CNN Based Diatom Identification: A Systematic Review of Methods, Performance, and Educational Implications

Muhammad Iqbal Najib Fahmi [1], Susriyati Mahanal [1], Siti Zubaidah [1], Ibrohim Ibrohim [1]

[1] *Department of Biology, Universitas Negeri Malang, Malang, Indonesia*

***Abstract* – Automated diatom identification using convolutional neural networks has attracted increasing attention for applications in environmental monitoring and microalgae learning, yet evidence remains fragmented across methodological and practical dimensions. The study conducted a systematic literature review of 27 articles related to diatom identification using convolutional neural networks. The review analyses dominant task formulations, model architectures, dataset characteristics, reported performance, technical challenges, and future research directions. The findings show that research is largely dominated by classification and detection tasks, relying on established architectures and microscopy derived datasets, while segmentation and regression-based approaches are comparatively limited. Although high performance is frequently reported, substantial heterogeneity in evaluation metrics, validation protocols, and dataset regimes constrains cross study comparability and limits generalizability. Persistent challenges include visual complexity, morphological similarity among taxa, class imbalance, limited interpretability, and reduced robustness beyond controlled settings. These issues are particularly critical for microalgae learning, where reliable and interpretable outputs are essential.**

Overall, this review highlights the need for data centred development, robustness-oriented evaluation, and educational alignment to advance scalable and pedagogically meaningful convolutional neural networks-based diatom identification systems.

***Keywords* – Convolutional neural networks, diatom identification, microalgae learning, systematic review.**

DOI: 10.18421/TEM153-99
https://doi.org/10.18421/TEM153-99

Corresponding author: Susriyati Mahanal,
Department of Biology, Universitas Negeri Malang, Malang, Indonesia
Email: susriyati.mahanal.fmipa@um.ac.id

Received: 27 December 2025.
Revised: 07 April 2026.
Accepted: 27 April 2026.
Published: 27 August 2026.

1. Introduction

Species identification is a fundamental process in biology that aims to determine the identity of organisms based on morphological characteristics and taxonomically relevant distinguishing features [1]. This ability plays an important role not only in biodiversity research, but also in biology education, environmental monitoring, and ecology based decision making [2], [3]. In the context of education and biomonitoring, species identification skills are a prerequisite for understanding the ecological functions of organisms, community dynamics, and ecosystem responses to environmental pressures [4], [5]. However, various studies show that the process of species identification, especially in microscopic organisms, still faces serious limitations due to morphological complexity, high taxonomic expertise requirements, and limited time and resources [6], [7].

In higher education contexts, diatom-based identification tasks are not only used for taxonomic training but also function as entry points for developing observational reasoning, image interpretation skills, and understanding of bioindicator-based ecological assessment [8]. Consequently, the reliability, transparency, and robustness of identification tools directly influence how students engage with microscopic data and construct biological meaning from visual evidence. Diatoms are especially well suited for this purpose due to their distinctive siliceous frustules, high morphological diversity, and intricate valve structures, which provide rich visual features for understanding cell architecture, morphological variation, and the relationship between form and biological function [9], [10].

Through direct engagement with microscopic observations, image analysis, and morphology based identification tasks, students interact with authentic biological data in a manner that promotes analytical and evidence based learning [11], [12].

Despite these advantages, diatom identification remains a cognitively demanding task that is traditionally performed through manual microscopic observation and consultation of taxonomic keys [13], [14]. This process requires extensive training, prolonged experience, and considerable time investment, making it challenging for students and novice learners to achieve consistent and accurate identification outcomes [15]. The high degree of interspecific morphological similarity, intraspecific variability, and dependence on image quality further complicate the learning process, often leading to misclassification and reduced learning efficiency in microscopic microalgae education [16]. These limitations highlight the need for computational approaches that can support image based learning by providing more objective, scalable, and reproducible identification assistance.

Recent advances in Artificial Intelligence, particularly in deep learning, have created new opportunities for automating image based species identification [17]. Convolutional Neural Networks (CNNs) have emerged as the dominant approach for biological image analysis due to their ability to automatically learn hierarchical visual features directly from raw image data, without the need for handcrafted feature extraction [18]. A wide range of CNN architectures, such as LeNet, AlexNet, VGGNet, ResNet, and EfficientNet, have been successfully applied to classification tasks involving microscopic organisms, demonstrating strong performance in recognizing complex morphological patterns [19]. In the context of diatom identification, multiple studies have reported high classification accuracy, indicating the potential of CNN based systems to support both research oriented analysis and image driven learning environments [20].

However, existing studies reveal substantial heterogeneity in the design and implementation of CNN based diatom identification systems [21]. Variations in network architectures, dataset size and composition, image acquisition protocols, and preprocessing techniques, such as segmentation, normalization, and data augmentation lead to considerable differences in reported performance outcomes [22], [23]. Moreover, several open technical challenges remain unresolved, including limited availability of well annotated datasets, class imbalance across species, reduced generalizability across microscopes and imaging conditions, and limited model interpretability [24], [25].

These challenges constrain the broader adoption of automatic identification systems, particularly in educational settings where robustness, transparency, and usability are essential.

Although the number of studies on CNN based automatic diatom identification has increased markedly in recent years, there is still a lack of comprehensive systematic reviews that synthesize architectural choices, dataset characteristics, preprocessing strategies, performance metrics, and open challenges in a unified framework [26], [27]. Existing reviews tend to focus on microscopic organisms in general or on technical aspects of image classification without in depth analysis of diatoms as a distinct and morphologically complex group [28], [29]. Therefore, this study conducts a systematic literature review to critically analyze and synthesize recent research on CNN based automatic diatom identification. By consolidating current evidence and identifying research gaps, this review aims to provide a structured reference for advancing the development of robust, interpretable, and scalable automatic identification systems that can support microscopic microalgae learning and related image based applications. The review is guided by the following research questions:

RQ1: *What types of tasks, architectures, and dataset characteristics are most commonly used in studies of automated diatom identification using CNN?*

RQ2: *How does the CNN model perform in diatom identification, and what are the most commonly applied pre-processing and data augmentation techniques?*

RQ3: *What technical challenges need to be considered in the application of CNN for diatom-based microalgae learning?*

RQ4: *What future research directions are proposed to advance CNN based automated diatom identification?*

2. Theoretical Background

This section links the theoretical foundations of species identification with recent technological advances in microalgae learning. It outlines the shift from manual identification toward artificial intelligence–based approaches and introduces key aspects of convolutional neural network implementation that guide the subsequent discussion.

2.1. Technological Evolution and AI in Mikroalgae Learning

Microscopic microalgae identification, particularly for diatoms, has traditionally relied on manual microscopy and morphology based taxonomic keys.

Although authoritative, this approach is limited by subjectivity, high expertise demands, and poor scalability, especially in educational and high throughput settings [16]. Early advances in digital microscopy enabled semi automated workflows using handcrafted morphometric features, but these methods showed limited robustness to morphological variability [11], [30]. The adoption of deep learning, especially CNNs, enabled fully automated image based identification by learning hierarchical visual features directly from raw microscopy data, achieving strong performance in recognising complex diatom morphologies that are difficult to encode using rule based approaches [17], [18], [20]. However, the rapid proliferation of CNN based approaches has produced a heterogeneous and fragmented research landscape. Existing studies differ substantially in task formulation, dataset composition, preprocessing strategies, and evaluation protocols, which limits cross study comparability and obscures generalisability [21], [29]. In microalgae learning contexts, this fragmentation hampers the translation of high laboratory performance into robust learning support systems that operate under realistic imaging conditions. These issues underscore the need for a systematic synthesis that examines CNN based diatom identification beyond isolated performance metrics, focusing instead on design choices, evaluation practices, and technical constraints.

2.2. Key Aspects of CNN Implementation in Microalgae Learning

This review analyses CNN based automated diatom identification through key implementation aspects that structure system design and reported outcomes across studies, namely task formulation, network architecture, dataset characteristics, performance evaluation, preprocessing and augmentation, and technical challenges [29]. Task formulation is tightly coupled with architectural choice: Classification tasks predominantly use established CNN backbones (e.g., AlexNet, VGG, Inception, ResNet) via transfer learning, whereas detection and segmentation tasks rely on mainstream object detection and instance-segmentation frameworks such as Faster R-CNN, Mask R-CNN, YOLO variants, and U-Net-based models, which are better suited to cluttered microscopy scenes [31], [32]. Dataset characteristics critically determine performance and generalisability. Most studies use microscopy derived datasets with substantial variation in taxonomic coverage, image quality, acquisition protocols, and scale [33].

The dominance of private or partially shared datasets, combined with class imbalance and limited representation of rare taxa, constrains reproducibility and limits generalisation, particularly in microalgae learning contexts where models may be applied to student generated or field collected images [34], [35].

Model performance is reported using task dependent metrics, including accuracy and F1-score for classification, mAP for detection, Dice or IoU for segmentation, and domain specific metrics for regression or mapping tasks [36], [37]. While preprocessing and augmentation practices show convergence most commonly patch extraction, background suppression, normalisation, and geometric transformations (e.g., rotation and flipping), differences in validation protocols and incomplete metric reporting limit cross study comparability [17], [38]. Recurring technical challenges, including visual complexity, morphological ambiguity, data imbalance, computational constraints, and limited interpretability, directly restrict applicability beyond controlled experimental settings [39], [40]. Framing the literature around these aspects provides a concise basis for identifying current limitations and future research priorities in CNN based automated diatom identification for microalgae learning.

3. Methodology

This study aims to answer research questions related to automatic diatom identification based on CNN in the context of microalgae learning. This study uses the SLR method, which refers to the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) guidelines. The SLR method was chosen because it provides a systematic and transparent framework for identifying, selecting, and synthesizing previous research findings, thereby providing a comprehensive and evidence based understanding of the topic under review [41]. The SLR process began with determining the search strategy, formulating keywords, selecting electronic databases, and establishing inclusion and exclusion criteria for articles relevant to the research focus.

3.1. Search Approaches

The search strategy began with defining the main keywords that represent the research focus, namely “artificial neural network” OR “deep neural networks” “convolutional neural network“ OR ‘cnn’ OR ”deep learning“ OR ”artificial intelligence" AND “diatom” OR “bacillariophyta” OR “algae” OR “plankton” OR “phytoplankton” AND “identification” OR ‘classification’ OR “image classification” OR “image analysis” AND "microalgae" OR “phycology” OR “algal biology”.

The search is limited to articles published in the last ten years, between 2015 and 2025. In addition, the search was conducted between May and December 2025. The search was conducted in major scientific databases, including Scopus, IEEE explore, Springer, and American Chemical Society (ACS).

3.2. Selection of Eligible Papers

The results obtained were analyzed to eliminate duplicate papers and papers not written in English. Then, several inclusion criteria were established to further review the papers.

The inclusion criteria included: (1) Papers discussing the identification of diatom species using Convolutional Neural Network (CNN); (2) Empirical research based articles, not reviews, books, book chapters, or proceedings; (3) Articles that present sufficient information about the data used (type and source of the dataset), data processing stages (data preprocessing), and the architecture and configuration of the model applied; and (4) Articles that report model evaluation results using standard performance metrics, such as accuracy, precision, recall, or F1-score. The inclusion/exclusion process and the number of articles that met the criteria at each level are illustrated in a flowchart [42] (Figure 1).

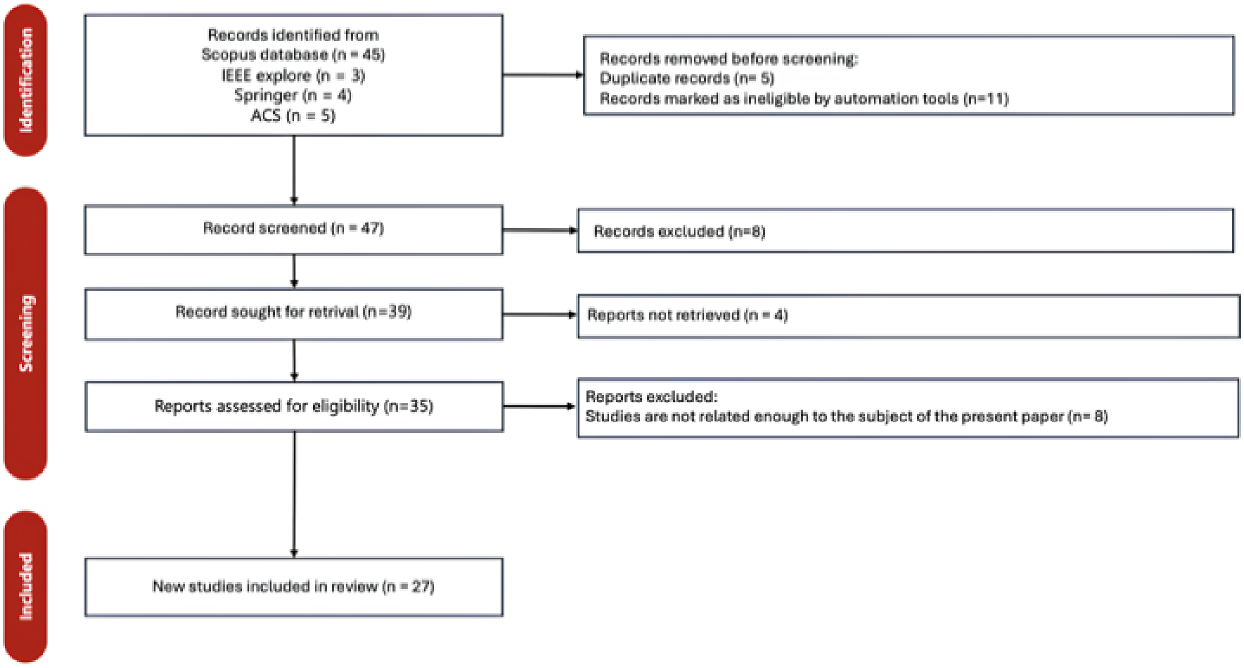

Figure 1. PRISMA flow diagram illustrating the study selection process

3.3. Data Extraction Strategy

Data from 27 articles were tabulated using Microsoft Excel in the form of extraction tables to obtain important information. The extraction table contained the architecture, dataset form, model performance, preprocessing techniques, and technical challenges in development in the form of learning media. The data was then organized and evaluated, then synthesized using a textual narrative approach. The textual narrative method was chosen because it could combine and summarize findings from various studies using text-based descriptions [43]. The team of authors and verifiers conducted focus group discussions (FGDs) to reduce potential bias in identifying key information and during the narrative synthesis process.

4. Results

This study provides a structured synthesis of current evidence on CNN-based diatom identification. The results are organised into four main parts: (1) Task types, model architectures, and dataset characteristics, (2) Model performance and commonly used preprocessing and augmentation techniques, (3) Technical challenges in applying CNN-based diatom identification for microalgae learning, and (4) Future research directions.

4.1. RQ1. Task Types, Architectures, and Dataset Characteristics in Automated Diatom Identification

This section summarises task types, model architectures, and dataset characteristics in CNN-based automated diatom identification.

It explains how these components are used in different studies and highlights the main patterns in current research. The synthesis of the reviewed studies (Table 1) shows that automated diatom identification research is predominantly structured around classification and detection-based tasks, whereas segmentation and regression or mapping tasks are less frequently addressed. Classification is the most common task, focusing on species or genus level identification from microscopy images or image patches [20], [27], [32], [41].

Detection combined with classification represents the second most prevalent task type, particularly in microscopy and SEM datasets where diatom objects must be localised within complex scenes prior to identification [38], [45], [46], [47]. Segmentation focused studies are comparatively limited and are primarily applied in high resolution imaging contexts that require precise object delineation [36], [48]. A smaller subset of studies applies CNNs to regression or mapping tasks, typically using spectral or multivariate data for environmental inference rather than direct species identification [37], [49], [50].

Table 1. Task type, architectures and dataset characteristics in automated diatom identification

Ref	Task Type	Architecture Used	Dataset
[36]	Segmentation	Mask R-CNN; U-Net	110+ taxa; ~3,000 specimens; grayscale virtual slides; private
[51]	Classification and Mapping	ParticleTrieur (CNN)	13 taxonomic groups; 16-year Zooscan time series; mixed
[52]	Classification	VGG16; InceptionV3; ResNet50V2; Xception	19 classes; >3,800 images; holographic microscopy; mixed
[46]	Detection and Classification	YOLOv8 (Nano)	2 species; 306 images; light microscopy; private
[32]	Classification	AlexNet	80 species; 11,000 images; brightfield microscopy; private
[53]	Detection and Classification	Modified ShuffleNet V2	10 classes; 10,507 segments; holography; private
[20]	Classification	VGG16; VGG19; Xception; DenseNet	10 classes; 3,319 specimens; virtual slides; private
[44]	Classification	Inception V3	8 species; 20,000 images; digested smears; private
[31]	Detection and Classification	R-FCN (ResNet-50)	8 genera; 2,606 images; SEM; private
[45]	Detection and Classification	Faster R-CNN (VGG16)	8 genera; ~2,600 images; SEM; private
[54]	Detection and Classification	SC-DiatomNet (YOLOv3 + CBAM, SPP)	6 species; 1,043 images; light microscopy; global
[55]	Detection	Faster R-CNN (ResNet50)	12 categories; 356 images; light microscopy; mixed
[49]	Regression and Mapping	1D-CNN with attention module.	5 algal pigments; 126 samples; drone hyperspectral (400–1000 nm); private
[27]	Classification	DiatomNet (Inception-based)	68 species; 3,027 images; microscopy; global
[38]	Detection and Classification	YOLOv4; AlexNet	80 species; ~8,000 images; microscopy; global
[34]	Classification	ResNet18; AlexNet; VGG11	14 species; 976 images; microscopy; global
[48]	Segmentation	Mask R-CNN; SegNet	10 taxa; 1,446 diatoms; microscopy; mixed
[39]	Classification	VGG-like CNN + metadata	27 categories; 350k images; Zooglider; mixed
[50]	Classification and Mapping	VGG16-based CNN	7,663 patches; Sentinel-2 MSI; mixed
[17]	Classification	GoogLeNet Inception V3	53 WSIs; lung tissues from 20 cases; 255×255 tiles; private
[56]	Detection	HDC-SSD	8 species; 2,600 images; SEM; private
[37]	Regression/Mapping	1D-CNN + Autoencoder	3 aquatic indices (TDI, BMI, FAI); 478 samples; 81 variables; global
[57]	Classification	DiatomNet v1.0 (Inception V3)	18 cases; 306 smears; microscopy; mixed
[58]	Classification	MobileNetV2 + GoogLeNet + SVM	6 species; 7,820 images; microscopy; private
[24]	Detection	YOLOv5	166 species; synthetic + atlas images; mixed
[47]	Detection and Classification	Faster R-CNN (Inception-ResNet V2)	52 species; forensic samples; microscopy; private
[40]	Detection	Fully Convolutional Neural Network	600 test images; 20 training images (4,046 patches); dual-channel microscopy; private

Across these task types, the most commonly used architectures are well established CNN backbones that are adapted rather than newly designed for diatom specific problems. For classification, studies predominantly employ AlexNet, VGG, Inception, and ResNet families, often through transfer learning [17], [20], [32], [52]. For detection and segmentation, the field consistently relies on mainstream object detection frameworks, including Faster R-CNN, Mask R-CNN, YOLO variants, and SSD or F-CNN based models [31], [38], [40], [54], [56]. Lightweight or hybrid architectures appear mainly in studies oriented toward deployment, such as ShuffleNet V2 for holography-based monitoring [53] , and MobileNetV2 based pipelines for efficient classification [58]. Overall, the evidence indicates that architectural choices are primarily driven by task requirements and computational constraints rather than by the development of novel CNN designs.

Dataset characteristics further reveal strong common patterns across studies. Most datasets are derived from microscopy modalities, particularly light microscopy and SEM, which remain the primary sources for species level identification and detection [32], [45], [47]. Holographic microscopy supports automated monitoring pipelines with large numbers of image segments [52], [53], while remote sensing or hyperspectral data are used mainly in regression or mapping tasks rather than in microscopy-based identification [49], [50]. Dataset size varies widely, ranging from hundreds of images in narrowly scoped detection or segmentation studies [46], [55] to thousands or tens of thousands of images in classification studies [32], [44], and up to hundreds of thousands of images in platform-based monitoring systems [39]. Notably, private or partially shared datasets dominate, indicating that current evidence remains largely dataset specific and that widely adopted open benchmarks for automated diatom identification are still limited.

4.2. RQ 2. CNN Model Performance in Diatom Identification, and the Most Commonly Applied Pre-Processing and Data Augmentation Techniques

This section examines the performance of CNN-based models for automatic diatom identification, along with the most commonly used data preprocessing and augmentation techniques. It highlights how performance is reported across various tasks and identifies methodological practices frequently employed to enhance the robustness and reliability of CNN. Table 2 shows that automated diatom identification studies generally report high model performance, but the reported levels are strongly task dependent and vary in how comprehensively they are documented.

For classification, most microscopy based studies report high accuracy and F1 scores, commonly above 0.95 [20], [27], [58], with several reaching or approaching 0.99 accuracy [32], [34]. More moderate yet still competitive performance is reported in non standard imaging settings, such as holography (accuracy = 0.90, [52]) and platform based monitoring (accuracy 0.923, [39]). For detection, performance is primarily reported using mAP, typically in the range of 0.89–0.95, across SEM and light microscopy studies [45], [47], [56]. Segmentation specific metrics are reported less consistently; when explicitly provided, results include Dice 0.775 for virtual slide segmentation [36]. In regression and mapping tasks, performance is reported using domain specific metrics rather than classification scores, such as $R^2 = 0.71$–0.87 for pigment estimation [49] and NSE 0.59–0.67 with corresponding RMSE values for aquatic health indices [37]. Overall, these results indicate that CNN based approaches achieve strong reported performance within their respective data regimes, however, differences in task type and metric selection limit direct cross study comparison.

Table 2. Summary performance preprocessing, and augmentation techniques

Ref	Performance (Reported Metrics)	Preprocessing Techniques	Augmentation Techniques
[36]	Dice 0.775; Precision 0.688–0.724; Recall 0.837–0.898	Tiling (512×512); grayscale conversion; padding	Object shifting; flipping; mirroring
[51]	NR (focus on MDS breakpoints)	Size fractioning; KDE-based size spectra	Not specified
[52]	Accuracy 90.1%; Precision 89.8%; Recall 90.7%; F1 >0.89	Reconstruction; Otsu thresholding; resizing	Rotation; translation
[46]	Detection mAP 0.895; Segmentation mAP 0.898	Manual annotation; resizing	Mosaic; color jitter; flipping
[32]	Accuracy 99.51%	Otsu segmentation; histogram normalization	Rotation; flipping
[53]	Accuracy 93.8–98%; F1 0.959	Background subtraction; filtering	Flipping; rotation; cropping
[20]	F1 = 0.97	Slide stitching; masking	Rotation; shift; zoom
[44]	Accuracy 97.67%; Inference accuracy 92.31%	Tiling; background removal	Rotation; contrast adjustment
[31]	mAP 0.897	Feature map optimization	OHEM; anchor tuning
[45]	Accuracy 0.864; mAP 0.878	ROI pooling; coarse-to-fine strategy	Training strategy optimization
[54]	Precision 0.942; Recall 0.950; F1 0.94; mAP 0.977	Resizing; adaptive anchors	Blur; flipping; translation
[55]	Precision 0.72; Recall 0.74	Focus stacking; settling	Flips; photometric distortion
[49]	R^2 = 0.71–0.87 (across pigments)	Radiometric calibration; ortho-rectification	Not specified
[27]	Accuracy 0.9895; F1 0.9892	Normalization; resizing	Flipping
[38]	Accuracy 99.51%; Precision 0.727	Background correction; CLAHE	Histogram matching
[34]	Accuracy = 0.993	Padding; contour descriptors	Morphological transforms
[48]	Precision 0.85; Recall 0.86; F1 0.85	Raw-image evaluation	Rotation; mirroring
[39]	Accuracy 0.923	Normalization; edge detection	Reflection; cyclic pooling
[50]	OA 0.927	Atmospheric correction; NDWI	Not specified
[17]	Accuracy 97.67%; AUC 99.51%	WSI tiling; background patch removal; histogram normalization	Rotation (90°–270°); contrast and brightness adjustment
[56]	mAP 0.948; F1 0.912	Resizing; dataset partitioning	Rotation; scaling
[37]	NSE 0.59–0.67; RMSE 12.28–17.25	Autoencoder-based dimensionality reduction	Not reported
[57]	Precision 0.969; Recall 0.968; F1 0.968	Tiling; ROI selection	Rotation; contrast
[58]	Accuracy 0.9879; F1 0.9883	Feature extraction	Not specified
[24]	Precision +25%; Recall +23%	PDF scraping; tiling	Copy–paste; rotation
[47]	mAP 0.902; mAR 0.906	Enzymatic digestion; resizing	Transfer learning
[40]	F1 0.82; Precision 0.77; Recall 0.87; AUC 0.88	Patch extraction (320×320); channel normalization; background under sampling	Rotation; flipping; intensity adjustment; Gaussian noise

Regarding methodological practices, the reviewed studies converge on a small set of recurring preprocessing techniques that address the dominant constraints of diatom imagery. The most common preprocessing step is tiling or patch extraction, applied to whole slide images or large fields of view to standardise input size and reduce computational cost [17], [36], [40], [44]. Background removal or suppression, often using thresholding or masking strategies, is also widely used to reduce non informative regions [17], [32], [52]. Normalisation and contrast enhancement, including histogram based methods and CLAHE, appear frequently to stabilise illumination differences across samples [17], [38].

Modality specific preprocessing is evident in holography, where numerical reconstruction precedes classification or detection [52], [53], and in remote sensing based studies, where preprocessing shifts toward radiometric or atmospheric correction [49], [50].

Augmentation strategies are more uniform than performance reporting and are dominated by simple geometric transformations. Rotation and flipping are the most frequently applied techniques across classification, detection, and segmentation tasks [20], [32], [40], [48], [56]. Several studies also apply photometric adjustments (brightness or contrast) to improve robustness to acquisition variability [17], [44], [55].

In detection oriented pipelines, stronger augmentation schemes such as mosaic or copy-paste are occasionally used to increase object density and background complexity [35], [46]. However, Table 2 also shows that augmentation is not always explicitly reported, and some studies rely primarily on transfer learning strategies (e.g., [47]). This pattern suggests that, while augmentation practices are relatively standardised, their reporting remains inconsistent, which constrains reproducibility and benchmarking across studies.

4.3. RQ3. Technical Challenges that Need to be Considered in the Application of CNN for Microalgae Learning

This section examines the main technical challenges in applying CNN-based automated diatom identification for microalgae learning.

It highlights key limitations related to data quality, model performance, and practical implementation that may affect the reliability and usability of these systems in real learning contexts. The reviewed literature identifies six recurring technical challenges that constrain the application of CNN-based diatom identification in microalgae learning (Table 3). First, limitations in image size and image quality, including large image dimensions, inconsistent illumination, and focus variability, reduce model reliability when applied to non-standardized microscopy images typical of educational settings. Second, complex visual backgrounds in real samples, such as overlapping diatoms and the presence of debris or non-target organisms, increase false detections.

Table 3. Main technical challenges in applying CNN-based diatom identification for microalgae learning

Main Challenge	**Brief Description of the Challenge**	**Ref**
Limitations in image size and image quality	Images are excessively large, poorly focused, inadequately illuminated, or difficult to capture under consistent conditions	[17], [35], [36], [38], [52], [55]
Complex visual backgrounds in real samples	Diatoms frequently overlap with each other and coexist with sand, debris, or other non-target organisms	[31], [35], [40], [45], [46], [47], [48], [54], [57]
High morphological similarity among diatoms	Different species, or different views of the same species, are difficult to distinguish based on visual features alone	[20], [32], [34], [47], [51]
Limited and uneven training data	Some species are highly represented in datasets, while others are rare or entirely absent	[17], [44], [45], [49], [52], [57], [58]
Limited model generalization and interpretability	Models perform well on controlled datasets but often fail under new conditions and provide limited explanation for their predictions	[17], [20], [37], [39], [50], [58]
High computational and technical requirements	Models demand substantial computational resources and are difficult to deploy in practical learning environments	[27], [45], [53], [56]

Third, morphological similarity among diatoms remains a persistent challenge, as visually similar species or different orientations of the same species are difficult to distinguish using image features alone. Fourth, limited and uneven training data, where common taxa dominate and rare taxa are underrepresented, restrict model robustness and bias performance toward frequently observed species. Fifth, studies consistently report limited model generalization and interpretability, with performance degradation under new conditions and limited explanatory insight into model decisions. Finally, high computational and technical requirements pose practical barriers to classroom deployment, particularly in resource constrained learning environments. Collectively, these findings indicate that successful educational adoption of CNN based diatom identification requires addressing data realism, morphological complexity, dataset representativeness, model transparency, and computational feasibility, rather than focusing on accuracy alone.

4.4. RQ4. Future Research Directions for Advancing CNN based Automated Diatom Identification

This section outlines key future research directions for advancing CNN based automated diatom identification. It highlights priority areas that address current limitations and support the development of more robust, scalable, and application-oriented systems. The reviewed studies identify five dominant directions for future research in CNN based automated diatom identification (Table 4). First, dataset expansion and representativeness are the most consistently reported priorities, emphasizing larger datasets, improved coverage of rare taxa, and external validation across institutions to reduce dataset bias. Second, advanced detection and segmentation strategies are required to better address overlapping diatoms, small targets, and complex visual scenes through improved instance segmentation and ROI extraction.

Third, studies highlight the need for improved model generalization and robustness, including explicit handling of domain shift and validation across different environments and acquisition conditions. Fourth, computational efficiency and deployability emerge as key concerns, with recommendations for streamlined pipelines, embedded or edge implementations, and architectures suitable for real time or resource limited use.

Finally, several studies stress integration with environmental monitoring applications, linking automated identification to water quality assessment, forensic analysis, and field-based monitoring. Overall, these findings indicate a shift from model centric optimization toward scalable, robust, and application ready systems.

Table 4. Major future research directions in CNN based automated diatom identification

Main Challenge	Brief Description of the Challenge	Ref
Dataset expansion and representativeness	Increasing dataset size, rare species coverage, and cross institutional validation	[42], [44], [52], [54], [48], [49], [31], [27]
Advanced detection and segmentation strategies	Improving handling of overlapping diatoms, small objects, and ROI extraction	[32], [40], [43], [45], [53], [36], [24], [39]
Model generalization and robustness	Addressing domain shift, adversarial robustness, and cross environment performance	[34], [44], [47], [41]
Computational efficiency and deployment	Reducing computational cost and enabling real time or embedded implementations	[38], [50], [55], [20]
Integration with environmental monitoring	Linking models to water quality assessment, forensic analysis, and real world monitoring	[17], [37], [46], [51]

5. Discussion

This study provides a clear overview of the latest research on CNN-based diatom identification by examining how task design, model architecture, and dataset characteristics influence model performance and practical applications. This is important to ensure that high performance results are not misinterpreted as evidence that these systems are ready for use in the real world or in educational settings. The analysis shows that (1) Most studies focus on classification tasks, (2) Reported performance demonstrates high accuracy in identifying diatoms, (3) Several technical challenges still limit reliable use in real-world conditions, and (4) Future research should focus on developing more robust and deployable systems to support classroom instruction.

5.1. Task Formulation, Architectural Choices, and Dataset Constraints in CNN based Diatom Identification

Task formulation, model architecture, and dataset characteristics are core components that fundamentally determine the performance and application of CNN-based diatom identification systems. In the current literature, the configuration of these elements shows a consistent pattern, where research is largely oriented toward image classification, while segmentation and regression remain relatively under-explored.

This imbalance reflects not only methodological preferences but also practical constraints, particularly the availability and structure of existing datasets. Most datasets are curated at the image or specimen level, which favors classification pipelines that are technically simpler to implement and evaluate. Consequently, current studies tend to prioritize species-level recognition accuracy over more complex analyses of diatom morphology and community structure.

The predominance of classification tasks indicates a continued emphasis on performance benchmarking under controlled conditions rather than comprehensive image understanding in realistic microscopy settings. Although classification models effectively demonstrate algorithmic capability, their reliance on cropped or pre-segmented inputs limits scalability to authentic microscopy workflows [59]. In practice, diatoms frequently appear overlapping, partially occluded, or embedded in complex backgrounds [60]. While detection based approaches are increasingly explored, they remain secondary, suggesting that fully automated diatom analysis has not yet become a central research objective [61].

Architecturally, the literature shows a strong reliance on established CNN backbones originally developed for general computer vision tasks, typically adapted through transfer learning. This pragmatic strategy prioritizes development efficiency and short term performance gains but limits exploration of architectures specifically designed to capture subtle, taxonomically relevant diatom features [62].

Consequently, reported performance improvements are often driven more by dataset composition and preprocessing strategies than by substantive architectural innovation [63].

Dataset characteristics further constrain task formulation and model generalization. Existing datasets vary widely in size and accessibility, with most remaining private, thereby limiting reproducibility and cross study validation. Light microscopy overwhelmingly dominates data sources, while alternative imaging modalities are used only in specialized contexts, reinforcing task and architecture choices tailored to narrow imaging conditions. High reported performance is therefore often dataset specific, reflecting controlled acquisition settings with limited variability [64], rather than robustness across real world educational or field based applications [65], [66].

5.2. Reported Performance, Evaluation Inconsistency, and the Role of Preprocessing and Augmentation

Reported model performance is often used as the primary indicator of progress in CNN-based diatom identification, with many studies reporting high results, particularly for classification tasks, followed by detection and more limited segmentation and regression applications. However, these performance claims need to be interpreted with caution, as they are strongly influenced by variations in evaluation metrics, validation strategies, and dataset conditions. Classification performance is typically expressed using accuracy, precision, recall, and F1-score, whereas detection studies predominantly rely on mean average precision, yet substantial heterogeneity in metric selection and evaluation protocols limits meaningful comparison across studies. As a result, high reported performance does not necessarily reflect comparable robustness or generalizability, but rather highlights underlying inconsistencies in how model effectiveness is assessed.

A key issue is the inconsistency of performance reporting. Many studies rely on incomplete metric sets, often emphasizing accuracy without adequately addressing class imbalance or error distribution. Validation strategies also vary widely, including differences in data splitting, cross-validation schemes, and the frequent absence of independent external test datasets. Consequently, high numerical performance does not necessarily indicate comparable robustness or transferability across experimental settings.

High reported performance is further shaped by controlled dataset conditions. Most datasets are acquired under standardized imaging protocols with limited variability in illumination, focus, and background composition, allowing CNN models to exploit stable visual patterns [25].

While this supports efficient within-dataset optimization, it does not adequately reflect the complexity of educational or field based microscopy, where image heterogeneity is substantially higher [67], [68]. Reported performance should therefore be interpreted as dataset-specific rather than as evidence of generalizable system readiness [69].

Clear patterns also emerge in preprocessing and augmentation practices. Tiling or patch extraction is widely used to standardize input size and reduce computational cost, while normalization, contrast enhancement, and background suppression are commonly applied to stabilize training. Data augmentation primarily relies on basic geometric transformations, particularly rotation and flipping, with occasional photometric adjustments and more complex strategies in detection-oriented studies [70], [71].

5.3. Technical Challenges and Implications for Microalgae Learning

The application of CNN based diatom identification in microalgae learning introduces challenges that extend beyond technical performance, requiring careful consideration of both biological complexity and real-world learning conditions. Evidence from the reviewed studies indicates that these challenges arise less from limitations in algorithmic capability than from mismatches between model assumptions and the realities of educational contexts. In practice, limitations in image size and quality, together with complex visual backgrounds, reduce model reliability when systems trained on curated datasets are applied to heterogeneous microscopy conditions typical of classroom settings [72]. Variability in illumination, focus, and sample preparation, combined with overlapping diatoms and non-target objects, increases visual complexity and reduces model reliability[13]. These findings suggest that performance reported under controlled settings does not readily translate to authentic learning environments, where visual noise and user dependent acquisition are unavoidable [73], [74]. As a result, high benchmark accuracy alone is insufficient to ensure robust or pedagogically meaningful deployment [75].

Beyond the issue of data realism, the high morphological similarity among diatoms represents a fundamental biological challenge in microalgae learning. Closely related species, as well as different orientations of the same species, are often difficult to distinguish visually, which limits the discrimination performance of CNNs [76]. This problem arises due to an unbalanced training dataset, where a small number of common taxa are dominant while rare species are underrepresented or even absent [77].

This imbalance not only influences model predictions toward frequently observed species but also risks narrowing the way diatom diversity is represented in educational applications [78], [79]. Therefore, dataset selection must be carefully considered to support CNN performance in microalgae learning implementations [80].

Further challenges stem from limited model generalization, low interpretability, and high computational requirements. Many models only perform well on familiar datasets and show a significant decline in performance when applied to new imaging conditions. This limits their usefulness as learning tools. In addition, the need for large computational resources and technical expertise reduces the feasibility of implementing these systems in typical classroom environments. Overall, these findings suggest that the effective integration of CNN-based diatom identification into microalgae learning requires a shift from accuracy-focused development toward an approach that emphasizes realistic data conditions, balanced representation of taxa, transparent model behavior, and practical applicability [34], [81]. Addressing these aspects is critical to ensure that AI systems support meaningful biological learning rather than oversimplifying biological complexity [82], [83].

5.4. *Future Research Directions: from Technical Optimisation to Educational Alignment (RQ4)*

Future research directions are critical for addressing current limitations in CNN-based diatom identification and improving its applicability in real-world contexts. The evidence synthesized in this review indicates a clear shift from model-centered optimization toward system readiness and practical deployment. In particular, limitations related to data availability and representativeness are consistently identified as fundamental barriers. As a result, future research increasingly emphasizes expanding datasets to cover a broader taxonomic range and validating models across institutions and diverse data collection settings [84], [85]. This emphasis reflects a growing awareness that high performance on curated datasets does not necessarily translate to reliable performance in diverse ecological or educational contexts, where variability in imaging conditions and community composition is unavoidable [86], [87].

In parallel, many studies highlight the need for more advanced detection and segmentation strategies to address the visual complexity of realistic microscopy data. Overlapping diatoms, small targets, and heterogeneous backgrounds remain major challenges when models are applied beyond cropped or pre processed images. As a result, future research increasingly prioritizes instance level localization and region extraction as prerequisites for meaningful classification.

Closely related to this challenge is the issue of model generalization and robustness [88], [89]. Differences in imaging devices, preparation methods, and environmental conditions introduce domain shifts that can substantially degrade performance [90]. Addressing these effects is framed not as an optional improvement, but as a requirement for ensuring stable and transferable model behavior across settings [91].

Computational efficiency and application level integration emerge as key directions shaping the next phase of development. Several studies stress the importance of reducing computational demands and enabling deployment on resource limited platforms, including real time or embedded systems. This concern is particularly relevant for educational and field based applications, where technical infrastructure is often constrained [92], [93]. At the same time, there is increasing emphasis on linking automated diatom identification directly to environmental monitoring tasks such as water quality assessment and operational analysis. Together, these directions signal a transition from proof of concept models toward scalable and application oriented systems, where performance is evaluated not only by accuracy metrics but also by robustness, feasibility, and practical relevance [94], [95].

This review presents a task-oriented synthesis of CNN-based automatic diatom identification by systematically linking task formulation, model architecture, dataset characteristics, performance reporting, technical challenges, and future research directions. Rather than focusing solely on model accuracy, this review places automated diatom identification in the context of microalgae learning and demonstrates that technical design choices directly influence model interpretability, scalability, and educational utility. Several limitations should be acknowledged. First, significant heterogeneity in datasets, imaging modalities, taxonomic coverage, and evaluation protocols limits direct comparisons between studies and reduces the reliability of aggregated performance conclusions. Second, some included studies addressed related but not equivalent tasks, such as pigment estimation or remote sensing-based mapping. While these studies provide methodological insights, they differ from microscopy-based species identification and therefore have limited relevance for evaluating learning-oriented diatom identification systems.

Overall, these findings suggest that future research should place greater emphasis on standardized evaluation practices, clear task definitions, and datasets that reflect authentic microscopy conditions. Advances in CNN-based diatom identification for educational use will depend more on an integrated approach that aligns technical development with biological validity and learning objectives, rather than incremental improvements in accuracy.

6. Conclusion

This systematic review synthesises recent research on CNN-based automated diatom identification by analysing task formulation, model architectures, dataset characteristics, performance reporting, technical challenges, and future research directions. The findings show that the field is largely dominated by classification and detection-centric approaches based on established CNN architectures and microscopy-derived datasets, while segmentation and regression-based tasks remain limited. Although high performance metrics are frequently reported, substantial heterogeneity in datasets, evaluation metrics, and validation protocols restrict cross-study comparability and weaken claims of generalisability.

The review further indicates a convergence toward pragmatic preprocessing and augmentation practices that address the visual complexity of diatom imagery, yet persistent technical challenges, such as morphological ambiguity, class imbalance, limited robustness, and low interpretability, continue to constrain reliable application. These limitations are particularly critical in microalgae learning contexts, where model predictions must support conceptual understanding rather than function as opaque classification outputs. Overall, the synthesis of future research directions highlights a shift toward data-centred development, robustness-oriented evaluation, and educational alignment. Expanding shared datasets, improving generalisation under realistic conditions, and integrating interpretable feedback mechanisms emerge as key priorities for advancing CNN-based diatom identification beyond proof-of-concept studies. This review therefore provides a concise analytical foundation for developing scalable, reliable, and educationally meaningful AI tools for diatom analysis and microalgae learning.

Acknowledgements

We would like to express our deepest gratitude to the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Education, Culture, Research, and Technology, INDONESIA, through the Grant for funding this research, specifically in the implementation of CNN in microalgae learning through contract numbers: 085/C3/DT.05.00/PL/2025 and 2.6.111/ UN32.14.1/LT/2025.

References:

[1]. Almeida, S. F., & Feio, M. J. (2024). Present and future perspectives for bioassessment of running water using diatoms. *Diatom Photosynthesis: From Primary Production to High-Value Molecules*, 351-382. Doi: 10.1002/9781119842156.ch12

[2]. Andreini, P., et al. (2020). Image generation by GAN and style transfer for agar plate image segmentation. *Computer methods and programs in biomedicine*, *184*, 105268. Doi: 10.1016/j.cmpb.2019.105268

[3]. Arnal Barbedo, J. G. (2013). Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, *2*(1). Doi: 10.1186/2193-1801-2-660

[4]. Avilés-Cruz, C., et al. (2019). Coarse-fine convolutional deep-learning strategy for human activity recognition. *Sensors*, *19*(7), 1556. Doi: 10.3390/s19071556

[5]. Baek, S. S., et al. (2020). Identification and enumeration of cyanobacteria species using a deep neural network. *Ecological Indicators*, *115*, 106395. Doi: 10.1016/j.ecolind.2020.106395

[6]. Buck, T., et al. (2019). Just a small bunch of flowers: The botanical knowledge of students and the positive effects of courses in plant identification at German universities. *PeerJ*, *7*, e6581. Doi: 10.7717/peerj.6581

[7]. Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks*, *106*, 249-259. Doi: 10.1016/j.neunet.2018.07.011

[8]. Serrano-Bueno, G., et al. (2025). Real-Time Edge Computing vs. GPU-Accelerated Pipelines for Low-Cost Microscopy Applications. *Electronics (Switzerland)*. Doi: 10.3390/electronics14050930

[9]. Burfeid-Castellanos, A. M., et al. (2022). A digital light microscopic method for diatom surveys using embedded acid-cleaned samples. *Water*, *14*(20), 3332. Doi: 10.3390/w14203332

[10]. Clark, A. T., et al. (2020). Predicting species abundances in a grassland biodiversity experiment: Trade-offs between model complexity and generality. *Journal of ecology*, *108*(2), 774-787. Doi: 10.1111/1365-2745.13316

[11]. Davidian, M., et al. (2024). Exploring the interplay of dataset size and imbalance on CNN performance in healthcare: Using X-rays to identify COVID-19 patients. *Diagnostics*, *14*(16), 1727. Doi: 10.3390/diagnostics14161727

[12]. de Benito Fernández, M., et al. (2023). Evaluation of XAI Models for Interpretation of Deep Learning Techniques' Results in Automated Plant Disease Diagnosis. *Lecture Notes in Networks and Systems*, *732 LNNS*. Doi: 10.1007/978-3-031-36957-5_36

[13]. Deng, J., et al. (2021). Identification of diatom taxonomy by a combination of region-based full convolutional network, online hard example mining, and shape priors of diatoms. *International Journal of Legal Medicine*, *135*(6), 2519-2530. Doi: 10.1007/s00414-021-02664-2

[14]. Deng, J., et al. (2022). A coarse to fine framework for recognizing and locating multiple diatoms with highly complex backgrounds in forensic investigation. *Multimedia Tools and Applications*, *81*(4), 4839-4857. Doi: 10.1007/s11042-021-11169-4

[15]. Devia, C., et al. (2024). Exploring biological challenges in building a thinking machine. *Cognitive Systems Research*, *87*. Doi: 10.1016/j.cogsys.2024.101260

[16]. Ding, W., & Li, C. (2024). Algal blooms forecasting with hybrid deep learning models from satellite data in the Zhoushan fishery. *Ecological Informatics*, *82*. Doi: 10.1016/j.ecoinf.2024.102664

[17]. Dourado-Filho, L. A., & Calumby, R. T. (2021). An experimental assessment of deep convolutional features for plant species recognition. *Ecological Informatics*, *65*, 101411. Doi: 10.1016/j.ecoinf.2021.101411

[18]. Ellen, J. S., Graff, C. A., & Ohman, M. D. (2019). Improving plankton image classification using context metadata. *Limnology and Oceanography: Methods*, *17*(8), 439-461. Doi: 10.1002/lom3.10324

[19]. Fatima, S., et al. (2025). Deep learning in dermatopathology: applications for skin disease diagnosis and classification. *Discover Applied Sciences*, *7*(9), 1006. Doi: 10.1007/s42452-025-07138-3

[20]. Garcia, T., et al. (2023). Temporal changes in zooplankton indicators highlight a bottom-up process in the Bay of Marseille (NW Mediterranean Sea). *Plos one*, *18*(10), e0292536. Doi: 10.1371/journal.pone.0292536

[21]. Godbillot, C., et al. (2024). A new method for the detection of siliceous microfossils on sediment microscope slides using convolutional neural networks. *Journal of Geophysical Research: Biogeosciences*, *129*(9), e2024JG008047. Doi: 10.1029/2024JG008047

[22]. Villa, A. G., Salazar, A., & Vargas, F. (2017). Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological informatics*, *41*, 24-32. Doi: 10.1016/j.ecoinf.2017.07.004

[23]. Gu, G., et al. (2022). Automated diatom detection in forensic drowning diagnosis using a single shot multibox detector with plump receptive field. *Applied soft computing*, *122*, 108885. Doi: 10.1016/j.asoc.2022.108885

[24]. Gunduz, H., & Gunal, S. (2024). A lightweight convolutional neural network (CNN) model for diatom classification: DiatomNet. *PeerJ Computer Science*, *10*, e1970. Doi: 10.7717/peerj-cs.1970

[25]. Guo, B., et al. (2021). Automated plankton classification from holographic imagery with deep convolutional neural networks. *Limnology and Oceanography: Methods*, *19*(1), 21-36. Doi: 10.1002/lom3.10402

[26]. Habib, S., et al. (2024). Advancing taxonomic classification through deep learning: a robust artificial intelligence framework for species identification using natural images. *IEEE Access*, *12*, 146718-146732. Doi: 10.1109/ACCESS.2024.3450016

[27]. Haddaway, N. R., & Vicente-Vicente, J. L. (2023). Fatal flaws in systematic review conduct, and how to fix them—A commentary to Rao. et al. (2022). *Sustainable Cities and Society*, *92*. Doi: 10.1016/j.scs.2023.104457

[28]. Hooykaas, M. J., et al. (2019). Identification skills in biodiversity professionals and laypeople: A gap in species literacy. *Biological Conservation*, *238*, 108202. Doi: 10.1016/j.biocon.2019.108202

[29]. Jahan, S. S., et al. (2022). Exploring the association between emotional intelligence and academic performance and stress factors among dental students: a scoping review. *Dentistry journal*, *10*(4), 67. Doi: 10.3390/dj10040067

[30]. Jaillard, B., et al. (2018). A combinatorial analysis using observational data identifies species that govern ecosystem functioning. *PloS one*, *13*(8), e0201135. Doi: 10.1371/journal.pone.0201135

[31]. Kaasinen, A., & Kervinen, A. (2026). Implementation of distance outdoor education in pre-service teacher education by using mobile interaction. *Environmental Education Research*, *32*(1), 211-231. Doi: 10.1080/13504622.2025.2512969

[32]. Kim, H. K., et al. (2024). Advancing river health assessments: Integrating microscopy and molecular techniques through diatom indices. *Water*, *16*(6), 853. Doi: 10.3390/w16060853

[33]. Kim, J. S. G., et al. (2025). Comparison of image preprocessing strategies for convolutional neural network-based growth stage classification of butterhead lettuce in industrial plant factories. *Applied Sciences*, *15*(11), 6278. Doi: 10.3390/app15116278

[34]. Kitano, H. (2016). Biological complexity and the need for computational approaches. *Philosophy of Systems Biology: Perspectives from Scientists and Philosophers*, 169-180. Springer International Publishing. Doi: 10.1007/978-3-319-47000-9_16

[35]. Kloster, M., et al. (2023). Improving deep learning-based segmentation of diatoms in gigapixel-sized virtual slides by object-based tile positioning and object integrity constraint. *Plos one*, *18*(2), e0272103. Doi: 10.1371/journal.pone.0272103

[36]. Kloster, M., et al. (2020). Deep learning-based diatom taxonomy on virtual slides. *Scientific reports*, *10*(1), 14416. Doi: 10.1038/s41598-020-71165-w

[37]. Kraft, K., et al. (2022). Towards operational phytoplankton recognition with automated high-throughput imaging, near-real-time data processing, and convolutional neural networks. *Frontiers in Marine Science*, *9*, 867695. Doi: 10.3389/fmars.2022.867695

[38]. Krause, L. M., et al. (2020). Fully convolutional neural network for detection and counting of diatoms on coatings after short-term field exposure. *Environmental science & technology*, *54*(16), 10022-10030. Doi: 10.1021/acs.est.0c01982

[39]. Kwon, D., et al. (2025). Morphological characteristic analysis of unrecorded freshwater diatoms in Korea. *Applied Microscopy*, *55*(1), 5. Doi: 10.1186/s42649-025-00112-8

[40]. Lakehal, A., et al. (2024). Multi-classification and Multi-labeling of UML Diagrams Using Transfer Learning and CNN. *International Conference on Artificial Intelligence in Renewable Energetic Systems*, 573-583. Doi: 10.1007/978-3-031-80301-7_62

[41]. Li, J., et al. (2024). SC-DiatomNet: An Efficient and Accurate Algorithm for Diatom Classification. *Journal of Marine Science and Engineering*, *12*(10), 1862. Doi: 10.3390/jmse12101862

[42]. Libreros, J., et al. (2018). Automated identification and classification of diatoms from water resources. In *Iberoamerican congress on pattern recognition*, 496-503. Springer International Publishing. Doi: 10.1007/978-3-030-13469-3_58

[43]. Liu, Q., He, D., & Xie, L. (2019). Prediction of off-target specificity and cell-specific fitness of CRISPR-Cas System using attention boosted deep learning and network-based gene feature. *PLoS computational biology*, *15*(10), e1007480. Doi: 10.1371/journal.pcbi.1007480

[44]. MacIsaac, J., et al. (2024). Improving acoustic species identification using data augmentation within a deep learning framework. *Ecological Informatics*, *83*, 102851. Doi: 10.1016/j.ecoinf.2024.102851

[45]. MacNeil, L., et al. (2021). Plankton classification with high-throughput submersible holographic microscopy and transfer learning. *BMC Ecology and Evolution*, *21*(1), 123. Doi: 10.1186/s12862-021-01839-0

[46]. Madkour, D. M., et al. (2023). A systematic review of deep learning microalgae classification and detection. *IEEE Access*, *11*, 57529-57555. Doi: 10.1109/ACCESS.2023.3280410

[47]. Mallawaarachchi, V., et al. (2018). Experiential learning in bioinformatics–learner support for complex workflow modelling and analysis. *International Journal of Emerging Technologies in Learning*, *13*(12), 19-34. Doi: 10.3991/ijet.v13i12.8608

[48]. Mane, D. T., & Kulkarni, U. V. (2017). A survey on supervised convolutional neural network and its major applications. *International Journal of Rough Sets and Data Analysis (IJRSDA)*, *4*(3), 71-82. Doi: 10.4018/ijrsda.2017070105

[49]. Marini, N., et al. (2021). Semi-supervised training of deep convolutional neural networks with heterogeneous data and few local annotations: An experiment on prostate histopathology image classification. *Medical image analysis*, *73*, 102165. Doi: 10.1016/j.media.2021.102165

[50]. Marzidovšek, M., et al. (2024). Computer Vision Techniques for Morphological Analysis and Identification of Two Pseudo-nitzschia Species. *Water*, *16*(15), 2160. Doi: 10.3390/w16152160

[51]. Nardelli, S. C., Gray, P. C., & Schofield, O. (2021). Developing a convolutional neural network to classify phytoplankton images collected with an Imaging FlowCytobot along the West Antarctic Peninsula. *OCEANS 2021: San Diego–Porto*, 1-7. Doi: 10.23919/OCEANS44145.2021.9706072

[52]. Nardelli, S. C., Gray, P. C., & Schofield, O. (2022). A convolutional neural network to classify phytoplankton images along the West Antarctic Peninsula. *Marine Technology Society Journal*, *56*(5), 45-57. Doi: 10.4031/MTSJ.56.5.8

[53]. Otálora, S., et al. (2021). Combining weakly and strongly supervised learning improves strong supervision in Gleason pattern classification. *BMC Medical Imaging*, *21*(1), 77. Doi: 10.1186/s12880-021-00609-0

[54]. Pacheco-Venegas, E. G., et al. (2018). Automatic identification of diatoms using descriptors obtained in the plane of frequencies. *Applications of Digital Image Processing XLI*, *10752*, 63-69. Doi: 10.1117/12.2320869

[55]. Pardeshi, R., & Deshmukh, P. (2023). Efficient microalgae species identification using compact convolutional neural network. *Int. J. Recent Innov. Trends Comput. Commun.*, *11*(7s), 8-15. Doi: 10.17762/ijritcc.v11i7s.6972

[56]. Parmar, T. K., Rawtani, D., & Agrawal, Y. K. (2016). Bioindicators: the natural indicator of environmental pollution. *Frontiers in life science*, *9*(2), 110-118. Doi: 10.1080/21553769.2016.1162753

[57]. Pedraza, A., et al. (2017). Automated diatom classification (Part B): a deep learning approach. *Applied Sciences*, *7*(5), 460. Doi: 10.3390/app7050460

[58]. Pooch, E. H. P., Ballester, P., & Barros, R. C. (2020). Can we trust deep learning based diagnosis? The impact of domain shift in chest radiograph classification. *International workshop on thoracic image analysis*, 74–83. Springer International Publishing. Doi: 10.1007/978-3-030-62469-9_7

[59]. Popkov, A., et al. (2022). Machine learning for expert-level image-based identification of very similar species in the hyperdiverse plant bug family Miridae (Hemiptera: Heteroptera). *Systematic Entomology*, *47*(3), 487-503. Doi: 10.1111/syen.12543

[60]. Pyo, J., Hong, S. M., Jang, J., Park, S., Park, J., Noh, J. H., & Cho, K. H. (2022). Drone-borne sensing of major and accessory pigments in algae using deep learning modeling. *GIScience & Remote Sensing*, *59*(1), 310-332. Doi: 10.1080/15481603.2022.2027120

[61]. Rani, P., et al. (2022). Machine learning and deep learning based computational approaches in automatic microorganisms image recognition: methodologies, challenges, and developments. *Archives of Computational Methods in Engineering*, *29*(3), 1801-1837. Doi: 10.1007/s11831-021-09639-x

[62]. Ruiz-Santaquiteria, J., et al. (2020). Semantic versus instance segmentation in microscopic algae detection. *Engineering Applications of Artificial Intelligence*, *87*, 103271. Doi: 10.1016/j.engappai.2019.103271

[63]. Saleem, N., et al. (2024). Class-specific data augmentation for plant stress classification. *The Plant Phenome Journal*, *7*(1), e20112. Doi: 10.1002/ppj2.20112

[64]. Salido, J., et al. (2020). A low-cost automated digital microscopy platform for automatic identification of diatoms. *Applied Sciences*, *10*(17), 6033. Doi: 10.3390/app10176033

[65]. Sánchez, C., et al. (2019). Diatom classification including morphological adaptations using CNNs. In *Iberian Conference on Pattern Recognition and Image Analysis*, 317-328. Springer International Publishing. Doi: 10.1007/978-3-030-31332-6_28

[66]. Schoolmaster Jr, D. R., Zirbel, C. R., & Cronin, J. P. (2020). A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations. *Ecology*, *101*(8), e03070. Doi: 10.1002/ecy.3070

[67]. Shafik, W., et al. (2023). A systematic literature review on plant disease detection: motivations, classification techniques, datasets, challenges, and future trends. *IEEE Access*, *11*, 59174-59203. Doi: 10.1109/ACCESS.2023.3284760

[68]. Sharma, S., Kumaraswamy, E., & Kumar, S. (2023). Retrospective study of convolutional neural network for medical image analysis and a deep insight through histopathological dataset. *Computational intelligence: Select proceedings of InCITe 2022* (*Lecture Notes in Electrical Engineering)*, *968*, 47–58. Doi: 10.1007/978-981-19-7346-8_5

[69]. Shrivas, V., et al. (2025). Machine learning and image analysis in diatom taxonomy: Progress and prospects. *AIP Conference Proceedings*, *3343*(1), 040036. AIP Publishing LLC. Doi: 10.1063/5.0292582

[70]. Simović, P., et al. (2024). The Effects of Data Quality on Deep Learning Performance for Aquatic Insect Identification: Advances for Biomonitoring Studies. *Water*, *17*(1), 21. Doi: 10.3390/w17010021

[71]. Skarstein, T. H., & Skarstein, F. (2020). Curious children and knowledgeable adults–early childhood student-teachers' species identification skills and their views on the importance of species knowledge. *International Journal of Science Education*, *42*(2), 310-328. Doi: 10.1080/09500693.2019.1710782

[72]. Soeprobowati, T. R., et al. (2023). Diatom index of Galela Lake, Halmahera, Indonesia in relation to human activities. *International Journal of Environmental Science and Technology*, *20*(7), 7707-7722. Doi: 10.1007/s13762-022-04463-7

[73]. Sonmez, M. E., et al. (2022). Convolutional neural network-Support vector machine based approach for classification of cyanobacteria and chlorophyta microalgae groups. *Algal Research*, *61*, 102568. Doi: 10.1016/j.algal.2021.102568

[74]. Sonmez, M. E., et al. (2024). Enhancing microalgae classification accuracy in marine ecosystems through convolutional neural networks and support vector machines. *Marine Pollution Bulletin*, *205*, 116616. Doi: 10.1016/j.marpolbul.2024.116616

[75]. Spaulding, S. A., et al. (2021). Diatoms. org: supporting taxonomists, connecting communities. *Diatom Research*, *36*(4), 291-304. Doi: 10.1080/0269249X.2021.2006790

[76]. Taurozzi, D., Cesarini, G., & Scalici, M. (2024). Diatoms as bioindicators for health assessments of ephemeral freshwater ecosystems: A comprehensive review. *Ecological Indicators*, *166*, 112309. Doi: 10.1016/j.ecolind.2024.112309

[77]. Tetard, M., et al. (2023). Merging databases for CNN image recognition, increasing bias or improving results?. *Marine Micropaleontology*, *185*, 102296. Doi: 10.1016/j.marmicro.2023.102296

[78]. Tiwary, P., & Bhattacharyya, K. (2024). Cycle consistent twin energy-based models for image-to-image translation. *Medical image analysis*, *91*, 103031. Doi: 10.1016/j.media.2023.103031

[79]. Tournois, L., et al. (2024). Automatic detection and identification of diatoms in complex background for suspected drowning cases through object detection models. *International journal of legal medicine*, *138*(2), 659-670. Doi: 10.1007/s00414-023-03096-w

[80]. Triest, L., et al. (2012). Epilithic diatoms as indicators in tropical African rivers (Lake Victoria catchment). *Hydrobiologia*, *695*(1), 343-360. Doi: 10.1007/s10750-012-1201-2

[81]. Ugurlu, D., et al. (2021). The impact of domain shift on left and right ventricle segmentation in short axis cardiac MR images. *International Workshop on Statistical Atlases and Computational Models of the Heart*, 57-65. Springer International Publishing. Doi: 10.1007/978-3-030-93722-5_7

[82]. Vallez, N., Bueno, G., Deniz, O., & Blanco, S. (2022). Diffeomorphic transforms for data augmentation of highly variable shape and texture objects. *Computer Methods and Programs in Biomedicine*, *219*, 106775. Doi: 10.1016/j.cmpb.2022.106775

[83]. Valova, I., Harris, C., Mai, T., & Gueorguieva, N. (2020). Optimization of convolutional neural networks for imbalanced set classification. *Procedia Computer Science*, *176*, 660-669. Doi: 10.1016/j.procs.2020.09.038

[84]. Venkatachalapathy, R., & Karthikeyan, P. (2013). Synedra species morphological observation at Bhavani region in Cauvery river, Tamil Nadu, India. *International Research Journal of Earth Sciences*, *1*(6), 1-5.

[85]. Venkataramanan, A., et al. (2024). "UDE DIATOMS in the Wild 2024": a new image dataset of freshwater diatoms for training deep learning models. *GigaScience*, *13*. Doi: 10.1093/gigascience/giae087

[86]. Wäldchen, J., Rzanny, M., Seeland, M., & Mäder, P. (2018). Automated plant species identification—Trends and future directions. *PLoS computational biology*, *14*(4), e1005993. Doi: 10.1371/journal.pcbi.1005993

[87]. Wang, R., Chaudhari, P., & Davatzikos, C. (2022). Embracing the disharmony in medical imaging: A simple and effective framework for domain adaptation. *Medical image analysis*, *76*, 102309. Doi: 10.1016/j.media.2021.102309

[88]. Warren, D. L., et al. (2014). Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 C alifornia vertebrate species of concern. *Diversity and distributions*, *20*(3), 334-343. Doi: 10.1111/ddi.12160

[89]. Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of planning education and research*, *39*(1), 93-112. Doi: 10.1177/0739456X17723971

[90]. Yue, L., et al. (2023). Detection of the status of diatom blooms in the tributaries of the Yangtze River based on Sentinel-2 images. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1-15. Doi: 10.1109/TGRS.2023.3326143

[91]. Zhang, J., et al. (2023). DiatomNet v1. 0: A novel approach for automatic diatom testing for drowning diagnosis in forensically biomedical application. *Computer methods and programs in biomedicine*, *232*, 107434. Doi: 10.1016/j.cmpb.2023.107434

[92]. Zhang, J., et al. (2021). An efficient method for building a database of diatom populations for drowning site inference using a deep learning algorithm. *International journal of legal medicine*, *135*(3), 817-827. Doi: 10.1007/s00414-020-02497-5

[93]. Zhou, H., Zhou, H., & Dong, J. (2024). Progress and challenges of optical convolutional computing (invited). *Acta Optica Sinica*, *44*(15), 1513022–1513022. Doi: 10.3788/AOS240782

[94]. Zhou, Y., et al. (2019). Digital whole-slide image analysis for automated diatom test in forensic cases of drowning using a convolutional neural network algorithm. *Forensic science international*, *302*, 109922. Doi: 10.1016/j.forsciint.2019.109922

[95]. Zhu, Y. Z., et al. (2022). Comparison among Four Deep Learning Image Classification Algorithms in AI-based Diatom Test. *Journal of Forensic Medicine*, *38*(1). Doi: 10.12116/j.issn.1004-5619.2021.411001