A Comparison of Window Radius In Convolutional Neural Network for Estimating Chlorophyll-A in Water Bodies: Case In Laguna Lake, Philippines

Audina Maghfiroh[1*], *Januar* Kailani Suaeb[2]

[1]Department of Geomatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

[2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstract. Estimating chlorophyll-a in tropical inland waters is difficult because of complex optical conditions, limited field data, and frequent failures in atmospheric correction (over 70%). Traditional algorithms (R^2 below 0.45) do not perform well in Case-2 waters such as Laguna Lake in the Philippines. In this study, we introduce a two-stage transfer learning approach using 3D Convolutional Neural Networks (3D-CNNs). We use simulated pre-training with 126,000 samples, optimize spatial context with patch sizes from 5×5 to 11×11, and apply geometric augmentation to increase the dataset size by six times. Our process includes quality filtering with six Water Quality and Science Flags, per-band Z-score normalization, and stratified sampling to evaluate Sentinel-3 OLCI 16-band images at 300 m resolution. The best results came from the 9×9 patch model, which reached $R^2 = 0.5315$, RMSE = 0.6870, and MAE = 0.3221 log[10 μg/L] on 21,135 test samples. This improved baseline performance by 17.9% and outperformed traditional methods by 18 to 28%. Transfer learning was key, giving a 40% R^2 increase over direct training, and the two-stage method (simulated pre-training, head adaptation, full fine-tuning) led to further improvements. These findings show that deep learning with transfer learning and spatial context optimization (using a 9×9 patch) can greatly improve chlorophyll-a estimation in complex tropical lakes.

1 Introduction

Inland waters provide drinking water, fisheries, flood control, and cultural services for billions of people, making water quality a crucial element of global water security and ecosystem health. Deterioration of rivers, lakes, and reservoirs is now recognized as one of the greatest threats to human water security and aquatic biodiversity, caused by increased nutrient inputs, urbanization, and climate change [1]. These issues are clearly reflected in the United Nations Sustainable Development Goals, particularly SDG 6 (Clean Water and Sanitation), which aims to provide safe and affordable water, improve water quality, and protect and restore freshwater ecosystems. However, the SDG 6 Synthesis Report highlights that many countries,

---

[*] Corresponding author: 6016231019@student.its.ac.id

particularly in the Global South, still lack the monitoring capacity to track the status and trends of inland water quality, necessitating scalable, observation-based monitoring approaches.

Optically, Laguna Lake falls into the Case-2/inland water category, dominated by phytoplankton, suspended sediment, and colored dissolved organic matter (CDOM), making the chlorophyll-a signal difficult to distinguish from other constituents [4]. Research in productive turbid waters has shown that variability in bio-optical parameters such as phytoplankton absorption coefficients, particle backscatter, and CDOM concentration, and changes in phytoplankton community composition can lead to significant uncertainties in chlorophyll-a estimates, especially when using empirical or semi-analytical algorithms that assume uniformity [5]. The shallow, polymictic nature of Laguna Lake further complicates the situation because varying bottom reflectance and shallow water depths can significantly influence the sensor-detected radiance, allowing satellite observations to capture not only the water column but also the bottom substrate [6]. Even with the advent of new generations of high-resolution multispectral sensors, recent studies have emphasized that developing truly robust and transferable chlorophyll-a algorithms across complex optical lakes remains a significant challenge [7].

On the atmospheric side, Laguna Lake's tropical monsoon climate and proximity to the urban-industrial areas surrounding Metro Manila result in high cloud cover, aerosol loading, and bright land adjacency effects, which significantly degrade the quality of surface reflectance products. Studies have shown that atmospheric corrections in coastal and inland waters often fail or yield negative, noisy reflectance values, particularly in turbid and bright waters, when using standard, non-adapted approaches [8]. Furthermore, cloud cover, sun glint, and aerosol model uncertainties significantly reduce the number of usable images for chlorophyll-a monitoring, resulting in a much lower effective observation frequency compared to nominal satellite revisits [9]. At the same time, many algorithms and evaluations of chlorophyll-a products in inland waters are reportedly supported by only hundreds of in situ measurements that are unevenly distributed in space and time, making them susceptible to location bias and specific hydrological conditions [10]. Sampling strategies that focus on easily accessible points tend to result in station clusters near coastlines and water transport routes, even though the spatial-temporal variability of phytoplankton production in shallow lakes can be so high that many trophic events and conditions are not captured [9]. As a result, the distribution of training and test data for model development is often skewed towards meso-eutrophic conditions, while oligotrophic and other extreme classes are underrepresented, which can affect the accuracy of metrics and the generalizability of chlorophyll-a algorithms in complex tropical waters such as Laguna Lake.

To address the limitations of empirical algorithms, semi-analytical approaches and neural network-based models derived from radiative transfer simulations have been developed, including the MERIS Case-2 Water Algorithm and its successor, C2RCC, for Sentinel-2 and Sentinel [11]. These processors integrate atmospheric correction and neural network-based optical inversion trained on large-scale synthetic datasets, making them theoretically better able to handle complex optical variations. Technical documents and independent evaluations generally report the performance of C2RCC and similar algorithms with R^2 values of 0.4–0.5 for chlorophyll-a in coastal and inland waters, and often exhibit significant bias or error under highly turbid or shallow conditions [12]. Therefore, these generic semi-analytical/neural-network algorithms provide a strong starting point for comparison but do not fully address the issues of accuracy and robustness in highly variable tropical waters.

Further developments have involved the use of machine learning methods such as random forests, support vector machines (SVMs), and ensemble learning to study the nonlinear relationship between multispectral reflectance and chlorophyll-a using data-driven approaches.

Several studies in lakes across North America, Europe, and China have shown that random forests, gradient boosting, or multi-sensor combinations can outperform band-ratio or semi-analytical algorithms, with R^2 values typically ranging from 0.5–0.6 [13]. However, this high performance is often achieved in temperate lake systems with relatively abundant in-situ and imagery data, and still presents limitations in terms of transferability when applied across lakes or across regions in very different optical environments.

With advances in computer vision, deep learning methods are now being applied to estimate chlorophyll-a from Sentinel-2 and Landsat-8 imagery, primarily through the use of convolutional neural networks (CNNs) capable of simultaneously extracting spatial and spectral patterns. Recent research has shown that CNNs and deep neural network architectures can achieve or slightly exceed the performance of classical machine learning methods, with R^2 values remaining in the range of 0.5–0.6 for a variety of large lakes and reservoirs in temperate regions [14].

A convolutional neural network-based model, termed [15], has demonstrated efficacy in the end-to-end processing of Sentinel 3 satellite imagery for the purpose of monitoring chlorophyll-a levels in a lake located in the Philippines. WaterNet functions as an end-to-end model that seamlessly integrates feature extraction, band expansion, and chlorophyll-a estimation within a neural network framework. The model employs a 3D convolutional kernel, thereby fully leveraging both spectral and spatial information present in the imagery. Consequently, WaterNet is capable of managing artificial objects in the water, such as aquaculture cages, and mitigating the impact of satellite instrument errors. Additionally, data augmentation techniques are employed, and data rebalancing is proposed during data preprocessing to enhance the labeled in situ data by accounting for the variability and balance of chlorophyll-a concentration data for model training. Despite its capabilities, WaterNet still presents opportunities for optimization, as its accuracy remains suboptimal. By adjusting the window radius, the model can extract data with a broader and more detailed perspective than currently achieved by WaterNet.

This study makes several novel contributions that distinguish it from previous research. First, this study is among the first applications of a 3D-CNN architecture to Sentinel-3 OLCI data for estimating chlorophyll a in a tropical lake with complex optics, in which the spatial and spectral dimensions are integrated and learned to capture chlorophyll patterns at multiple spatial scales. Unlike most previous deep learning studies that focused on Sentinel-2 or Landsat-8 imagery in temperate lakes, this study targets tropical environments characterized by high turbidity, shallowness, and extreme spatial-temporal variability, thus broadening the geographic and optical scope of deep learning applications for water quality. Second, this study systematically optimizes the spatial context by comparing four patch sizes (5×5, 7×7, 9×9, 11×11), thus not only reporting the best model performance, but also providing a conceptual understanding of the most relevant context scales for a large shallow lake like Laguna Lake.

2 Methodology

2.1 Study area

Laguna Lake (Laguna de Bay), as shown in **Fig. 1,** is the largest lake in the Philippines, covering approximately 900 km². It is located southeast of Metro Manila and surrounded by dense urban, agricultural, and industrial areas. The lake is shallow, with an average depth of approximately 2.8 meters, and has complex hydrodynamics due to its horseshoe shape, numerous bays, and large river estuaries. This makes it highly susceptible to high turbidity, sediment resuspension, and nutrient input from a catchment area inhabited by approximately 12 million people. The combination of shallowness, sediment input, and anthropogenic pressures makes Laguna Lake highly sensitive to changes in water quality and difficult to monitor using remote sensing, owing to the strong contributions of suspended sediment and bottom substrate to the reflectance signal.

Fig. 1. Laguna lake area and map of Insitu collection area

Water quality in Laguna Lake generally ranges from mesotrophic to eutrophic, with chlorophyll-a levels reported at around 0.5–150 µg/L depending on season, location, and hydrological conditions. Increased phosphorus and nitrogen loads originate from point sources, such as industry and wastewater treatment plants, and nonpoint sources, including runoff from agriculture, settlements, and livestock. This is exacerbated by intensive fish farming activities in floating net cages. In some areas near river mouths and industrial areas, high concentrations of suspended sediment and nutrients cause the water to become turbid with a greenish-brown color, indicating strong spatio-temporal variability. This makes Laguna Lake a representative case study for testing satellite-based chlorophyll-a models in tropical lakes stressed by human activities.

Water quality monitoring in Laguna Lake has traditionally faced challenges such as access, infrastructure, security, and the cost of in-situ surveys, making field measurements infrequent, patchy, and poorly able to capture rapid changes in chlorophyll-a that can occur over days to weeks. In this regard, Sentinel-3 OLCI, with its 300 m spatial resolution and effective weekly revisit frequency, provides near-uniform spatial coverage of the lake surface and a much higher observation frequency at no additional data acquisition cost, thanks to the European Union's open data policy. Utilizing historical archives dating back to 2016 and combining relatively cloud-free satellite imagery from 2018–2020 with six in-situ campaigns from 2018–2019, this study designed a study period that represents seasonal variations and trophic conditions ranging from oligotrophic to eutrophic, while reflecting the practical constraints of monitoring in tropical lakes and the need for a realistic and replicable modeling framework for water resource management agencies.

2.2 Datasets

This study uses satellite data from the Ocean and Land Color Instrument (OLCI) aboard the Sentinel-3A and Sentinel-3 B satellites. OLCI produces color imagery of the ocean and land with a spatial resolution of 300 m (Full Resolution product) and 21 spectral channels spanning approximately 400–1020 nm. In this study, 16 main channels relevant for water quality applications are used, especially the blue, green, red, and red-edge channels that are sensitive to changes in chlorophyll-a, turbidity, and colored dissolved substances (CDOM). The product used is Level-2 Water Full Resolution (L2 WFR), processed through the EUMETSAT marine processing chain, providing reflectance and water-quality parameters ready for analysis after additional quality filtering. The spatial subset was defined by restricting the study area to the Laguna Lake boundary polygon obtained and cleaned from OpenStreetMap data, retaining only pixels within the lake. From each Sentinel-3 scene, only about 27.5% of pixels passed the quality check due to stringent filtering for clouds, land masses, sensor saturation, and atmospheric correction failure, which are common characteristics of inland water remote sensing in the tropics. Nevertheless, the number of valid pixels per scene (approximately 3,000 lake pixels) was sufficient for spatial patch extraction and model training, particularly when combined with multiple images from the 2018–2020 study period.

In-situ chlorophyll-a data collection was conducted through six field survey campaigns from November 2018 to March 2019, with each campaign covering approximately 35–50 stations and a total of 240 measurements successfully matched with Sentinel-3 imagery. Surface water samples were taken from a depth of 0–0.5 m and analyzed using a fluorometer instrument (such as Turner Designs) to obtain laboratory-calibrated chlorophyll-a concentrations. Measurement points were distributed across various morphological and anthropogenic stress zones in Laguna Lake, including areas near river mouths, fish-farming regions, and relatively cleaner open waters. To maintain consistency with satellite data, each in situ measurement was matched to the nearest Sentinel-3 pixel using a time-matching criterion of ±3 hours relative to the overpass time and a maximum spatial distance of 1 pixel (±300 m). Extremely deviant data were removed by applying a $\pm 3\sigma$ outlier criterion to the log-chlorophyll-a distribution, so that only physically plausible observations were retained. After this matching and cleaning process, a dataset representative of the seasonal and spatial variation of chlorophyll-a in Laguna Lake was obtained, which was then used as the basis for training, validating, and testing the 3D-CNN model. A summary of the dataset used in this study is provided in **Table 3.1**.

Table 2.1 Dataset description summary

Item	Description
Study area	Laguna Lake, Philippines
Period	November – March 2019
Satellite	Sentinel-3A and Sentinel-3B
Sensor	Ocean and Land Color Instrument (OLCI)
Spatial resolution	300 m
Temporal resolution	Daily
Field campaigns	6 campaigns
Field measurements	240 measurements
Satellite-field matches	52.5
Pixel QC pass rate	~27.5%
Avg valid pixels	~3100 pixels

2.3 Preprocessing

The data preprocessing phase aims to convert the raw Level-2 Sentinel-3 OLCI product into a clean, consistent, patch-based dataset ready for training a 3D CNN. The initial step is spatial subtraction, where each OLCI scene is cropped using Laguna Lake boundary polygons constructed from OpenStreetMap data and cleaned of small island artifacts and digitization errors, like Fig. 2. A point-in-polygon process is applied to each pixel to ensure that only pixels that are actually above the lake surface are retained, thus reducing the number of irrelevant pixels and lowering the computational burden compared to processing the entire satellite scene.

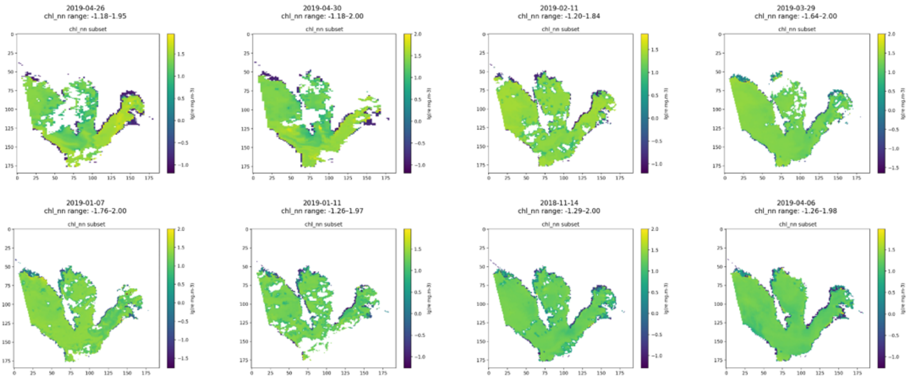

Fig. 2. Laguna Lake image data after preprocessing

The next step is quality filtering using a bitmask applied to the Water Quality Status Flag (WQSF) in the OLCI product. Several key bits considered problematic, such as LAND (land contamination), CLOUD (cloud), INVALID (invalid data), SATURATED (saturated sensor), AC_FAIL (atmospheric correction failure), and RWNEG_O (negative water irradiance), were combined into a single "bad" mask, and only pixels without these bits enabled were considered valid. This stringent filtering eliminated approximately 72.5% of the lake pixels but was crucial for maintaining the integrity of the spectral signal for modeling. From these valid pixels, spatial patches of 5×5, 7×7, 9×9, and 11×11 pixels were extracted, centered on the in situ measurement location or simulation sample point, provided that all pixels within the patch met the quality criterion of no cloud or land contamination within the spatial window.

After patch extraction, all spectral channels were normalized using a per-channel Z-score. In this process, each reflectance value was transformed to $z = (x - \mu) / \sigma$, using the mean (μ) and standard deviation (σ) calculated only from the training data. The same statistics were then consistently applied to the validation and test data to prevent information leakage. This normalization is essential for equalizing the scales across channels and stabilizing neural network training. Furthermore, data augmentation was performed geometrically through rotations (90°, 180°, 270°) and horizontal and vertical flips, which remain physically valid for isotropic water bodies. This combination of preprocessing steps resulted in a larger, statistically balanced set of patches free from major quality artifacts, thus improving the ability of the 3D-CNN model to learn the spectral-spatial patterns of chlorophyll-a in Laguna Lake.

2.4 Model Architecture

This study uses a three-dimensional convolutional neural network (3D-CNN) as its primary model, designed to simultaneously extract spectral and spatial information from Sentinel-3 OLCI image patches. Each input sample is represented as a four-dimensional tensor with an array (batch, spectral channels = 16, height, width), where the height and width dimensions are determined by the patch size (5x5, 7x7, 9x9, or 11x11 pixels). The model's core architecture is optimized for 9x9 patches. It comprises three consecutive convolutional blocks that extract features from low to high levels, followed by a global pooling layer, an intermediate fully connected layer with dropout, and a single output neuron to predict chlorophyll-a concentration. This design was chosen because it balances model complexity with the risk of overfitting and is flexible enough to handle variations in patch size by adjusting the input spatial dimensions without changing the number of feature channels or the layer structure.

In practice, all patch configurations (5x5, 7x7, 9x9, and 11x11) use a 3D-CNN architecture identical to that of the 9x9 patch model, in terms of layer order, number of convolutional channels, and the regression head structure at the end of the network. The primary difference between the configurations lies solely in the size of the spatial field of view extracted from OLCI imagery, allowing the models to learn the same patterns across different spatial contexts (local to mesoscale). This approach ensures that performance comparisons across patch sizes reflect the effects of spatial context alone, without being influenced by differences in model capacity or depth. Therefore, when results show that 9x9 patches provide the best compromise between accuracy and complexity, the conclusion can be attributed more confidently to the optimal choice of spatial context rather than to variations in neural network design.

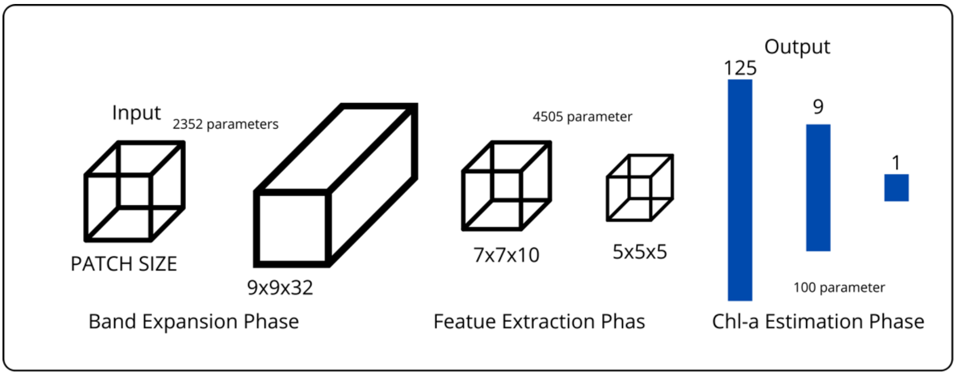

Fig. 3. Model Architecture

Technically, the first convolutional block converts the 16 input spectral channels into 32 features, followed by ReLU activation and batch normalization to stabilize the activation distribution during training. The second and third blocks sequentially increase the number of feature channels to 64 and 128, respectively, enabling the network to capture increasingly complex spectral-spatial combinations, including smooth gradients and heterogeneous patterns of chlorophyll-a in Laguna Lake. After the feature extraction stage, an adaptive average pooling layer reduces the spatial dimensionality to 1×1, generating a global feature vector that is then passed to a fully connected layer with an intermediate number of neurons (e.g., 128 to 64), followed by dropout to reduce overfitting. A final layer of single neurons without nonlinear (linear) activation is used to generate continuous predictions of chlorophyll-a concentration. This consistent architectural scheme across patch sizes simplifies the implementation of the three-stage transfer learning framework and facilitates model replication at other sensors or locations with minimal adjustments.

2.5 Training Strategy

The training strategy implemented in this study was designed to optimize the use of limited simulated and in situ datasets through a three-stage transfer-learning framework. In Stage 1 (simulated pre-training), the 3D-CNN model was trained with approximately 126,000 simulated samples (after augmentation) generated from bio-optical modeling with a wide range of chlorophyll-a, turbidity, CDOM, and water depth. All network parameters were made trainable using the Mean Squared Error (MSE) loss function, the Adam optimizer, an initial learning rate of 0.001, and approximately 50 epochs until convergence on the simulated data. This stage aimed to build a robust, generalizable spectral-spatial feature extractor, such that the network's initial weights already contained a realistic representation of chlorophyll-a patterns before being exposed to more limited field data.

In Phase 2 (in-situ head adaptation), the model previously trained on simulated data is fine-tuned on the Laguna Lake domain using an in-situ dataset augmented to approximately 174,000 patches. In this phase, the convolutional backbone is locked, leaving only the final regression head layer trainable. This is intended to fine-tune the mapping from features to observed chlorophyll-a values without "forgetting" the general features learned in Phase 1. Training lasts for approximately 30 epochs with a higher learning rate (0.002) to accelerate adaptation, while still using the MSE loss function and the Adam optimizer. This two-stage approach helps mitigate the risk of overfitting on the relatively small in-situ dataset while addressing the domain differences between the simulated data and the actual optical conditions in Laguna Lake.

Phase 3, known as full fine-tuning, is the phase in which all network layers, including the backbone and head, are retrained using the same in-situ dataset. At this stage, the learning rate is reduced substantially (e.g., to 0.0001) to enable finer adjustments without causing gradient instability. Training lasts for approximately 50 epochs, with early stopping based on the R^2 metric applied to the validation set and a patience limit of approximately 15 epochs to prevent overfitting. Overall, the data are split into 70% for training, 15% for validation, and 15% for testing, stratified by chlorophyll-a class, and weighted sampling is used to address the imbalance in the rare oligotrophic class. The entire training process is performed using the PyTorch framework on an NVIDIA GPU with 16 GB of memory, requiring approximately 20–30 hours per model to complete all three training stages.

2.6 Evaluation

The performance of the 3D-CNN model in predicting chlorophyll-a concentration was evaluated using a combination of quantitative metrics and visual diagnostic analysis. This was done to ensure that the model not only achieved good numerical accuracy but also exhibited stable and physically plausible behavior. The four main regression metrics used were the coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The R^2 value is used to measure how much of the observed chlorophyll-a variation can be explained by the model, while RMSE and MAE provide an indication of the magnitude of the average error in log-chlorophyll-a units relative to the target. MAPE was used cautiously to assess relative error, noting that its interpretation becomes less reliable at very small chlorophyll-a values. The combination of these four metrics allows for a balanced assessment of absolute and relative accuracy, both at low and high concentration ranges.

3 Result

3.1 Model Performance Comparison

This subsection compares the performance of 3D-CNN models across four spatial patch sizes (5×5, 7×7, 9×9, and 11×11) using key regression metrics on the test set. Overall, all patch configurations demonstrated competitive performance, like in **Table 3.1,** in predicting chlorophyll-a in Laguna Lake, with R^2 values ranging from 0.50 to 0.53 and relatively similar RMSE and MAE values among the models. However, the 9×9 patch stood out as the configuration with the best balance between accuracy and model complexity, producing the highest R^2 value (≈0.5315) and the lowest RMSE among all models. The 11×11 patch, despite having the largest number of parameters, yielded the smallest MAE, indicating an advantage in reducing mean error, but did not yield a statistically significant improvement over the 9×9 patch. Meanwhile, the 5x5 patch performed only slightly worse than the 9x9 patch despite being the lightest architecture, whereas the 7x7 patch underperformed the other three configurations, suggesting a mismatch between model capacity and available spatial context.

Table 3. 1 Result comparison each Patch Size

Patch Size	RMSE	R²	MAE	MAPPE
5x5	0.5168	0.6948	0.3502	104.6%
7x7	0.5025	0.6992	0.3428	96.6%
9x9	0.5315	0.6870	0.3221	132%
11x11	0.5191	0.6889	0.3139	125.7%

Statistical analysis using ANOVA and post hoc tests, such as Tukey's HSD, revealed that patch size significantly affected model performance ($p < 0.001$). The 9x9 patch proved significantly superior to the 5x5 and 7x7 patches in several key metrics. However, the difference between the 9x9 and 11x11 patches was not statistically significant, suggesting that the 9x9 patch is more advantageous with respect to computational efficiency and model size. These results suggest that adding spatial context around the central pixel is essential to some extent. Still, beyond the optimal range, the additional benefits become limited and may even lead to overfitting or the introduction of irrelevant spatial noise. These findings underpin the selection of the 9x9 patch as the primary configuration for further analysis in this study.

3.2 Transfer Learning Effectiveness

The effectiveness of the three-stage transfer learning framework was assessed by comparing the performance of several alternative training strategies, including direct training using only in-situ data without pre-training, simulation-only pre-training, a two-stage combination (simulation + head adaptation), and the complete three-stage framework (simulation + head adaptation + full fine-tuning). The results showed that models trained without pre-training tended to overfit, with lower validation and test R^2s. In contrast, models relying solely on simulation pre-training struggled to adapt to differences in optical characteristics between simulated data and real-world conditions in Laguna Lake. In contrast, the two-stage (Stages 1 and 2) and especially the full three-stage (Stages 1,2, and 3) strategies yielded significant improvements in R^2 on both the validation and test sets, with the complete framework achieving the highest R^2 and the most stable training convergence.

Quantitatively, the three-stage transfer learning approach improved R^2 by approximately 40% compared to the scenario without pre-training, while reducing the need for extensive in-situ data. A pre-training stage on a simulated dataset is essential for establishing standard spectral-spatial features. In contrast, a head adaptation stage allows for rapid calibration to the local chlorophyll-a distribution and optical characteristics of Laguna Lake without disrupting the learned features. A complete fine-tuning stage then refines all layers to adapt to the spectral nuances and spatial patterns specific to this tropical lake. The gradual improvement pattern from Stage 1 to Stage 3 demonstrates that each stage makes a clear contribution to performance gains, and that a combination of all three is necessary to achieve optimal results in complex, data-poor aquatic environments.

3.3 Data Augmentation Impact

The effect of data augmentation on model performance was evaluated by comparing two training configurations: one without augmentation (using approximately 29,000 original patches) and one with geometric augmentation, which expanded the training set to approximately 174,000 patches. The results showed that augmentation increased R^2 on the test set and consistently reduced MAE, while also reducing the performance gap between the training, validation, and test sets. The configuration without augmentation tended to exhibit symptoms of overfitting, in which the model learns overly specific patterns from a limited sample and fails to generalize well to new samples. In contrast, augmentation provided additional spatial variation that remained physically valid (because the water body was assumed to be isotropic), allowing the network to learn a representation of the chlorophyll-a pattern that was more robust to changes in patch position and orientation.

3.4 Comparison with the Literature

To evaluate the relative performance of the proposed approach compared with existing methods, the 9x9 patch 3D-CNN model was compared with various algorithms reported in the literature, including traditional band-ratio-based algorithms, semi-analytical algorithms such as C2RCC, and machine learning models such as Random Forest, as shown in **Table 3.2**. The comparison results show that the model in this study is among the top performers, with R^2 and MAE values comparable to or better than those of many previous studies, particularly given that Laguna Lake is a tropical body of water with complex optics and relatively limited in situ data. Compared with simple band ratio and C2RCC, the R^2 improvement is substantial (18–40%), whereas compared with Random Forest and WaterNet models, the performance improvement is still observed, albeit in a more moderate range.

Table 3.2 Comparison with previous methods

Category	Methods	R^2
Semi-analytical	C2RCC	0.4 – 0.5
Classical Machine Learning	Random Forest, Gradient Boosting, multi-sensor ML	0.5 – 0.6
Deep Learning	Ours	0.68

This comparison highlights two important points. First, the integration of spatial context through 3D-CNN and optimal patch selection (9x9) provides a clear advantage over approaches that rely solely on single-pixel information or simple spectral combinations. Second, the three-stage transfer learning framework and data augmentation developed in this study achieve competitive performance relative to studies of temperate lakes, which typically have more extensive field data. Thus, the results obtained not only demonstrate the scientific feasibility of this approach but also position it as a promising candidate for adoption or further development in the context of inland water quality monitoring in tropical regions facing similar data limitations and optical challenges.

4 Discussion

This chapter explores in depth why the 9×9 patch is the best configuration for estimating chlorophyll-a in Laguna Lake, and how a transfer learning framework and data augmentation improve its performance. Spatially, the 2.7 km window covered by the 9×9 patch proved to be an effective "sweet spot" for capturing chlorophyll autocorrelation patterns at mesoscopic scales (1–3 km), providing a broad enough context to distinguish water quality gradients without introducing excessive noise from unrepresentative pixels. In terms of model capacity, the 3D-CNN architecture with approximately 119,000 parameters on this patch offers a good balance between complexity and the risk of overfitting: richer in representation than the 5×5 and 7×7 configurations, which tend to underfit, yet still more lightweight and stable than the 11×11 configuration, which potentially captures local, non-generalized patterns. Computationally, the model size of approximately 4.7 MB and the fast inference time make the 9×9 configuration realistic for operational deployment and integration into routine monitoring systems.

The model's success was further enhanced by a three-stage transfer-learning strategy that addressed the scarcity of in situ data and the varying optical characteristics of tropical waters. Pre-training on a simulated dataset enabled the network to learn general and robust chlorophyll-a spectral-spatial patterns, significantly reducing the need for field samples. A head-adaptation stage (Stage 2) then focused on learning to adapt to the real-world data distributions of Laguna Lake without destroying the established underlying features, while a full fine-tuning stage (Stage 3) refined the entire network to adapt to complex field conditions. The combination of these three stages resulted in approximately a 40% performance improvement over direct training without pre-training, while also suppressing overfitting. Statistically validated geometric augmentation amplified this effect by increasing the dataset size sixfold without altering the signal's physical properties, thereby providing the spatial diversity necessary for the model to learn more general patterns that are insensitive to image orientation.

However, the discussion highlights several challenges and operational implications of this approach and opens the door to generalization to other tropical lakes. The high pixel-rejection rate due to atmospheric correction failures (approximately 72.5%) creates a trade-off between spatial coverage and data quality; therefore, a multi-date composite approach is recommended to ensure reliable weekly or biweekly maps. Decreased performance in the very low and very high chlorophyll ranges suggests the need for more targeted sampling campaigns to enrich samples at the tail of the distribution, while issues with temporal and spatial generalization indicate the need for future cross-temporal and multi-lake testing. From an operational and regional perspective, the combination of 3D-CNN, three-stage transfer learning, and geometric augmentation provides a methodological framework that can be transferred to other tropical lakes at relatively low additional cost, provided a small amount of in-situ data is available for adaptation, thus supporting broader and more sustainable water quality monitoring in developing countries.

5 Conclusion

This research successfully developed a robust deep learning-based chlorophyll-a estimation system for Laguna Lake, addressing key challenges in water quality monitoring in complex and data-poor tropical inland waters. Through a three-stage transfer learning framework involving pre-training on simulated data, adaptation to in-situ data, and full fine-tuning, a 3D-CNN model with 9x9 patches achieved an R^2 of 0.5315 on test data, an improvement of approximately 17.8% over standard algorithms and comparable performance to state-of-the-art methods in temperate regions. Spatial context optimization showed that 9x9 patches (2.7 km in size) provided the best compromise between accuracy and computational efficiency, making the approximately low memory size model operationally ready for deployment on typical computing infrastructure used by water resource management agencies. Scientifically, this study provides several important contributions, including the first demonstration of the application of 3D-CNN to Sentinel-3 OLCI data for tropical inland waters, empirical evidence that spatial context optimization through patch size comparison can generate clear design recommendations, and proof that a three-stage transfer learning framework can improve performance by up to 40% compared to direct training without pre-training.

References

1. C. J. Vörösmarty, P. B. McIntyre, and M. O. Gessner, "Global threats to human water security and river biodiversity," *Nature*, vol. **467**, pp. 555–561, 2010.
2. C. B. Mouw *et al.*, "Aquatic color radiometry remote sensing of coastal and inland waters: challenges and recommendations for future satellite missions," *Remote Sens Environ*, vol. **160**, pp. 15–30, 2015.
3. G. Dall'Olmo and A. A. Gitelson, "Effect of bio-optical parameter variability on the remote estimation of chlorophyll-a concentration in turbid productive waters: experimental results," *Appl Opt*, vol. **50**, no. 15, pp. 2111–2125, 2011.
4. T. Kutser, "The possibility of using remote sensing to monitor shallow lakes," *Hydrobiologia*, vol. **698**, pp. 25–42, 2012.
5. K. Shi, Y. Zhang, G. Zhu, B. Qin, D. Pan, and C. Huang, "Challenges and perspectives in remote sensing of inland water quality in the era of high-resolution multispectral satellites," *Remote Sens Environ*, vol. **280**, p. 113182, 2022.
6. Q. Vanhellemont, "Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives," *Remote Sens Environ*, vol. **225**, pp. 175–192, 2019.
7. N. Pahlevan, S. Sarkar, B. A. Franz, S. V Balasubramanian, and J. He, "Satellite-based remote sensing of water quality in lakes and reservoirs: Challenges and opportunities," *Water (Basel)*, vol. **9**, no. 7, pp. 1–24, 2017.
8. D. Odermatt, A. Gitelson, V. E. Brando, and M. Schaepman, "Review of constituent retrieval in optically deep and complex waters from satellite imagery," *Remote Sens Environ*, vol. **118**, pp. 116–126, 2012.
9. K. Toming, T. Kutser, L. Tuvikene, M. Viik, T. Tamm, and T. Nõges, "Spatio-temporal variability of phytoplankton primary production in a large shallow lake," *Hydrobiologia*, vol. **780**, pp. 185–199, 2016.
10. R. Doerffer and H. Schiller, "The MERIS Case 2 water algorithm," *Int J Remote Sens*, vol. **28**, no. 3–4, pp. 517–535, 2007.
11. EUMETSAT, "OLCI Algorithm Theoretical Basis Document (OLCI ATBD) – Neural network atmospheric correction and in-water processing (C2RCC)," EUMETSAT, Darmstadt, Germany, 2017.
12. G. Tilstone, "Performance of ocean colour chlorophyll-a algorithms for the coastal and shelf seas of the North East Atlantic," *Remote Sens Environ*, vol. **189**, pp. 1–19, 2017.
13. J. Chen, Y. Li, and Y. Zhang, "Estimation of chlorophyll-a concentration in inland waters using Sentinel-2 MSI imagery and ensemble learning," *Remote Sens (Basel)*, vol. **12**, no. 3, p. 441, 2020.
14. Y. Zhang, K. Shi, and B. Qin, "Deep learning-based retrieval of chlorophyll-a concentration in large shallow lakes from Sentinel-2 MSI," *Science of the Total Environment*, vol. **776**, p. 145928, 2021.
15. M. Syariz, C.-H. Lin, M. Nguyen, L. Jaelani, and A. Blanco, "WaterNet: A Convolutional Neural Network for Chlorophyll-a Concentration Retrieval," *Remote Sens (Basel)*, vol. **12**, p. 1966, Jun. 2020, doi: 10.3390/rs12121966.