Journal of Information Systems Engineering and Management
2025, 10(25s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Review on Deep Learning Approaches to Address Multi-Class Imbalance: An Emphasis on Water Quality Data

Manjusha Nambiar PV[1] *, Arpita Gupta [2]

Computer Science Engineering Department, Koneru Lakshmaiah Education Foundation, KL Deemed to be University, Hyderabad-500075, Telangana, India.

*Correspondence to: manjupnambiar@gmail.com

ARTICLE INFO

Received: 05 Jan 2025

Revised: 24 Feb 2025

Accepted: 04 Mar 2025

ABSTRACT

The deep learning methods for dealing with multi-class imbalance in water quality data are comprehensively examined in this paper. The findings show a notable increase in publications, especially since 2021, and a clear preference for deep learning techniques when classifying imbalanced data.

Background: A thorough review of the body of literature included articles from significant digital libraries that were published between January 2012 and December 2024. Based on several important factors, such as commonly used datasets and types, years of publication, different sources, research and empirical types, assessment measures, and development tools, 59 articles were chosen and examined.

Objective: This study examines how deep learning techniques respond to data that is unbalanced when there are multiple classes and how the deep learning models' large capacity and intricate structures make them appropriate for these kinds of tasks, with an emphasis on water datasets.

Conclusion: The paper emphasizes the difficulties with data diversity and computing efficiency, along with possible solutions for reliable real-world applications. It also discusses innovative solutions that enhance the reliability of real-world applications.

Keywords: Water Quality Classification, Deep Neural Networks, Multiclass Data Imbalance, Ensemble

1. INTRODUCTION

Water quality is essential to maintaining the well-being of populations and ecosystems. Water quality data has become increasingly complicated due to the rapid generation of massive data during the creation and use of smart water quality monitoring devices based on the Internet of Things (IoT). Water quality factors such as pH, dissolved oxygen, chemical concentrations, and contaminants must be accurately classified to maintain safe and sustainable environments. Monitoring systems frequently depend on machine learning (ML) and deep learning (DL) approaches to automate classification. However, class imbalance is a common problem in multi-class classification utilizing real-world water datasets because some quality levels (like major pollution incidents) are significantly less frequent than others (like normal quality circumstances). Some pollution levels or conditions in these datasets are underrepresented, creating an imbalance. The minority classes, often the most important from an ecological or health standpoint, are frequently misclassified. In the past decade, deep learning techniques have gained popularity because of the advancements in speech recognition, computer vision, and various other fields [1-3]. Their recent success can be ascribed to enhanced data availability, advancements in hardware and software, and numerous algorithmic innovations that expedite training and enhance generalization to novel data. Notwithstanding these advancements, less statistical research has been conducted to adequately assess methods for addressing class imbalance through deep learning and their associated structures, namely deep neural networks.

Deep learning models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), can be extensively applied to analyze time series such as water quality monitoring and spatial data, effectively capturing complex patterns and trends for improved predictive accuracy. In the study [4], using the powerful performance of long short-term memory (LSTM) deep neural networks in time-series prediction, a drinking-water quality model was designed to forecast massive amounts of water quality data using the advanced deep learning theory. However, these models struggle with imbalanced class distributions. Deep learning models are known to be data-hungry, requiring large, balanced datasets to achieve optimal performance. When the data distribution is skewed, deep learning models, especially those with large capacities, may overfit the majority classes, ignoring the underrepresented ones. In imbalanced datasets, training deep models may require more time or even fail to converge to a satisfactory solution due to the rarity of minority class instances. While deep learning can learn intricate patterns in data, it still requires careful handling of imbalances, especially in multi-class scenarios where each class should be learned equally well [5]. In many applications, the minority classes are critically important, even though they are underrepresented in the dataset. Detecting rare diseases may be far more critical than identifying common health conditions. Missing a rare fraudulent transaction can have significant financial implications, whereas misclassifying a legitimate transaction as fraud has fewer consequences. In cybersecurity, recognizing uncommon types of cyberattacks is crucial to prevent breaches. To address unbalanced human activities from smart homes and increase the learning algorithms' sensitivity to the minority class, the research in [6] suggests a data-level perspective in conjunction with a temporal window technique. As deep learning systems are increasingly applied in high-stakes environments, it becomes essential to

ensure that these models perform well across all classes, particularly the minority ones. Solving the imbalanced classification problem in deep learning will lead to more robust models that generalize to diverse real-world data, ensuring more reliable performance. The study in [7] seeks to summarize the most recent developments in this field by exploring the real-world application of imbalanced learning and presenting a thorough analysis of current approaches to address unbalanced learning. This paper provides a timeline between 2012 and 2024 illustrating the development of Deep Neural Networks (DNN) and the increasing prevalence of multi-class imbalance in section 5. The review focuses on strategies that address the imbalance problem in multi-class classification tasks, specifically within the context of water quality data.

1.1 Motivation and Contribution

Many researchers concur that the topic of deep learning with class-imbalanced data is not widely studied [8,9]. The multiclass imbalance problem is not as established as the binary class imbalance problem. The situation is more complicated in multiclass imbalance. There can be multiple majority and minority classes. Suppose an imbalanced dataset has three classes: A, B, and C. Here, class A can be the majority class concerning class B and also can be the minority class with respect to class C. So, in multiple classes, it is hard to identify the majority and minority classes concerning other classes. According to the study in [10], the type of class size configuration and the degree of class overlap significantly impact the difficulty in multiclass unbalanced data.

This review paper aims to provide a comprehensive overview of the methods, challenges, and recent advancements related to handling imbalanced data in multi-class classification using deep learning, primarily in water quality data sets. Here are the key objectives that should guide the structure and content of the paper:

- Highlight the unique challenges posed by multi-class classification problems compared to binary classification and investigate how class confusion between minority classes affects learning in deep neural networks and the lack of multi-class imbalance studies focused on water quality

- Include advanced strategies such as cost-sensitive learning, ensemble methods, and generative models like GANs to combat imbalanced class distributions.
- Provide insight into why deep learning models struggle with imbalanced classes despite their power in representation learning and lack of multi-class imbalance studies focused on water quality
- Analyse the role of architectural modifications in deep learning models that help mitigate imbalance, such as attention mechanisms, transfer learning, and meta-learning.

The following sections categorize the methods into data-level, algorithm-level approaches, and hybrid techniques. Discuss the merits and demerits of these methods, comparing their effectiveness in improving the performance of minority classes while maintaining good overall classification. The paper highlights the need for integrating advanced deep-learning techniques into water quality monitoring systems.

2. DATA-LEVEL SOLUTIONS

To resolve class imbalance and improve the representation of minority classes, data-level techniques focus on modifying the training set. A detailed analysis of data types reveals various modalities, with medical imaging and ecological datasets being particularly affected by imbalance issues. Techniques such as data augmentation, resampling, and the creation of synthetic data points have been successful in reducing these problems. Fig. (1) summarizes the data-level solutions.

2.1 Resampling Techniques

Two primary approaches to address the class imbalance problem are random under-sampling and random oversampling. This work [11] examined and confirmed the value of several sampling techniques over non-sampling techniques to produce a machine-learning model with a well-balanced sensitivity and specificity on unbalanced chemical data. This study [12] examined popular techniques from both categories assessed in this work for their capacity to improve the unbalanced ratio of five extremely unbalanced datasets from various application fields. To balance the classes, these methods add or remove samples at random. Samples eliminated in a random order can eliminate valuable information. The scenario is also the same for random oversampling, where performance matches up to under-sampling but requires more processing power to increase samples randomly in the minority class. These methods take extensive time, particularly when dealing with massive data sets. According to Chawla et al. (2002), [13] the two most popular oversampling techniques are the random replication of a small number of samples and the synthetic minority over-sampling technique (SMOTE). While SMOTE effectively balances classes, it can introduce noise and duplicate samples, reducing its reliability in high-dimensional datasets. Consequently, there are numerous variations of techniques that enhance SMOTE. ADASYN modifies the generation of synthetic samples based on minority class density [14]. Safe-Level-SMOTE [15] balances class distribution using a safe-level technique, thereby minimizing misclassification risks. A variation of SMOTE called borderline-SMOTE only generates synthetic samples on the boundary between two classes. Borderline-SMOTE [16] finds and synthesizes samples close to the decision boundary. SDSMOTE [17], a spatial distribution-based SMOTE, primarily creates an entirely new data set by obtaining raw spatial data distribution. When choosing majority-class sample subsets from various angles, two sample selection techniques—more precisely, a top-down and a bottom-up strategy—are suggested to preserve the original data distribution pattern. Using the SMOTE technique with a post-processing strategy to modify the datasets is another method of resolving the class imbalance issue. The investigation [18] uses the SMOTE method, the Tomek link, and the combination of these two resampling techniques for fault classification with simulated and experimentally imbalanced data. To address the class disparity in a non-noisy manner, A simple and effective oversampling strategy called k-means SMOTE was presented by Douzas et al. [19]. It combines SMOTE with k-means clustering.

The main target of these techniques is to identify the dataset's outlier, redundant, and noisy patterns and remove them for better classifier performance. The techniques Tomek-link and Edited Nearest-Neighbor rule are based on this. According to the study [20], the effect of class imbalance is significantly accelerated by redundant borderline occurrences and outliers in the data set. Cluster-based instance selection (CBIS), a unique under sampling technique that blends instance selection and clustering analysis, is presented in the study [21]. While the instance selection component removes unrepresentative data samples from each "subclasses," the clustering analysis component clusters similar data samples of the majority class dataset into "subclasses." Before completing a learning task, this work in [22] presents a novel package for recent relevant oversampling approaches to enhance data quality in imbalanced datasets. However, oversampling can introduce noise into the dataset, the risk of overfitting, increase computing costs, and reduce information gain. While using the oversampling technique, it is necessary to consider the distinct features of the dataset and the specific oversampling procedure. These methods lessen the issues brought on by class imbalance while enhancing the models' general quality and dependability. As a result, decisions based on the model's predictions are eventually better. Some researchers have been investigating sampling strategies that integrate spatial information between data classes, deep learning, and clustering algorithms using intrinsic features to improve the data further [23]. Examine how DL-based clustering algorithms are trained, highlight several clustering quality indicators, and assess various DL-based methods using three bioinformatics cases: biomedical text mining, cancer genomics, and bioimaging. Of course, different sampling methods can be categorized based on the complexity of the corresponding sampling method.

2.2 Data Augmentation

To determine the optimal combination to improve classification performance on unbalanced datasets, the study in [24] provides a generic framework for evaluating nine ensembles learning and nine data augmentation strategies for class imbalance challenges. Synthetic data points will be generated for underrepresented classes using the technique. Data augmentation for multiclass imbalance can take several forms. Data augmentation techniques are frequently used with image and time series data. Sensor-based data, such as water quality monitoring and environmental data, can benefit from its application. Adding random noise to existing data is known as noise injection, and it is a frequently used form of data augmentation that can produce new samples [25]. Minor alterations to the sensor readings of parameters such as pH, nitrate levels, or residual chlorine levels, for example, can provide new data points while preserving the data's original structure.

2.3 Generative Adversarial Networks (GANs)

By integrating a variational autoencoder (VAE) into a multi-head graph attention network (GAT), this

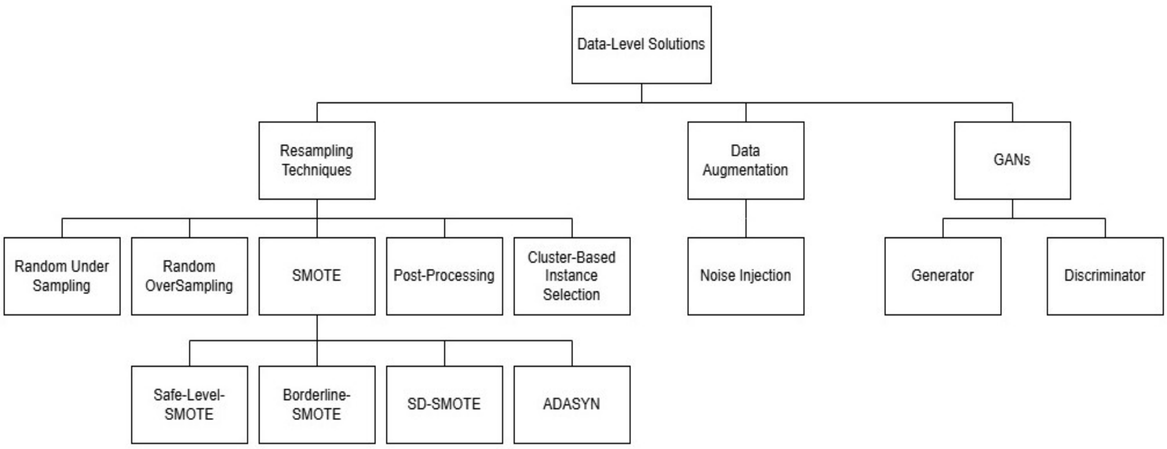

Fig. (1). Data level Solutions

study [26] gathered a thorough dataset from various projects about various construction activities, each with a unique structure and degree of class imbalance. Adding a generative model that enhances

the dataset for the underrepresented class enhanced the performance of building prediction models. The Deep Learning Generative Adversarial Random Neural Network, which has the same properties and capabilities as a GAN, is presented in this article. These models learn the underlying data distribution and generate new samples that mimic real-world data. A discriminator and a generator neural network make up a GAN. The discriminator tries to differentiate between synthetic and real data, while the generator produces artificial data points. Genetic algorithms based on biologically influenced agents like mutation, crossover, and selection have been mimicked by GANs, a neural technique that creates populations of individuals [27]. GAN can generate unusual event conditions, such as chemical pollution, in the context of water quality indicators. Deep learning models that require balanced datasets to produce accurate predictions may find these generated samples particularly useful.

3. ALGORITHM-LEVEL SOLUTIONS

These techniques work by modifying the algorithm by giving the highest priority to enhancing the algorithms' capacity to identify instances belonging to the minority class correctly. Generating algorithms to be more responsive to the underrepresented groups helps improve models' overall performance and generalization when presented with imbalanced data. Among the most prevalent algorithm-level strategies is cost-sensitive learning, wherein the classification performance is enhanced by modifying the algorithm's objective function. This alteration ensures that the model receives greater emphasis on learning from underrepresented classes. This study suggests using these measures as cost functions, frequently obtained using the confusion matrix [28]. Mienye et al. [29] focus on building robust cost-sensitive classifiers by altering the objective functions. Since the improved algorithms consider the unequal class distribution during training, altering the original data's distribution is not essential. This leads to more dependable performance than when the data is resampled.Fig.(2) summarizes algorithm-level solutions.

3.1 Cost-Sensitive Learning

Cost-sensitive learning takes into account the different costs associated with different types of

misclassification and aims to optimize the model for scenarios where the consequences of errors are uneven. Cost-sensitive cross-entropy is a common choice for multi-class imbalance in water quality data. Analyzing the performance of data-level approaches against algorithm-level approaches that emphasize cost-sensitive models and against a hybrid approach that combines those two approaches is the main purpose of the work [30]. To emphasize accurate minority class classification, cost-sensitive learning introduces penalty terms to the loss function [31]. By assigning higher costs to misclassified minority samples, models must improve their performance in these underrepresented categories. The study to improve the classification performance of SVMs [32] proposed a method that automatically adjusts the error cost between class samples to identify a preferred hyperplane effectively. In classification accuracy, it can evaluate the efficiency of an error cost, and, should it proven ineffective, it modifies

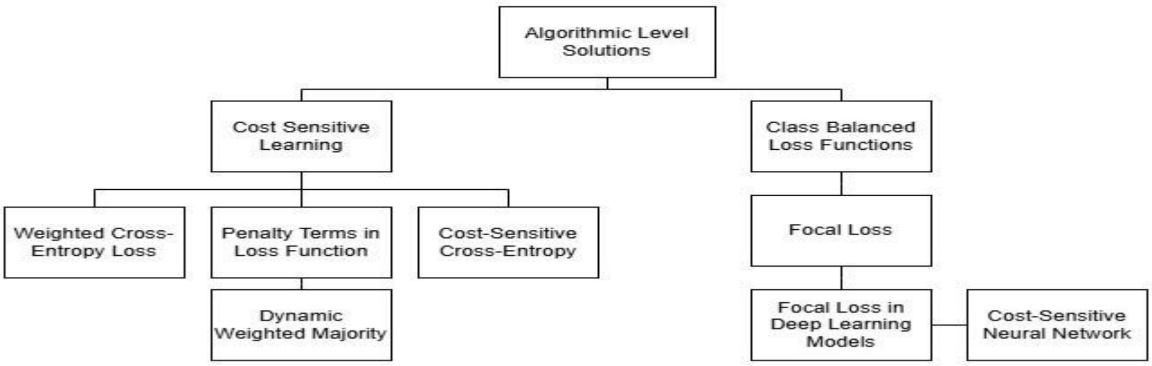

Fig. (2). Algorithmic level Solutions

the error cost in the correct direction. To handle imbalanced streaming data involving concept drift, the work in [33] presents a chunk-based incremental learning method termed adaptive chunk-based dynamic weighted majority. The loss function is modified to balance misclassifications of minority classes using cost-sensitive learning. Every class has a weight that is determined by inversely proportioning to how frequently it occurs in the training set in the Weighted Cross-Entropy Loss method. The weight W_c for a class C is calculated as

W_c = 1/frequency of class C.

Thus, classes with lower frequencies are given higher weights. This technique improves model performance on minority classes, which is particularly helpful for deep learning models that deal with multiclass imbalance issues by lowering the bias toward majority classes.[34] Suggested a class rebalancing technique based on a class-balanced dynamically weighted loss function, where weights are allocated according to class frequency and the predicted probability of ground truth class, to solve the imbalance in the class distribution in deep learning.

3.2 Class-Balanced Loss Functions:

Hard-to-classify examples can be focused more on using focal loss, an extension of cross-entropy loss. By lowering the weights of examples that are simple to categorize, this approach enhances the model's attention to uncommon occurrences in water quality datasets and successfully handles unbalanced datasets. Focal loss is a sophisticated function created to deal with the problem of class imbalance, especially when a model has to manage a major class difference. Facebook AI researchers first proposed the idea of focal loss for object detection in computer vision; it dynamically scales the cross-entropy loss and concentrates the model's training on samples that are difficult to categorize or incorrectly identified. The focal loss network intrusion detection system, a cost-sensitive neural network based on focal loss, is suggested to solve the unbalanced data issue in the study [35].

4. HYBRID APPROACHES

Hybrid approaches effectively address multi-class imbalance problems, particularly in intricate, high-dimensional data like water quality. To improve classification performance for minority classes, hybrid approaches combine algorithmic and data-level techniques to create a fairer representation of classes. For example, Hybrid Data Augmentation with a Loss Adjustment approach utilizes synthetic data generation techniques, such as SMOTE, alongside algorithmic changes like class-reweighting in the loss function. In water quality data, for instance, generating synthetic samples of rare contaminant levels while adjusting the loss function to penalize misclassification of these rare classes can significantly enhance the model's sensitivity to minority classes. The data-level augmentation increases the representation of minority classes in the training data, while loss adjustment methods—such as focal loss or class-balanced loss—bias the model to pay greater attention to these underrepresented classes. The study [36] combines SMOTE for augmenting minority class samples with algorithmic adjustments, enhancing classification performance in imbalanced water quality datasets. It details the impact of this hybrid approach on minority class detection accuracy in water monitoring contexts, showing improved results through a combined data and algorithmic methodology. One key limitation is the increased computational cost due to the hybrid approach, especially when applying SMOTE alongside algorithmic tuning for enhanced performance. This computational demand can make

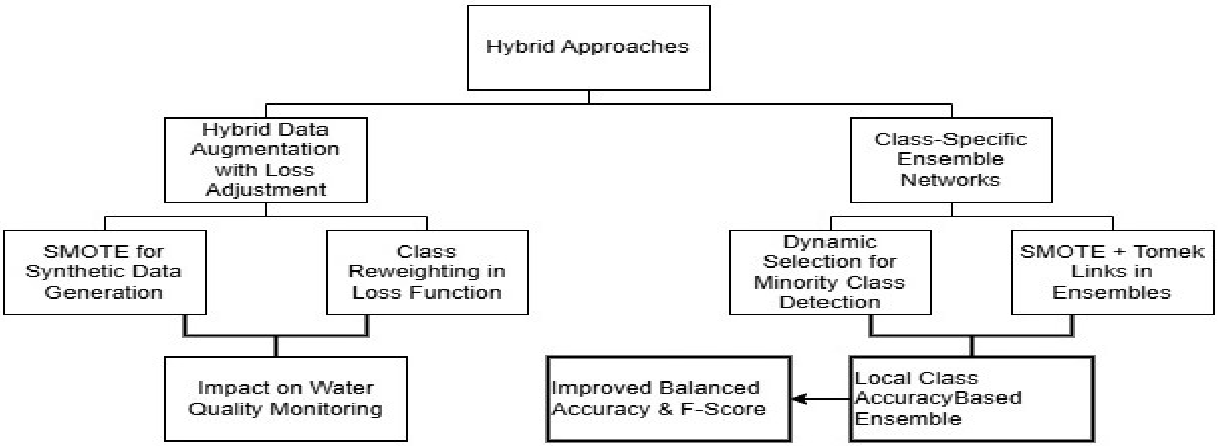

Fig. (3). Hybrid Approaches

real-time processing challenging. Class-specific ensemble networks are a powerful approach to addressing class imbalance in water quality data by tailoring ensemble components to enhance the representation of minority classes. The research study [37] demonstrates a dynamic selection approach to improve water quality anomaly detection, combining multiple ensemble methods with resampling techniques, including SMOTE and Tomek Links. When developers combine large amounts of heterogeneous data, which is structured, semi-structured, and unstructured data, a significant barrier for applications is presented [38].By selecting specific models or components based on each class's characteristics, this approach adapts dynamically to boost classification performance on imbalanced data. The study utilized techniques like Local Class Accuracy (LCA), decision trees, and random forests within a dynamic ensemble setup. This method showed improved balanced accuracy, F-score, and G-mean, with the LCA-based ensemble showing the best results for minority class detection. However, these approaches' complexity and computational cost are higher, making real-time applications challenging in some scenarios.Fig.(3) summarizes these solutions.

5. DEEP LEARNING SOLUTIONS FOR MULTICLASS IMBALANCE

Deep learning, the subfield of machine learning, is employed with artificial neural networks (ANNs) with two or more hidden layers. A completely connected feedforward neural network with at least one hidden layer is called a multilayer perceptron (MLP). A particular kind of feedforward neural network, the CNN, is designed to process multi-dimensional data, including images [39]. The MLP and CNN are only two of the alternate DNN designs that have been created over time. Autoencoders and RNNs are described in detail in [40–42]. They also provide advanced optimization tactics such as

increased regularization, parameter initialization, normalization, activation functions, and optimizers that have been shown to reduce training times and boost efficiency. While deep learning models such as CNNs and RNNs have shown remarkable success in various domains, handling multi-class imbalance remains a significant challenge, particularly in water quality classification. Representation learning is mapping raw input data features into a new representation or a new feature space using machine learning to improve detection and classification tasks. Non-linear input data transformations accomplish this process of translating unprocessed input data to new representations. This automated feature development saves a great deal of time by eliminating the need for specialists to manually hand-engineer features. It enhances performance in difficult issues, like picture and speech, which are otherwise challenging to determine. By assembling several hidden layers, DNNs can learn high-level feature representations of inputs when given enough data. These learned representations decrease irrelevant input components while strengthening those crucial for discrimination. Deep learning architectures are made of more sophisticated abstract representations, giving them power.This section discusses developments in deep learning methods that tackle imbalance problems and demonstrates how well they perform in real-time water monitoring systems.

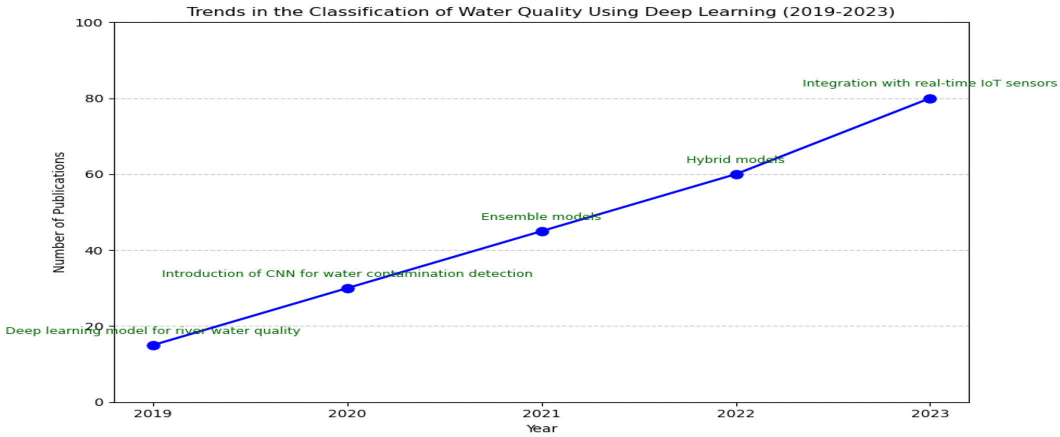

Fig. (4). Advancement in the classification of water data sets

5.1 DNN Advancements and Multi-Class Imbalance (2012-2023)

The evolution of deep learning models has significantly influenced the ability to manage imbalanced datasets. Below is a timeline of key advancements and their implications for handling multi-class imbalance, which is graphically shown in Fig. (4).

2012-2014: Emergence of Deep Learning in Classification.

Alexnet's [43] demonstration of the efficacy of deep learning in image classification sparked interest in DNN research in the period 2012-2014. Although it is not the main focus, the inherent imbalance in datasets such as ImageNet is started to be recognized as a potential issue. The study in [44] examined how the depth of the convolutional network affects its accuracy when used for large-scale image recognition. To meet the new state of the art for detection and classification in the ImageNet Large-Scale Visual Recognition Challenge, the study in [45] proposed a deep convolutional neural network architecture codenamed Inception. This study also implemented GoogLeNet, a 22-layer deep network whose quality is evaluated in detection and classification. The study in [46] investigates a technique the stack model classifier for multiclass classification for water quality prediction, which integrates machine learning with practical and simple water quality measurements. The introduction of GANs made it possible to create artificial data.[47] suggested a novel framework for using an adversarial approach to estimate generative models, enabling synthetic data generation, which became useful in addressing class imbalances by augmenting datasets. Early applications of machine learning for water quality focus primarily on regression-based models predicting specific parameters (e.g., pH, dissolved oxygen), with limited exploration of deep learning.

2013-2015: Deeper Networks and Architectures

Deeper networks like VGGNet [48] and GoogleNet [49] emerge, pushing the boundaries of accuracy in image classification. RNNs gain traction for sequential data such as time-series analysis. Researchers begin exploring techniques like data augmentation and cost-sensitive learning to address imbalances in specific applications. The introduction of RNNs sparks interest in using time-series data for water quality monitoring. However, studies remain sparse, and multi-class imbalance is still underexplored.

2016-2018: Generative Models and Attention Mechanisms

GANs are introduced, enabling the generation of synthetic data. Attention mechanisms [50] improve performance in tasks like machine translation and image captioning. The use of GANs for data augmentation to address class imbalance is explored. The problem of long-tailed distributions, a specific type of multi-class imbalance, gains attention. Early studies began experimenting with deep learning for classifying water quality indices, but imbalanced class distributions (e.g., more 'safe' than 'unsafe' samples) pose challenges. GAN-based data augmentation starts being considered for underrepresented water quality classes.

2019-2021: Transformers and Self-Supervised Learning

Transformer networks [51] revolutionize natural language processing and begin to influence other domains like vision and time-series analysis. Self-supervised learning techniques [52] allow models to learn from unlabelled data, reducing dependence on large labelled datasets.Research on specialized loss functions (e.g., focal loss) and training strategies (e.g., re-weighting, resampling) for imbalanced datasets intensifies. New metrics for evaluating performance under imbalance, such as balanced accuracy and macro-averaged F1 scores, are proposed.Researchers apply Transformer models to time-series water quality data, improving classification accuracy for multi-class water quality indices. Studies highlight that class imbalance—particularly in real-world datasets from sensor networks—adversely affects model performance, leading to hybrid approaches combining data augmentation and algorithmic solutions.

2022-2024: Focus on Robustness and Efficiency

Emphasis shifts towards improving the robustness and efficiency of DNNs through techniques like model pruning, quantization, and federated learning, allowing deployment in resource-constrained environments. The connection between class imbalance and model robustness is further investigated. Methods for detecting and mitigating the effects of imbalance in real-world applications, including imbalanced environmental monitoring data, gain importance.*2022:* The paper *"A Stacked Ensemble Deep Learning Approach for Imbalanced Multi-Class Water Quality Index Prediction"* [Wen Yee Wong et al.] introduces a novel ensemble strategy combining deep neural networks to tackle multi-class imbalance in water quality datasets. The study demonstrates how ensemble methods can significantly improve prediction accuracy across imbalanced classes. Increased interest in integrating self-supervised learning and domain adaptation for water quality classification, enabling models to handle data scarcity and class imbalance. Researchers focus on real-time monitoring systems incorporating class imbalance mitigation techniques, improving water quality predictions in diverse environmental conditions.

5.2 Advantages of Deep Learning in Handling Class Imbalance for Water Quality Data

Deep learning neural networks have several inherent advantages when handling multiclass data imbalance. Still, they may also be improved and modified with certain methods to deal with the problem more successfully. [53] suggests a novel framework for extracting hybrid characteristics that emphasize the combination and ideal choice of high- and low-level attributes. The suggested method achieves scalability and dependability by automatically adjusting the final ideal features based on real-time scenarios while producing an accurate and effective disease classification of medical images. Deep learning models, particularly CNNs and RNNs, excel at extracting intricate features directly from raw data without requiring extensive manual feature engineering. In [54], the most used DL network type CNNs are introduced, and their architecture development and key characteristics are discussed. This is essential for classifying water quality because it enables models to pick up on small differences between classes, like chemical compositions or contamination levels, that conventional models can miss. Layers in the deep neural network catch increasingly complicated information; the deeper layers concentrate on particular, in-depth aspects that distinguish one water quality level from another, while the first layers record fundamental patterns. This adaptability is highly advantageous in multi-class classification tasks with imbalanced data, as the model can focus on learning key traits of minority classes through feature extraction.

5.2.1 Scalability and Flexibility

Water quality datasets, which may include pH levels and other chemical indicators, are examples of huge, complicated datasets with high-dimensional features that deep learning models can manage. In contrast to traditional algorithms, which frequently require pre-processed or reduced input for effective training, architectures such as CNNs and transformers are naturally scalable; they can analyze large data volumes and complicated characteristics without requiring significant restructuring. This scalability is essential for water quality monitoring systems that gather data at high frequencies and from numerous sources to ensure that the model retains accuracy over various expanding datasets.

5.3 Transfer Learning and Pretrained Models

Transfer learning allows training models on larger, balanced datasets and then fine-tuning them on a specific water quality dataset, which is particularly useful for minority class detection. Using pre-trained models like any domain-specific datasets, deep learning architectures can leverage the general feature representations learned from balanced data and apply these to smaller, imbalanced water quality datasets. This technique enhances the model's ability to detect rare pollutants or contaminants, as the pre-trained layers retain generalizable knowledge that only needs fine-tuning, often resulting in improved accuracy for minority classes

5.4 Innovative Imbalance-Specific Architectures

Recent deep learning advancements include architectures specifically tailored for imbalanced data, like attention mechanisms and class-aware learning.[55] developed an Enhanced Long Short-Term Memory (E-LSTM) based on the feature attention mechanism that employs word-feature seizing and sequential modelling to identify and classify the stress polarity. By focusing on the most significant parts of an input, attention mechanisms enable models to enhance attributes from underrepresented classes. Attention methods can draw attention to important pollutant indicators in water quality categorization, even if those indicators only appear in just a fraction of the data. By producing embeddings emphasizing class differences, class-aware representation learning further enhances model performance by guaranteeing that minority classes have a separate representation space inside the model.

Summary of Case Studies

The case studies in Table 1 show the effectiveness of various deep learning strategies for addressing multi-class imbalances in water quality data. Techniques such as SMOTE, cost-sensitive learning, GANs, and ensemble methods have shown promise in improving minority class detection, although each method comes with trade-offs. Combining these methods in real-world water quality monitoring systems is often necessary to achieve robust performance.

6.CONCLUSION AND FUTURE DIRECTIONS

Water quality datasets often suffer from severe class imbalances, where certain pollution levels are significantly underrepresented. This limits the ability of deep learning models to generalize effectively. Many available datasets are small or collected under specific conditions, making it difficult to develop models with broad applicability. Future research should focus on establishing standardized, well-balanced water quality datasets by leveraging data augmentation techniques and synthetic data generation. This paper highlights the need for integrating advanced deep-learning techniques into water quality monitoring systems GANs and hybrid deep learning approaches are powerful tools for handling class imbalance but require significant computational resources. Research should explore lightweight GAN architectures and efficient training strategies to make these methods more accessible for real-world applications.

LIST OF ABBREVIATIONS

IoT = Internet of Things

ML = Machine Learning

DL = Deep Learning

CNN = Convolutional Neural Networks

RNN = Recurrent Neural Networks

LSTM = Long Short-Term Memory

DNN = Deep Neural Networks

SMOTE = Synthetic Minority Over-sampling Technique

SDSMOTE = Spatial Distribution-based SMOTE

CBIS = Cluster-Based Instance Selection

GAN= Generative Adversarial Networks

LCA = Local Class Accuracy

ANN = Artificial Neural Networks

MLP = MultiLayer Perceptron

E-LSTM = Enhanced Long Short-Term Memory

CONSENT FOR PUBLICATION

NOT APPLICABLE

FUNDING

None

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none

Case Study	Authors	Title	Year	Technique	Dataset Characteristics	Results	Limitations
1	Zhang, H., and Wang, M.[56]	"Cost-Sensitive LSTM for Imbalanced Industrial Wastewater Quality Prediction"	2021	Cost-sensitive learning + LSTM	Severe class imbalance; only 2% critical contamination samples	AUC-PR increased by 15%; improved sensitivity to minority class	Struggled with false positives; suggested combining with under sampling techniques.
2	Nicholaus et al.,[57]	"Anomaly Detection in Coastal Water Quality Using Autoencoders"	2019	Autoencoders	Highly imbalanced data; rare pollution events	Recall for rare events increased by 30%; effectively identified rare pollution events	Flagged natural fluctuations as anomalies, leading to false positives; proposed integrating domain knowledge.
3	Karami Lawal et al.,[58]	Optimized Ensemble Methods for Classifying Imbalanced Water Quality Index Data	2020	Ensemble learning	Resampled subsets of imbalanced data	Overall accuracy: 92%; improved recall for minority class	None specified; demonstrated improved performance through aggregation of predictions.
4	Wong, W.Y., et al.[59]	"A Stacked Ensemble Deep Learning Approach for Imbalanced Multi-Class Water Quality	2023	Stacked ensemble learning	Imbalanced datasets with rare events	Superior balanced performance across accuracy, precision,	None specified; applied balanced bagging and RUSBoost techniques for improved classification.

		Index Prediction"				recall, and F1 score	
5	Karami Lawal et al.,[60]	"A Deep Learning Strategy for Water Quality Monitoring"	2023	Deep learning + SMOTE	Key water quality indicators data	Improved classification accuracy for minority classes	None specified; suitable for real-time assessments.
6	K, K., Krishnan et al.,[61]	"Water Quality Prediction: A Data-Driven Approach with Data Augmentation "	2023	Data augment ation	Imbalanced water quality datasets	Significant improvement in prediction accuracy for minority classes	None specified; evaluated multiple models.

Table 1. Case Studies

REFERENCES

[1] Mo, S., Wang, Y., Zhi, J., He, J., & Guo, L. (2022). Research on water quality classification of unbalanced river image data under deep learning framework. *2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 272–275. https://doi.org/10.1109/ICFTIC57696.2022.10075153

[2] Junyi Chai, Hao Zeng, Anming Li, Eric W.T. Ngai,Deep learning in computer vision: A critical review of emerging techniques and application scenarios,Machine Learning with Applications,Volume 6,2021,100134,ISSN 2666-8270, https://doi.org/10.1016/j.mlwa.2021.100134.

[3] Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, *2*(6), 420. https://doi.org/10.1007/s42979-021-00815-1

[4] Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., & Poria, S. (2023). A review of deep learning techniques for speech processing. *Information Fusion*, *99*, 101869. https://doi.org/10.1016/j.inffus.2023.101869

[5] Ghosh, K., Bellinger, C., Corizzo, R., Branco, P., Krawczyk, B., & Japkowicz, N. (2024). The class imbalance problem in deep learning. *Machine Learning*, *113*(7), 4845–4901. https://doi.org/10.1007/s10994-022-06268-8

[6] Lango, M., & Stefanowski, J. (2022). What makes multi-class imbalanced problems difficult? An experimental study. *Expert Systems with Applications*, *199*, 116962. https://doi.org/10.1016/j.eswa.2022.116962

[7] Hamad, R. A., Kimura, M., & Lundström, J. (2020). Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments. *SN Computer Science*, *1*(4), 204. https://doi.org/10.1007/s42979-020-00211-1Hellín, C. J., Olmedo, A. A., Valledor, A., Gómez, J., López-Benítez, M., & Tayebi, A. (2024). Unraveling the Impact of Class Imbalance on Deep-Learning Models for Medical Image Classification. *Applied Sciences*, *14*(8), 3419. https://doi.org/10.3390/app14083419

[8] Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, *6*(1), 27. https://doi.org/10.1186/s40537-019-0192-5

[9] Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. P. (2024). A survey on imbalanced learning: latest research, applications and future directions. *Artificial Intelligence Review*, *57*(6), 137. https://doi.org/10.1007/s10462-024-10759-6

[10] Banerjee, P., Dehnbostel, F. O., & Preissner, R. (2018). Prediction is a balancing act: Importance of sampling methods to balance sensitivity and specificity of predictive models based on imbalanced chemical data sets. *Frontiers in Chemistry*, *6*(AUG). https://doi.org/10.3389/fchem.2018.00362

[11] Tyagi, S., & Mittal, S. (2020). *Sampling Approaches for Imbalanced Data Classification Problem in Machine Learning* (pp. 209–221). https://doi.org/10.1007/978-3-030-29407-6_17

[12] Chawla, N. v, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).

[13] Haibo He, Yang Bai, Garcia, E. A., & Shutao Li. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322–1328. https://doi.org/10.1109/IJCNN.2008.4633969

[14] Bunkhumpornpat, C., Sinapiromsaran, K., & Lursinsap, C. (2009). *Safe-Level-SMOTE: Safe-Level-Synthetic Minority Over-Sampling TEchnique for Handling the Class Imbalanced Problem* (pp. 475–482). https://doi.org/10.1007/978-3-642-01307-2_43

[15] Han, H., Wang, W.-Y., & Mao, B.-H. (2005). Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. In *LNCS* (Vol. 3644).

[16] Yan, Y., Zhu, Y., Liu, R., Zhang, Y., Zhang, Y., & Zhang, L. (2022). Spatial Distribution-based Imbalanced Undersampling. *IEEE Transactions on Knowledge and Data Engineering*, 1–1. https://doi.org/10.1109/TKDE.2022.3161537

[17] Swana, E. F., Doorsamy, W., & Bokoro, P. (2022). Tomek Link and SMOTE Approaches for Machine Fault Classification with an Imbalanced Dataset. *Sensors*, *22*(9), 3246. https://doi.org/10.3390/s22093246

[18] Douzas, G., Bacao, F., & Last, F. (2018). Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Information Sciences*, *465*, 1–20. https://doi.org/10.1016/j.ins.2018.06.056

[19] Devi, Debashree & Biswas, Saroj & Purkayastha, Biswajit. (2017). Redundancy-driven modified Tomek-link based under sampling: A solution to class imbalance. Pattern Recognition Letters. 93. 3-12.

[20] Tsai, C.-F., Lin, W.-C., Hu, Y.-H., & Yao, G.-T. (2019). Under-sampling class imbalanced datasets by combining clustering analysis and instance selection. *Information Sciences*, *477*, 47–54. https://doi.org/10.1016/j.ins.2018.10.029

[21] Cordón, I., García, S., Fernández, A., & Herrera, F. (2018). Imbalance: Oversampling algorithms for imbalanced classification in R. *Knowledge-Based Systems*, *161*, 329–341. https://doi.org/10.1016/j.knosys.2018.07.035

[22] Karim, M. R., Beyan, O., Zappa, A., Costa, I. G., Rebholz-Schuhmann, D., Cochez, M., & Decker, S. (2021). Deep learning-based clustering approaches for bioinformatics. *Briefings in Bioinformatics*, *22*(1), 393–415. https://doi.org/10.1093/bib/bbz170

[23] Khan, A. A., Chaudhari, O., & Chandra, R. (2024). A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation, and evaluation. *Expert Systems with Applications*, *244*, 122778. https://doi.org/10.1016/j.eswa.2023.122778

[24] Akbiyik, M. E. (2023). *Data Augmentation in Training CNNs: Injecting Noise to Images*. http://arxiv.org/abs/2307.06855

[25] Muhammad Saqib. (2025). Optimizing Spot Instance Reliability and Security Using Cloud-Native Data and Tools. In Journal of Information Systems Engineering and Management (Vol. 10, Issue 14s, pp. 720–731). Science Research Society. https://doi.org/10.52783/jisem.v10i14s.2387

[26] Mostofi, F., Behzat Tokdemir, O., & Toğan, V. (2024). Generating synthetic data with variational autoencoder to address the class imbalance of graph attention network prediction model for construction management. *Advanced Engineering Informatics*, *62*, 102606. https://doi.org/10.1016/j.aei.2024.102606

[27] Serrano, W. (2023). The Deep Learning Generative Adversarial Random Neural Network in data marketplaces: The digital creative. *Neural Networks*, *165*, 420–434. https://doi.org/10.1016/j.neunet.2023.05.028

[28] Aurelio, Y. S., de Almeida, G. M., de Castro, C. L., & Braga, A. P. (2022). Cost-Sensitive Learning based on Performance Metric for Imbalanced Data. *Neural Processing Letters*, *54*(4), 3097–3114. https://doi.org/10.1007/s11063-022-10756-2

[29] Mienye, I. D., & Sun, Y. (2021). Performance analysis of cost-sensitive learning methods with application to imbalanced medical data. *Informatics in Medicine Unlocked*, *25*, 100690. https://doi.org/10.1016/j.imu.2021.100690

[30] López, V., Fernández, A., Moreno-Torres, J. G., & Herrera, F. (2012). Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification. Open problems on intrinsic data characteristics. *Expert Systems with Applications*, *39*(7), 6585–6608. https://doi.org/10.1016/j.eswa.2011.12.043

[31] Tang, J., Li, Y., Hou, Z., Fu, S., & Tian, Y. (2024). Robust two-stage instance-level cost-sensitive learning method for class imbalance problem. *Knowledge-Based Systems*, *300*, 112143. https://doi.org/10.1016/j.knosys.2024.112143

[32] Cao, B., Liu, Y., Hou, C., Fan, J., Zheng, B., & Yin, J. (2021). Expediting the Accuracy-Improving Process of SVMs for Class Imbalance Learning. *IEEE Transactions on Knowledge and Data Engineering*, *33*(11), 3550–3567. https://doi.org/10.1109/TKDE.2020.2974949

[33] Lu, Y., Cheung, Y.-M., & Yan Tang, Y. (2020). Adaptive Chunk-Based Dynamic Weighted Majority for Imbalanced Data Streams with Concept Drift. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(8), 2764–2778. https://doi.org/10.1109/TNNLS.2019.2951814

[34] Fernando, K. R. M., & Tsokos, C. P. (2022). Dynamically Weighted Balanced Loss: Class Imbalanced Learning and Confidence Calibration of Deep Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(7), 2940–2951. https://doi.org/10.1109/TNNLS.2020.3047335

[35] Mulyanto, M., Faisal, M., Prakosa, S. W., & Leu, J.-S. (2020). Effectiveness of Focal Loss for Minority Classification in Network Intrusion Detection Systems. *Symmetry*, *13*(1), 4. https://doi.org/10.3390/sym13010004

[36] K, K., Krishnan, S., & Manikandan, R. (2024). Water quality prediction: a data-driven approach exploiting advanced machine learning algorithms with data augmentation. *Journal of Water and Climate Change*, *15*(2), 431–452. https://doi.org/10.2166/wcc.2023.403

[37] Dogo, E. M., Nwulu, N. I., Twala, B., & Aigbavboa, C. (2021). Accessing Imbalance Learning Using Dynamic Selection Approach in Water Quality Anomaly Detection. *Symmetry*, *13*(5), 818. https://doi.org/10.3390/sym13050818

[38] Manjusha Nambiar, P. v., & Anupriya, E. (2023). Heterogeneous BigData Analysis in IoT Cloud Environment. In *Lecture Notes in Networks and Systems* (Vol. 491, pp. 387–397). https://doi.org/10.1007/978-981-19-4193-1_37

[39] Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*, *11*(3), 52. https://doi.org/10.3390/computation11030052

[40] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

[41] Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, *2*(6), 420. https://doi.org/10.1007/s42979-021-00815-1

[42] Chen, S., & Guo, W. (2023). Auto-Encoders in Deep Learning—A Review with New Perspectives. *Mathematics*, *11*(8), 1777. https://doi.org/10.3390/math11081777

[43] Loukil, Z., Mirza, Q. K. A., Sayers, W., & Awan, I. (2024). A Deep Learning based Scalable and Adaptive Feature Extraction Framework for Medical Images. *Information Systems Frontiers*, *26*(4), 1279–1305. https://doi.org/10.1007/s10796-023-10391-9

[44] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

[45] Dogo, E. M., Nwulu, N. I., Twala, B., & Aigbavboa, C. (2021). Accessing Imbalance Learning Using Dynamic Selection Approach in Water Quality Anomaly Detection. *Symmetry*, *13*(5), 818. https://doi.org/10.3390/sym13050818

[46] Nambiar PV, M., & Urkude, G. (2024). An Approach to Forecast Quality of Water Effectively Using Machine Learning Algorithms. *Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering)*, *17*. https://doi.org/10.2174/0123520965267326231115071849

[47] Zulqarnain, M., Shah, H., Ghazali, R., Alqahtani, O., Sheikh, R., & Asadullah, M. (2023). Attention Aware Deep Learning Approaches for an Efficient Stress Classification Model. *Brain Sciences*, *13*(7), 994. https://doi.org/10.3390/brainsci13070994

[48] Hussain, M., Thaher, T., Almourad, M. B., & Mafarja, M. (2024). Optimizing VGG16 deep learning model with enhanced hunger games search for logo classification. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-82022-5

[49] Sa'idah, S., Fany, A., & Suparta, I. P. Y. N. (2022). Convolutional Neural Network GoogleNet Architecture for Detecting the Defect Tire. *2022 International Conference on Computer Science and Software Engineering (CSASE)*, 331–336. https://doi.org/10.1109/CSASE51777.2022.9759694

[50] Al-Malla, M. A., Jafar, A., & Ghneim, N. (2022). Image captioning model using attention and object features to mimic human image understanding. *Journal of Big Data*, *9*(1). https://doi.org/10.1186/s40537-022-00571-w

[51] Thirunavukarasu, R., & Kotei, E. (2024). A comprehensive review on transformer network for natural and medical image analysis. *Computer Science Review*, *53*, 100648. https://doi.org/10.1016/j.cosrev.2024.100648

[52] Azlaan, H. (2024). *Self-Supervised Learning:Techniques for leveraging vast amounts of unlabeled data to improve model performance*. https://www.researchgate.net/publication/383463877

[53] Jerome, M. S., & Mohamed, A. (2024). A Novel Framework for Extracting and Recognizing Text in Scene Images through Self-Attention CNN Enhanced with Fuzzy DNN. In *Journal of Information Systems Engineering and Management* (Vol. 2025, Issue 5s). https://www.jisem-journal.com/

[54] Wai, K. P., Chia, M. Y., Koo, C. H., Huang, Y. F., & Chong, W. C. (2022). Applications of deep learning in water quality management: A state-of-the-art review. *Journal of Hydrology*, *613*, 128332. https://doi.org/10.1016/j.jhydrol.2022.128332

[55] Xin, Z., Liu, X., Zhang, H., Wang, Q., An, Z., & Liu, H. (2024). An enhanced feature extraction based long short-term memory neural network for wind power forecasting via considering the missing data reconstruction. *Energy Reports*, *11*, 97–114. https://doi.org/10.1016/j.egyr.2023.11.040

[56] Zhang, X., Peng, H., Zhang, J., & Wang, Y. (2023). A cost-sensitive attention temporal convolutional network based on adaptive top-k differential evolution for imbalanced time-series classification. *Expert Systems with Applications*, *213*, 119073. https://doi.org/10.1016/j.eswa.2022.119073

[57] Nicholaus, I. T., Park, J. R., Jung, K., Lee, J. S., & Kang, D.-K. (2021). Anomaly Detection of Water Level Using Deep Autoencoder. *Sensors*, *21*(19), 6679. https://doi.org/10.3390/s21196679

[58] Karami Lawal, Z., Aldrees, A., Yassin, H., Dan'azumi, S., Raghavendra Naganna, S., Abba, S. I., & Sammen, S. Sh. (2024). Optimized Ensemble Methods for Classifying Imbalanced Water Quality Index Data. *IEEE Access*, *12*, 178536–178551. https://doi.org/10.1109/ACCESS.2024.3502361

[59] Yee Wong, W., Hasikin, K., Salwa Mohd Khairuddin, A., Abdul Razak, S., Farzana Hizaddin, H., Istajib Mokhtar, M., & Mokhzaini Azizan, M. (2023). A Stacked Ensemble Deep Learning Approach for Imbalanced Multi-Class Water Quality Index Prediction. *Computers, Materials & Continua*, *0*(0), 1–24. https://doi.org/10.32604/cmc.2023.038045

[60] Karami Lawal, Z., Aldrees, A., Yassin, H., Dan'azumi, S., Raghavendra Naganna, S., Abba, S. I., & Sammen, S. Sh. (2024). Optimized Ensemble Methods for Classifying Imbalanced Water Quality Index Data. *IEEE Access*, *12*, 178536–178551. https://doi.org/10.1109/ACCESS.2024.3502361

[61] K, K., Krishnan, S., & Manikandan, R. (2024). Water quality prediction: a data-driven approach exploiting advanced machine learning algorithms with data augmentation. *Journal of Water and Climate Change*, *15*(2), 431–452. https://doi.org/10.2166/wcc.2023.403