Journal of Engineering Science 16(1), 2025, 01-10
DOI: https://doi.org/10.3329/jes.v16i1.82661

JES
an international Journal

DEVELOPMENT OF A CONVOLUTIONAL NEURAL NETWORK IMAGE DETECTION MODEL TO IDENTIFY THE TROPHIC STATUS OF LAKES

Sinha Mahira Sultana[*1], Mehnaz Sultana[1] and Nabil Ahmad Durjoy[1]

[1]*Department of Civil Engineering, Military Institute of Science & Technology, Dhaka-1216, Bangladesh*

Received: 12 March 2024 Accepted: 09 December 2024

ABSTRACT

The eutrophication of lakes is a major concern in Bangladesh since it is largely responsible for water quality degradation in lakes throughout the country. As per previous research, lake trophic status may be identified using different methods. This study was conducted to seek a new method for determining trophic status across Dhaka City. A total of 110 lake water samples were collected from Banani, Gulshan, Hatirjheel, Ramna, Dhanmondi, Uttara, Nirjhor, Shaheed Sharani I & II, and Dhaka Cantonment Lakes, respectively. Lake surface images were captured at different angles, as part of the sampling process. The samples were tested for nitrate, phosphate, and chlorophyll-a concentration. These parameters were used to compute a trophic state index for each sample. The results indicated that 55% of the lake samples were hypereutrophic, 25% were eutrophic, 11% were mesotrophic, and only 9% were oligotrophic. The indices were used to label the corresponding lake images into the four categories. The images were divided into a training batch and a test batch. A Convolutional Neural Network (CNN) image detection model was designed on a Kaggle platform using Python and fed the images from the training batch. After the training, the model was tested using the other batch of images and displayed an overall accuracy of 79.4% in identifying the trophic states of the lake samples based on their images. The image detection method thus demonstrated higher efficiency than manual computation methods of identifying trophic status.

Keywords: *Trophic Status, Neural Network, Image Detection, Classification*.

1. INTRODUCTION

Eutrophication refers to a state in aquatic systems characterized by elevated nutrient levels, primarily nitrogen and phosphorus, which lead to algal blooms and a decline in water quality (Bali and Gueddari, 2019). Anthropogenic disruptions and climate change have led to organic pollution in numerous lakes, deteriorating water quality, harming species within the food chain, and endangering the stability of lake ecosystems (Gullian-Klanian *et al*., 2021; Nosrati *et al*., 2012). Over recent decades, global lakes have undergone significant changes due to the escalating influence of human activities and climate warming. The occurrence of algal blooms in lakes has increased notably, with 8.8% of global lakes exhibiting a heightened risk of bloom formation between the 1980s and 2010s (Ho *et al*., 2019).

Einar Naumann, a Swedish limnologist at the University of Lund, Sweden in 1919, first introduced the concept of the trophic status of lakes. Trophic states are generally defined per his classification system. Naumann viewed the classification as an "artificial outgrowth of a biological reality" (Naumann, 1919). Phosphorus, nitrate, chlorophyll-a, dissolved oxygen, and water clarity are the most significant parameters used to assess the trophic status of surface waters (El-Serehy *et al*., 2018). There were a variety of methods used across the world to measure the quality of water bodies, such as rivers and lakes. In 1965, the water quality index (WQI) was established in the United States and was extensively used and approved in Europe, Africa, and Asia (Horton, 1965). Robert E. Carlson established the trophic state index (TSI) as a numerical classification of lake trophic status (Carlson, 1977). The trophic level index (TLI) was developed by modifying Carlson's TSI (Burns *et al*., 2009; Wojtkowska and Bojanowski, 2021).

Although lakes transition from oligotrophic to eutrophic, the transitions do not take place at clearly defined points. The changes do not occur at the same rate or in the same place. Carlson proposed that the Trophic State Index (TSI) value, which may be categorized into four main classes: oligotrophic, mesotrophic, eutrophic, and hypereutrophic, be used to characterize the trophic state of a water body. As a result, trophic status could not be determined by looking at only one or two variables. Multiparameter indices may have developed due to such reasoning (Brezonik and Shannon, 1971; Michalski and Conroy, 1972). The number of parameters to be measured limits the effectiveness of a multiparameter index. Indices relying on one criterion can be transparent and more responsive to changes. However, there is no agreement on what the sole criterion of trophic status should be. Landsat satellite imagery has been effectively used to estimate a Trophic State Index (eTSI) based on Secchi disk

*Corresponding Author: *mahirasinha180@gmail.com*

transparency (SDT) measurements for inland lakes in several areas before, such as Minnesota (Olmanson *et al*., 2007), Wisconsin (Chipman *et al*., 2004), Michigan (Fuller *et al*., 2004) and so on.

Machine learning refers to the ability of systems to utilize problem-specific training data to automate the creation of analytical models and address related tasks (Janiesch *et al*., 2021). Deep learning models have been carefully patterned after the human brain. So, they provide a more sophisticated approach to machine learning and are geared to tackle issues that machine learning cannot. Data is transmitted between layers of interconnected neurons known as "deep neural networks." As a result, the data undergoes a transformation that is not linear and becomes progressively abstract (Middleton, 2021). Convolutional Neural Networks are a subset of deep learning. They consist of a stack of layers that takes in an input image, perform a mathematical operation, and predict the class or label probabilities at the output. Instead of standard manual feature extraction methods, CNN performs feature extraction automatically by trial and error.

In a reservoir in central Taiwan, a back-propagation neural network was utilized to link the significant parameters that influence a variety of water quality indicators such as dissolved oxygen (D.O.), total phosphorus (T.P.), chlorophyll-a (Chl-a), and Secchi disk depth (S.D.). The study's findings revealed that the neural network could predict these indications with a reasonable degree of accuracy, implying that a neural network is a valuable tool for reservoir management (Kuo *et al*., 2006). A new back propagation artificial neural network model for evaluating eutrophication in China's eastern provinces was built, and a new training method was developed. The model was tested on four eastern lakes, with encouraging results, indicating the method was appropriate for assessing the eutrophication of lakes (Jiang *et al*., 2006).

Neural networks have long been applied to identify and assess water quality parameters, including eutrophication. These methods mainly involved parameter modelling, such as chlorophyll-a, total phosphate, and Secchi depth. Convolutional neural networks (CNN) specialize in image processing and detection. The prime objective of this study was to build a multiclass CNN model for image detection of lakes. The study also aims to identify the difference in trophic states of lakes located in different areas of Dhaka City.

2. MATERIALS AND METHODS

This research involved collecting water samples from different lakes to identify their trophic states to aid the development of an image detection model using convolutional neural networks (CNN). The following sections briefly summarise the research scheme, tests conducted, and procedure followed in this study.

2.1 Study Area and Sample Collection

A total of 110 water samples were collected from the lakes in the following locations: Gulshan, Banani, Hatirjheel, Dhanmondi, Ramna, Uttara, Nirjhor, Shaheed Sharani Lake and BUP Lake at Mirpur Cantonment.

Generally, sample collection was done from the banks of the lakes. Where possible, boats were used to collect samples from different points along a lake. This process ensures thorough lake coverage since multiple point sources discharge into one lake. The bottles were sealed with caps and labelled immediately after collection.

Images of the lake's surface were captured at each point of sample collection using a high-resolution camera. Multiple images were taken for every sample from different angles since a large number of images are required to create a dataset that can be used to train a CNN model.

2.2 Governing Equations

The samples were tested for total nitrate, total phosphate and chlorophyll-a concentrations. Trophic state index calculations were performed based on a water quality assessment report from the Florida Department of Environmental Protection (Paulic *et al*., 1996). First, it was determined whether the lakes were “phosphorous limited” or “nitrogen limited” lakes, using the ratio of Total Nitrogen to Total Phosphorous.

Table 1: Type of Lake Based on the Ratio of Nitrogen and Phosphorous

Nutrient Ratio	Lake Type
$TN/TP < 10$	Nitrogen Limited Lake
$10 < TN/TP < 30$	Nutrient Balanced Lake
$TN/TP > 30$	Phosphorous Limited Lake

Based on the test results for this study, the lakes were Nitrogen Limited Lakes, i.e., $TN/TP < 10$, as denoted in Table 1. Since nitrogen was the limiting nutrient, the nitrogen trophic state index was calculated using equation (1) (Paulic *et al*., 1996). Thus, no phosphorous index was required. Equation (2) would be used for the phosphorous trophic state index if the limiting nutrient had been phosphorous (Paulic *et al*., 1996).

$$TSI\ (TN) = 10^*\ [\ 5.96 + 2.15\ \ln\ (TN + 0.001)] \quad (1)$$

$$TSI\ (TP) = 10^*\ [\ 2.36\ \ln\ (TP\ ^*1000) - 2.38] \quad (2)$$

Next, another index was calculated using the chlorophyll-a concentration, as per Equation (3) (Paulic *et al*., 1996).

$$TSI\ (Chl\text{-}a) = 16.8+\ [\ 14.4\ ^*\ \ln\ (Chl\text{-}a)] \quad (3)$$

Thus, the limiting nutrient concentration was used for one index and the chlorophyll-a concentration for another. The average of these two indices was then taken as the final trophic state index, as per Equation (4).

$$TSI\ (average)= [TSI\ (TN)+TSI\ (Chl\text{-}a)]\ /2 \quad (4)$$

Next, the sample images were sorted into four classes based on their TSI value: Oligotrophic, Mesotrophic, Eutrophic, and Hypereutrophic, as shown in Table 2 (Carlson and Simpson, 1996).

Table 2: Trophic State Classification

Carlson TSI	Lake Trophic State
<30- 40	Oligotrophic
40-50	Mesotrophic
50-70	Eutrophic
> 70	Hypertrophic

2.3 Image Dataset Preparation and CNN Modelling

The lake images captured were transferred to a computer and processed to adjust the dimensions, light, and shadow and to crop out parts that may cause interference during the image detection process. The processed images were sorted into a group of folders. Around 20% of the images (160) were kept aside for testing. The remaining 80% (700) were used for training the CNN model. The images were uploaded to the "Kaggle" database online and the convolutional neural network model was designed using Python on a Kaggle Notebook.

A total of 860 images have been used for training the model, with a batch size of 16. Thus, 16 images were processed at a time. The number of iterations set to be performed on the whole training dataset was 30. However, if five iterations produced the same result, the process was automatically terminated using Keras's "Early Stopping" function. After the convolutional layer processing, 10% of the extracted features were discarded randomly to eliminate unnecessary information. 20% was discarded after the deep layer processing. These were done using the built-in "Dropout" feature of the Keras Library for Python. Once the training was complete, the model was tested using images from the test batch to determine its accuracy. This time all 160 images were tested in one batch. The overall research methodology is illustrated in Fig. 1.

3. RESULTS AND DISCUSSION

3.1 Findings from Laboratory Test

Figure 2 shows the proportion of samples found for each trophic state. As can be seen from the figure, over half the lake water samples were hypereutrophic. One-fourth of the samples were eutrophic; a little over 10% were mesotrophic, while the oligotrophic portion was less than 10%.

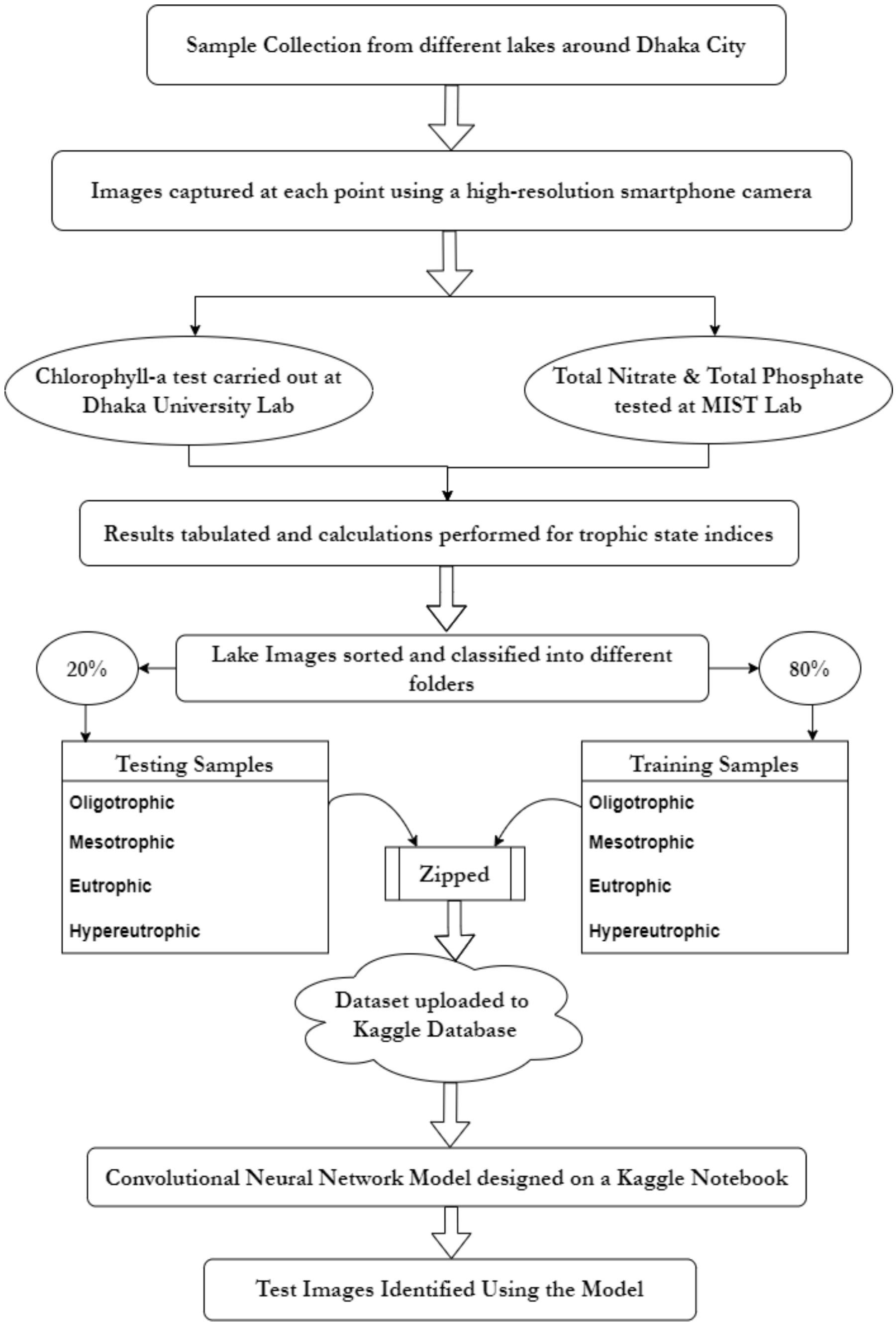

Figure 1: Research Outline

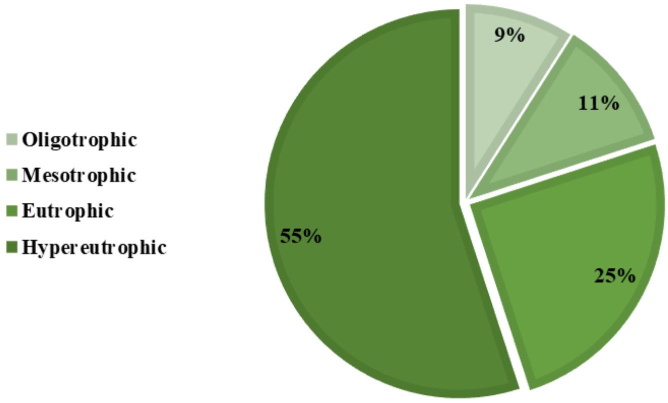

Figure 2: Trophic Status of the Selected Lakes

The highest concentration of chlorophyll-a found in each lake is shown in Fig. 3, while the highest concentrations for nitrate and phosphate are shown in Fig. 4.

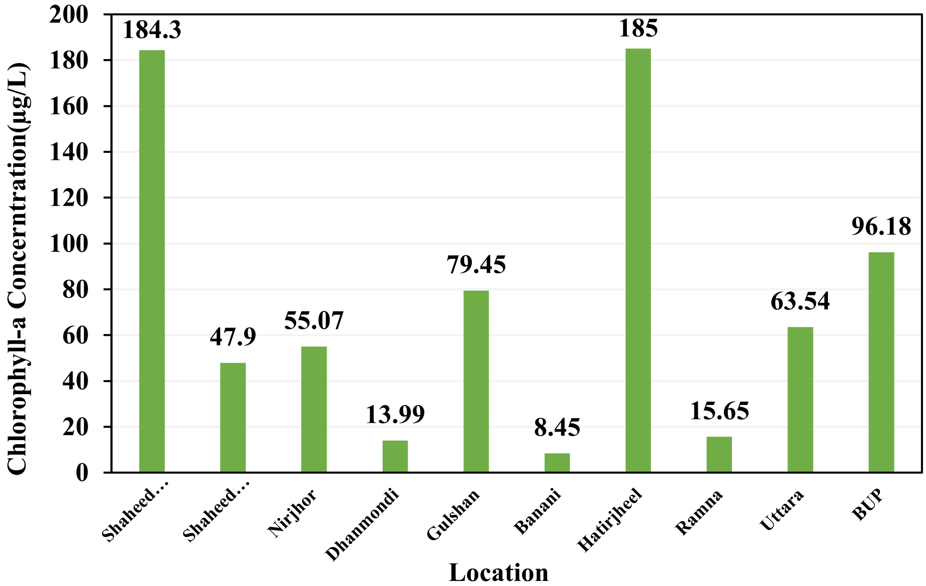

Figure 3: Highest Chlorophyll-a Levels in the Lakes

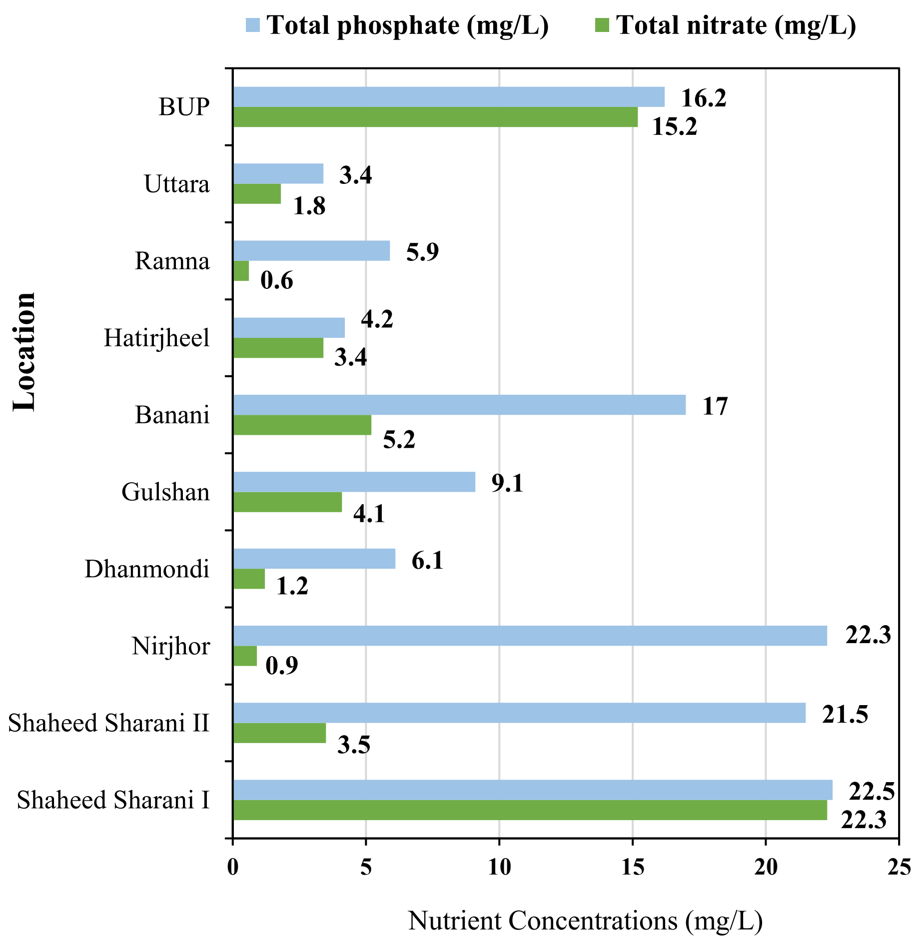

Figure 4: Highest Nutrient Levels in the Lakes

Shaheed Sharani I, Shaheed Sharani II, and Nirjhor (Lakes of Dhaka Cantonment)

As shown in Fig. 4, these lakes exhibited the highest phosphate concentrations, with Shaheed Sharani I at 22.5 mg/L, Shaheed Sharani II at 21.5 mg/L, and Nirjhor at 22.3 mg/L. This is likely due to intensive fish farming which can contribute to nutrient loading by means of fish feed and organic waste. Nitrate levels were significant in Shaheed Sharani I (22.3 mg/L), while Nirjhor exhibited lower nitrate levels (0.9 mg/L). According to Fig. 3, chlorophyll-a concentrations were highest in Shaheed Sharani I (184.3 µg/L), which could be a result of the stagnant water and favourable conditions for algal blooms driven by nutrient availability. Shaheed Sharani II (47.9 µg/L) and Nirjhor (55.07 µg/L) exhibited lower concentrations, suggesting less favorable conditions. The controlled access to these lakes limits external pollution sources, suggesting that internal nutrient cycling and anthropogenic activities are the primary eutrophication drivers (Carvalho *et al*., 2006; Solimini *et al*., 2006). The results indicate that these lakes range from eutrophic to hypertrophic zones.

Dhanmondi, Gulshan, and Banani

As demonstrated in Figure 4, Dhanmondi Lake had low nitrate (1.2 mg/L) and phosphate (6.1 mg/L) levels, along with chlorophyll-a (13.99 µg/L) as seen in figure 3. This reflects reduced pollution inputs and effective nutrient management. Gulshan Lake exhibited low nitrate (4.1 mg/L) and moderate phosphate (9.1 mg/L) levels (figure 4), and higher chlorophyll-a levels (79.45 µg/L) as illustrated in figure 3, possibly due to urban runoff, supporting algal growth.

Analysing the Figure 4, Banani Lake's moderate phosphate levels (17 mg/L) may come from urban runoff and sewage containing detergents and fertilizers, while low nitrate levels (5.2 mg/L) may be due to pollutant toxicity affecting nitrate production (Carvalho *et al*., 2006; Solimini *et al*., 2006). It has been observed that Banani lake has faecal matter floating all over the surface. This observation is supported by the fact that there is a sewerage line nearby discharging its waste into the lake. The sewage disposal could contribute to the high level of nitrates in the lake. These lakes range from a mid-mesotrophic to a eutrophic zone and are detected as such by the CNN model.

Hatirjheel, Ramna, and Uttara Lakes

From the Figure 4, it can be seen that Hatirjheel exhibited moderate nitrate (3.4 mg/L) and phosphate (4.2 mg/L) levels, yet showed high chlorophyll-a concentrations (185 µg/L) in Figure 3, likely due to internal nutrient recycling, high sunlight exposure, and stagnant water conditions (Carvalho *et al*., 2006; Solimini *et al*., 2006). Ramna Lake recorded low nitrate (0.6 mg/L) and phosphate (5.9 mg/L) levels as shown in Figure 4, along with the low chlorophyll-a concentration (15.65 µg/L) shown in figure 3, indicative of effective urban maintenance practices that reduce nutrient accumulation and algal growth. Uttara Lake displayed low nitrate (1.8 mg/L) and phosphate (3.4 mg/L) levels as depicted in Figure 4, with moderate chlorophyll-a concentrations (63.54 µg/L) which can be seen in Figure 3. This state may be influenced by occasional nutrient runoff and sediment recycling (Carvalho *et al*., 2006; Solimini *et al*., 2006). These lakes range from the oligotrophic to the mesotrophic zone and are detected as such by the CNN model.

Ramna and Uttara are relatively protected zones, with good maintenance and supervision by city authorities, which explains the lack of eutrophication in those lakes. However, Hatirjheel has always been under great public exposure. Even though the exposure pollutes Hatirjheel Lake, there is no point source discharging a significant amount of nutrients into the lake. At the same time, no fish farming takes place in these lakes.

3.2 Image Detection

Figure 5 demonstrates the CNN Model running for individual images on the Kaggle Notebook after completing the training.

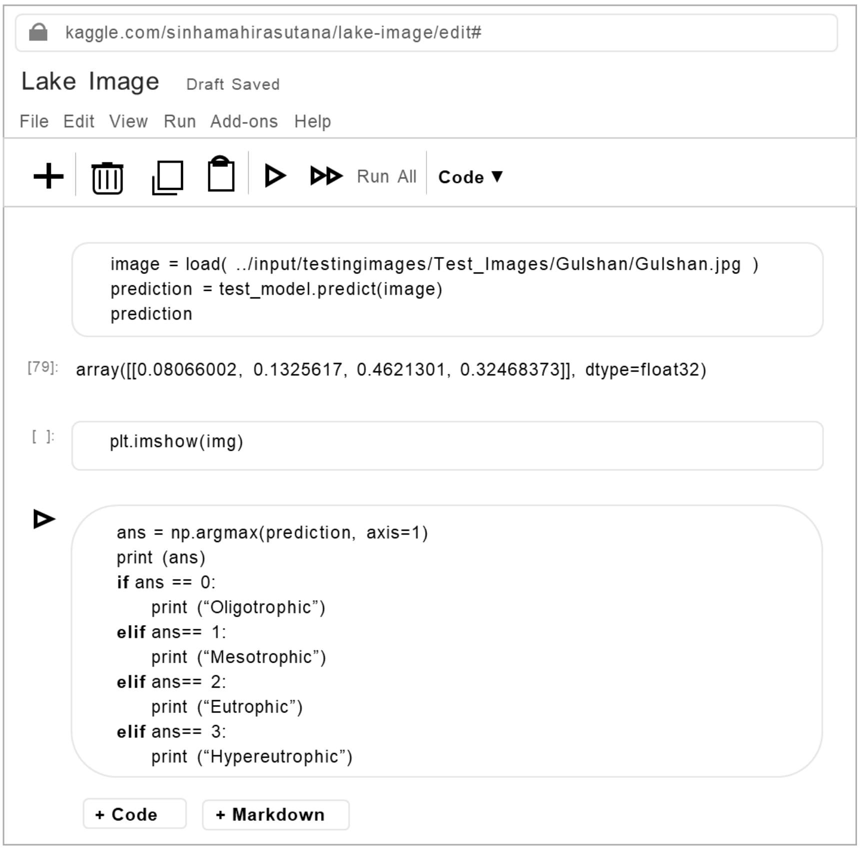

Figure 5: CNN Model running for individual images on the Kaggle Notebook

Figures 6 and 7 illustrate the CNN Model running for individual images on the Kaggle Notebook after completing the training.

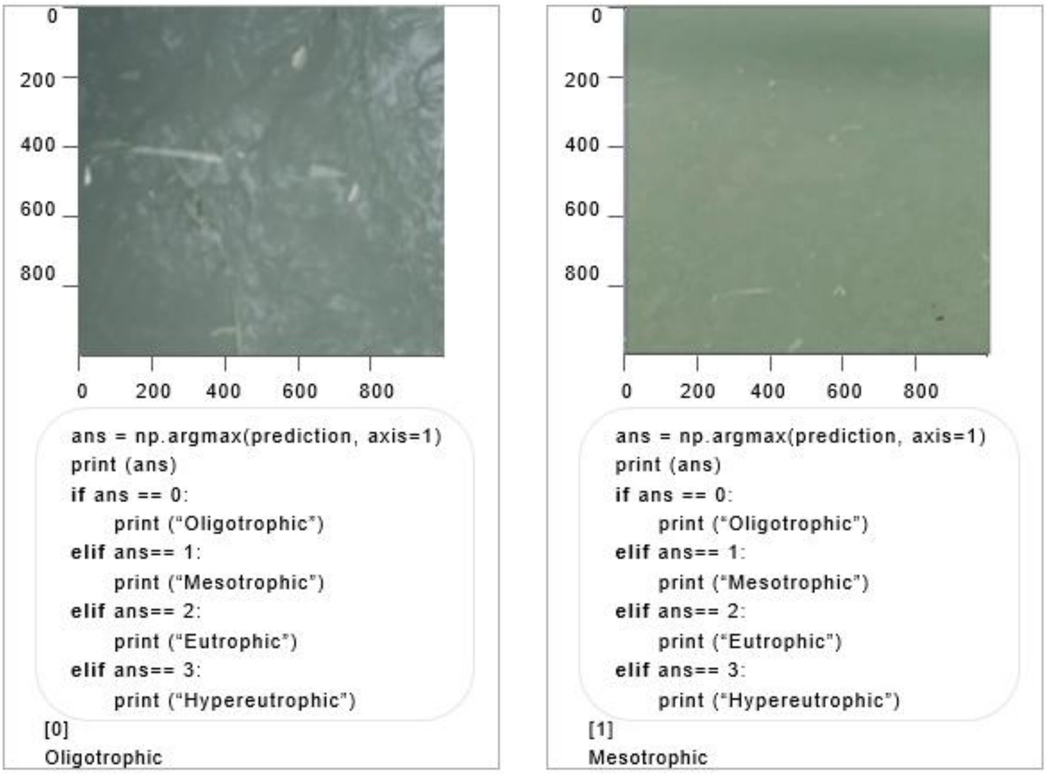

Figure 6: Uttara and Ramna Lake Images Being Detected as Oligotrophic and Mesotrophic, respectively.

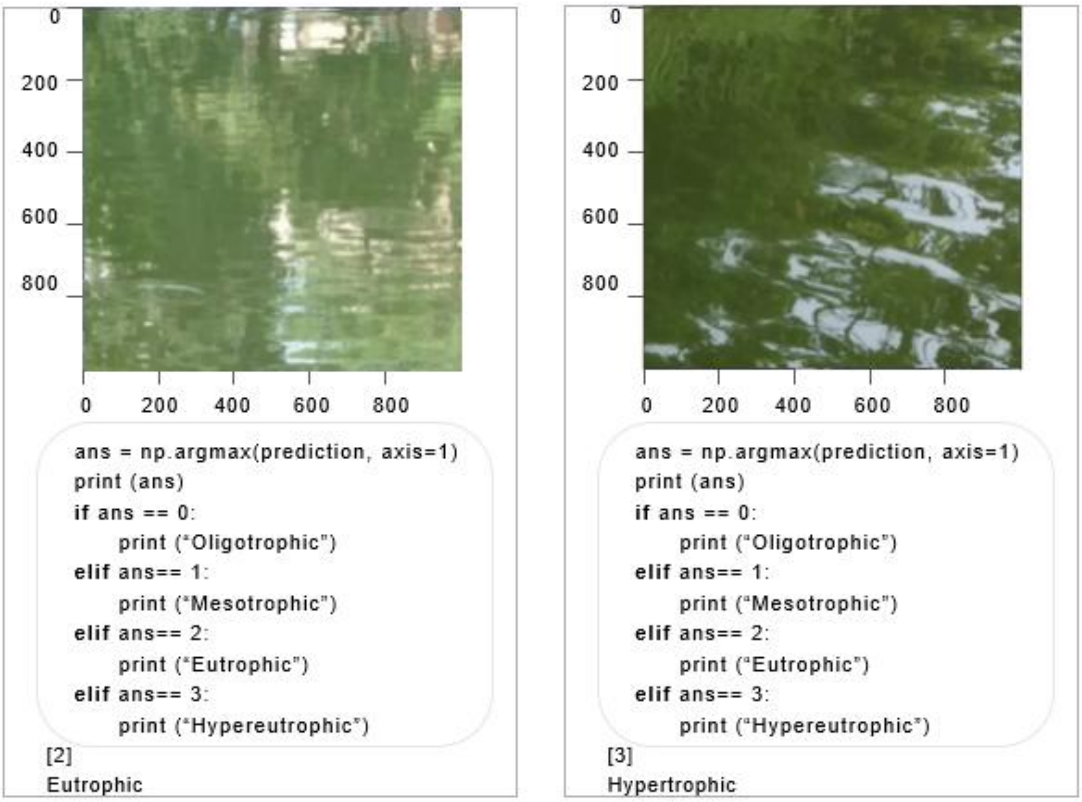

Figure 7: Gulshan and Shaheed Sharani Lake Images being Detected as Eutrophic and Hypereutrophic, respectively.

The confusion matrix shows the number of images identified in each category, correctly or incorrectly, by the CNN model. This is illustrated in Figure 8. This matrix was generated using the 160 images in the test dataset to determine the model's accuracy. Out of the 40 oligotrophic test images, 31 were correctly identified. Eight were detected as mesotrophic and one as eutrophic. Similarly, 32 mesotrophic samples were detected correctly, while 2 were marked oligotrophic and 6 were marked eutrophic. For the eutrophic samples, only three were incorrectly labelled as mesotrophic. The remaining images were detected correctly. Finally, 27 hypereutrophic samples were correctly identified, while the other 13 were not. These results show a 79.4% accuracy for the model.

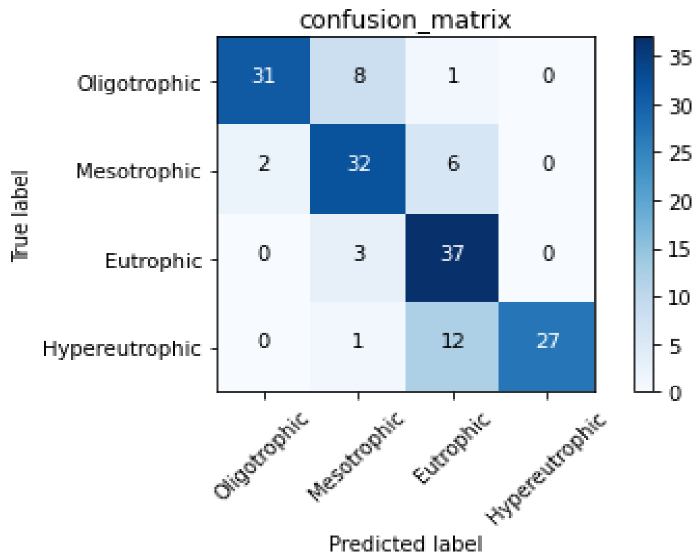

Figure 8: Confusion Matrix Showing Result of Image Detection

4. CONCLUSIONS

In this study, a multiclass convolutional neural network model was designed to identify the trophic status of selected lakes. Water samples were tested for total nitrate, total phosphate, and chlorophyll-a. The results were used to divide the sample images into four categories to train and test the CNN model. As per this study, the lakes cannot be strictly identified into four categories. This is because the level of nutrients and chlorophyll-a vary significantly throughout each lake, owing to the different point sources discharging into the same lake. Thus, a trophic range may be defined for each lake. The CNN model correctly detected most of the samples. However, it incorrectly identified some images, resulting in an overall accuracy of 79.4%. Thus, the identification of lake samples using CNN image detection has shown to be quicker and more efficient than testing all the parameters in the laboratory and performing calculations every time.

In Dhaka Cantonment, Shaheed Sharani I and Nirjhor lakes exhibit high phosphate concentrations (22.5 mg/L and 22.3 mg/L) due to intensive fish farming, with Shaheed Sharani I showing a high chlorophyll-a level (184.3 µg/L) driven by nutrient availability. These lakes range from eutrophic to hypertrophic zones. Banani, Gulshan, and Dhanmondi lakes fall in the mid-mesotrophic to eutrophic range. Banani's moderate phosphate levels (17 mg/L) stem from urban runoff and sewage inputs, while low nitrate levels (5.2 mg/L) reflect pollutant toxicity. Gulshan and Dhanmondi lakes show lower nutrient levels due to better management practices. Ramna, and Uttara lakes, in the oligotrophic to mesotrophic range, have lower nutrient levels due to effective maintenance. Ramna and Uttara lakes benefit from reduced external inputs. Hatirjheel Lake's high chlorophyll-a (185 µg/L) results from internal nutrient recycling and stagnant water.

The study has scope for improvement since trophic states do not have one fixed or standard classification system. They can be computed in many ways, per studies and experiments conducted by different researchers. The lake classification also depends on multiple parameters, which do not always have a consistent correlation. For determining trophic states, the index values calculated, or parameters tested are not always consistent with the colour of lakes. Based on the calculation, some lakes may be identified as eutrophic, whereas the colour of the lake is very light green or bluish. This problem arises because eutrophication is not the only factor controlling the colour of lakes. Lake colour may vary for several other reasons. A greater number of samples may be used to train the CNN model further to improve its accuracy. Trophic state index (TSI) values may be included in the coding, along with images, to train the CNN model. This approach could provide an approximate TSI value for each image, thus widening the model's scope.

ACKNOWLEDGEMENT

The authors extend their heartfelt gratitude to M Tauhid Ur Rahman, Professor of Civil Engineering, for his invaluable guidance. Special thanks are extended to the Department of Civil Engineering at the Military Institute of Science and Technology for providing access to exceptional testing facilities, enhancing the quality and scope of the research.

REFERENCES

Bali, M., Gueddari, M. (2019). Removal of phosphorus from secondary effluents using infiltration–percolation process. *Applied Water Science* 9, 54. https://doi.org/10.1007/s13201-019-0945-5

Brezonik, P. L., & Shannon, E. E. (1971). Trophic State of Lakes in North Central Florida, Research Project Technical Completion Report. *Florida Water Resources Research Centre*.

Burns, N., McIntosh, J., & Scholes, P. (2009). Strategies for Managing the Lakes of the Rotorua District, New Zealand. *Lake and Reservoir Management*, 21(1), 61-72.

Carlson, R.E. and Simpson, J. (1996). A Coordinator's Guide to Volunteer Lake Monitoring Methods. *North American Lake Management Society*. 96 pp.

Carlson, R. E. (1977). A trophic state index for lakes. *Limnology and Oceanography*, 22(2).

Carvalho, L., Lepistö, L., Rissanen, J., Pietiläinen, O.-P., Rekolainen, S., Torok, L., Solheim, A. L., Saloranta, T., Ptacnik, R., Tartari, G., Cardoso, A. C., Premazzi, G., Gunn, I., Penning, E., Hanganu, J., Hellsten, S., Orhan, I., & Navodaru, I. (2006). Nutrients and eutrophication in lakes. *Eutrophication and organic pollution* (Chapter 6).

Chipman, J. W., Lillesand, T. M., Schmaltz, J. E., Leale, J. E., & Nordheim, M. J. (2004). Mapping lake water clarity with Landsat images in Wisconsin, USA. *Canadian Journal of Remote Sensing*, 30(1), 1-7.

El-Serehy, H. A., Abdallah, H. S., Al-Misned, F. A., Al-Farraj, S. A., Al-Rasheid, K. A. (2018). Assessing water quality and classifying trophic status for scientifically based managing the water resources of Lake Timsah, the lake with salinity stratification along the Suez Canal. *Saudi Journal of Biological Sciences*, 25(7), 1247-1256.

Fuller, L. M., Aichele, S. S., & Minnerick, R. J. (2004). *U.S. Geological Survey Scientific Investigations Report* 2004-5086.

Gullian-Klanian, M., Gold-Bouchot, G., Delgadillo-Díaz, M., Aranda, J., & Sánchez-Solís, M. J. (2021). Effect of the use of *Bacillus spp.* on the characteristics of dissolved fluorescent organic matter and the phytochemical quality of *Stevia rebaudiana* grown in a recirculating aquaponic system. *Environmental Science and Pollution Research, 28*(36), 36326–36343. https://doi.org/10.1007/s11356-021-13148-6

Ho, J. C., Michalak, A. M., & Pahlevan, N. (2019). Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature*, 574(7780), 667–670. https://doi.org/10.1038/s41586-019-1648-7

Horton, R. K. (1965). An Index Number System for Rating Water Quality. *Journal of the Water Pollution Control Federation*, 37, 300-306.

Janiesch, C., Zschech, P. & Heinrich, K. (2021). Machine learning and deep learning. *Electron Markets* 31, 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jiang, Y., Xu, Z., & Yin, H. (2006). Study on improved B.P. artificial neural networks in eutrophication assessment of China eastern lakes. *Journal of Hydrodynamics*, 18(3), 528-532.

Kuo, J.-T., Hsieh, M.-H., Lung, W.-S., & She, N. (2006). Using artificial neural network for reservoir eutrophication prediction. *Ecological Modelling*, 200(1), 171-177.

Michalski, M., & Conroy, N. (1972). Water quality evaluation-Lake Alert study, Ontario *Ministry of the Environment Report*. p. 23.

Middleton, M. (2021). Deep Learning vs. Machine Learning — What's the Difference? [Online]. Available at: https://flatironschool.com/blog/deep-learning-vs-machine-learning/.

Naumann, E. (1919). Some aspects of the ecology of the limnoplankton, with special reference to the phytoplankton. [Translation from: Svensk Botanisk Tidskrift 13(2) 129-163, 1919.]. *Freshwater Biological Association*.

Nosrati, K., Govers, G., & Smolders, E. (2012). Dissolved organic carbon concentrations and fluxes correlate with land use and catchment characteristics in a semi-arid drainage basin of Iran. *Catena, 95*, 177–183. https://doi.org/10.1016/j.catena.2012.02.019

Olmanson, L. G., Bauer, M. E., & Brezonik, P. L. (2007). A 20-year Landsat Water Clarity Census of Minnesota's 10,000 Lakes. *Remote Sensing of Environment*.

Paulic, M., Hand, J., & Lord, L. (1996). Water-quality assessment for the state of Florida. *Florida Department of Environmental Protection*.

Solimini, A. G., Cardoso, A. C., & Heiskanen, A.-S. (2006). Indicators and methods for the ecological status assessment under the Water Framework Directive: Linkages between chemical and biological quality of surface waters. *European Commission Joint Research Centre*. Luxembourg: Office for Official Publications of the European Communities.

Wojtkowska, M., & Bojanowski, D. (2021). Assessing trophic state of surface waters of Służewiecki Stream (Warsaw). *Applied Water Science*, 11, 118.