International Journal of Multidisciplinary and Current Educational Research (IJMCER)
ISSN: 2581-7027 ||Volume|| 6 ||Issue|| 6 ||Pages 73-81||2024||

Model Design of LSTM Based on Enhanced NAdamClip Algorithm in Prediction of Ammonia in Aquaculture

[1,]Tu Jun, [2,]Azman Yasin , [3,]Nur Suhaili Mansor
[1,2,3,]School of computer, University Utara Malaysia, Kedah, Malaysia

ABSTRACT : The prediction of ammonia concentration in aquaculture is of great significance for maintaining water quality and protecting the health of aquaculture species. However, due to the complexity of water quality parameters and measurement errors, the traditional prediction methods are faced with many challenges. In order to improve the stability and prediction accuracy of the model training, an LSTM model based on the improved NAdamClip optimization algorithm was proposed for the prediction of ammonia concentration in aquaculture. By combining gradient clipping technique with NAdam optimization algorithm, the problems of gradient explosion and model instability which are easy to appear in the training of deep neural networks by traditional NAdam algorithm are solved. By analyzing the time series characteristics of ammonia concentration, the training flow of the improved algorithm is designed to overcome the instability of the traditional NAdam algorithm in the LSTM network. The model design provides a new technical path for aquaculture water quality management in the future, and lays a foundation for sustainable development of aquaculture industry.

KEYWORDS - Nadam, Gradient Clipping Technique, LSTM Model

I. INTRODUCTION

A. Background: Aquaculture is an economic activity that uses artificial means to cultivate and raise various aquatic organisms in water bodies. As the population grows and the demand for high-quality protein increases, aquaculture has become an effective way to compensate for the inability of traditional fisheries to meet demand [1]. In general, in most farming environments, concentrations of free ammonia above 1.9 milligram/liter may pose a threat to fish [2]. In practical problems, it is impossible to determine when to predict ammonia levels due to the variability of ammonia sources, diversity of water quality parameters, measurement errors, and challenging chemical interactions, making accurate prediction a complex task. Ammonia, often abbreviated as ammonia-N or NH3-N, is a compound that contains both ammonia (NH3) and nitrogen (N) [3]. In aquaculture ammonia (NH3) also is a critical water quality parameter, as it can be toxic to aquatic organisms at elevated concentrations [4]. In general, in most farming environments, concentrations of free ammonia above 1.9 mg/L may pose a threat to fish [2]. The breakdown of ammonia from fish waste is vital for fish health. Fish defecate, releasing ammonia (NH3) into the water. This ammonia undergoes nitrification, facilitated by specific bacteria. Firstly, Nitrosomonas bacteria convert ammonia into nitrites (NO2-), and then Nitrobacter bacteria turn nitrites into nitrates (NO3-). High ammonia and nitrite levels are toxic to fish, causing gill damage, stress, and in severe cases, death. Nitrites also impair blood's oxygen-carrying capacity, leading to brown blood disease. While nitrates are less toxic, they can still harm young or sensitive fish species [5]. Accurate prediction of ammonia levels in aquaculture practices is essential for maintaining water quality, ensuring the health and productivity of cultured species, and promoting the sustainability of the industry [6]. The process is shown in Fig. 1.

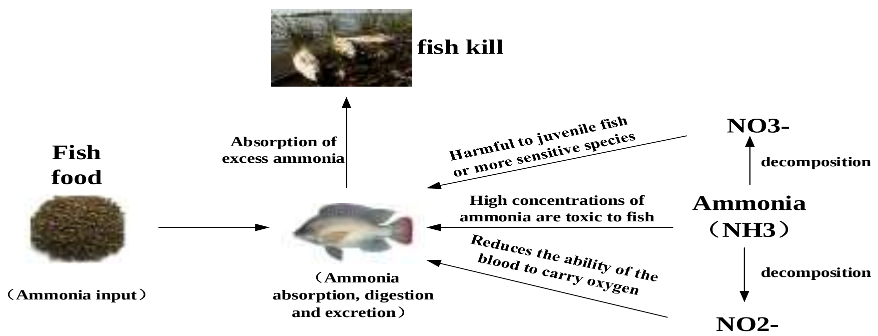

Fig. 1. Flowchart of Ammonia harms fish

Given the complexity and importance of maintaining optimal water quality in aquaculture, accurate prediction of ammonia levels is crucial. Traditional methods may not suffice due to the variability and interactions of ammonia sources, water quality parameters, and measurement errors. To address the challenge of predicting ammonia levels, various machine learning models like ANN, CNN, and RNN have been developed. However, these models lack memory cells, making them unsuitable for tasks involving sequences or time-series data, such as predicting ammonia in aquaculture. To address this limitation, LSTM (Long Short-Term Memory) models are chosen for ammonia prediction. Advanced machine learning techniques, particularly Long Short-Term Memory (LSTM) networks, offer promising solutions by capturing dependencies and patterns over time, making them suitable for tasks that involve time-series data [7].LSTM models are designed to handle the sequential nature of ammonia concentration data, thereby improving prediction accuracy and supporting the sustainability of aquaculture practices[8]. LSTM models excel in handling complex nonlinear relationships and adapt well over time. Additionally, for efficient training of deep neural networks, the stochastic gradient descent (SGD) algorithm is widely used.

To address this limitation, LSTM (Long Short-Term Memory) models are chosen for ammonia prediction. LSTM models excel in handling complex nonlinear relationships and adapt well over time. Additionally, for efficient training of deep neural networks, the stochastic gradient descent (SGD) algorithm is widely used.

B. Problem statement : NAdam are often used with artificial neural network models to solve specific challenges in training deep learning models [6]. However, these studies have found the complexity and potential limitations of NAdam training LSTM models, such as gradient explosion during the training process, resulting in fluctuations in the loss curve and instability problems in the model [9]. Therefore, when training very deep or very long time-step LSTM networks, where NAdam does not clearly define the gradient, the gradient may accumulate during backpropagation, resulting in the model weight values becoming very large. This cumulative effect is called gradient explosion, and it causes the model to easily fall into local optimal solutions [10]. To overcome this limitation, the gradient clipping was early recognized as standard technique to control the size of gradient updates and serve as a straightforward potent technique [11].

Previous studies have proposed various gradient clipping approaches such as clipping by norm and clipping by value [12]. To date, some of the clipped techniques employed to stochastic gradient algorithms include the sub gradient-based clipping techniques [10], autoclip [13], and restarted clipped-SSTM first order technique[14], batch clipping and adaptive layer wise clipping techniques [15], and Elesedy & Hutter proposed the U-Clip, which was trained using dataset CIFAR10 and ImageNet (combined with SGD, momentum or Adam), CIFAR10 training results arrive at 99% training accuracy [16].

The disadvantage of these gradient clipping techniques in combination with stochastic gradient descent algorithms is that SGD lacks an adaptive learning rate, making it less desirable for solving instability. Moreover, the aforementioned empirical studies on the combining SGD, Adam with clipped techniques to stabilized the training of RNN models, literatures indicate that no studies have looked at the effects of combining clipping method with Nadam optimization algorithm to RNN such as LSTM. The gaps that exist in highlighted prior works have paved the way for this proposed study NAdamClip by to enhancing the Nadam using the clipping technique which may ultimately addressed the said limitations. This can help the model adapt to the characteristics of the data. To address the problem of NAdam optimization algorithm being unstable during model training, an effective approach is to augment the NAdam algorithm with gradient clipping techniques and adjust the gradient clipping range. This adjustment helps to set a threshold for reducing the gradient, preventing problems caused by too large a gradient. By experimenting with different cropping ranges, a balance can be found between preventing gradient explosion and ensuring that the model learns effectively. Suitable gradient clipping ranges are usually determined through experimentation and tuning for specific models and datasets, which not only avoids too small or too large gradient values, but also optimizes the stability of training.

II. RESEARCH METHOD

This chapter outlines the methodology to be employed in this proposed research. The study will adhere to a structured approach consisting of four stages, as illustrated in Figure 3. These stages encompass key four tasks: data collection & preparation, enhancement design, implementation and evaluation. The simplified form of proposed methodology is illustrated as process to explain the proposed steps to enhance the NAdam optimizer in Fig 2.

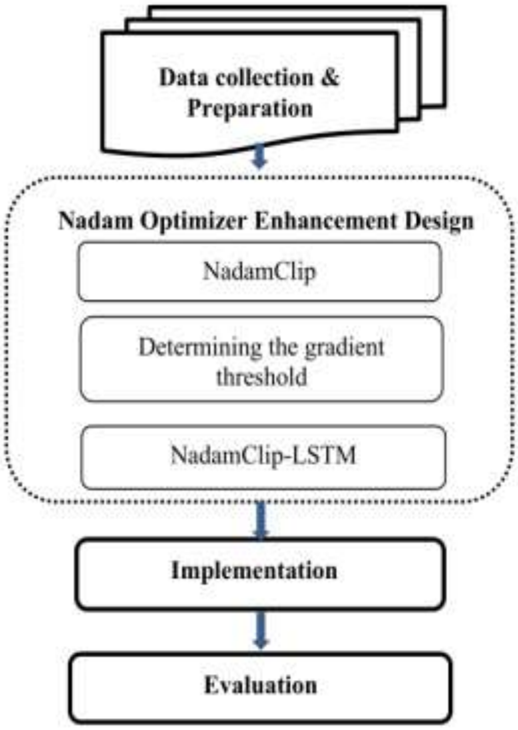

Fig. 2. Research Process

A. Data Collection and Preparation : Two sets of time series data—Data sets P1 and P2, respectively—will be used in this proposal. The section contains a description of these data sets. The initial time series dataset proposed for use in this proposal study corresponds to the datasets of three different ponds, such as P1 and P2, which were initially used in the open-source data sensor-based aquaponics fish ponds dataset from Kaggle. Details and characteristics of the two different pond datasets with sensors are tabulated in Table 1.

TABLE I. DATASETS P1, P2 OF TWO DIFFERENT PONDS

TABLE II.

Datasets	Historical	Input	Symbol
P1&P2	19 June 2021 to 31 October 2021	Temperature	TP
		Dissolved Oxygen	DO
		potential of hydrogen	PH
		Ammonia	Ammonia

Data set P1 consist of 83,127, and P2 consist of 76,388, and is available at Sensor based aquaponics Fish Pond Datasets.

The dataset for this proposed study is indeed a time series dataset, as it contains measurements of water quality parameters at different timestamps. This Time series data is often suitable for models like LSTM networks, because there is a temporal dependency in the data. LSTM networks indeed a RNN designed to capture patterns and dependencies over time. They are particularly effective in handling sequences and time series data due to their ability to remember information for long periods. They are particularly effective in dealing with long-term dependent sequences and large datasets because of their complex structure and ability to remember information over long periods of time. Figure 3 shows the ammonia visualization of the P1 dataset. Figure 4 shows the ammonia visualization for the P1 dataset and Figure 4 shows the ammonia visualization for the P2 dataset.

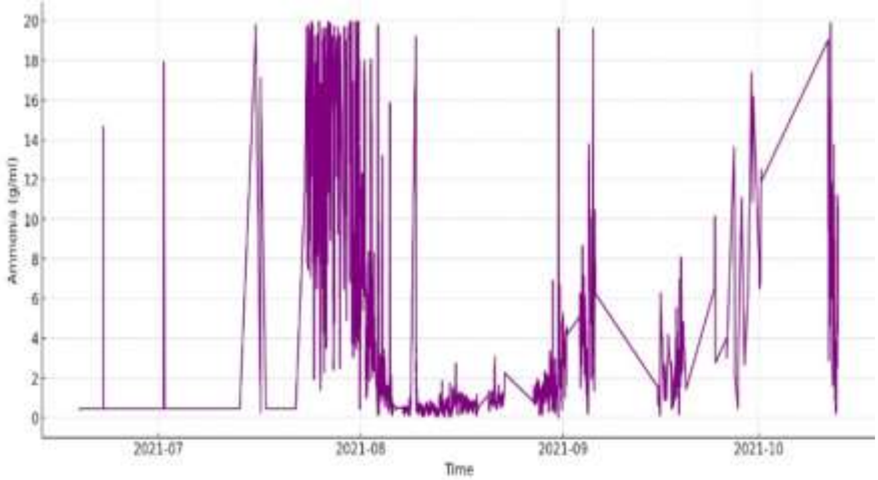

Fig. 3. P1 data set ammonia visualization

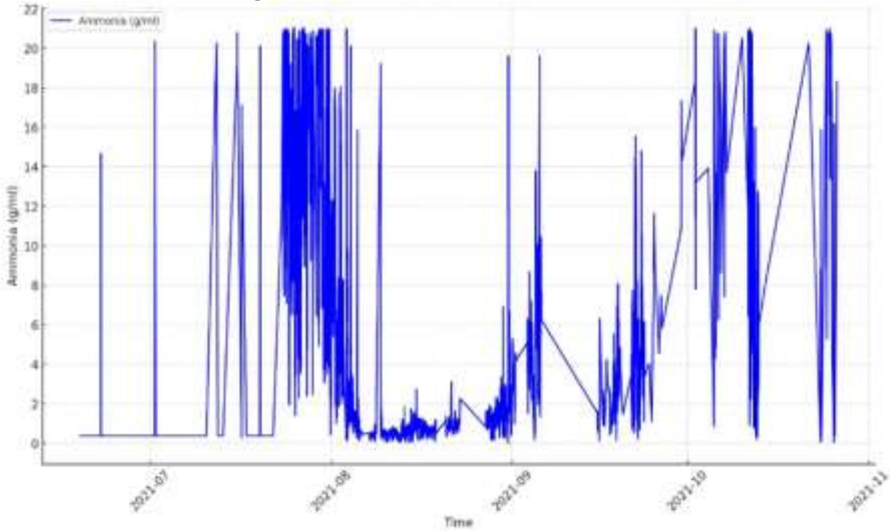

Fig. 4. P2 data set ammonia visualization

This graph is a visualization of ammonia for the P1 and P2 dataset and shows that ammonia level does not increase or decrease uniformly with time. Instead, there are spikes, decreases, and plateaus, indicating a complex relationship between time and ammonia level that cannot be captured in a straight line. The visualized ammonia data results indicate that some data points deviate significantly from the overall trend, suggesting that there are other factors or complex interactions affecting ammonia levels. These fluctuations, as well as the presence of outliers and non-constant rates of change, suggest that linear models are not sufficient. Instead, the data require more sophisticated non-linear methods to accurately model underlying patterns and predict future values, which are characteristic of non-linear systems.

Based on the two datasets P1 and P2 the calculation of the Hurst Index for ammonia in the two datasets can be argued that ammonia exhibits long-term dependence. Ammonia was determined to be a long-term dependent formula 1 using the Hurst index method:

$$H = \frac{log\left(\frac{R}{S}\right)}{log\ (seg\ ment\ size)} \tag{1}$$

In this example, the Hurst exponent values for ammonia in both P1 and P2 are less than 0.5, specifically 0.284 and 0.267, which indicates that there is a long-term negative correlation, and the smaller the value, the stronger the negative correlation. If a time series exhibits long-term negative correlation, then higher values tend to follow lower values and vice versa, even if the lag time is long. This is the opposite of what one might observe in short-term dependence, thus indicating the degree of long-term dependence in the ammonia time series. The LSTM network is particularly effective when dealing with sequential and time series data due to its ability to remember information over long periods of time. In order to improve the quality and applicability of the prediction model in this study, we considered pre-processing two samples from different ponds (P1, P2) with the following data pre-processing procedure:

Fig. 5. Data preprocessing process

- The standard deviation method [17] was implemented in python for data cleaning and the simple moving average (SMA) method [18] was used for data denoising.
- The multivariate scatterplot matrix method was used to demonstrate the non-linear relationship of multiple variables using python software tool.
- The min-max normalization method will be used to normalize the values of the input and output variables to a specific range.

B. Research Model Design and Evaluation : This stage describes the process of training an LSTM model using the NAdamClip algorithm, and then describes the way to optimize the NAdam algorithm using the U-Clip algorithm, as well as presenting the gradient range table format. Finally, the initialization process of the LSTM model is presented.

Flow NAdamClip-LSTM : The proposed NAdamClip-LSTM algorithm consists of four main phases: the initialization phase, the data phase, the measurement phase, and the validation phase. The prediction model will be fed normalized training and validation data before initialization.

- Initialize the LSTM model, specifying the architecture, hyperparameters, and other relevant settings. Define the NAdamClip optimizer with appropriate hyperparameters

- Prepare the dataset for training, including data preprocessing, sequence padding, and splitting into training, validation, and test sets. Use the NAdamClip optimizer in the training loop. The LSTM model will be trained on the dataset using the proposed NAdamClip optimizer.

- The use metrics like mean absolute error (MAE) and mean squared error (MSE) will be use to track training and validation performance. Can visualize training curves to spot possible explosions early in the training process. It can also allow adjusting the clip value and other characteristics unique to the LSTM in light of the validation set's performance.

After training, the LSTM model is evaluated on a separate test set using NAdamClip to evaluate its generalization performance. Compare SGDClip, AdamClip, and NAdamClip-LSTM models with NAdam-LSTM models, hyperparameters, and details of key findings. Report final performance metrics and any insights gained during the experiment. Draw the proposal model according to the above steps, as shown in figure 6.

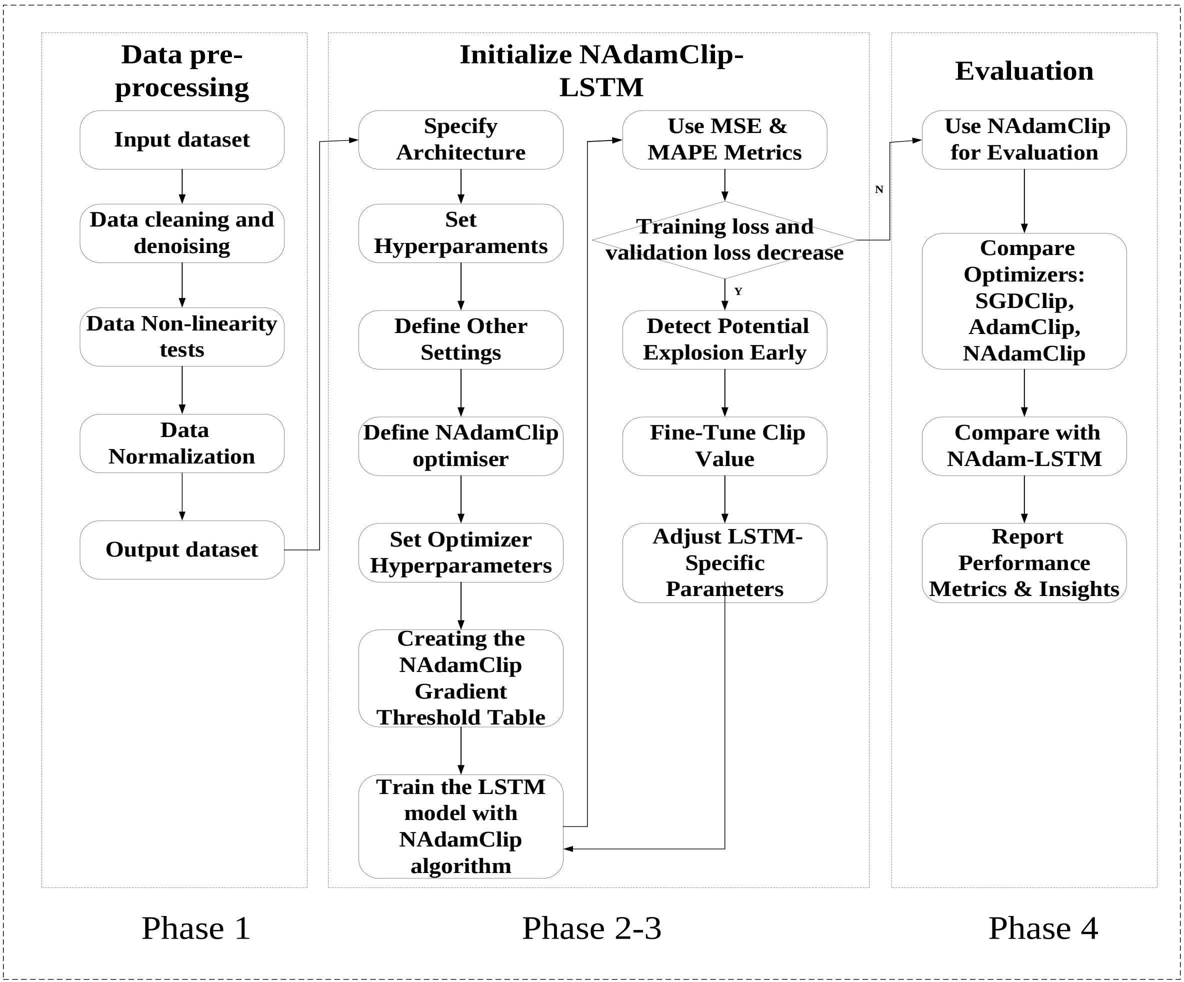

Fig. 6. Proposed Model

❖ **Initialization:** First, the task, such as time series prediction, must be clearly defined, along with the type of data involved. The dataset needs to be prepared, ensuring its suitability for the LSTM task. This involves splitting the data into training, validation, and test sets. Subsequently, the architecture of the LSTM network is designed, considering aspects like the number of layers, hidden units, and the dimensions of both input and output.

❖ **Evaluation :** The NAdamClip algorithm will be evaluated using model performance evaluation metrics and the performance of NAdam-LSTM and NAdamClip-LSTM will be compared to SGDClip and AdamClip, which serve as benchmarks.

❖ **Performance Metrics :** This proposed study will use two performance evaluation metrics to determine the performance of the enhanced NAdamClip algorithms on different dataset in this proposed study as presented in Chapter 3. As this proposed study relies on the analysis of time series data, it is imperative to emphasize the criticality of employing a suitable evaluation metric. The selection of an appropriate evaluation metric assumes paramount importance as it serves the crucial purpose of justifying the expected results. The metrics used for performance evaluation in this study are mean square error (MSE) and root mean square error (RMSE). MSE squares to emphasize large errors, making it sensitive to outliers. On this basis, RMSE provides the same and more intuitive and interpretable error measure as the predicted target unit.

III. RESULT

The expected results of this study include providing an efficient and accurate ammonia nitrogen prediction model to help the aquaculture industry better manage water quality. By comparing the performance of the NadamClip optimized LSTM model with the AdamClip, SGDClip and Nadadam models, we expect to demonstrate the significant advantages of the NadamClip optimized model for ammonia concentration prediction. First, we expect that the optimized LSTM model with NadamClip can significantly improve the accuracy of ammonia nitrogen concentration prediction and reduce the prediction error. This increase in accuracy will help the aquaculture industry to detect and deal with water quality problems in a timely manner, ensuring the health and growth of aquatic organisms.

Secondly, by introducing NadamClip optimization algorithm, we expect that the training and prediction time of the model will be effectively shortened, and the overall computing efficiency will be improved. The efficient computational performance not only saves resources, but also enables the model to be applied to real-world scenarios more quickly, providing timely water quality management support. In addition, we expect that the results of this study will provide practical guidance for water quality management in aquaculture industry and promote the progress and application of water quality monitoring technology. Specifically, the model developed in this study can more accurately predict changes in ammonia nitrogen concentration, take appropriate management measures, prevent water quality deterioration, and improve production efficiency.

This study is also expected to provide new ideas and methods for future research in the field of water quality prediction. The successful application of the NadamClip optimization algorithm in time series prediction is expected to demonstrate its potential in other similar tasks, driving the development and innovation of related technologies. Through comparative analysis, we not only validate the superiority of NadamClip in ammonia nitrogen prediction, but also provide valuable experience for future model optimization and application. Through the exploration and validation of this study, we expect to achieve breakthroughs in water quality monitoring and management, contributing to the sustainable development of the aquaculture industry. As technology continues to advance, we believe that such research will continue to optimize and improve, enabling the aquaculture industry to make greater progress in scientific management and environmental protection. In conclusion, this study not only proves the advantages of NadamClip optimization LSTM model in theory, but also demonstrates its application value in practice, providing a solid foundation for future related research and practical applications.

IV. CONCLUSION

The findings of this study hold significant implications for the water quality management of the aquaculture industry. By developing an efficient and accurate model for predicting ammonia nitrogen levels, this study offers a robust tool for improving the scientific rigor and effectiveness of water quality monitoring and management within aquaculture. This enhancement not only facilitates timely detection and response to water quality issues, ensuring the health and growth of aquatic organisms, but also advances the sustainable development of aquaculture. Furthermore, this study presents novel directions and insights for future research in environmental monitoring, showcasing the considerable potential of advanced optimization algorithms in practical applications. Through comparative analysis of different optimization algorithms, this study not only validates specific algorithmic advantages but also provides valuable references for other prediction tasks. Overall, this study makes innovative theoretical contributions while demonstrating substantial practical value. It promotes the advancement of scientific management and environmental protection technology within the aquaculture industry while providing a solid foundation and useful guidance for future research and practice.

REFERENCES

[1] C. E. Boyd *et al.*, “Achieving sustainable aquaculture: Historical and current perspectives and future needs and challenges,” *J. World Aquac. Soc.*, vol. 51, no. 3, pp. 578–633, 2020.

[2] J. Heise, H. Müller, A. J. Probst, and R. U. Meckenstock, “Ammonium Removal in Aquaponics Indicates Participation of Comammox Nitrospira,” *Curr. Microbiol.*, vol. 78, no. 3, pp. 894–903, 2021.

[3] A. A. Akindele and M. Sartaj, “The toxicity effects of ammonia on anaerobic digestion of organic fraction of municipal solid waste,” *Waste Manag.*, vol. 71, pp. 757–766, 2018.

[4] C. E. Boyd, *General relationship between water quality and aquaculture performance in ponds*. Elsevier, 2017.

[5] Z. Xu, J. Cao, X. Qin, W. Qiu, J. Mei, and J. Xie, “Oxidative Stress , Immune Responses and Tissue Structure in Fish Exposed to Ammonia Nitrogen : A Review,” *Anim. Rev.*, vol. 11, pp. 1–19, 2021.

[6] R. R. Karri, J. N. Sahu, and V. Chimmiri, “Critical review of abatement of ammonia from wastewater,” *J. Mol. Liq.*, vol. 261, pp. 21–31, 2018.

[7] S. S. Baek, J. Pyo, and J. A. Chun, "Prediction of water level and water quality using a cnn-lstm combined deep learning approach," *Water (Switzerland)*, vol. 12, no. 12, 2020.
[8] T. Awoke, M. Rout, L. Mohanty, and S. C. Satapathy, "Bitcoin price prediction and analysis using deep learning models," in *Communication software and networks*, Springer Singapore, 2021, pp. 631–640.
[9] X. Wang, J. Wu, C. Liu, H. Yang, Y. Du, and W. Niu, "Exploring LSTM based recurrent neural network for failure time series prediction," *Beijing Hangkong Hangtian Daxue Xuebao/Journal Beijing Univ. Aeronaut. Astronaut.*, vol. 44, no. 4, pp. 772–784, 2018.
[10] V. V. Mai and M. Johansson, "Stability and Convergence of Stochastic Gradient Clipping: Beyond Lipschitz Continuity and Smoothness," in *Proceedings of Machine Learning Research*, 2021, pp. 7325–7335.
[11] G. Black, *Museums and the challenge of change: Old institutions in a new world*, Old instit. Old institutions in a new world. Routledge., 2020.
[12] Z. H. Liu *et al.*, "A Regularized LSTM Method for Predicting Remaining Useful Life of Rolling Bearings," *Int. J. Autom. Comput.*, vol. 18, no. 4, pp. 581–593, 2021.
[13] P. Seetharaman, G. Wichern, B. Pardo, and J. Le Roux, "Autoclip: Adaptive gradient clipping for source separation networks," *IEEE Int. Work. Mach. Learn. Signal Process. MLSP*, vol. 2020-Septe, no. 1, 2020.
[14] E. Gorbunov, M. Danilova, and A. Gasnikov, "Stochastic optimization with heavy-tailed noise via accelerated gradient clipping," *Adv. Neural Inf. Process. Syst.*, vol. 2020-Decem, no. NeurIPS, 2020.
[15] T. N. Nguyen, P. H. Nguyen, L. M. Nguyen, and M. Van Dijk, "Batch Clipping and Adaptive Layerwise Clipping for Differential Private Stochastic Gradient Descent," pp. 1–20, 2023.
[16] B. Elesedy and M. Hutter, "U-Clip: On-Average Unbiased Stochastic Gradient Clipping," 2023, [Online]. Available: http://arxiv.org/abs/2302.02971
[17] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, 2013.
[18] V. Pandey and V. K. Giri, "High frequency noise removal from ECG using moving average filters," *Int. Conf. Emerg. Trends Electr. Electron. Sustain. Energy Syst. ICETEESES 2016*, pp. 191–195, 2016.

Biographies and Photographs

	*Tu Jun* is a lecturer and she is currently a Ph.D. student in Computer Science in the School of Computer Science, Faculty of Arts and Science, University of Northern Malaysia (UKM) Malaysia. Her main research interests are artificial intelligence, machine learning, and Internet of Things.
	*Dr. Azman Yasin* is An Associate Professor at the School of Computing, College of Arts and Sciences, University Utara Malaysia. He holds Ph.D. in Science dan System Management from University Kebangsaan Malaysia (UKM) Malaysia. His main research focuses are artificial intelligence, information system, and software engineering.

	*Dr. Nur Suhaili Binti Mansor* is a Senior Lecturer at the School of Computing, College of Arts and Sciences, University Utara Malaysia. He holds Ph.D. in Science dan System Management from University Pertanian Malaysia (UPM) Malaysia. Her main research focuses are Machine Learning, Remote Sensing, Other Geographical Information System (GIS)