iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82747

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Hybrid Swin-CNN Autoencoder for Arctic Sea Ice Anomaly Detection Using Monthly Passive Microwave Satellite Observations

Ouadghiri Ouafae[1], Jingming Xia[2], Ibrahim Khaled Q.F[3]
School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, China

Abstract: Arctic sea ice is important for the global climate system, polar ecosystems, and polar regions' navigation. Precise recognition of abnormal patterns of Arctic sea ice concentration is significant to observe climate variability and detect the unusual change in polar regions. In this study, we propose a Hybrid Swin-CNN Autoencoder for Arctic sea ice anomaly detection based on monthly satellite observations of the NOAA/NSIDC Climate Data Record of passive microwave sea ice concentration. Monthly sea ice concentration in the northern hemisphere from 2010 to 2024 are used, data in 2010–2020 are used as the training set and data in 2021–2024 are used as the test set. Our model has the combination of convolutional encoder-decoder layer to do the local spatial feature reconstruction and Swin-style window attention layer for discovering the global spatial feature in the sea ice concentration distribution. Comparison of the proposed method To verify our method, four comparison models were evaluated: statistical climatology baseline model, CNN Autoencoder, Light Swin Autoencoder, and Swin-CNN refinement variant. Experimental results show that the proposed Hybrid Swin-CNN Autoencoder outperforms the other models in every evaluation measure, for example, the error measure (MSE = 0.000143, MAE = 0.003391, RMSE = 0.011673, SSIM = 0.994328), indicating that our proposed model is able to better reconstruct the pixel-level value and preserve the spatial structure of sea ice than the baseline models. Error maps also verify that high-error zones mainly happen at or near sea ice boundaries and marginal ice areas, showing that the reconstruction-error-related anomaly maps are meaningful for localizing abnormal sea ice patterns.

Keywords: Arctic sea ice, anomaly detection, autoencoder, Swin Transformer, convolutional neural network, sea ice concentration, NSIDC, deep learning.

I. INTRODUCTION

Arctic sea ice is one of the most sensitive indicators of climate variability and change to environmental conditions in the Arctic. Variations in sea ice concentration influence ocean–atmosphere heat exchange, surface albedo, the polar environment and ecosystems, and human activities (such as shipping and resource development). Given the fast environmental changes in the Arctic, continuous observations of the sea-ice conditions are also very important. Satellite passive microwave measurements are available to provide long-term and consistent sea-ice concentration records, making them very attractive and useful for climate monitoring and data-driven modelling. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration provides daily and monthly observations of sea ice concentration in both polar regions on a 25 km polar stereographic grid, for temporal coverage starting October 1978 and ongoing to the present [1]. Traditional sea ice monitoring techniques are often based on climatology comparison, thresholding, and statistic analysis of anomalies. These techniques are useful because they are interpretable and implementable; however, they cannot fully characterize the complex spatial structures, nonlinear variation and local changes of the sea ice concentration field. In particular, Arctic sea ice anomaly detection needs to be sensitive to spatial morphology features such as displacement of ice edges, change of marginal ice zone, irregular sea ice concentrations, etc. Thus, deep learning models are becoming attractive for sea ice study as they can directly learn the spatial feature from gridded satellite observation. Recent studies have demonstrated the potential of deep learning for sea ice forecasting and analysis. For instance, IceNet demonstrated the use of probabilistic deep learning for the seasonal Arctic sea ice forecasting and was trained using climate simulations and observations to forecast monthly maps of sea ice concentration up to six months ahead [2]. More recent studies also attempted to use U-Net-based and Transformer-based deep learning models for Arctic sea ice concentration prediction, demonstrating that deep learning architectures can capture the spatial dependence and are able to improve the forecasting skill relative to traditional approaches [3],[5]. These demonstrations demonstrate that deep neural architectures can be useful not only for forecasting but also for reconstruction-based anomaly detection in sea ice concentration fields.

Autoencoders are particularly suitable for unsupervised anomaly detection. The basic way that an autoencoder learns to reconstruct normal input patterns is by using an encoder-decoder structure. If the model is primarily trained based on historical data, the reconstruction errors for abnormal models or spatial models that occur less frequently will be larger. These errors can then be used as anomaly scores. Classical autoencoders have been extensively applied for representation learning and dimensionality reduction [6] and convolutional autoencoders further develop this concept for image-like spatially structured data to learn local spatial features. CNN-based Autoencoders, on the other hand, can be very good at reconstructing the local texture, but may not be very well suited to modelling the broader spatial dependencies across the Arctic sea ice field.

Another approach to spatial relationship modelling is used by Vision Transformers. The Swin Transformer proposed shifted-window self-attention that it uses hierarchical shifted windows for local window attention and shifted windows for cross-window attention, but it limits the computation of attention to local windows [7]. It is more efficient for image-like data than global self-attention, and is widely used as a good vision backbone. For sea ice anomaly detection, the concept is appealing because maps of sea ice concentration in the Arctic also feature the local sea ice boundary patterns as well as spatial patterns on a larger scale. But an AE based solely on the Transformer could lose fine local details if the feature representation is compressed too much.

In order to overcome these shortcomings, this study introduces a Hybrid Swin-CNN Autoencoder for Arctic sea ice anomaly detection. The idea behind the proposed model is to use convolutional layers to reconstruct the local space and use window attention blocks in the bottleneck layer with the model of Swin. CNN encoder and decoder captures the local sea ice boundary details, and Swin-style attention blocks enhance spatial dependency modelling. Skip connections are also gated to recover spatial detail while minimizing unwanted direct input to output copying. This model is trained with a combination of MSE, L1 and SSIM based reconstruction losses to optimize for both pixel accuracy and structural similarity.

This paper has the following main contributions:

1) Monthly satellite data obtained by the National Snow and Ice Data Center (NSIDC) is suggested for the detection of Arctic sea ice anomalies from a non-supervised Hybrid Swin-CNN Autoencoder.
2) The proposed architecture consists of CNN-based local reconstruction and the window attention mechanism proposed in Swin, which enhances the spatial representation of the sea ice concentration fields.
3) A thorough comparison is carried out with the statistical climatology, CNN Autoencoder, Lightweight Swin Autoencoder, and Swin-CNN refinement bases.
4) The experimental results of the monthly Arctic sea ice concentration data from 2010 to 2024 demonstrate that the proposed model performs well on the mean square error, mean absolute error, root mean square error and SSIM.
5) Reconstruction error visualization is also consistent with the high reconstruction error in areas with meaningful sea ice boundary and marginal ice zone features.

The rest of this paper is organized as follows. Section II covers the previous research efforts on sea ice monitoring, deep learning in sea ice analysis, autoencoder based anomaly detection and vision Transformers. In Section III, the dataset description, pre-processing procedure and anomaly detection method based on reconstruction are provided. In Section IV, the proposed Hybrid Swin-CNN Autoencoder architecture is presented. The results of the experiments, a comparison with the baseline methods and the discussions about the anomaly visualisations are presented in Section V. Lastly, Section VI ends the paper and provides directions to future research.

II. RELATED WORK

The use of satellite passive microwave observations for monitoring Arctic and Antarctic sea ice has been extensive since it can cover the entire ice-covered region with continuous footprint over wide areas under most weather and illumination conditions. Sea ice concentration is one of the important parameters for the description of the sea ice conditions in the polar areas, and long-term sea ice concentration data are critical for monitoring sea ice conditions and performing anomaly analysis. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration: This data set includes daily and monthly sea ice concentration data for the polar regions at 25 km resolution measured on a polar stereographic grid from 1978 until the present day [1]. The algorithm used in the product is based on a well-established passive microwave retrieval algorithm, the NASA Team algorithm, and the Bootstrap algorithm, which have been used for sea ice concentration estimation for many years [8], [9]. The conventional method of sea ice anomaly analysis is to compare the current observations with the climatology of the past. This method is easy to interpret and straightforward, but it may not be sufficient to represent nonlinear and spatially complex patterns at sea ice boundaries and marginal ice zones.

The ability of deep learning to learn spatial patterns directly from gridded satellite observations has made it recently an interesting possibility for sea ice prediction and remote sensing analysis. CNNs are well suited for image-like geophysical fields, since they are able to learn local spatial features with the convolutional filters. U-Net like encoder-decoder models have been extensively employed in spatial reconstruction tasks, due to their ability to extract the contextual features along with spatial recovery [10]. IceNet was found to be effective in seasonal Arctic sea ice forecasting in the field of sea ice research [2]. Recent research includes CNN-based and U-Net-based models as well as Transformer-based models for forecasting Arctic sea ice concentration. For instance, Kim et al. explored prediction of Arctic sea ice concentration (ASIC) in a long time series using deep learning and Ren et al. presented SICNetseason, a Transformer-based model for sea ice concentration prediction in seasonal timescales within the Arctic region [11], [5]. In these studies, deep learning models were found to be better at learning the spatiotemporal patterns in sea ice data than traditional statistical models.

Autoencoders are widely used in unsupervised anomaly detection as they have an encoder-decoder architecture for data reconstruction. If it is trained by historical or normal patterns, an autoencoder will typically reconstruct common patterns successfully, and patterns that are unusual, will result in higher reconstructed errors. Such reconstruction errors can then be used as anomaly scores. CNN Autoencoders are good for the spatial reconstruction, as local details like edges and textures are preserved. Although CNNs are predominantly based on local receptive fields, they may be less effective at capturing the spatial scale of the Arctic sea ice domain, broader than that captured by a local receptive field. This constraint is important because sea ice anomalies can be local and spatially structured.

Alternative method for modelling spatial relationship using self-attention is the Vision Transformers. Swin Transformer proposed a hierarchical vision Transformer with shifted window attention that significantly decreases computational complexity by calculating the attention in local windows with respect to the cross window attention computation. This design applies to data that also has "image-like" characteristics, in that it is efficient and good for spatial modelling. But, an autoencoder based on Transformer can miss out on the fine local details if the compression is quite heavy. Hybrid CNN-Transformer models have become attractive since CNNs can reconstruct local features effectively and Transformers can model spatial dependency. This idea is adopted in this study by introducing the proposed Hybrid Swin-CNN Autoencoder that utilizes CNN encoder-decoder layers for local sea ice reconstruction and Swin-style attention blocks in the bottleneck to boost the spatial representation.

Anomaly detection based on reconstruction should take into account numerical error and spatial structure for evaluation. MSE, MAE and RMSE values are used to assess pixel level reconstruction accuracy and the Structural Similarity Index Measure is used to assess the structural similarity between original and reconstructed images [12]. This study uses all four of those metrics while the detection of sea ice anomaly is not only based on the error of each pixel, but also on maintaining the spatial morphology.

III.DATASET AND PREPROCESSING

A. Dataset and Experimental Split

The NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 6, is used for this study in order to detect sea ice anomalies in the Arctic [8]. Only the Northern Hemisphere monthly sea ice concentration product was made use of. Sea ice concentration maps are selected for the period from January 2010 to December 2024 to provide a total of 180 maps per month. A sea ice concentration field is stored in each NetCDF file as the variable cdr_seaice_conc_monthly.

The size of the original monthly maps is 448 x 304 on the polar stereographic grid. Since the overheads would be small on the machine, the size of each monthly field was reduced to 64 × 64 pixels by area based interpolation. All sea ice concentration (SIC) values were capped between 0 and 1 (where 0 indicates open water or no sea ice and 1 indicates ice cover) and missing and invalid values were filled with zero. The final dataset was represented as:

$$X \in R^{180\times1\times64\times64}$$

Monthly samples of 180, input channel of 1 and a resized spatial resolution of 64×64.

The data was split chronologically to create a realistic future-testing situation. The data period of 2010 to 2020 were used for training and 2021 to 2024 were used for testing, providing 132 and 48 respective samples. A split was chosen based on the chronological order of the years since the sea ice concentration is a time-dependent climate variable and because the test years are closer to the present climate conditions than the years in the random split.

B. *Reconstruction-Based Anomaly Detection*

The approach of this research is anomaly detection by reconstruction. Each model is trained to reconstruct historical time series of monthly sea ice concentration maps. In testing, the difference between the input map provided and the reconstructed output is used as the anomaly score. A greater reconstruction error will be made if the monthly pattern is unusual or hard for the model to reconstruct.

The Absolute Reconstruction Error (ARE) is computed pixel-wise as:

$$E = |X - \hat{X}|$$

Where X is the original sea ice concentration map and $\hat{X}$ is the sea ice concentration mask. High-error regions are considered as anomaly-sensitive areas. To fairly evaluate the proposed model, it was compared with four different methods: the statistical monthly climatology baseline, CNN Autoencoder, Lightweight Swin Autoencoder, Swin-CNN Refine Autoencoder. All models used the same training and testing split.

IV. PROPOSED METHODOLOGY

A. *Overview of the Proposed Model*

The method proposed in this study is a Hybrid Swin-CNN Autoencoder for Arctic sea ice anomaly detection. This model brings together the best of convolutional neural networks and Swin-style attention. CNN layers are employed to extract and reconstruct local spatial details (sea ice edges and marginal ice zone structures) and Swin-style window attention blocks are used in the bottleneck to model spatial dependency in the sea ice concentration field.

The proposed model is not restricted to local convolutional receptive fields as is the case with an autoencoder based purely on CNN. It is not as strongly compressed as a Swin Autoencoder into a small latent representation. The model instead learns the intermediate level features, thus being able to capture both local spatial details and wider spatial relationships.

The overall architecture is made up of four main components: CNN encoder, Swin-style attention bottleneck, Gated skip connections and CNN decoder.

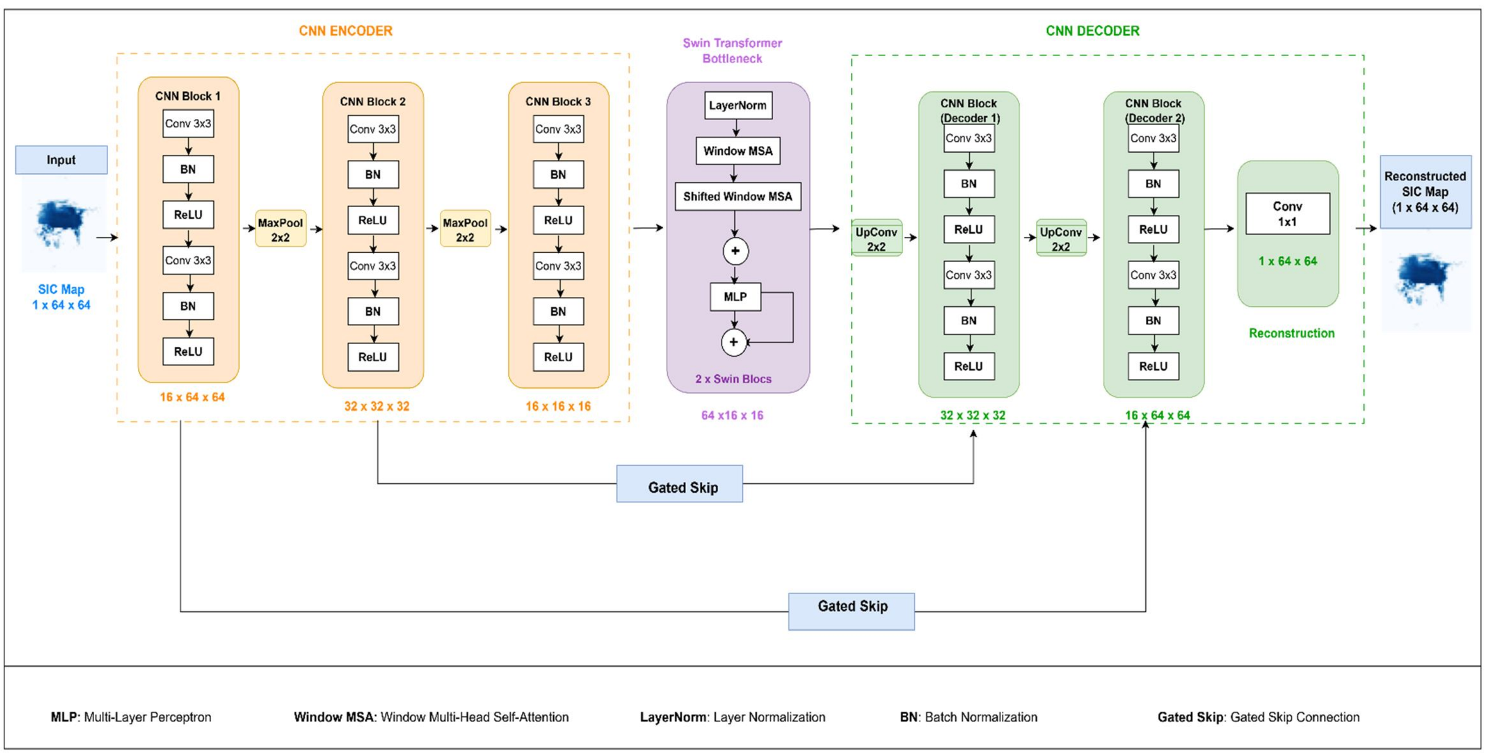


Figure 1 Architecture overview of the proposed Hybrid Swin-CNN Autoencoder for Arctic sea ice concentration reconstruction

B. CNN Encoder
The input sea ice concentration map is first processed to extract local spatial features by the encoder. The shape of the input image is:

$$X \in R^{1\times 64\times 64}$$

The encoder has convolutional blocks as well as downsampling layers. The first convolutional block has a resolution of the original image and identifies low-level spatial features. Then, there are two downsampling stages that decrease feature map size and increase channels. This enables the model to acquire a more abstract spatial representation.

The feature maps generated by the encoder are 64×64, 32×32 and 16×16. The feature map of size 16×16 is then input into the Swin-style attention bottleneck.

C. Swin-Style Attention Bottleneck
The main attention component of the proposed model is the bottleneck. The model employs window-based attention inspired by the Swin Transformer instead of using the full global self-attention. The feature map is divided into local windows, and the self-attention is computed within local windows. Shifted-window attention is also used to allow information exchange between neighbouring windows.

This design is appropriate for sea ice maps in the Arctic region because the Arctic has local structures in its sea ice patterns as well as spatial relationships. Local attention facilitates focus on the sea ice boundaries and marginal ice regions and shifted windows enable interaction among nearby regions.

D. Gated Skip Connections
A popular technique in encoder-decoder systems for reconstructing spatial information is skip connections. But, in the case of anomaly detection, the direct skip connections can enable the model to mimic the input too well, and this may result in a loss of anomaly sensitivity. To tackle with this issue, the proposed model introduces gated skip connections.

The gated skip connection is a learnable attention-like gating of the encoder features before they are fed to the decoder. This will help the model to recover useful spatial information while minimizing uncontrolable copying from the source image. The skip operation can be represented with the following:

$$S_g = S.\sigma(W_s(S))$$

In the above equation, S is the encoder skip feature, W_s is a 1x1 convolution, σ is the sigmoid activation function and S_g is the gated skip feature.

E. CNN Decoder
The decoder reconstructs the sea ice concentration map from the bottleneck representation. It uses transposed convolution layers to upsample the feature maps from 16×16 back to 64×64. The decoder adds upsampled feature map and gated skip feature for the same stage of the decoder at every decoding stage.

The last layer is an output layer with a sigmoid activation function to obtain sea ice concentration reconstructed values in the range [0, 1]. The output reconstruction is described as:

$$\hat{X} \in R^{1\times 64\times 64}$$

where $\hat{X}$ is the reconstructed sea ice concentration map.

F. Reconstruction Loss Function
A hybrid reconstruction loss was adopted to train the proposed model. The loss is a sum of Mean Squared Error, Mean Absolute Error and a structural loss based on SSIM. MSE helps reduce pixel-wise squared reconstruction errors, MAE improves average absolute accuracy, and the SSIM-based term encourages structural similarity between the original and reconstructed maps.

The total loss is the sum of:

$$\mathcal{L} = \mathcal{L}_{MSE} + \mathcal{L}_{L1} + \mathcal{L}_{SSIM}$$

The mean squared error loss function, mean absolute error loss function, and structural similarity loss function are presented, respectively, as $\mathcal{L}_{MSE}$, $\mathcal{L}_{L1}$, and $\mathcal{L}_{SSIM}$. This loss functions jointly to force the model to reproduce both the pixel level values and the structure of sea ice morphology well.

G. Baseline and Ablation Models

Four comparison methods were put in place to assess the effectiveness of the proposed model. The statistical baseline is based on monthly climatology for the training period as the reference for reconstruction. CNN Autoencoder: the performance of a pure convolutional reconstruction model is evaluated. The Lightweight Swin Autoencoder is a test of a reconstruction model based on Transformer. To investigate if the Swin encoder can provide a better reconstruction with the addition of a convolutional refinement, an ablation variant, the Swin-CNN Refine Autoencoder is used.

The proposed Hybrid Swin-CNN Autoencoder is different from these autoencoders in that it incorporates Swin attention into a CNN bottleneck alongside gated skip connections. This design enables the model to maintain the local detail and also to reflect the spatial dependency.

V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section we show the experimental results on the proposed Hybrid Swin-CNN Autoencoder and compare the performance of the proposed Hybrid Swin-CNN Autoencoder with a simple statistical baseline, a regular CNN Autoencoder, a Lightweight Swin Autoencoder, and the Swin-CNN Refine Autoencoder. each method was tested using data from the same stretch of time, specifically from 2021 through 2024. For our measurements, we relied on MSE, MAE, RMSE, and SSIM.

A. Quantitative Results

The reconstruction results of all the tested methods are presented in Tabel I. The statistical baseline is a simple climatology reference for each month, CNN Autoencoder, Lightweight Swin Autoencoder, and Swin-CNN Refine Autoencoder are deep learning baselines and ablation models. The proposed Hybrid Swin-CNN Autoencoder was the best performer on all the four measures.

TABLE 1
PERFORMANCE COMPARISON OF THE PROPOSED MODEL AND BASELINE METHODS

Model	MSE	MAE	RMSE	SSIM
Statistical Baseline	0.002840	0.011950	0.052120	0.931780
CNN Autoencoder	0.000732	0.011115	0.026807	0.912180
Lightweight Swin Autoencoder	0.001245	0.008856	0.034718	0.949300
Swin-CNN Refine Autoencoder	0.001591	0.009018	0.038946	0.956530
Hybrid Swin-CNN Autoencoder	0.000143	0.003391	0.011673	0.994328

The proposed Hybrid Swin-CNN Autoencoder achieved an MSE of 0.000143, MAE of 0.003391, RMSE of 0.011673, and SSIM of 0.994328.

The values indicate that the proposed model gives the best reconstruction out of all the models that were evaluated. Compared with the CNN Autoencoder, the proposed model reduced MSE from 0.000732 to 0.000143, MAE from 0.011115 to 0.003391, and RMSE from 0.026807 to 0.011673. In addition, it also increased SSIM from 0.912180 to 0.994328, which demonstrates that it preserved the spatial sea ice structure much better.

A relatively high SSIM was obtained with the statistical baseline, due to the fact that the monthly climatology can reflect the general seasonal variations of Arctic sea ice. On the other hand, the MSE and RMSE values were substantially larger than those of the deep learning models indicating that it is unable to perfectly replicate the spatial changes from month to month. The CNN Autoencoder gives the best results in terms of MSE and RMSE but the lowest result for SSIM, indicating less preservation of the spatial morphology. The Lightweight Swin Autoencoder showed better MAE and SSIM than the CNN Autoencoder with a higher MSE and RMSE. This means that the Transformer-only model managed to capture spatial structure to a larger extent, but some details of the image were lost. The Swin-CNN Refine Autoencoder had a high SSIM and did not have a positive impact on the pixel-level error metrics. By contrast, the proposed Hybrid Swin-CNN Autoencoder model had the best balance between pixel-level accuracy and structural preservation.

B. Effectiveness of the Hybrid Architecture

The results showed that the CNN layers and Swin-style attention are more effective than using the CNN or Swin layers separately. CNN layers can help recover the details of the sea ice boundaries and the patterns in the sea ice at a per-pixel level. Yet, CNNs might not be able to model broader spatial correlations throughout the Arctic domain. Attention in Swin style using models of relationships within local windows and shifted windows to improve spatial representation. The proposed model benefits from local reconstruction ability and spatial dependency modelling by placing Swin-style attention blocks in the CNN's bottleneck of the encoder-decoder structure.

By comparing the proposed model with the Lightweight Swin Autoencoder, the effectiveness of the model is demonstrated, indicating that using the attention mechanism after heavy spatial compression is not enough. The attention introduced at intermediate feature resolution, such as Hybrid Swin-CNN Autoencoder, helps to retain key local features while enabling attention-based spatial interaction.

C. Anomaly Visualization

To assess the ability of the proposed model to detect meaningful anomaly-sensitive regions, in addition to quantitative evaluation, reconstruction error maps were produced. The absolute reconstruction error between the original sea ice concentration map and the reconstructed output was used to calculate the anomaly map.

It is seen that the overall sea ice concentration field is reconstructed correctly, while the visualization results still contain visible error patterns around important spatial structures, which indicates that the proposed model has the ability to reconstruct the sea ice concentration field. The high error areas are primarily located in the area bordering with the sea ice and marginal ice zone. This is relevant since these areas tend to be more variable and are more likely to show variations in sea ice that are not normal. The error maps are not arbitrary and do not contain a lot of zeros, so the model does not just reproduce the input image, but it learns a meaningful representation of the spatial structure of sea ice.

Examples of the sea ice concentration maps used and generated by the proposed model, as well as the reconstruction error heatmaps, are shown in figure 2. The reconstructed maps are very similar to the original sea ice extent, and the error heatmaps show the areas where reconstruction was more challenging.

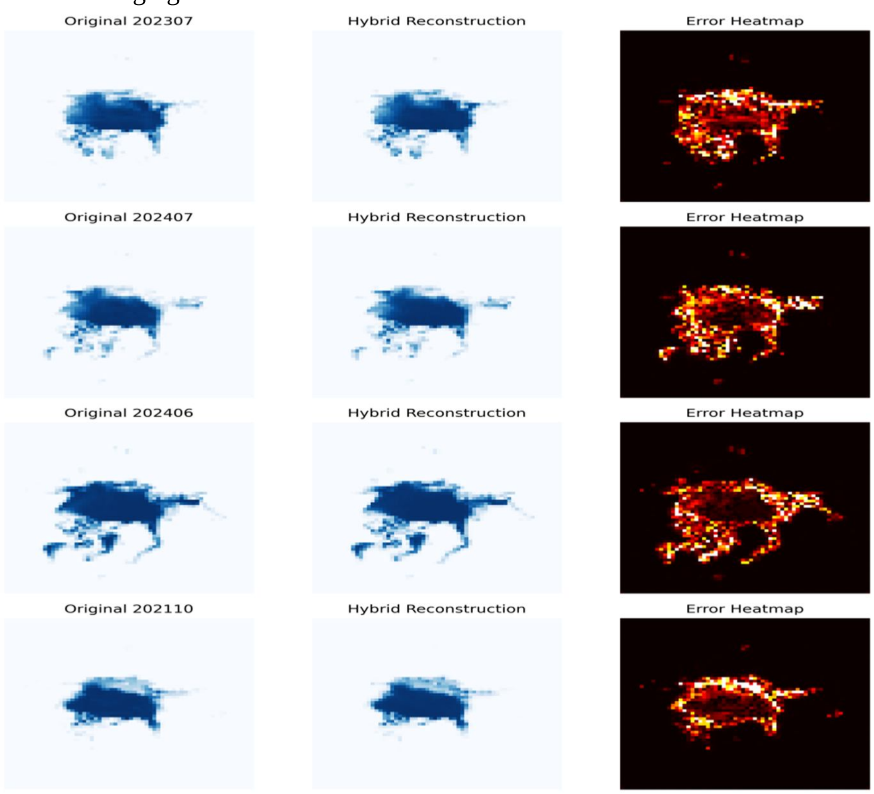


Figure 2 Original Arctic sea ice concentration maps, Hybrid Swin-CNN reconstructions, and reconstruction error heatmaps for selected test months.

To check the real-world relevance of the error areas detected, we placed the sea ice edge outline on top of the error maps. The spots with high errors mostly appeared near the ice edge, which shows that the model is responding to actual sea ice features instead of just random noise in the pixels.

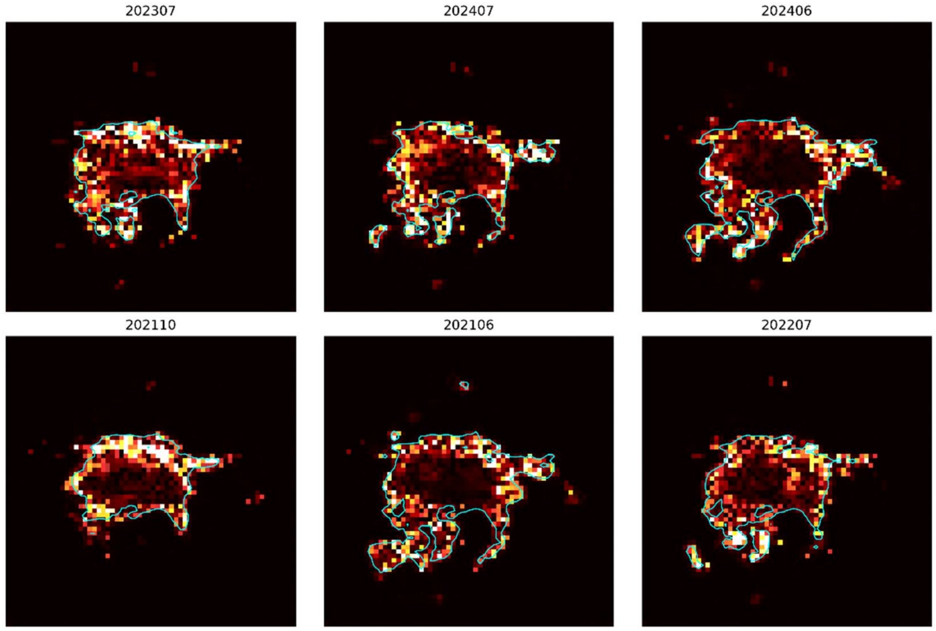


Figure 3 Reconstruction error maps with sea ice edge contours. High-error regions are mainly concentrated near the marginal ice zone and sea ice boundary

D. Discussion

The experimental results show that the Hybrid Swin-CNN Autoencoder works well for detecting anomalies in Arctic sea ice. The model performed best across all four evaluation metrics, indicating it can recreate monthly sea ice concentration maps with good numerical accuracy and structural similarity. This matters because spotting anomalies in sea ice requires not just low pixel errors but also keeping the shape and features like ice edges and border zones intact.

Comparing with the CNN Autoencoder, purely convolutional models can reconstruct pixels well, but they might not keep the spatial structure as intact. Looking at the Lightweight Swin Autoencoder, Transformer-based attention helps capture the structure better, but if the spatial info is compressed too much, it could lose fine local details. The hybrid design we propose addresses these issues by mixing CNNs for local reconstruction with Swin-style attention to handle spatial information at the bottleneck.

The results suggest that combining CNN and Transformer models works well for detecting sea ice anomalies through reconstruction. The model is simple and efficient enough to be trained with limited computing resources, yet still delivers good reconstruction results. This balance makes it a good option for real-world climate monitoring, where both speed and accuracy matter.

VI. CONCLUSION AND FUTURE WORK

This study introduced a Hybrid Swin-CNN Autoencoder to detect anomalies in Arctic sea ice by reconstructing monthly sea ice concentration data from the NSIDC, covering 2010 to 2024. The model combines CNNs for capturing local features with Swin-style window attention in the bottleneck, which helps keep local sea ice patterns intact while also considering broader spatial relationships. Gated skip connections were used to bring back important spatial details without simply copying the input to the output.

The new model was tested against four others: a statistical monthly climatology baseline, a CNN Autoencoder, a Lightweight Swin Autoencoder, and a Swin-CNN Refine Autoencoder. Results from 2021 to 2024 showed that the Hybrid Swin-CNN Autoencoder outperformed them all based on measures like MSE, MAE, RMSE, and SSIM. This suggests the method does better at both reconstructing sea ice concentration at the pixel level and keeping the overall structure.

When looking at where reconstruction errors occurred, they mainly appeared around sea ice edges and marginal zones. This means the model is picking up meaningful anomalies rather than random noise. So, the Hybrid Swin-CNN Autoencoder offers a relatively simple but effective way to spot anomalies in Arctic sea ice using satellite data.

Looking ahead, there are several ways to improve this approach. One option is to use sea ice data with higher resolution to better identify small-scale anomalies. Another is to include other environmental factors like sea surface temperature, air temperature, wind, or ice thickness to create a system that detects anomalies using multiple data types. It might also help to analyze longer time periods to track how anomalies change over time. Finally, testing the method on Antarctic sea ice or other climate anomaly detection problems could show how well it works in different settings.

REFERENCES

[1] W. N. Meier, F. Fetterer, A. K. Windnagel, J. S. Stewart, and T. Stafford, "NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 6," National Snow and Ice Data Center, Boulder, CO, USA, 2026, doi: 10.7265/b18j-z797.

[2] T. R. Andersson et al., "Seasonal Arctic sea ice forecasting with probabilistic deep learning," Nature Communications, vol. 12, no. 1, Art. no. 5124, 2021, doi: 10.1038/s41467-021-25257-4.

[3] C. Palerme *et al*., "Improving short-term sea ice concentration forecasts using deep learning," *The Cryosphere*, vol. 18, pp. 2161–2182, 2024, doi: 10.5194/tc-18-2161-2024.

[4] J. Park, Y. Cho, J.-J. Jeon, J. Park, H.-C. Kim, and S. Hong, "Unicorn: U-Net for sea ice forecasting with convolutional neural ordinary differential equations," Scientific Reports, vol. 15, Art. no. 36330, 2025, doi: 10.1038/s41598-025-20097-4.

[5] Y. Ren, X. Li, and Y. Wang, "SICNetseason V1.0: A transformer-based deep learning model for seasonal Arctic sea ice prediction by incorporating sea ice thickness data," Geoscientific Model Development, vol. 18, pp. 2665–2678, 2025, doi: 10.5194/gmd-18-2665-2025.

[6] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," Science, vol. 313, no. 5786, pp. 504–507, 2006, doi: 10.1126/science.1127647.

[7] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin Transformer: Hierarchical vision transformer using shifted windows," in Proc. IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 9992–10002, doi: 10.1109/ICCV48922.2021.00986.

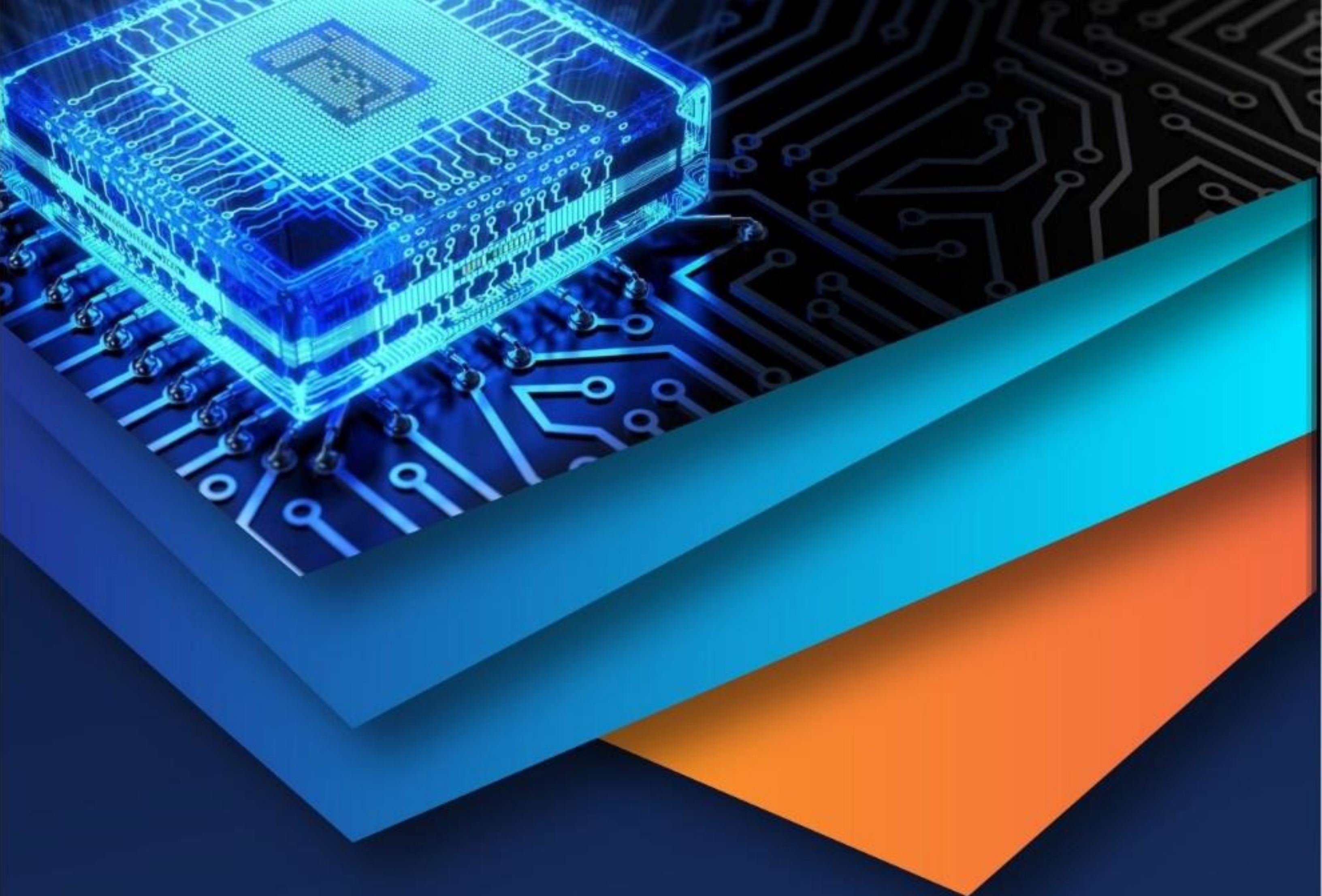
[8] D. J. Cavalieri, P. Gloersen, and W. J. Campbell, "Determination of sea ice parameters with the Nimbus 7 SMMR," Journal of Geophysical Research, vol. 89, no. D4, pp. 5355–5369, 1984, doi: 10.1029/JD089iD04p05355.

[9] J. C. Comiso, "Characteristics of Arctic winter sea ice from satellite multispectral microwave observations," *Journal of Geophysical Research*, vol. 91, no. C1, pp. 975–994, 1986, doi: 10.1029/JC091iC01p00975.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[11] Y. J. Kim, H. Kim, D. Han, J. Stroeve, and J. Im, "Long-term prediction of Arctic sea ice concentrations using deep learning: Effects of surface temperature, radiation, and wind conditions," Remote Sensing of Environment, vol. 318, Art. no. 114568, 2025, doi: 10.1016/j.rse.2024.114568.

[12] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY