Deep Learning-Based Sea Surface Temperature Prediction in Korean Coastal Waters

line 1: 1st Ji-Yeon Kim
line 2: *Department of Artificial Intelligence*
line 3: *Dong-Eui University*
line 4: Busan, Republic of Korea
line 5: jnsp0907@naver.com

line 1: 2nd Ki-Hwan Kim
line 2: *AI Grand ICT Research Center*
line 3: *Dong-Eui University*
line 4: Busan, Republic of Korea
line 5: 15541@deu.ac.kr

line 1: 3rd Young-Jin Kang
line 2: *AI Grand ICT Research Center*
line 3: *Dong-Eui University*
line 4: Busan, Republic of Korea
line 5: 15073@deu.ac.kr

line 1: 4th Seok Chan Jeong
line 2: *Dept. of e-Business, AI Grand ICT Research Center*
line 3: *Dong-Eui University*
line 4: Busan, Republic of Korea
line 5: scjeong@deu.ac.kr

***Abstract*—Sea surface temperature (SST) is a critical variable for monitoring ecosystems, managing fisheries resources, and analyzing climate variability. This study used NOAA OISST data to predict the next 7-day SST from the past 7 days in the Korean coastal waters and compared the prediction performance of various deep learning models. Preprocessing was applied to account for the land-sea boundary, and prediction performance was evaluated using RMSE, MAE, and MSE. Experimental results showed that the U-Net and ConvLSTM models achieved the best prediction performance, highlighting the importance of preserving spatial structure and learning spatiotemporal continuity in SST prediction.**

Keywords—Sea Surface Temperature, Deep Learning, Climate Change, Korean Coastal Waters

I. Introduction

Sea Surface Temperature (SST) is a key variable governing heat exchange between the ocean and the atmosphere[1]. It directly impacts not only climate system variability but also marine ecosystems, coastal environments, and marine industries. SST fluctuations are closely linked to marine heat waves, coastal ecosystem fluctuations, fishing activities, and marine hazards. Therefore, accurate SST predictions play a crucial role in marine environmental management and policy development.

The coastal waters of South Korea comprise a complex marine environment encompassing the West Sea, South Sea, and East Sea, where the Tsushima Warm Current, a tributary of the Kuroshio Warm Current, intersects with the North Korean Cold Current [2]. This results in pronounced seasonal and spatial variability in sea surface temperature, along with frequent SST changes. These characteristics make the waters off South Korea a challenging area for SST forecasting, and reliable predictions of SST fluctuations are essential for ensuring the stability of coastal management and marine activities.

However, in real-world marine environments, rapid fluctuations in SST can lead to various problems. Unexpected changes in water temperature may reduce productivity in aquaculture and fisheries, disrupt coastal ecosystems, and complicate responses to marine hazards. These problems are exacerbated when SST fluctuations cannot be predicted in advance, negatively impacting the entire marine management system.

To effectively mitigate these issues, a systematic approach utilizing SST prediction information based on observational data is required. Specifically, predicting SST changes on a multi-day scale in advance would enable more proactive responses in marine environmental management and decision-making processes. Accordingly, interest in data-driven SST prediction techniques utilizing observational data has recently been increasing.

Research utilizing data-driven approaches to SST prediction has already been conducted in various forms. Studies analyzing SST fluctuations using statistical time series models or regression-based methods have been reported [3, 4], and attempts have been made to learn spatial patterns of SST using convolutional neural networks (CNNs) [5]. Furthermore, studies utilizing recurrent neural network structures such as ConvLSTM to simultaneously consider temporal variations and spatial distributions of SST have been reported [6], and attempts to apply Transformer-based models utilizing self-attention mechanisms to predict SST and similar ocean variables are also increasing [7]. These prior studies demonstrate that data-driven approaches can achieve significant results in SST prediction problems.

Building on this research trend, this study focuses on SST prediction in the coastal waters of South Korea using actual observational data. In real-world marine operational contexts, such as coastal management, fisheries activities, and marine hazard response, it is essential to capture continuous SST variations over multiple days. Accordingly, time-series forecasting over a specified period is required rather than single-point prediction.

In this study, a prediction problem was formulated in which SST variations over a given period are forecast using SST observations from a preceding period, based on the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) dataset. Specifically, SST over the subsequent 7 days was predicted using SST observations from the previous 7 days as input. This prediction setting reflects the characteristics of SST variability over multiple days while considering a time scale that is practically relevant for real-world marine environmental applications. Daily SST data from 2020 to 2024 were utilized, and all data were preprocessed with a focus on oceanic regions prior to model training. Furthermore, multiple deep learning models were applied, and their performances under the defined prediction setting were systematically compared to evaluate the applicability of data-driven approaches based on actual observational data.

II. METHODOLGY

A. Data and Preprocessing

This study used the OISST data provided by the NOAA. OISST is a high-quality sea surface temperature reanalysis dataset that combines satellite, buoy, and ship observations using optimal interpolation techniques. It offers daily temporal resolution and a spatial resolution of 0.25° × 0.25°. This data set has been continuously constructed for the entire global ocean since 1981, making it widely used for analyzing long-term ocean environmental changes.

The study area is the waters surrounding the Korean Peninsula, spanning latitudes 32°N–40°N and longitudes 124°E–132°E. The dataset consists of daily SST data for a total of five years, from January 1, 2020, to December 31, 2024.

OISST may contain missing or anomalous values due to satellite observation limitations, cloud influence, and observation gaps. To address these issues, this study performed a step-by-step preprocessing process. First, a binary ocean mask was created to distinguish between ocean and land areas, and only valid ocean grids were selected. This process prevents physically meaningless SST values occurring in land areas from affecting the learning and evaluation processes.

Missing data handling was designed to preserve the temporal continuity of sea surface temperature as much as possible. First, linear interpolation was applied along the temporal axis at each grid location to correct missing values. Subsequently, any remaining missing values were imputed using the spatial mean of adjacent grid points to ensure data completeness. This missing data treatment was intended to minimize abrupt discontinuities in the SST time series.

During the normalization process, to prevent pixel values over land areas from distorting the global statistical distribution, the global mean and standard deviation were computed using only pixel values corresponding to ocean regions. Z-score normalization was then performed using these statistics, and normalization is defined as follows.

$$\tilde{x} = \frac{x-\mu}{\sigma} \tag{1}$$

Where x denotes the raw SST value, and μ and σ represent the global mean and standard deviation computed using only ocean-region data from the training dataset. Land regions were masked with a fixed constant value to eliminate their influence during model training.

Time-series samples for model training were generated using a sliding window approach. Each sample used SST observations from the previous 7 days as input, while SST values over the subsequent 7 days served as the prediction target.

B. Model Architectures

In this study, various deep learning architectures were compared to analyze the effects of different spatiotemporal modeling strategies on prediction performance in a sea surface temperature forecasting task. The models considered were categorized into convolution-based models, recurrent neural network-based models, hybrid models, and Transformer-based models. All models were designed to use the same input/output configuration (7-day input → 7-day prediction) to ensure a fair comparison between model architectures.

CNN-based models have the advantage of effectively learning local spatial patterns of SST fields through convolutional operations. This study included a CNN architecture based on 2D convolutions and a 3D CNN model capable of simultaneously processing spatiotemporal information. By applying 3D convolutions that include the temporal dimension, the 3D CNN can learn SST change patterns over short time periods along with spatial information. These models offer a relatively simple structure and high computational efficiency, providing a baseline performance for SST prediction problems.

The ConvLSTM model combines convolutional operations with recurrent neural network structures, enabling it to effectively process spatiotemporal time-series data such as SST. ConvLSTM simultaneously learns temporal dependencies and spatial patterns by replacing fully connected operations within the LSTM structure with convolutional operations. In this study, ConvLSTM was utilized to explicitly model temporal changes from past SST sequences and, based on this, to predict future SST changes.

Hybrid models combine spatial feature extraction using CNNs with temporal dependency learning using RNN-based architectures. This study included a model combining a ResNet-based CNN encoder and an LSTM architecture. This model extracts complex spatial patterns through a deep convolutional network, then inputs the extracted feature sequences into an LSTM to learn temporal variations. This architecture offers the advantage of simultaneously securing spatial expressiveness and temporal modeling capabilities.

The U-Net model effectively utilizes multi-resolution spatial information through an encoder-decoder architecture and skip connections. In this study, we combined input SST time series by channel dimension and fed them into the U-Net. This allowed us to spatially integrate and process past SST information and simultaneously predict continuous future SST fields. U-Net-based models excel at preserving spatial structure and serve as a valuable comparison target for evaluating the performance of SST spatial pattern reconstruction.

Transformer-based models have been applied to various geoscience problems due to their ability to effectively learn global spatiotemporal dependencies by leveraging self-attention mechanisms. This study included a Vision Transformer (ViT)-based model and a lightweight Transformer model. The Transformer model segments the input time series into patches and then learns features by considering both time series and spatial information. This structure can simultaneously model long-range spatial correlations and temporal dependencies, making it suitable for SST prediction in complex ocean environments.

C. Evaluation Metrics

In this study, we used Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Square Error (MSE) to quantitatively evaluate the performance of SST forecasts. All evaluation indices were calculated for SST values restored to actual temperature units (°C) by inversely normalizing them. An ocean mask was applied to calculate these indices only for the ocean grid, excluding land areas.

RMSE is a metric calculated by taking the square root of the mean squared prediction error. It is relatively sensitive to large errors and can effectively reflect prediction failures. MAE represents the average absolute error between the predicted value and the observed value, and is useful for intuitively assessing overall average prediction accuracy. MSE is defined as the mean squared prediction error and is reported together with RMSE and MAE to maintain consistency with the loss function used during model training.

Each metric was computed by averaging the values calculated over the entire forecast period (7 days), enabling a comprehensive comparison of SST prediction performance across models. Through this set of evaluation metrics, this study aims to provide a balanced assessment of both the magnitude and characteristics of prediction errors. The evaluation metrics are defined as follows:

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

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \tag{3}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{4}$$

III. Results

In this study, we compared the predictive performance of various deep learning models for the problem of predicting SST for the next seven days in the Korean coastal waters using the past seven days as input. Model performance was evaluated using a test dataset, and all evaluation indices were calculated based on the RMSE, MAE, and MSE in real temperature units (°C) after inverse normalization.

TABLE I. Comparison of SST Prediction Performance Based on RMSE, MAE, and MSE

Model	RMSE	MAE	MSE
ConvLSTM	0.658	0.481	0.433
3D-CNN	0.965	0.733	0.931
U-Net	0.654	0.487	0.427
ResNet-LSTM	1.423	1.125	2.026
ViT	1.133	0.871	1.284
Light Transformer	1.031	0.796	1.063

Table 1 presents the results of a quantitative comparison of the SST prediction performance of each model. The results showed that the U-Net and ConvLSTM models achieved the best prediction performance. The U-Net model achieved an RMSE of 0.654°C and a MAE of 0.487°C, while the ConvLSTM model also achieved similar performance, with an RMSE of 0.658°C and a MAE of 0.481°C. This suggests that the strong spatial structure preservation capability of U-Net and the effective learning of spatiotemporal continuity by ConvLSTM are well suited for predicting SST variability in the coastal waters of Korea.

In contrast, the 3D CNN model showed relatively lower performance compared to the top two models, with an RMSE of 0.965°C and an MAE of 0.733°C. This suggests that the convolutional structure, with its limited temporal dimension, has limitations in adequately reflecting the complex temporal variations of SST. The ResNet-LSTM model showed the largest prediction error, with an RMSE of 1.423°C, which can be interpreted as being due to the input time series length and data scale being insufficient compared to the complexity of the model structure.

For Transformer-based models, ViT and Light Transformer achieved RMSEs of 1.133°C and 1.031°C, respectively. This suggests that while self-attention mechanisms are effective at learning long-range dependencies, their advantages may not be fully utilized in forecasting problems with short input time series lengths, such as this study. In particular, Light Transformer, with its lightweight architecture, exhibited relatively lower errors compared to ViT, demonstrating that a balance between model complexity and data scale is a critical factor in SST prediction performance.

The MSE metric also showed a consistent trend with the RMSE and MAE results, with models showing smaller prediction errors yielding lower MSE values. This indicates that the evaluation metrics used in this study provide consistent information, supporting the reliability of performance comparisons across models.

In summary, the results of this study demonstrate that a model structure that simultaneously preserves spatial patterns and learns spatiotemporal continuity is effective for predicting SST in the Korean coastal waters. Furthermore, complex models do not always guarantee better performance in all cases, highlighting the importance of selecting a model that is appropriate for the time scale and data characteristics of the prediction problem.

IV. Conclusion

This study addressed a data-driven SST prediction problem for the coastal waters of Korea, in which SST over the next seven days was predicted using SST observations from the previous seven days as input. Daily SST data from 2020 to 2024 were obtained from the NOAA OISST dataset. Using preprocessing and evaluation procedures restricted to oceanic regions, the predictive performance of various deep learning models was compared and analyzed.

Experimental results showed that the U-Net and ConvLSTM models outperformed other models in SST prediction performance. This suggests that a convolutional encoder-decoder architecture that effectively preserves spatial structure and a recurrent neural network architecture that explicitly learns spatiotemporal continuity are effective in predicting SST fluctuations in complex marine environments such as the Korean coast. In contrast, the Transformer-based model showed relatively high prediction errors under the input time series length and limited data scale specified in this study, demonstrating that the advantages of the model structure can manifest differently depending on the time scale and data characteristics of the prediction problem.

The results of this study demonstrate that model complexity does not always lead to improved prediction performance in SST prediction problems, highlighting the importance of selecting a model structure appropriate to the target time scale and ocean environmental characteristics. Furthermore, this study confirms that a data-driven approach based on actual observations has practical applicability in SST predictions off the Korean coast.

However, this study has several limitations. It used only a single variable (SST) as input and did not consider external factors such as atmospheric variables or ocean circulation information. Furthermore, because the forecast period was limited to seven days, the model's performance for longer-term SST variability was not evaluated. Future studies may apply a multivariate forecasting model that considers various physical variables, such as wind, sea surface height, and ocean currents, and extend the forecast period to comprehensively analyze short-, mid-, and long-term SST forecasting performance.

In summary, this study quantitatively analyzed the applicability of data-driven deep learning models and characteristics of different model architectures for SST prediction in the Korean coastal waters. The findings provide fundamental insights that can be utilized in future marine environmental forecasting and management.

ACKNOWLEDGMENT

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2026-RS-2020-II201791)

REFERENCES

[1] A. G. O'Carroll, E. M. Armstrong, H. M. Beggs, M. Bouali, K. S. Casey, G. K. Corlett, et al., "Observational needs of sea surface temperature," Frontiers in Marine Science, vol. 6, Art. no. 420, 2019.

[2] K. I. Chang, W. J. Teague, S. J. Lyu, H. T. Perkins, D. K. Lee, D. R. Watts, et al., "Circulation and currents in the southwestern East/Japan Sea: Overview and review," Progress in Oceanography, vol. 61, no. 2–4, pp. 105–156, 2004.

[3] S. M. Barbosa, "Testing for deterministic trends in global sea surface temperature," Journal of Climate, vol. 24, no. 10, pp. 2516–2522, 2011.

[4] S. M. Barbosa, "Testing for deterministic trends in global sea surface temperature," J. Climate, vol. 24, no. 10, pp. 2516–2522, May 2011.

[5] Y. Miao, C. Zhang, X. Zhang, and L. Zhang, "A multivariable convolutional neural network for forecasting synoptic-scale sea surface temperature anomalies in the South China Sea," Weather and Forecasting, vol. 38, no. 6, pp. 849–863, 2023.

[6] L. Wei and L. Guan, "Seven-day sea surface temperature prediction using a 3DConv-LSTM model," Frontiers in Marine Science, vol. 9, Art. no. 905848, 2022.

[7] H. Dai, Z. He, G. Wei, F. Lei, X. Zhang, W. Zhang, and S. Shang, "Long-term prediction of sea surface temperature by temporal embedding transformer with attention distilling and partial stacked connection," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 17, pp. 4280–4293, 2024.