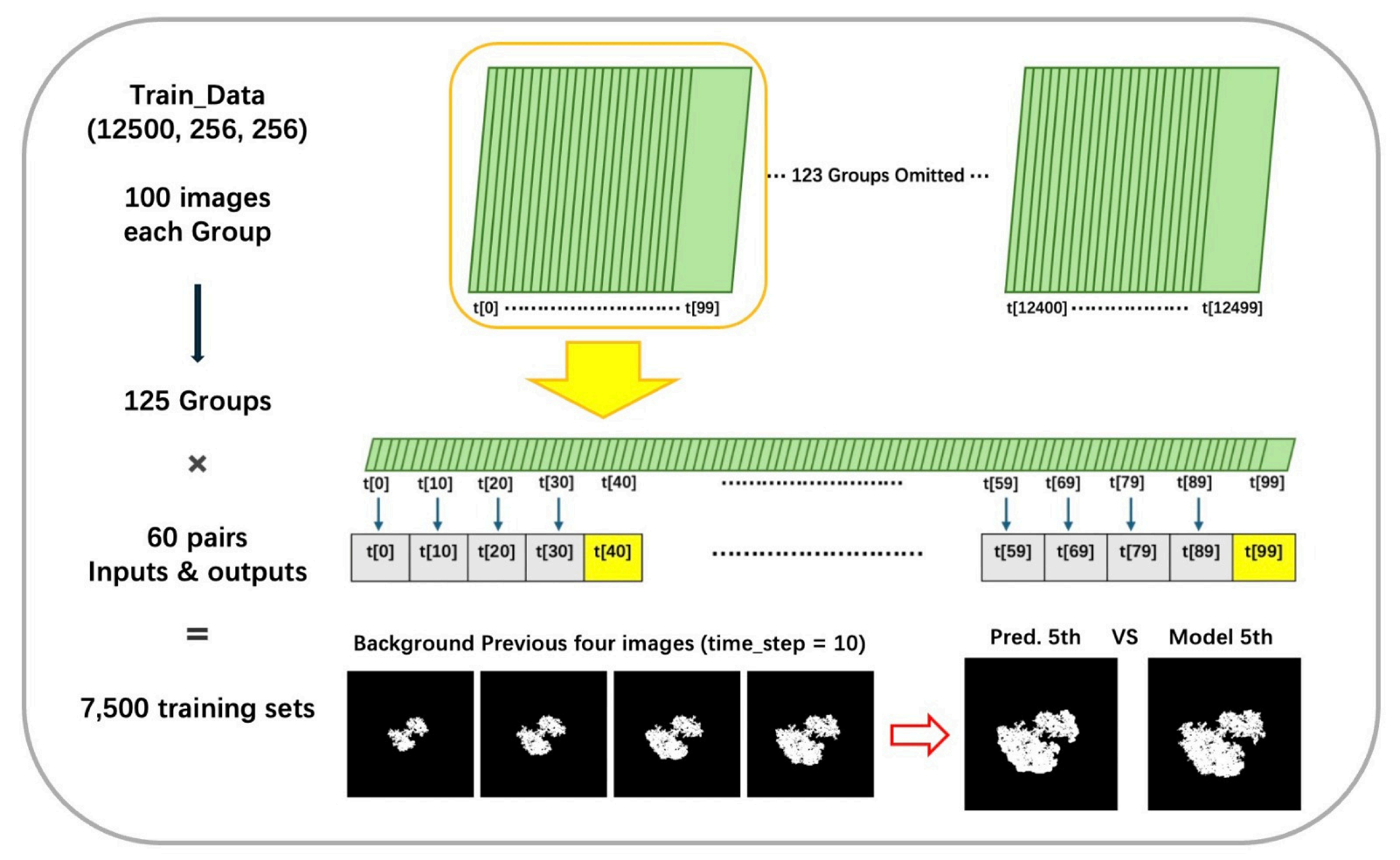
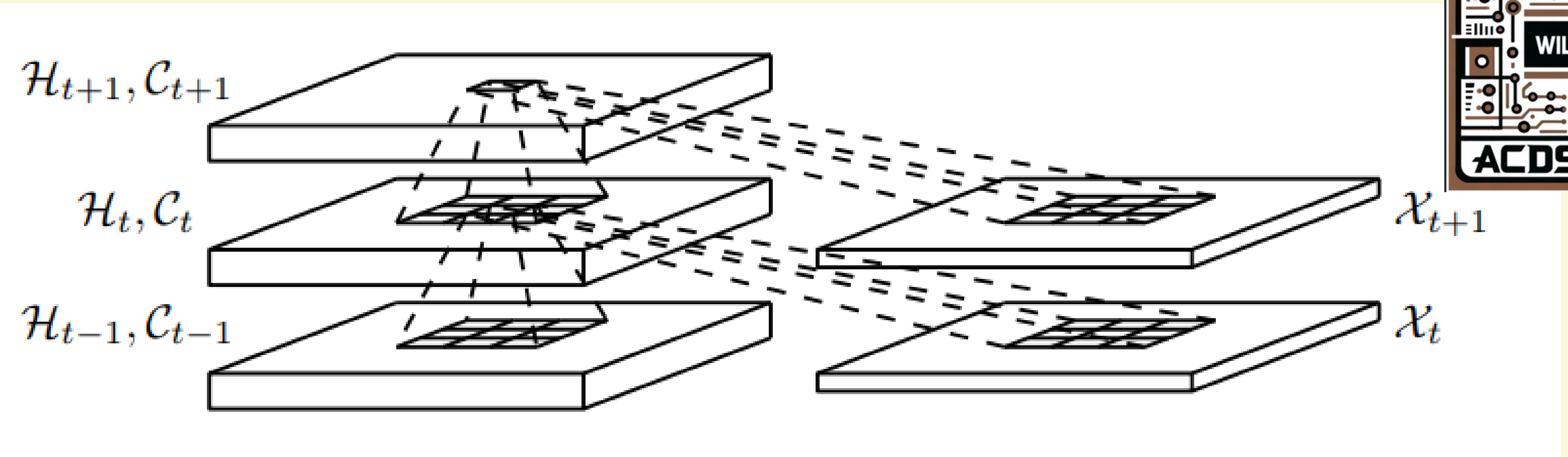
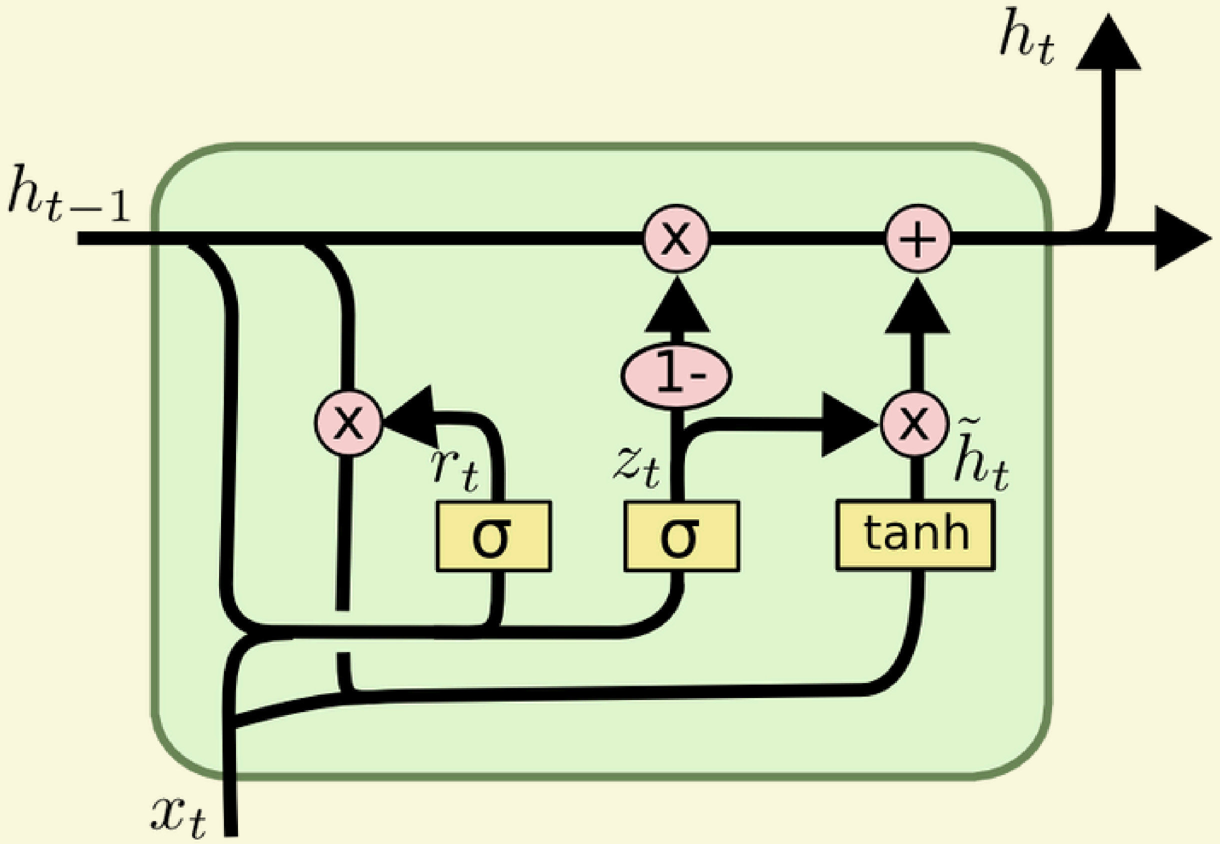


Image produced with BioRender

Final Model: ConvLSTM



- Overcomes the limitations of traditional RNNs and CNNs by combining CNNs and LSTMs
- Spatial convolution and temporal loop operation
- Takes long-term dependence of spatial and temporal correlation into account

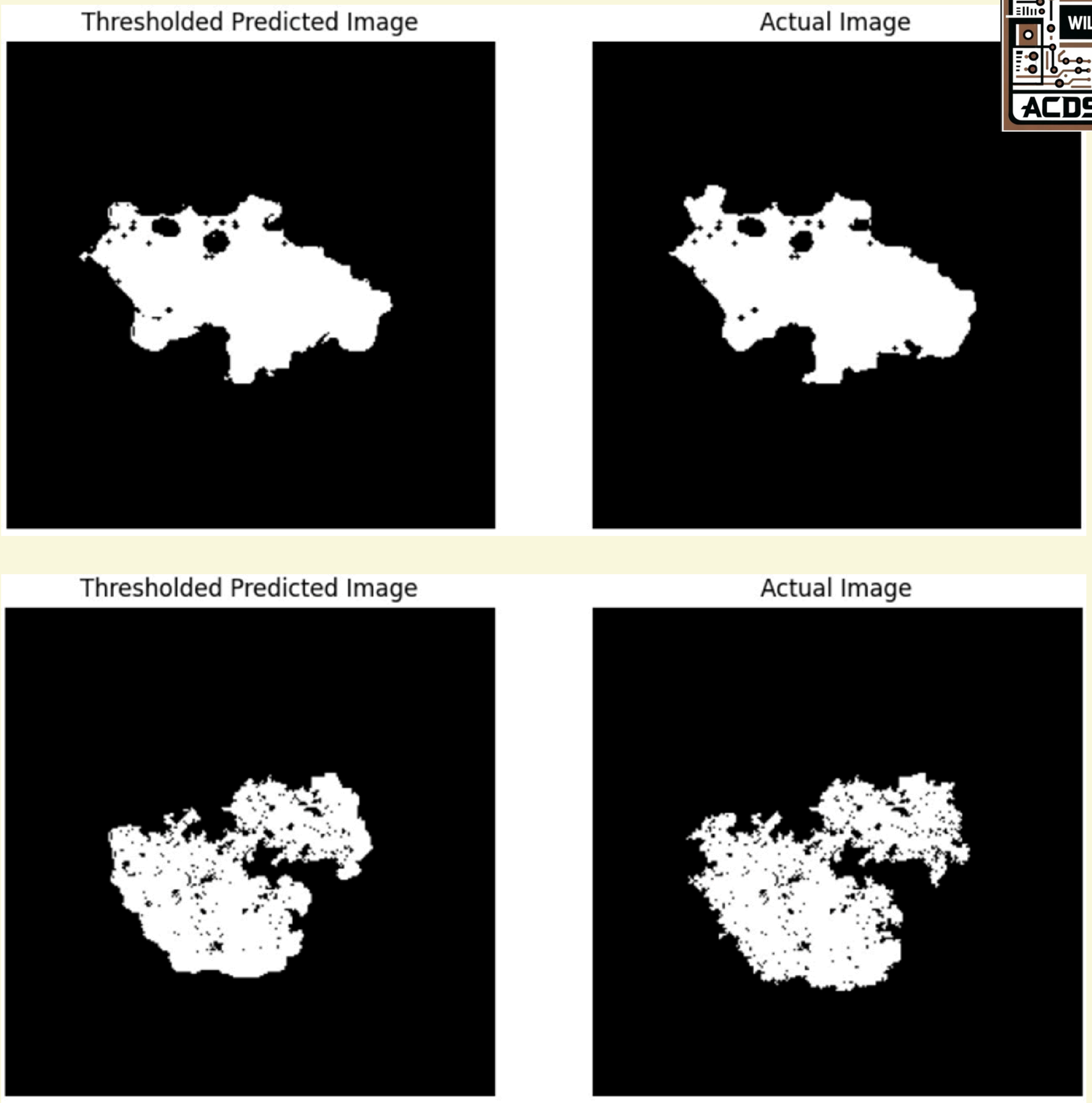


Shi, X et al, 2015

https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Results

- Compute MSE between the predicted and real model images
- **MSE:** 0.01759
- Indicates that the model is excellently trained and generates predicted images that are very similar.



Objective 1 - Conclusion

Challenges

- Requires large data sizes
- Difficulty in data pre-processing
- High consumption of computing resources

Model -Strengths

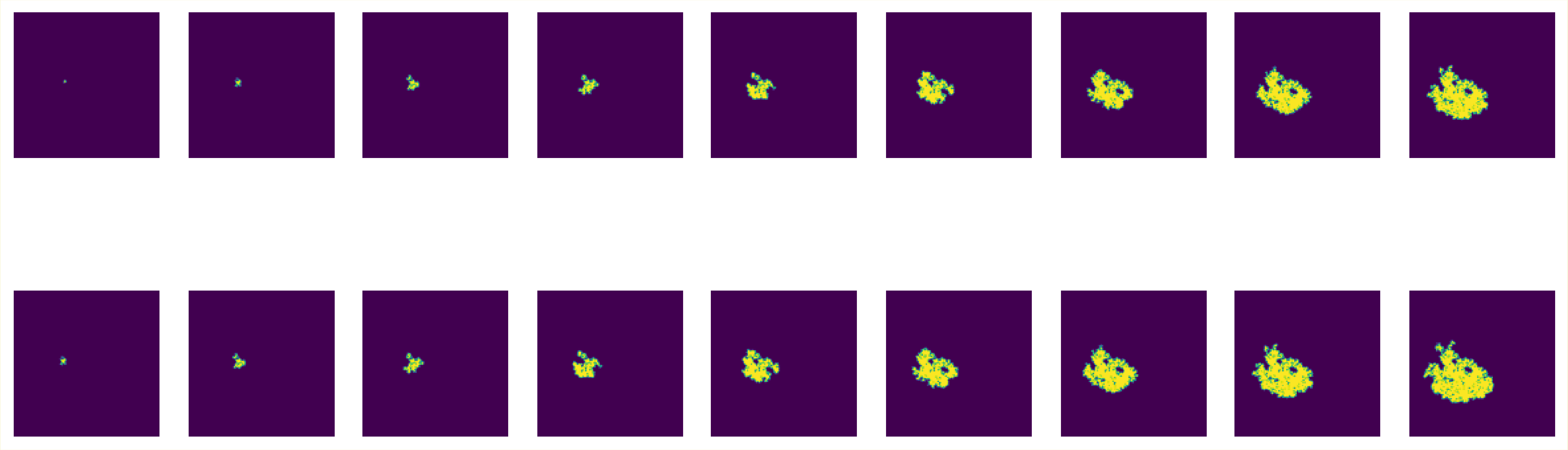
- Combines spatial and temporal features
- Effective handling of spatial and temporal data
- Convolution reduces the parameter counts

Model -Weaknesses

- Complex structure difficult to debug
- Unstable training process

Objective 2

Build a **Generative AI** surrogate wildfire generative model

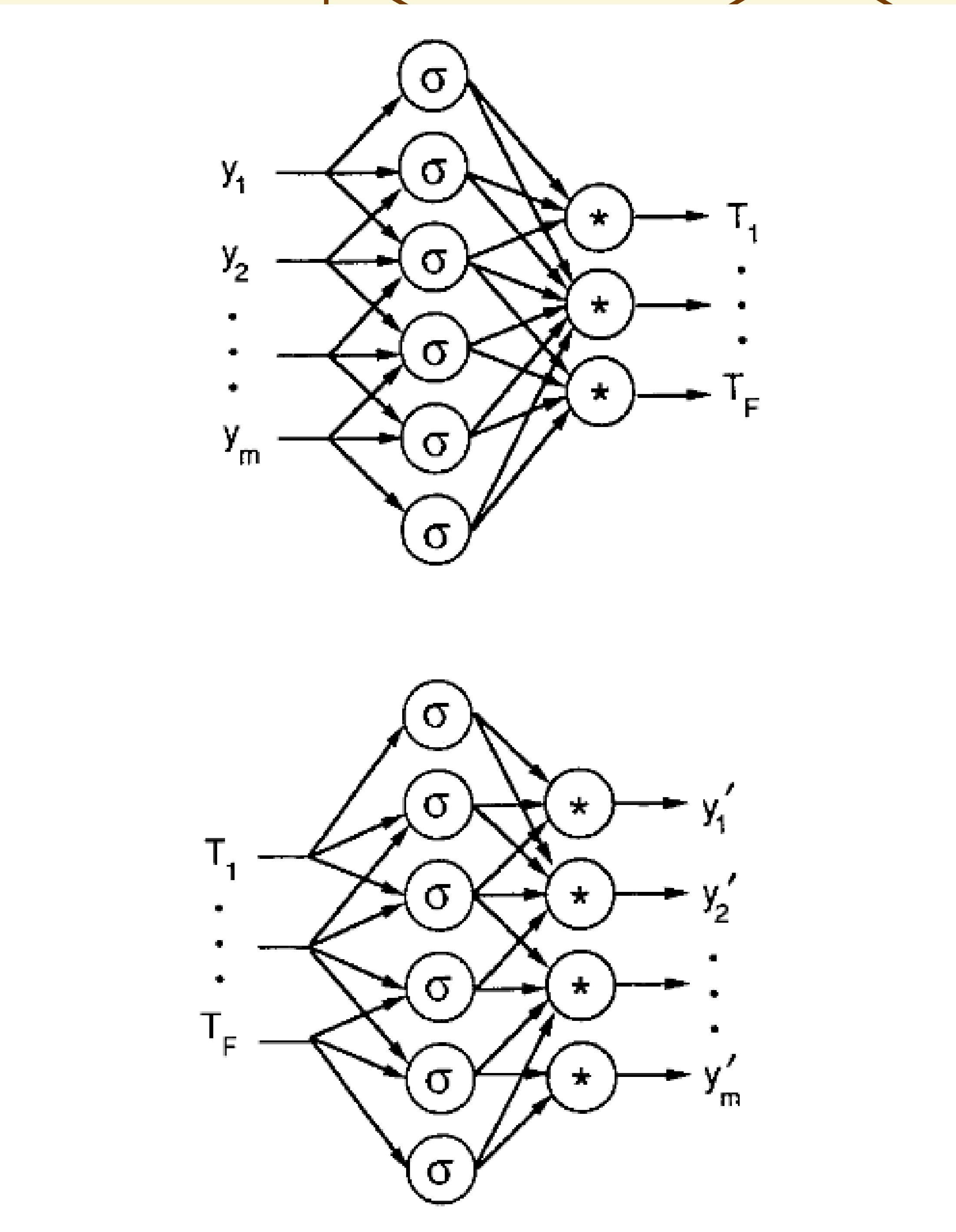


Our plan

Build a surrogate model using generative AI

Forecast with Ferguson fire background data

Compare forecast results with satellite data



Kramer, M.A. (1991)

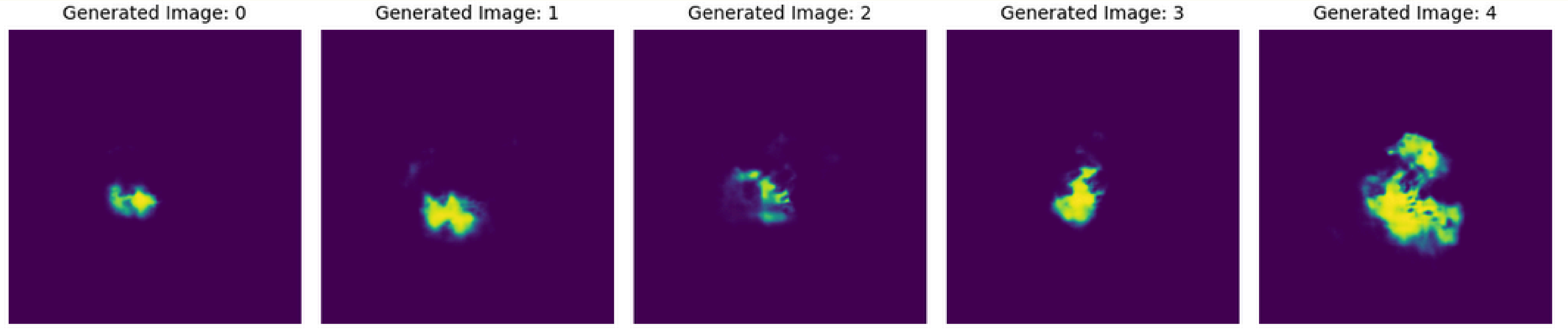
Latent space considerations

Creating latent space that captures temporal dynamics of the system

- In order to better capture temporal features we tried a number of strategies:
  - Using a sequence of images in order to predict the next sequence
  - Using only every 10 timestamps from training data
  - Conditioning with timestep by including a time encoder directly in model.

Creating 'well behaved' latent space

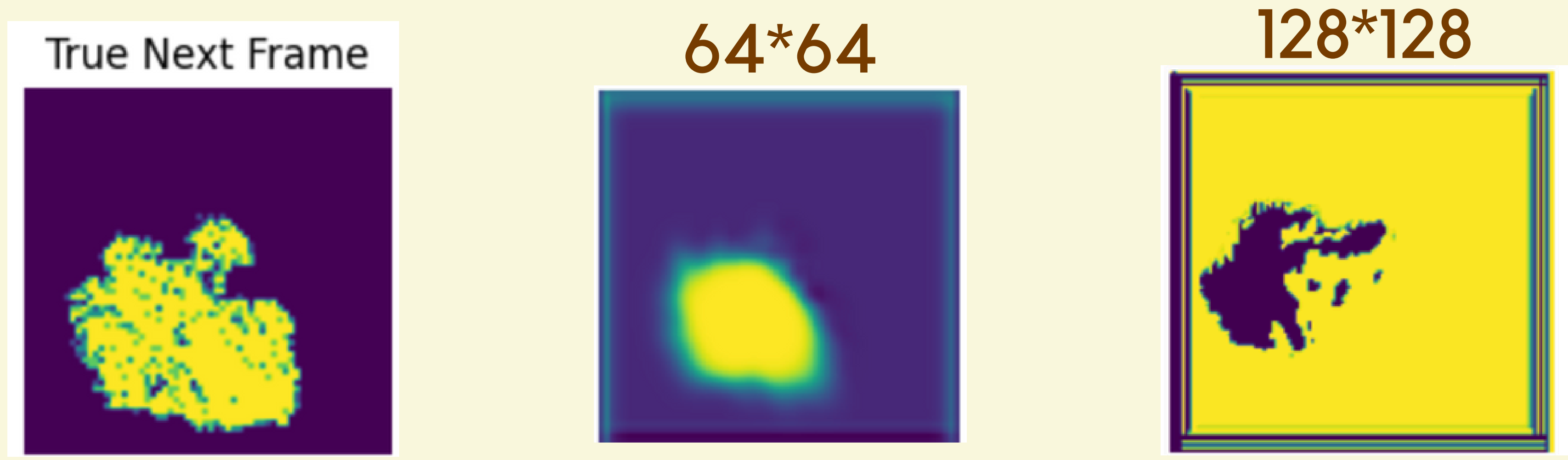
- Creating a good reconstruction model is not too difficult with Autoencoder
- Task relies on a latent space that you can sample from
- Experimented with VQ-VAE to create more 'real' discretized latent space
- Tried using weighting schedules for kl divergence loss



Trialled methods

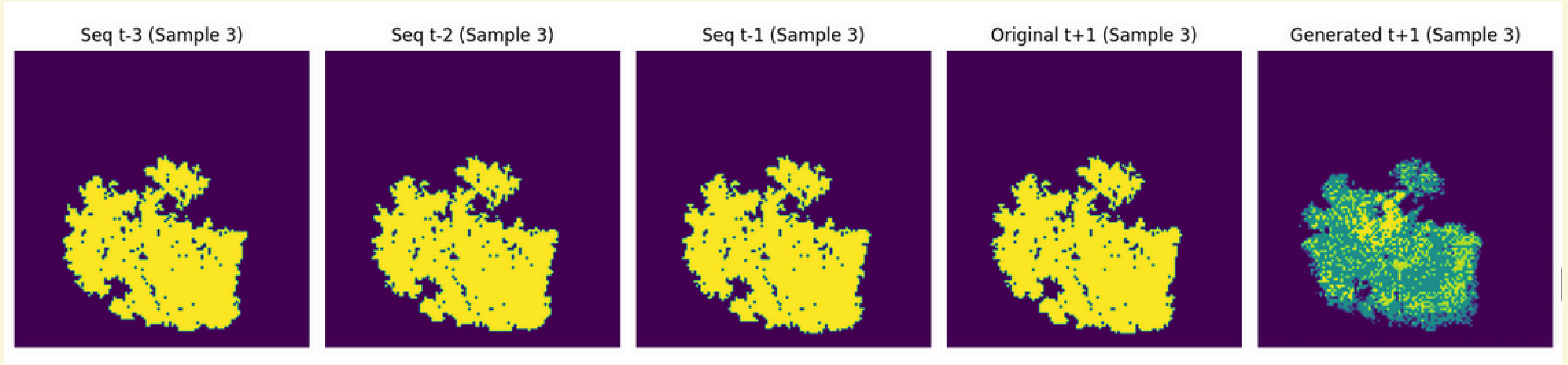
Temporal GAN

- Designed to deal with sequential data
- MSE for 64*64 was 0.259
- MSE for 128*128 was 0.530



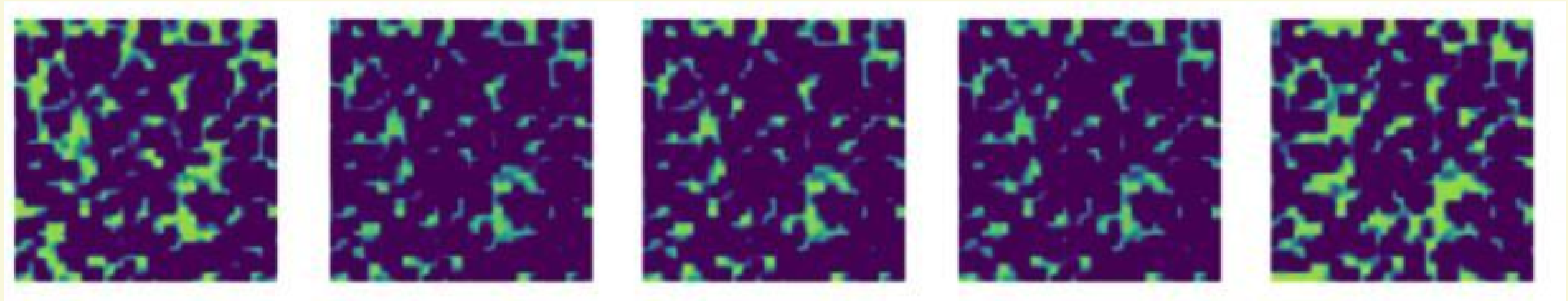
LSTM VAE

- Uses LSTM layers to manage sequential data
- MSE: 0.096



VQ-VAE

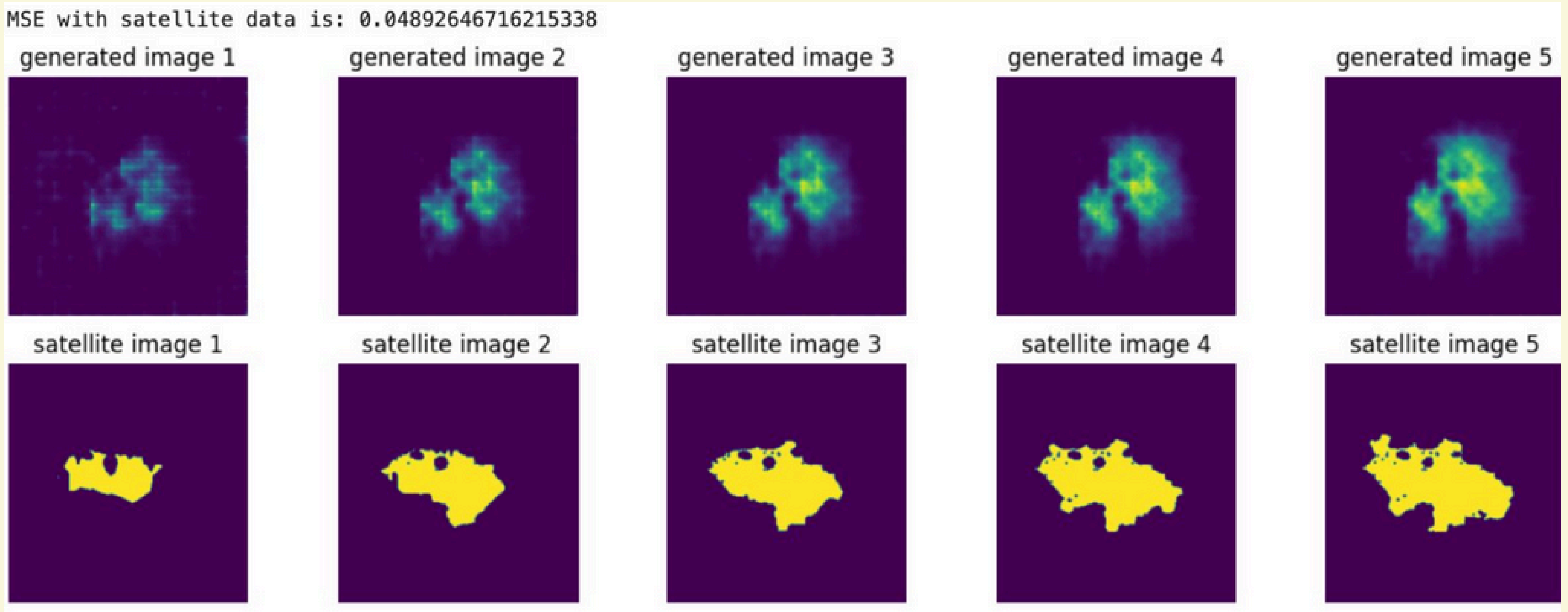
- Uses a discrete latent space with vector quantisation
- Effective in high-quality reconstructions
- MSE: much higher than other methods



Trialled methods continued..

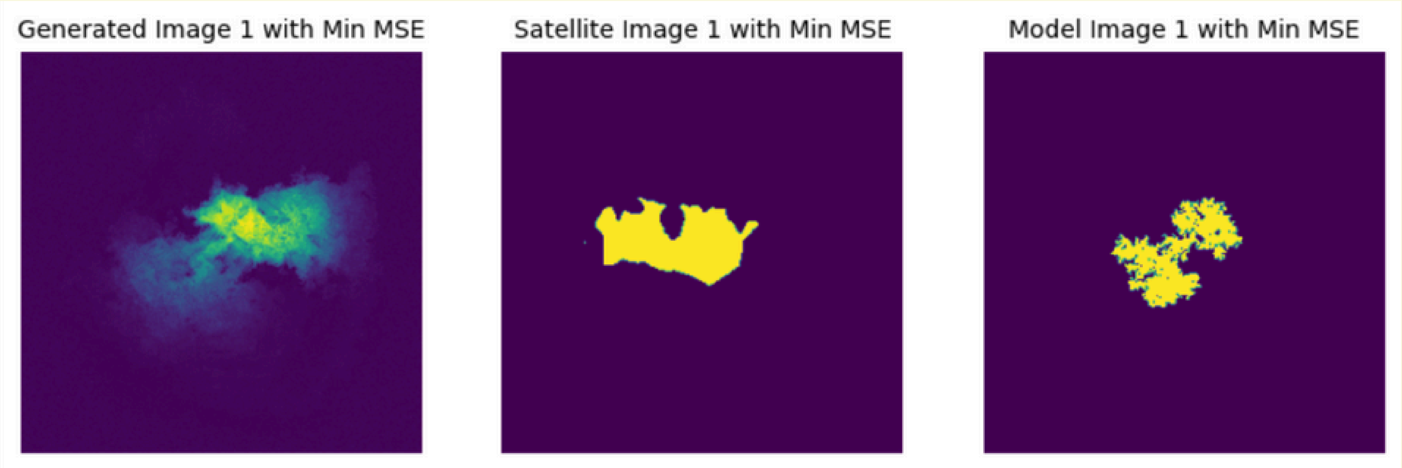
Convolutional (2D&3D) VAE

- Uses convolutional layers which benefits images
- Captures spatial hierarchies and features
- 2D MSE 0.051
- 3D MSE: 0.049



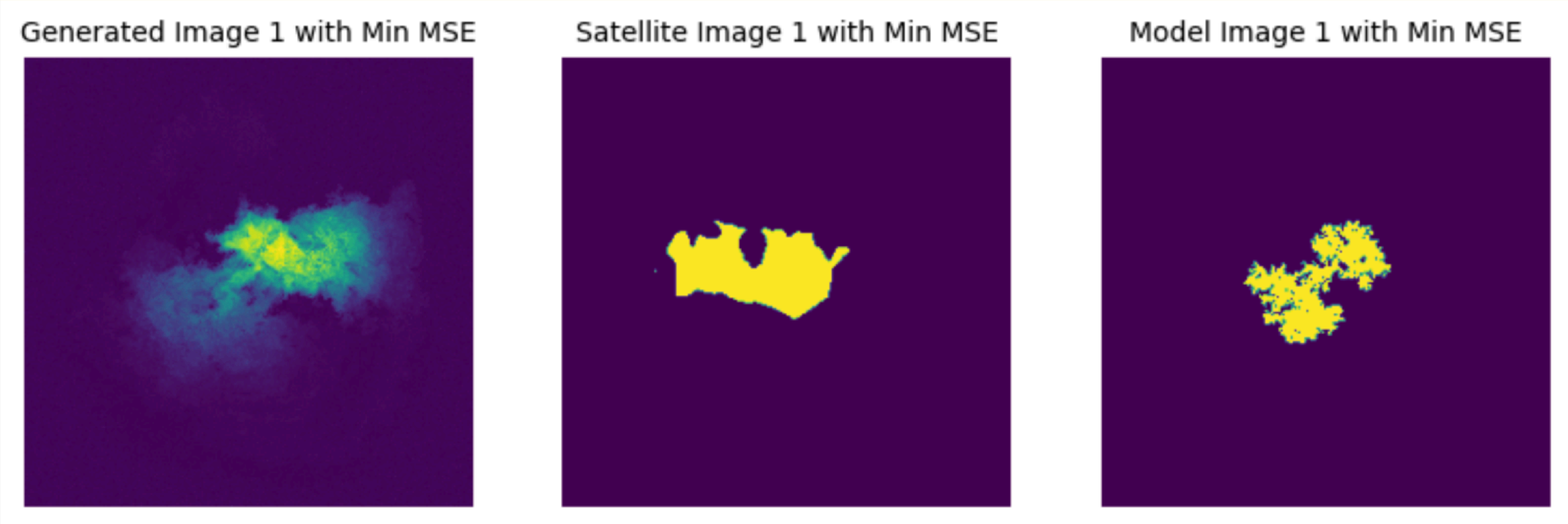
Linear VAE

- Basic concept
- May not perform well on complex datasets
- MSE: 0.0401



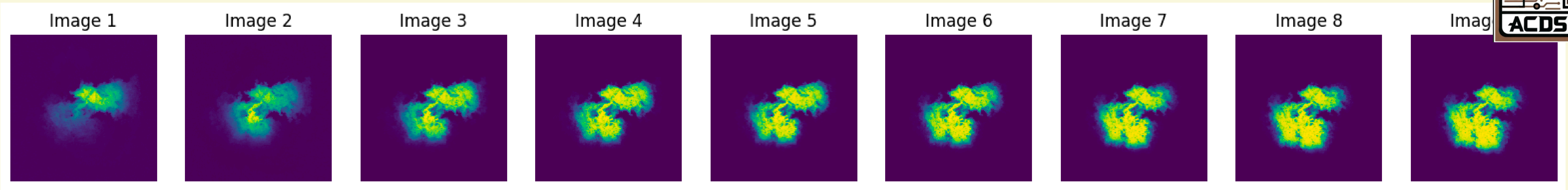
Final Model - Linear VAE

- Sufficient to capture the necessary features
- High Resource Efficiency
- Quicker experimentation and iteration



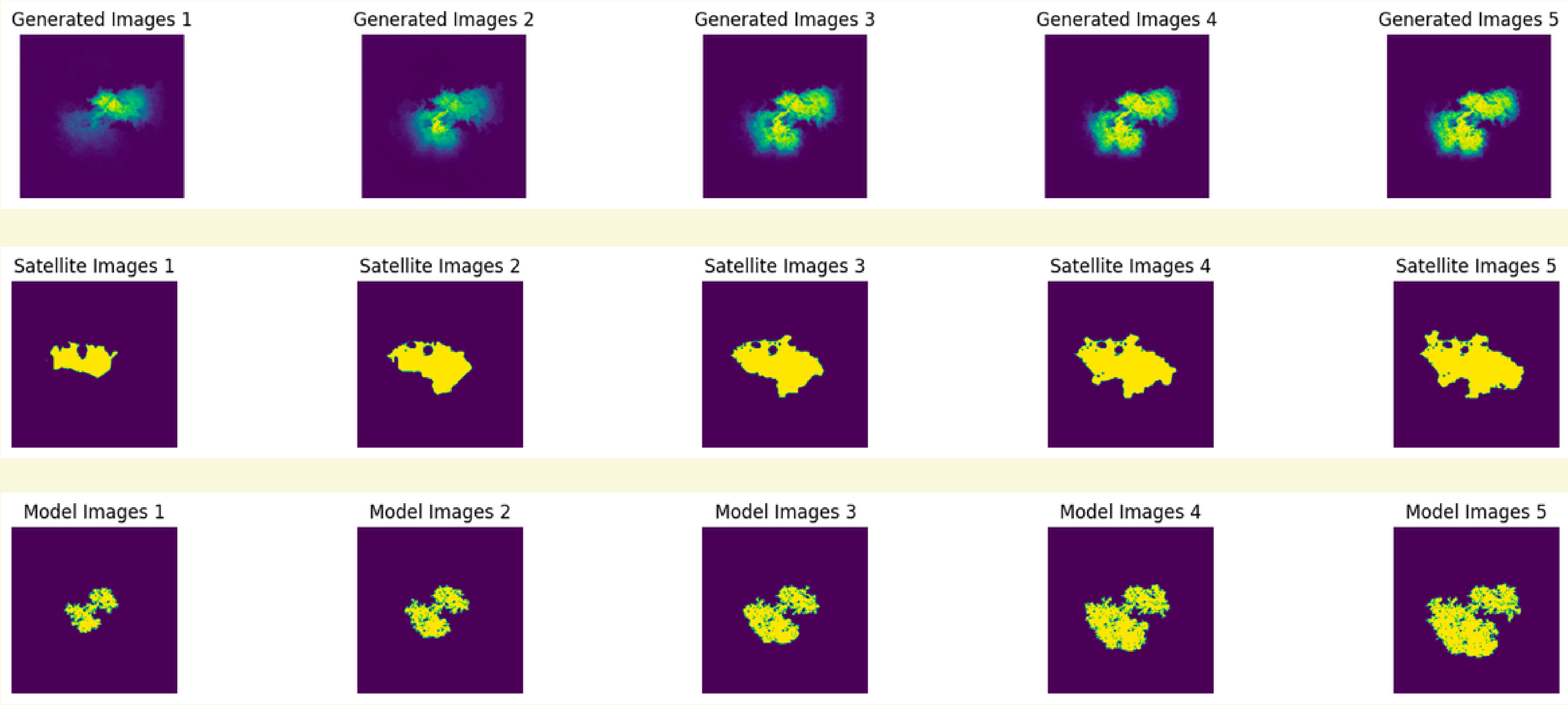
MSE(with satellite data): 0.04

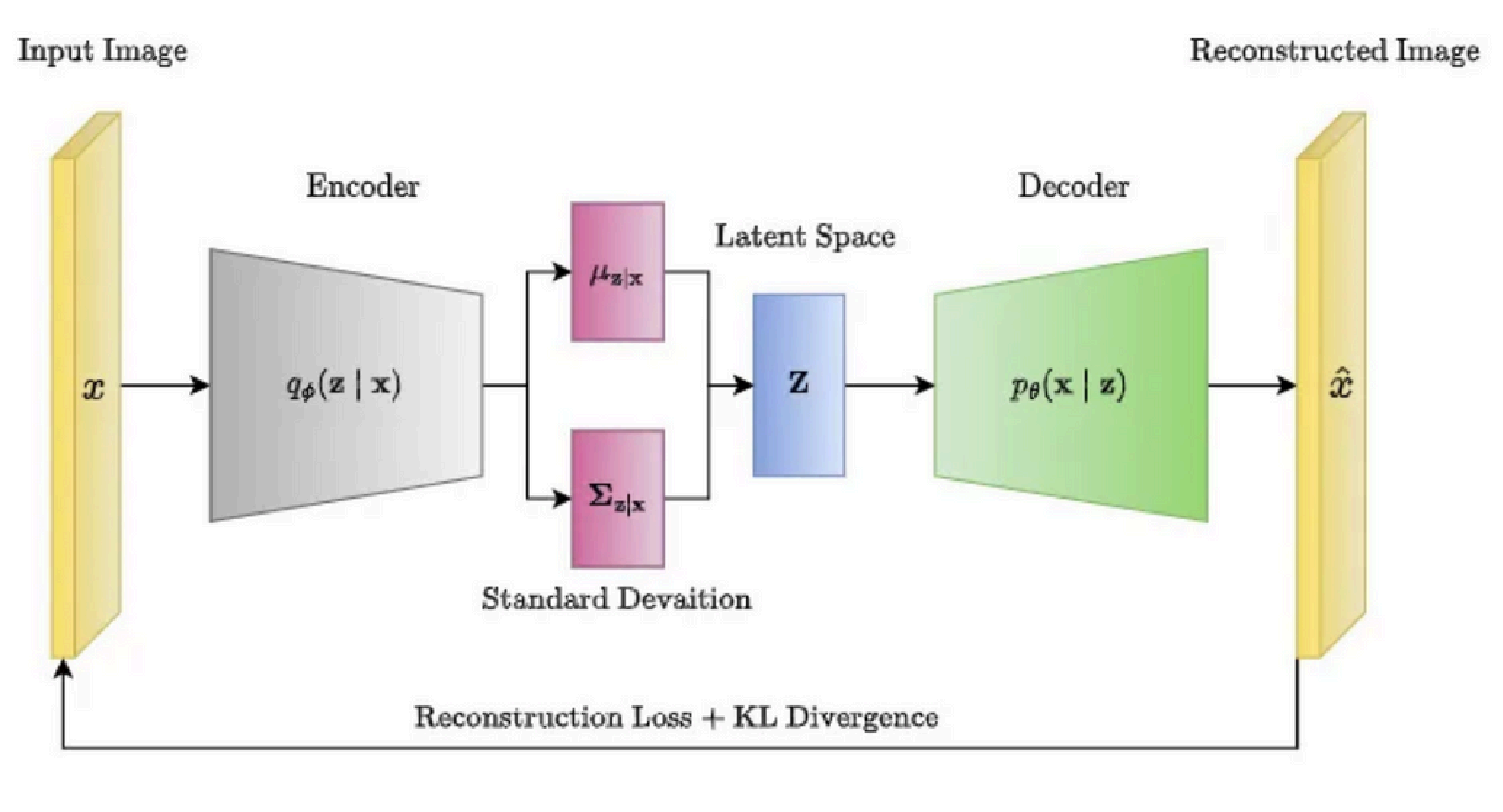
Generate Data with Time Steps



1st is decoded by **randomly sampling** an initial latent vector **from the latent space**

Each subsequent image is generated based on the **re-encoded latent vector** of the **previous image.**





https://viso.ai/deep-learning/synthetic-data-ai-training-solution/

Objective 2 - Conclusion

Challenges

- Handling Temporal Dynamics and Sequence Prediction
- Model Complexity and Computation Time

Model -Strengths

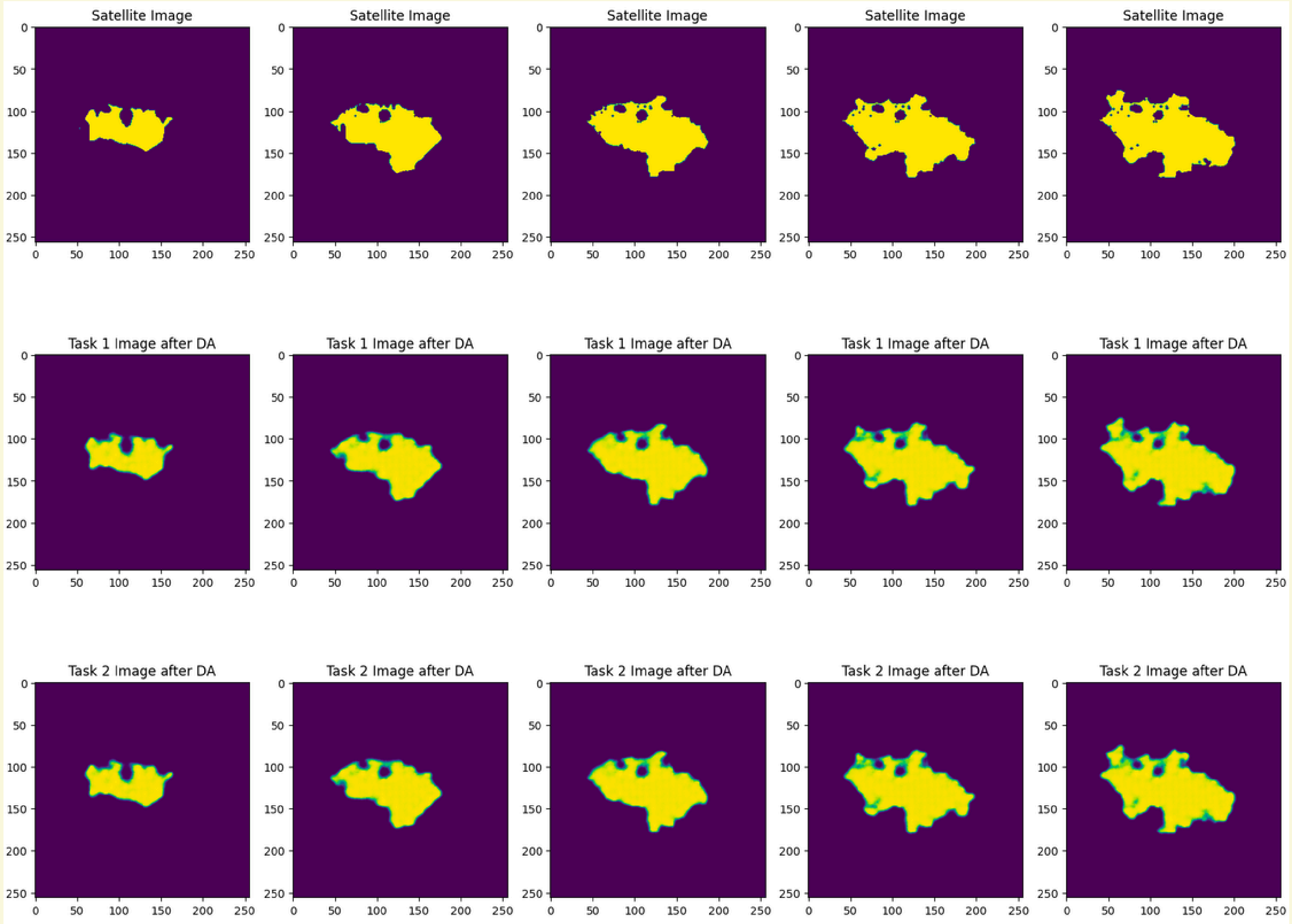
- Fast and simple
- Effective forecasting
- Low computational cost
- Scaleable
- Interpretability

Model -Weaknesses

- Limited complexity
- Incapable of learning from distant past patterns.
- Limited Capacity for Complex Patterns

Objective 3

Perform **data assimilation** on objectives 1 & 2



Our Plan

Compute error covariance matrices

Background data (Matrix B)

Satellite data (Matrix R)

Perform DA in a reduced space with satellite data

Compare pre- and post-DA results

Trialled methods

Compression Method

PCA

- Limited to linear transformations
- Might not capture non-linear relationships

CAE

- Generally provides higher reconstruction quality for image data
- More flexibility in terms of architecture

Data Assimilation Methods

Kalman Filter

- Effective for small datasets

Ensemble Kalman Filter

- Performance might be sensitive to the size of the ensemble

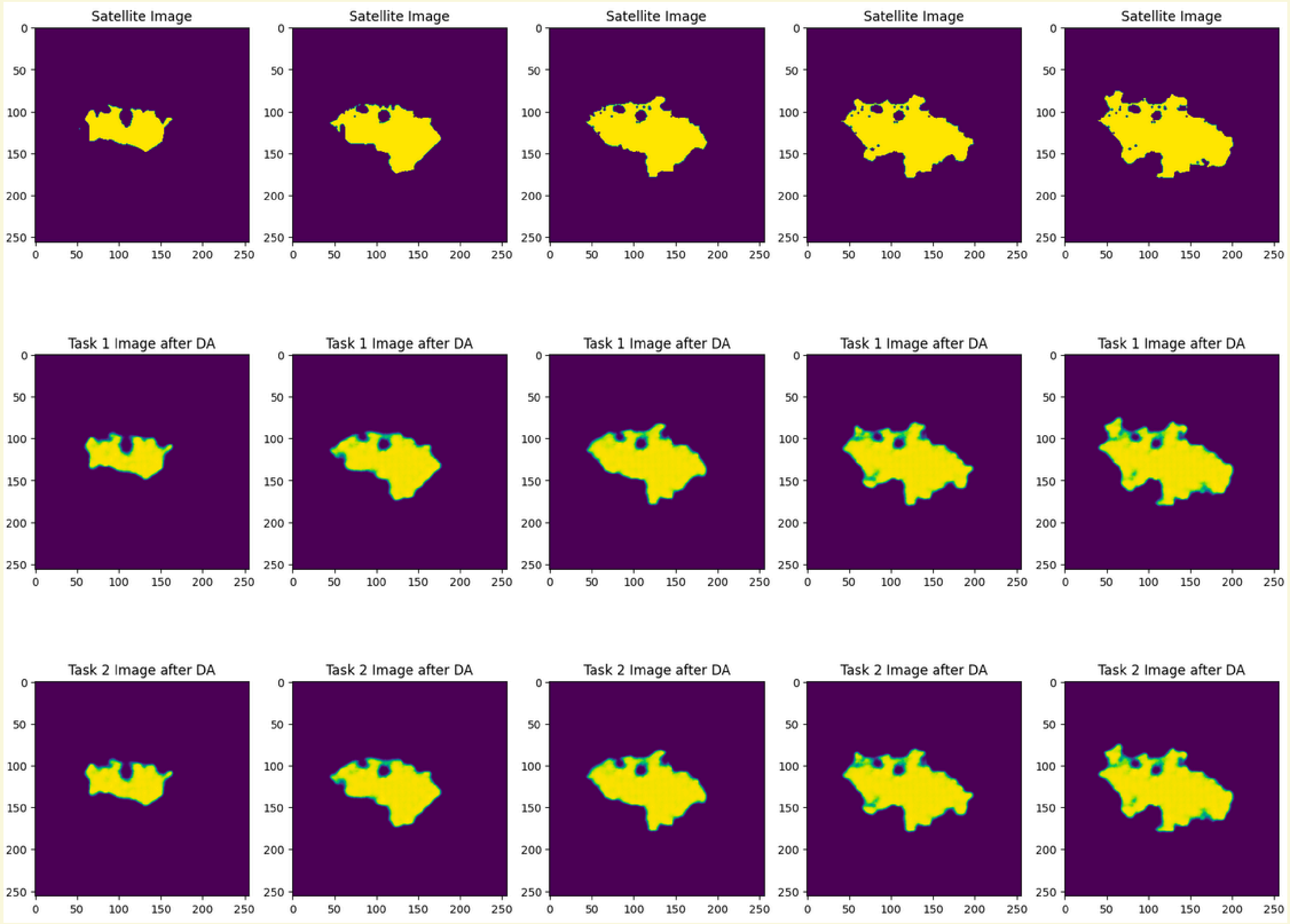
3DVar and 4DVar

- Computational inefficiency
- Complex error structures

Final Choice

- CAE for Data Compression
- Calculate R-matrix using covariance of all 5 images
- Grid search to find optimal R and B weighting factors
- Apply Kalman Filter for Data Assimilation

	MSE before DA	MSE after DA
Objective 1 (RNN)	0.073	0.0039
Objective 2 (Gen AI)	0.066	0.0039



Objective 3 - Conclusion

Challenges

- Difficulty in determining the R and B matrices and their weights
- Not sure if satellite data is ‘ground truth’

DA -Strengths

- Simple to implement
- Suitable for small dataset
- Easily extended to other complex (EnKF)

DA -Weaknesses

- Limits applicabilities in complex systems

Sustainability and Good Practices

- Packaging
- Classes and Functions
- Test Cases and Code Coverage
    - GitHub actions
    - pytest
- Documentation
    - Sphinx
    - README.md
    - Doc-strings, lint
- Pull Requests and Code Review
- License
- Agile methodology

References

Kramer, M.A. (1991). Nonlinear principal component analysis using autoassociative neural networks. AIChE Journal, 37(2), pp.233–243. doi:https://doi.org/10.1002/aic.690370209.

Olah, C. (2015). *Understanding LSTM Networks.* [online] Github.io. Available at: https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

S. Cheng, Y. Guo and R. Arcucci, "A Generative Model for Surrogates of Spatial-Temporal Wildfire Nowcasting," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 7, no. 5, pp. 1420-1430, Oct. 2023, doi: 10.1109/TETCI.2023.3298535.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W. and WOO, W. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. [online] *Neural Information Processing Systems.*

Klingler, N. (2024). *Synthetic Data: A Model Training Solution.* [online] viso.ai. Available at: https://viso.ai/deep-learning/synthetic-data-ai-training-solution/ [Accessed 24 May 2024].

Thank you

Sarah Nesti, Aoife Robertson, Shrreya Behll, Hongxuan Jiang,
Robert Smith, Yixin Li, Yibing Chen, Peiyi Leng