VOL. 23, NO. 2, 2024, 10-22
elektrika.utm.my
ISSN 0128-4428

ELEKTRIKA
Journal of Electrical Engineering

Contemporary Methodologies for Identifying and Categorizing Microalgae: A Comprehensive Review and Future Perspectives

Pok Wei Han[1] and **Mohd Ridzuan Bin Ahmad**[1*]

[1]Department of Control and Mechatronics Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.

[*]Corresponding author: mdridzuan@utm.my

Abstract: As more people become connected to the grid and the energy demand continues to increase, non-renewable energy sources are being consumed at a faster rate than they are being replaced. Microalgae, often referred to as "green gold" have shown great potential as a renewable energy source due to their unique characteristics. However, not all microalgae are suitable to use as replacements for traditional fossil fuels. Therefore, it is essential for researchers to accurately identify and classify microalgae based on their species and energy-producing capabilities. The main objective of this paper is to present a detailed summary of the current technologies employed in the classification and detection of microalgae. Both traditional manual microscopy and advanced artificial intelligence techniques such as machine learning and deep learning are covered in this overview. Furthermore, this paper offers a critical analysis of these technologies and provides suggestions for enhancing their effectiveness. Despite deep learning being the most advanced technology for microalgae classification and detection, there is still significant potential for future improvements that could further increase the accuracy.

Keywords: microalgae, microscopy, machine learning, deep learning

Article History: received 7 May 2023; accepted 15 July 2024; published 29 August 2024

1. INTRODUCTION

The emerging field of microalgae-derived biodiesel offers a promising alternative to fossil fuels. Microalgae, microscopic photosynthetic organisms found in diverse aquatic and terrestrial environments, are dubbed "green gold" due to their exceptional productivity, carbon capture capabilities, and minimal environmental footprint. They represent a renewable resource with potential applications in food, bioactive compounds, and notably, biofuels [3].

Microalgae stand out as a sustainable biofuel source due to their rapid growth rate and ability to convert sunlight and CO2 into organic matter through photosynthesis, effectively reducing atmospheric carbon dioxide levels [4]. However, significant challenges persist in harnessing their full potential:

1. **Identification of High-Energy Species:** Despite their abundance, not all microalgae species possess suitable lipid content for efficient biofuel production. Further research is essential to identify and characterize species with optimal lipid yields across diverse environmental conditions [6, 60].
2. **Environmental and Sustainability Assessments:** While microalgae are touted as eco-friendly alternatives, comprehensive life cycle assessments are needed to evaluate their overall environmental impact, including land use, water consumption, and nutrient requirements [4, 61].
3. **Technological Advancements in Detection and Classification:** Current methods for identifying and classifying microalgae are limited by their small size and morphological similarities. Advancements in microscopy, image processing techniques, and machine learning algorithms are critical to enhancing accuracy and efficiency in species identification [8, 62].
4. **Optimization of Cultivation Techniques:** To maximize biomass and lipid productivity, there is a pressing need to optimize cultivation techniques. Research should focus on refining nutrient supply, light exposure, temperature control, and CO2 utilization efficiency [5, 63].
5. **Integration into Energy Systems:** Effective integration of microalgae-based biofuels into existing energy infrastructures requires further exploration. This includes compatibility studies with fuel distribution systems, blending strategies with conventional fuels, and economic viability assessments for scaled production [7, 64].
6. **Policy and Regulatory Frameworks:** Development of supportive policies and regulatory frameworks is essential to incentivize investment and commercialization of microalgae-based biofuels. Research can contribute by identifying policy gaps,

evaluating economic incentives, and addressing barriers to market adoption [5].

Addressing these research gaps is crucial for advancing microalgae as a sustainable renewable energy source. By overcoming technological, environmental, and economic challenges, microalgae-based biofuels can play a pivotal role in achieving global energy sustainability goals.

To address all these challenges requires a lot of effort. This review paper focuses on one of the challenges associated with technology in detection and classification of microalgae. Technological advancements are required to facilitate the detection and classification of microalgae species, particularly for those samples found in a wild environment that have a mixture of impurities and microalgae. Researchers have employed several advanced methods for microalgae applications that have yielded significant results. The focus of this paper is to review the current technologies employed by researchers for microalgae detection and classification, including manual inspection and artificial intelligence techniques such as machine learning and deep learning. A brief overview and introduction of the technologies utilized by researchers as well as the author's own perspective will be discussed.

2. TECHNOLOGIES USED IN MICROALGAE DETECTION AND CLASSIFICATION

Visualizing microalgae is the initial step in commencing research on them because algae have been estimated to include anything from 30,000 to more than 1 million species [9]. Microalgae exhibit unique features and distinct colors that differentiate one species from another. In most cases, when a sample is taken from a given environment, such as a pond, it is likely to contain multiple microalgal species. Therefore, researchers need to accurately identify and classify the specific microalgae that possess the desired characteristics for their potential use in renewable energy applications. Various methods are available for the identification and characterization of microalgae, including manual inspection using a light microscope and artificial intelligence. These methods are crucial for enabling researchers to isolate and study the microalgae species that hold the most potential for use as "green gold" in renewable energy production. The next session will briefly introduce some common technologies use in microalgae detection and classification.

2.1 Microscopy

The observation of microalgae by researchers typically involves the use of a light microscope. Although, with the availability of various types of microscopes, each with unique features and capabilities, selecting an appropriate microscope is crucial for optimal observation. Different species of microalgae exhibit diverse sizes and structures, thus necessitating careful consideration of the most suitable microscope.

Microscopy enables close observation of microalgae's morphological structures and colors, which yield significant insights into their biological functions such as health, lipid content, and adaptation to different environmental conditions. To achieve the highest level of detail and accuracy, a microscope with advanced resolution and magnification capabilities is vital. This allows for precise observations of even the smallest features of the microalgae's structures, and real-time monitoring of their behavior.

2.2 Artificial Intelligence

AI, which stands for Artificial Intelligence, is a broad field of computer science focused on the development of intelligent machines that can perform tasks typically requiring human intelligence [10], [11]. AI, is broadly and generally used to refer to any sort of machine learning program [12]. Computers with artificial intelligence are designed to perform various activities, including speech recognition, learning, planning, and problem-solving which are not programmed in a machine [13],[14]. These capabilities allow AI systems to understand and interpret human language, improve their performance over time through data analysis, create and execute plans to achieve specific goals, and solve complex problems by identifying patterns and making predictions. AI involves the use of various techniques, including Machine Learning (ML), Deep Learning (DL), natural language processing, robotics, and expert systems. Figure 1 shows the relationship between AI, ML, and DL. ML is the subset of AI and DL is the subset of ML [15]. The ultimate goal of AI is to create intelligent machines that can not only perform human-like tasks but also surpass human capabilities in areas such as speed, accuracy, and capacity.

2.2.1 Machine Learning

Machine Learning (ML) is a subset of artificial intelligence that focuses on developing algorithms that can learn from data and improve their performance over time. It is used to teach machines how to handle the data more efficiently [16] and make predictions or decisions based on that learning. A framework with many parameters is first built, and then the prepared data is fed into the model. The parameters are continuously adjusted until they match or are close to the correct result [17].

There are three common machine learning algorithms, which are supervised learning, unsupervised and reinforcement learning. Supervised learning is a Machine Learning technique that involves training a model using a

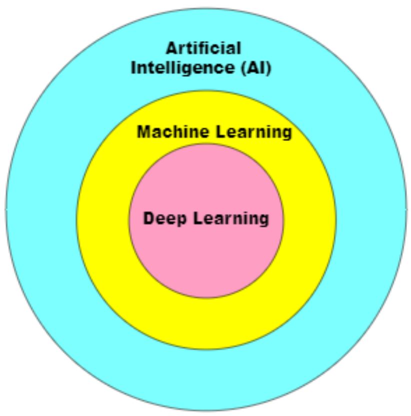

Figure 1. The relationship between AI, ML, and DL

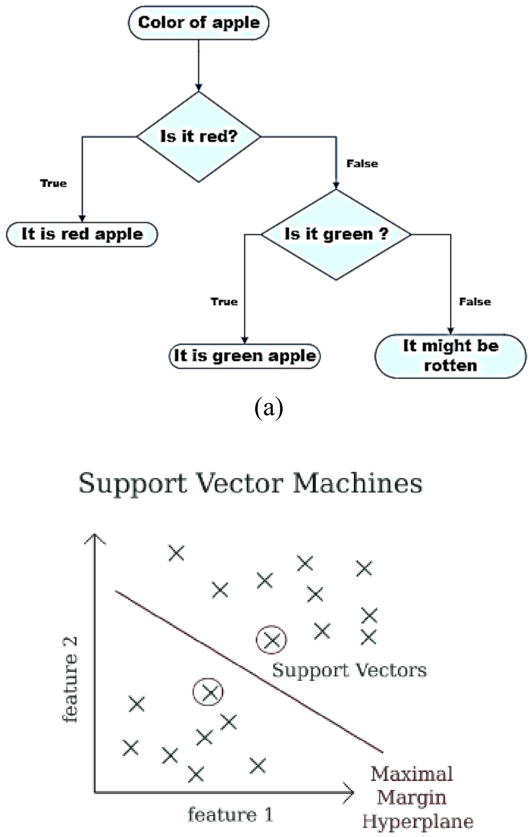

Figure 2. Supervise learning model (a) decision tree; (b) SVM

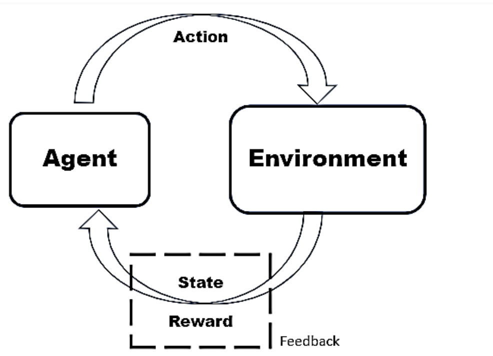

Figure 3. Framework for reinforcement learning

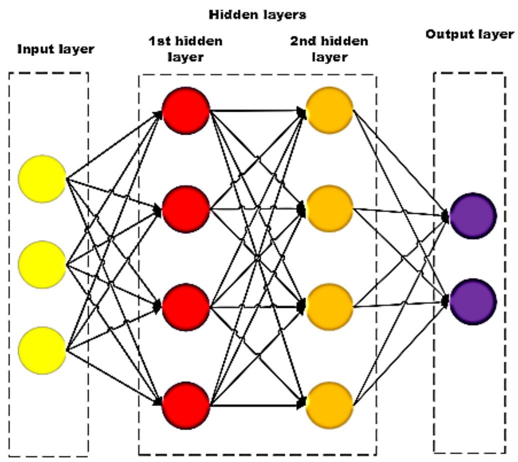

Figure 4. Common architecture of neural networks

labeled dataset, where each example is assigned a corresponding output value. This enables the model to learn from the labeled examples and make predictions on new, unseen data. Supervised learning is commonly used in applications where historical data predicts likely future events [13]. Some common examples of supervised learning are decision tree and Support Vector Machine (SVM).

The decision tree is an algorithm used in machine learning for addressing classification and regression problems, displaying a visual illustration of potential outcomes from a decision based on specific conditions. It is a tree-like graph which each of them consists of nodes and branches. Each node represents attributes in a group that is to be classified and each branch represents a value that the node can take [16]. The accuracy of a tree model is significantly influenced by its complexity, which can be managed by adjusting the stopping criteria and selecting an appropriate pruning method [18].

SVM is also a common technique used in classification and regression. In classification problems, nonlinear kernel functions are frequently utilized to convert input data to a high-dimensional feature space. Maximum-margin hyperplanes are then established. The model thus produced depends on only a subset of the training data near the class boundaries [19]. Figure 2 show an example of supervised learning model.

In contrast, unsupervised learning does not involve labeled data. Instead, the model must find patterns or structure in the data on its own. The k-means clustering algorithm is the common method in this learning. This clustering model is either based on centroids or distances, and it determines the assignment of data points to each cluster by measuring their distances. Initially, a value of K is selected, and the data is partitioned into K categories to enable better differentiation between data points within the same cluster. The main goal of the K-Means algorithm is to minimize the sum of the distances between the points and their respective cluster centroids, and cluster them in an iterative manner [20]. It is best suited for data mining because of its efficiency in processing large datasets [21].

Reinforcement learning is another type of Machine Learning technique where the model learns by interacting with an environment and receiving feedback in the form of rewards or penalties. Figure 3 shows the basic framework for this algorithm. Reinforcement Learning problems are related to learning which is the best action to perform, situation-by-situation, to maximize the aggregated reward [22].

2.2.2 Deep Learning

Deep Learning (DL) is considered a subset of ML and AI, and thus DL can be seen as an AI function that mimics the human brain's processing of data [23]. It differs from the machine learning model in terms of complexity and data

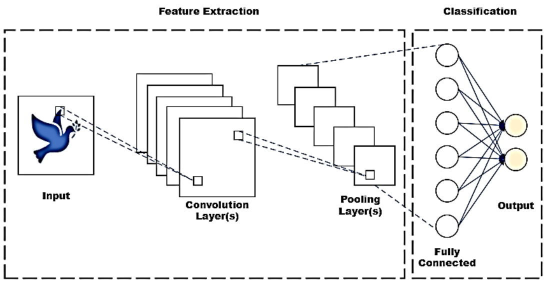

Figure 5. CNN model architecture

training ability. Deep learning operates based on the concept of extracting features from raw data using a multi-layered approach to identify various aspects that are pertinent to the input data. Its main objective is to train artificial neural networks to detect patterns within data. These neural networks are composed of multiple layers of interconnected nodes, which enable the network to learn more complex features and representations of the data. The nodes of neural networks are stacked next to each other in three layers, which are the input layer, hidden layers and output layer as shown in Figure 4. DL is able to achieve higher power and flexibility due to its ability to process a large number of features when it deals with unstructured data [24], [25].

The term "deep" in deep learning refers to the depth of the neural network, which can have many layers, sometimes up to hundreds or even thousands. The more layers there are, the more complex the features that the network can learn. Among all types of DL techniques, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and YOLO (You Only Look Once) are the most popular to use in image classification and object detection.

CNN techniques are based on the idea of convolution, which involves applying a filter or kernel to an image to extract features or patterns. A traditional convolutional neural network is made up of single or multiple blocks of convolution and pooling layers, followed by one or multiple fully connected (FC) layers, and an output layer [26], [27] as shown in Figure 5. The function of convolutional layers is to apply filters to the image and identify specific features, while the pooling layers are responsible for reducing the output size by down sampling the feature maps. The most commonly used pooling technique is Max Pooling. Fully connected layers are utilized for classification or regression purposes. CNN model is mainly used in image processing applications. The most common CNN architectures are ZFNet, GoogLeNet, VGGNet, AlexNet, and ResNet [28].

RNN stands for Recurrent Neural Network, which is a type of deep learning algorithm that is commonly used in processing natural language, speech recognition, and other sequential data analysis tasks. RNNs are a class of supervised machine-learning models, made of artificial neurons with one or more feedback loops [29]. A basic RNN consists of three layers, which are input, recurrent hidden, and output layers. An RNN is essentially a neural network that has been expanded across time by linking edges to the subsequent time step instead of the subsequent layer in the same time step [28]. The hidden units of the RNN preserve the previous input data in a state vector, which is used to calculate the outputs. The nodes in different layers of the neural network are compressed to form a single layer of recurrent neural networks [30]. Figure 6 shows an architecture of a simple RNN. Some popular architectures of RNN include Bidirectional RNN, Long-Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) [29].

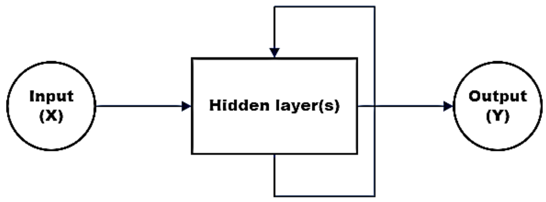

Figure 6. Simple RNN architecture

YOLO (You Only Look Once) is commonly used for real-time object detection due to its faster detection speed. This algorithm uses a single network for both the classification and localization of objects in an image [31]. It is simple to construct and can train directly on full images [32]. YOLO divides the input image into a grid and predicts the objectness score and bounding box coordinates for each grid cell. This allows YOLO to detect multiple objects in a single forward pass, making it extremely fast and efficient. Yet, it has limitations in terms of accuracy because it cannot detect small objects [33]. Currently, the YOLOv7, published in 2022 proved to be the fastest and most accurate real-time object detection model for computer vision tasks [34] compare to others YOLO versions. It used Extended Efficient Layer Aggregation Network (E-ELAN) as its backbone which enables the framework to learn better [34], [35].

3. CRITICAL ANALYSIS OF TECHNOLOGIES EMPLOYED FOR MICROAALGAE DETECTION

In this section, a thorough review and discussion of various technologies employed by researchers for microalgae detection will be conducted. The method used by the researchers and the respective results obtained will be discussed. Additionally, a critical analysis of the technologies used to identify their strengths, weaknesses, and limitations will be done. The purpose is to provide valuable insights into the latest advancements in microalgae detection technologies, as well as the potential for further improvements in this field. Ultimately, the information gathered from this section will aid in the development of more effective and efficient microalgae detection techniques, which can have significant implications for microalgae research in renewable energy.

3.1 Manual Observation Using Microscope

As microalgae are extremely small in size, they are not able to view using the humans' naked eyes. Consequently, to detect their presence, a conventional microscope is utilized as an optical instrument to facilitate human observation of these microorganisms. Abate Ayele et al. identifies the morphological structure of microalgae by using a light microscope, (model Labomed USA). From the experiment result, they were able to classify the 12 microalgae genera into two common groups, which are prokaryotes and eukaryotes. Out of the twelve identified genera, 8 belong to the category of eukaryotic protist microalgae, while the remaining four falls under prokaryotic cyanobacteria, with three of them being filamentous algae. Several of the identified isolates hold industrial significance, including Chlorella, Anabaena, Diatoms, and Scenedesmus [36].

The review paper produced by Fuguo Liu et al. stated that the microscopic examination technique is a traditional way that is commonly used to identify the microalgae species which cause Harmful Algae Bloom (HAB) [37]. Light microscopy can be utilized to do species identification among the various type of HAB [38]. Besides, this technique was also used by Rahmadi Tambaru et al. to investigate the quality of phytoplankton to sustain the health of seafood. The observation of 16 genera from the Class Bacillariophyceae, one genus from the Class Cyanophyceae, and five genera from the Class Dinophyceae were successfully conducted using a binocular microscope model, specifically the Olympus CX21. These genera were accurately identified and classified through microscopic examination [39].

The use of the microscopy method for identifying microalgae based on their morphological features is limited in its effectiveness. This method is suitable only for the determination of small numbers of microalgae under specific conditions, for example, microalgae species with totally different sizes and structures. However, it is essential to note that microscopic examination is a time-consuming and labor-intensive process that requires specialized training and expertise [40],[41],[42]. The accuracy of microalgae identification is heavily relied on the knowledge and experience of trained professionals, emphasizing the need for expertise in this field. Given the vast diversity of microalgae species across the globe, it is impractical for an individual to possess comprehensive knowledge of every species. Slight variations in the morphology of the same microalgae species can occur due to various factors such as environmental changes, genetic mutations, and other factors. These changes can affect the accuracy of results obtained through morphological identification. Furthermore, it is important to consider the cost-effectiveness of different microscopy methods for microalgae detection in renewable energy applications. One key factor to consider is the resolution of the microscope, as higher resolutions typically come with higher costs. In the specific context of microalgae detection, manual microscopy may not be the most suitable option due to its limitations and higher potential for error. Therefore, it is crucial to carefully evaluate and select the most appropriate light microscope model based on its capabilities, accuracy, and cost-effectiveness.

3.2 Machine Learning for Microalgae Detection and Classification

The process of machine learning for microalgae classification and detection typically involves several stages, beginning with data collection. Researchers will gather a large quantity of microalgae datasets either from microscopic observations or online data repositories. An online dataset can use to prove the accuracy of the model however it is advised to use primary data instead of online data when training the model because sometime the real microalgae input image might have slightly different in morphological structure, size and color due to different environments. The size of the dataset require is based on the model complexity. Some of the models will need a longer time to process the datasets, especially for color images. Next, the collected data must undergo preprocessing to eliminate unwanted noise or background from the images. The presence of impurities in the data can negatively affect the accuracy of the trained model.

Following data preprocessing, the most relevant features are selected from the dataset. The unique features of the microalgae sample in terms of physical characteristics such as color, shape, and size, should be able to extract by the model. Comparison and modification of available machine learning models can be done to produce an appropriate machine learning algorithm based on the research objectives. Testing and validation of the potential models should be done using the same dataset. After the training of the model, evaluation and optimization is then performed to further enhance the accuracy of the results.

Promdaen et al. developed a method to accurately segment and compute texture descriptors of 12 types of microalgae from image backgrounds by categorizing them into two shape groups and using single and multi-resolution edge detection [43]. The authors applied a feature combination approach to handle the variation in algae shapes within the same genus. The proposed method averaged texture descriptors extracted from an input image with different levels of edge enhancement, yielding better classification accuracy. They used the Sequential Minimal Optimization (SMO) algorithm to train a support vector classifier for image classification and achieved a high accuracy of 97.22%. The authors of the study faced several challenges in their research, including the fact that some microalgae from different genera have almost identical shapes which made it difficult to differentiate. As a result, the authors had to rely on texture features to distinguish between microalgae. Additionally, some algae in the microscopic images did not have clear boundaries, which made it difficult to accurately segment them.

A generalized segmentation algorithm (GSA) combining a convolution filter (Kirsch) and pixel clustering algorithm (Otsu) is proposed by Anaahat Dhindsa et al. to accurately extract microorganisms. 25 features are identified and mutual information-based models are used for feature selection [44]. After evaluating five ML algorithms, which are multi-layer perceptron (MLP), K-Nearest Neighbors (KNN), Quadratic Discriminant Analysis (QDA), Logistic Regression (LR) and Support Vector Machine (SVM), the authors select SVM as their ML model due to its good performance compare to others model. Despite that, to further improvisation the SVM model, an improvised SVM (ISVM) model was proposed by modifying its radial basis function (rbf) Kernel, in which first inter quartile range

(IQR) is computed for each feature, and then the rbf equation is applied. The rbf kernel was chosen because it is most frequently used by contemporary researchers [45], [46]. This results in better accuracy (98.2%), precision, recall, and F1 score than the traditional SVM, achieving 2% higher accuracy as compared to the SVM radial. The advantage of ISVM is that it removes outliers and extreme values by computing the difference between the first and third quartiles.

Zhan Peng Xu et al. introduced a transmission hyperspectral microscopic imager (THMI) that uses machine learning algorithms to detect microalgae through hyperspectral analysis [40]. The researchers utilized Principal Component Analysis (PCA) and peak ratio algorithms for feature extraction and dimensionality reduction of transmission spectra. The classification was done using a SVM model. The study compared the classification outcomes of two methods: PCA-SVM and peak ratio-SVM. The results showed that the two methods were nearly identical, achieving an average accuracy, sensitivity, and specificity of 94.4%, 94.4%, and 97.2%, respectively. The THMI system has a significant limitation in that it is very complex to build.

The accuracy of Linear Discrimination Analysis (LDA), linear and non-linear SVM are compared to classify polarized light scattering data of 35 categories of marine microalgae by Zepeng Zhuo et al. in their research. The result shows that non-linear SVM achieves higher accuracy, more than 80% as compared to others [47].

Hui Huang et al. introduced a micro-hyperspectral imaging (MHSI) approach to differentiate between spherical engineered microplastics (polyethylene) and microalgae (*Isochrysis galbana*) [48]. The VNIR MHSI system (Hyperspec VNIR-A, Headwall Photonics Inc, USA, spectral resolution 3 nm) and a microscope (OLYMPUS, magnification 40×) were used to measure the MHSI of the samples. The researchers employed several classifiers, including Support Vector Machine with Radial Basis Function (SVM-RBF), Least Squares Support Vector Machine, k-Nearest Neighbors, and others, which the performance was then compared with the MHSI system. However, the findings suggest that SVM (RBF) is the most effective classifier for identifying microplastic and microalgae, with a recall and precision of over 0.86. The lower accuracy of the result decreased due to the model not being able to detect the smaller size of microalgae and the imbalanced data set used while training.

From the review, Support Vector Machine (SVM) is one of the most widely used machine learning models in microalgae detection and classification due to its ability to handle high-dimensional data and nonlinear relationships between features. However, it is recommended to thoroughly test and compare various models to determine the most effective approach for microalgae classification. By doing so, researchers can ensure that the selected model provides the most accurate and reliable results, and can help to advance the field of microalgae detection for renewable energy production. SVM has limitations in processing a large dataset which will increase the model complexity. Modification of the model can also be done to further improve the accuracy of the result. One possible approach to improve SVM's performance is by optimizing the kernel function [44], [49], [50]. The accuracy of results can also be tested by combining multiple machine-learning techniques. Besides, the clearer image should be used as input data for easier feature extraction.

3.3 Deep Learning for Microalgae Detection and Classification

Even though machine learning and deep learning are a subset of AI, deep learning have some differences compared to machine learning in term of data representation, feature extraction, and model complexity. Deep learning is more complex and requires more input data for training. Deep learning is able to extract features from raw, unstructured data making it more convenient to use in several industries. The following section summarized several deep learning models proposed and used by the researchers to perform the classification and detection of microalgae.

Two Artificial Neural Network (ANN) models were developed by P. Otálora et al. to classify two well-known species of microalgae, *Scenedesmus almeriensis* and *Chlorella vulgari*, using feature variables and images as inputs. FlowCAM device is used to capture images [51]–[53]. The study employed the AlexNet architecture, which consists of 25 convolutional layers and an input layer size of 227x227x3 for color images, and an output layer size of two. The training dataset comprised 80% of the images, with the remaining 20% being used for validation purposes and no testing set to reduce the required training time. MATLAB's "Deep Learning Toolbox" has been used for the training and validation of both ANNs models. The research discovered that the ANN model using color images as input attained higher accuracy levels compare to feature variables as input but necessitated additional time for training and classification due to the complexity of the data and the denser network [53]. Besides, the accuracy of the result is also affected by the clustering of unwanted particles in the sample.

Mesut Ersin Sonmez et al. conducted data augmentation on microalgae images captured by a Nikon Eclipse TS100 inverted microscope. These images were classified using seven different CNN models including AlexNet, ResNet18, MobileNet, ResNet50, GoogleNet, DenseNet, and Inceptionv3, with a maximum of five epochs, an initial learning rate of 0.001, a 0.1 learning rate drop factor, and learning rate drop period of 20. The mini-batch size was set at 32, and the Stochastic Gradient Descent (SGDM) optimizer was used. Results showed that AlexNet had the lowest classification accuracy of 98.81%, while Inceptionv3 was among the models with the highest accuracy of 99.66%. In the second method, the researchers used SVM to improve the classification accuracy of AlexNet. The deep features extracted from the AlexNet model were classified with SVM using four different kernel functions: Gaussian, Linear, Cubic, and Quadratic. The result showed that the accuracy of AlexNet, which was initially 98.81%, increased to 99.66% [54]. This approach requires knowledge or expertise in culturing the microalgae sample.

Iago Corrêa et al. proposed an eight-layer deep learning model consisting of five convolutional and three fully connected layers, with the first layer comprising 16 convolutions of size 7×7 [51]. They aimed to minimize image preprocessing steps and maintain the accuracy of microalgae classification results. Because of the low resolution of the FlowCAM images due to the small size of the microalgae samples (1um), 64 x 64-pixel images were used to maintain the aspect ratio. Data augmentation techniques were used to increase the amount of data available for model training, resulting in an approximately 17% improvement in accuracy. The model achieved a classification accuracy of 88.59%, indicating its success.

Peisheng Qian et al. has introduced a new deep-learning approach for detecting and classifying algae using a modified version of the Faster R-CNN model [55]. This modified framework includes an additional classification branch for multi-task learning, where the first branch predicts the genus of the algae, the second branch detects and localizes the algae, and the third branch predicts the class of the algae. To train this framework, a ResNet-50-based FPN network that was pre-trained on ImageNet was used, with a batch size of 32 and a gradient descent optimizer with a momentum of 0.9. The initial learning rate was set to 0.02, which was reduced by a factor of 10 at steps 6000 and 7000. Cross-entropy loss function was used for algal classification at both genus and class levels. The proposed framework achieved a mean Average Precision (mAP) of 74.64% and 81.17% for algal detection based on genera and classes, respectively. This model is unable to detect some microalgae that are transparent and overlap each other.

In addition to the commonly used CNN model, the YOLO model is also frequently employed for the classification and detection of microalgae. One major advantage of the YOLO model is its ability to perform high-speed detection while maintaining reasonable accuracy. The real-time analysis capability is one of the key features of YOLO. Despite having multiple available models, YOLO technology is continuously being improved to develop the most efficient model for object detection. However, the available YOLO model has certain limitations when it comes to detecting microalgae that are very small in size. Hence. to enhance the accuracy of the YOLO model, researchers have implemented various modifications such as incorporating machine learning or CNN models, as well as adjusting the backbone of the YOLO network. Effective image preprocessing methods are also crucial for maximizing the model's ability to identify relevant features for classification. The upcoming section will discuss the research work carried out by the researchers and the specific parameters used by them for training the model.

Abdullah et al. used YOLOv3, YOLOv4, and YOLOv5 to detect and classify harmful algae bloom (HAB) on a custom microscopic image dataset [56]. A deep-learning-based data augmentation technique called DC-GAN, which is an advanced version of the GAN model, increases the number of images in the dataset. The Adam optimizer was used to train YOLOv3 and YOLOv4 on 80% of the dataset with 100 epochs and a batch size of 32, while the stochastic gradient descent (SGD) optimizer was used to train YOLOv5 on the same percentage of the dataset with 100 epochs and a batch size of 16. The learning rate was set to 0.01 for all three models. YOLOv5 shows better classification results compared with others YOLO models with a mAP score of 90.1%. The model proposed by the authors outperformed the regular YOLO model due to the use of DC-GAN based generated data for training. However, one of the biggest disadvantages in this model is the lack of real environmental data.

Mengying Cao et al. replace the Darknet-53 backbone network with a lightweight network called MobileNet in order to reduce the model's parameter requirements [57]. They also present an improved Spatial Pyramid Pooling (SPP) technique for pooling and concatenating multi-scale region features, reducing position error during small object detection. The Complete Intersection over Union (CIoU) algorithm is used to optimize the YOLOv3 model's loss function by taking into account the overlap area of the bounding box, central point distance, and aspect ratio. When the authors compared the improved YOLOv3 to the standard YOLOv3, they discovered that the improved model achieved an average accuracy of 98.90% and a detection efficiency that was 8.59% higher than the original model. Besides, it can detect small microalgae accurately as compared to other models.

The YOLOv3 algae image detection model was developed using a total of 1,114 microscope-generated images of 30 genera of algae by Jungsu Park et al [8]. Their study aimed to compare the performance of YOLOv3 with different input image types, namely, color and grayscale. The algae images were divided into four groups with 5, 10, 20, and 30 genera for training and testing the model. Darknet-53 was used as the primary YOLOv3 network, with a batch size of 64 and a learning rate of 0.001. After comparing the mAP and precision, the authors concluded that grayscale images are suitable for detecting a small number of genera, while color images are more appropriate for detecting a large number of genera, as they provide more useful information for microalgae detection. This model may likely encounter classification errors because of the presence of overlapping microalgae images, as well as the similarities in the morphological structures of different microalgae species, particularly in the case of small and densely packed microalgae.

Jesús Salido et al. developed an automated microscope that integrates an algorithm for stage and focuses control, image acquisition, and diatom detection and classification [58]. To detect and classify microalgae, they used a combination of YOLO and AlexNet, where YOLO perform live detection of microalgae and AlexNet classified it. AlexNet was pre-trained with ImageNet data, and a learning rate of 0.001 was used, decreasing by a factor of 0.1 every eight epochs. The SGD optimizer with L2-regularization of 0.004 was selected to avoid overfitting. The microalgae identification using YOLO achieved a maximum precision of 86%, while microalgae classification using AlexNet achieved an accuracy of 99.51% for the combination of the normalized and original datasets.

Based on the literature review, it has been found that many researchers collect microalgae samples and culture them prior to conducting classification and detection of the microalgae. Microalgae cultivation requires specialized knowledge and expertise as each type of microalgae has its own specific living requirements. The culture medium used must provide the necessary conditions for the microalgae to thrive. Furthermore, it is important to note that results obtained from cultured microalgae samples may not necessarily reflect the conditions found in their natural environment during detection. The presence of unwanted impurities or particles may affect the detection result. Besides, classification errors will occur due to the behavior of some microalgae that have a high potential to overlap each other.

In the context of microalgae classification, it is suggested that the use of CNN may be more suitable due to its demonstrated ability to achieve high levels of accuracy. YOLO is a suitable option for real-time microalgae analysis and detection due to its high speed, although its accuracy may be less stable compared to other models like CNNs. In real-time microalgae monitoring, YOLO faced more challenges due to the microalgae characteristic and impurities in the sample. The YOLO model has limitations in detecting the small-size microalgae.

The utilization of FlowCAM for capturing microalgae images can be comparatively costly [59] and may result in images with reduced clarity when dealing with smaller microalgae sizes. Low-resolution cameras can also result in blurry training datasets, further impacting the accuracy of the classification model. It is important to note that increasing the number of training epochs does not necessarily lead to improved classification outcomes, as overfitting might be occurred. Additionally, the accuracy of the model can be influenced by the size and behavior of the microalgae, such as their tendency to cluster together, as well as the morphological similarities between some microalgae species. Combining machine learning and deep learning models can improve the accuracy of the classification and identification process. The use of YOLO combined with CNN models has also shown promise in further enhancing the accuracy of the result.

Table 1 presents a summary of the author's perspectives on the limitations of manual microscopy and its associated findings. In contrast, Tables 2 and 3 provide a summary of the use of AI technologies for microalgae detection and classification, which include machine learning and deep learning techniques.

Table 1. Usage of microscopy in microalgae detection and classification with its limitations

Year	Application	Microscope model	Result	Reference
2022	Harmful Algae Bloom detection	-	Able to obtained some unique morphological microalgae cells accurately using light microscope for further analysis	[37], [38]
2021	Quality of phytoplankton to sustain the health of seafood	Binocular microscope model- Olympus CX21.		[39]
2019	Isolate and identify the potential native microalgae for local application	Labomed US		[36]
Limitations: 1. Manual microscopy requires an expert to identify some specific microalgae. 2. The accuracy of the result is highly affected by the expertise and experience of the researchers. 3. Time-consuming 4. To obtain a good result, a high-performance light microscope is needed which is costly. 5. Only suitable to investigate a small number of morphologically different microalgae species.				

Table 2. Machine Learning techniques for microalgae detection and classification with their limitation

Year	Method	Result	Limitation	Reference
2022	Compare the performance of LDA and SVM in microalgae classification.	Non-linear SVM achieves higher accuracy, 80%.	The measurement of polarized light scattering of individual microalgae may not be applicable for each condition and time-consuming.	[47]
2021	Generalized segmentation mechanism (Kirsch + Otsu) Proposed ISVM model which modifies the rbf kernel of SVM with IQR	2% higher performance compare to the original SVM radial.	The proposed segmentation mechanism may not work well for images with low contrast or high noise levels.	[44]

2021	Micro-hyperspectral imaging (MHSI) Compare ML classifier	SVM RBF achieves recall and precision of over 0.86.	Microplastic and microalgae have too many shapes and size. The smaller size of microalgae and imbalance data set cause decease in accuracy. Expensive hardware is needed.	[48]
2020	Introduce transmission hyperspectral microscopic imager (THMI) Compare PCA-SVM and peak ratio-SVM	PCA-SVM and peak ratio SSVM achieve the almost same result of 94.4% accuracy.	Complex hardware. Only 3 microalgae species are used to train the model so the accuracy might differ when handling more data.	[40]
2014	Feature combination approach uses to handle different algae shape Proposed single-resolution edge detection for segmenting images Use the Sequential Minimal Optimization (SMO) algorithm to train a support vector	Accuracy of 97.22%. Able to get clear microalgae image boundary and minimize noise.	Fewer microalgae image dataset is used which might affect the accuracy of the result. The paper does not compare the proposed method with other existing methods for microalgae recognition.	[43]

Table 3. Limitations of Deep Learning techniques for microalgae detection and classification

Year	Model use	Result	Limitations	Reference
2022	YOLOs with DC-GAN	The mAP score of YOLOv5 is the highest, 90.1%.	Lack of real environment microalgae data set.	[56]
2022	AlexNet and SVM	Improve the classification accuracy of AlexNet from 98.81% to 99.66%.	Need knowledge in culturing the microalgae sample. Limited microalgae training data.	[54]
2021	YOLOv3 with SPP and CIoU algorithm	Improved YOLOv3 model achieve 98.90% accuracy. Fast detection speed.	The performance of the model may vary on other datasets	[57]
2021	ANN model with AlexNet architecture	ANN model using color images as input attained higher accuracy levels compare to feature variables.	Longer training time for color image Clustering of unwanted particles affects the accuracy	[53]
2021	YOLOv3	Greyscale image is suitable for a small number of genera whereas the color image is for a large number of genera.	Classification error due to overlap images, same morphological structure of microalgae, small and crowded microalgae	[8]
2020	Modified Faster R-CNN	Average mAP of 74.64% and 81.17% for algal detection based on genera and classes.	Not able to detect transparent algae, overlapping algae and algae with almost similar structure	[55]
2020	YOLO and AlexNet	YOLO achieved a maximum precision of 86%, while microalgae classification using AlexNet achieved an accuracy of 99.51%.	Complex hardware Using online datasets for model training which sometimes does not reflect a real condition	[58]

2017	8-layer deep learning model consisting of 5 convolutional and 3 fully connected layers	Achieve 88.59 % accuracy with fewer image preprocessing steps needed. Fully automatic system No need for expert	Less resolution input image from FlowCAM	[51]

4. FUTURE PERSPECTIVE

As a microalgae researcher, it is important to acknowledge that there are still a number of challenges associated with identifying microalgae. Given the small size of these organisms, there is a risk of detection errors arising from their behavior under a microscope. Some microalgae species may overlap, change color or shape due to environmental factors and transparent will lead to detection error. While deep learning is the most advanced technology available for microalgae detection and classification, there is still room for improvement. To obtain a deep learning model that is both accurate and efficient, researchers can explore the option of generating a lighter-weight model. Having a model that can achieve high accuracy with limited data and also has a high detection speed would be more convenient. Sometimes it is almost impossible to obtain a large scale of raw data for model training purposes. However, less amount of training data will lead to the decrease in accuracy. Thus, achieving high classification accuracy with a small amount of data is only possible if the training data is both clear and highly representative of the problem domain. Additionally, the proposed model architecture must be capable of capturing relevant microalgae features from the input data during training.

Researchers are advised to generate their own microalgae datasets from the cultivation tank or microalgae sample to make sure that the training dataset reflects the real morphological structure in that particular environment. While secondary data may be useful in verifying model accuracy, primary data sources are typically preferred.

Image preprocessing is an essential step in analyzing images of microalgae. The decision to convert the image to greyscale or remain in color is based on the dataset and the characteristics of the microalgae being studied. If the microalgae are all green in color, a color image may introduce noise and affect the accuracy of the detection algorithm. Therefore, converting the image to greyscale can reduce the likelihood of detection errors and improve the accuracy of the algorithm. Greyscale images contain only shades of gray, ranging from black to white, and can reduce the complexity of the image, making it easier to detect and analyze features. On the other hand, color images can contain more information, but can also introduce unnecessary complexity that can negatively impact the analysis.

Granted that technology can aid in detecting microalgae, there is no necessity to rely on a specific approach. As a researcher, it is important to consider the various technological approaches available for detecting microalgae, and to carefully select the most suitable model for a given purpose. The choice of model will depend on the specific needs of the application, such as whether high precision or rapid detection capabilities are required. Employing a suitable detection model can significantly enhance the effectiveness of identifying the most suitable microalgae for energy production purposes.

5. CONCLUSION

This paper presents an overview and analysis of recent techniques used for detecting and classifying microalgae. The literature review identified three common methods for microalgae detection and classification, namely manual microscopy, machine learning, and deep learning. However, manual microscopy techniques are not very effective as they require skilled experts and are time-consuming. The study suggests that SVM is the most effective machine learning classification model for microalgae classification compared to other algorithms. Although deep learning is considered more advanced than machine learning and can perform more complex classification tasks, it requires more training data. While normal CNN models can achieve high classification accuracy, they require more computation time. In contrast, YOLO is a faster detection model but with lower accuracy. It is essential to note that relying solely on manual classification methods can lead to classification errors, which can be reduced by using automatic classification models such as machine learning or deep learning. By implementing these models, researchers can improve the accuracy and reliability of microalgae detection for renewable energy production.

ACKNOWLEDGMENT

This work has been supported by Universiti Teknologi Malaysia through Hi-Tech (F4) Q.J130000.4623.00Q15 grant in the project F4 3.3: Edge Artificial Intelligence for Suspicious Human Activity Recognition.

REFERENCES

[1] A. Z. A. Saifullah, A. Karim, and A. Ahmad-yazid, "Microalgae: An Alternative Source of Renewable Energy," *American Journal of Engineering Research*, vol. 03, no. 03, pp. 330–338, 2014.

[2] Y. M. Sani, W. M. A. W. Daud, and A. R. Abdul Aziz, "Solid acid-catalyzed biodiesel production from microalgal oil - The dual advantage," *J Environ Chem Eng*, vol. 1, no. 3, pp. 113–121, 2013, doi: 10.1016/j.jece.2013.04.006.

[3] F. Aksoy, E. Koru, and M. Alparslan, “Microalgae for Renewable Energy: Biodiesel Production and other Practies,” pp. 167–174, 2014.
[4] M. S. Ghayal and M. T. Pandya, “Microalgae biomass: A renewable source of energy,” *Energy Procedia*, vol. 32, pp. 242–250, 2013, doi: 10.1016/j.egypro.2013.05.031.
[5] Philip T. Pienkos; Al Darzins, “The promise and challenges of microalgal-derived biofuels,” *Biofuels, Bioproducts and Biorefining*, vol. 3, pp. 431–440, 2009, doi: 10.1002/BBB.
[6] S. Hemaiswarya, R. Raja, R. Ravikumar, and I. S. Carvalho, “Microalgae taxonomy and breeding,” *Biofuel Crops: Production, Physiology and Genetics*, no. June, pp. 44–53, 2013, doi: 10.1079/9781845938857.0044.
[7] M. R. Djoru, R. Gimin, and Suwari, “Phytoplankton (microalgae) as an alternative of renewable energy sources,” *IOP Conf Ser Mater Sci Eng*, vol. 823, no. 1, 2020, doi: 10.1088/1757-899X/823/1/012019.
[8] J. Park, “Microalgae Detection Using a Deep Learning Object Detection Algorithm, YOLOv3,” *Journal of Korean Society on Water Environment*, vol. 37, no. 4, p. 2021, 2021, [Online]. Available: https://doi.org/10.15681/KSWE.2021.37.4.275
[9] M. D. Guiry, “How many species of algae are there?” *J Phycol*, vol. 48, no. 5, pp. 1057–1063, 2012, doi: 10.1111/j.1529-8817.2012.01222.x.
[10] S. Mannam, “Artificial Intelligence, Machine Learning, and Deep Learning: Are They All the Same?” pp. 1–3, 2019.
[11] P. Ongsulee, “Artificial intelligence, machine learning and deep learning,” *International Conference on ICT and Knowledge Engineering*, pp. 1–6, 2018, doi: 10.1109/ICTKE.2017.8259629.
[12] H. Wehle, “Machine Learning, Deep Learning, and AI: What’s the Difference?” no. August, 2017.
[13] A. Habeeb, “Artificial intelligence,” *Research Gate*, vol. 7, no. 2, 2017, doi: 10.13140/RG.2.2.25350.88645/1.
[14] K. Aggarwal *et al.*, “Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning,” *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 1, pp. 115–123, 2022, doi: 10.52866/ijcsm.2022.01.01.013.
[15] Seema Singh, “Cousins of Artificial Intelligence,” no. Ml, pp. 1–13, 2018.
[16] B. Mahesh, “Machine Learning Algorithms - A Review,” *International Journal of Science and Research (IJSR)*, 2018, doi: 10.21275/ART20203995.
[17] H. Ning, R. Li, and T. Zhou, “Machine learning for microalgae detection and utilization,” *Front Mar Sci*, vol. 9, no. July, pp. 1–22, 2022, doi: 10.3389/fmars.2022.947394.
[18] Oded. okach, Lior; Maimon, “Decision Trees,” 2005, doi: 10.1007/0-387-25465-X_9.
[19] W. Lipo, “Support Vector Machines: Theory and Applications,” *Springer Science & Business Media*, vol. 177, pp. 1–47, 2005.
[20] M. Cui, “Introduction to the K-Means Clustering Algorithm Based on the Elbow Method,” pp. 5–8, 2020, doi: 10.23977/accaf.2020.010102.
[21] H. Hamdan Ali and L. Emad Kadhum, “K-Means Clustering Algorithm Applications in Data Mining and Pattern Recognition,” *International Journal of Science and Research (IJSR) ISSN*, vol. 6, no. 8, pp. 1577–1584, 2017, doi: 10.21275/ART20176024.
[22] M. Naeem, S. T. H. Rizvi, and A. Coronato, “A Gentle Introduction to Reinforcement Learning and its Application in Different Fields,” *IEEE Access*, vol. 8, pp. 209320–209344, 2020, doi: 10.1109/ACCESS.2020.3038605.
[23] I. H. Sarker, “Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions,” *SN Computer Science*, vol. 2, no. 6. Springer, Nov. 01, 2021. doi: 10.1007/s42979-021-00815-1.
[24] A. Mathew, P. Amudha, and S. Sivakumari, “Deep learning techniques: an overview,” in *Advances in Intelligent Systems and Computing*, Springer, 2021, pp. 599–608. doi: 10.1007/978-981-15-3383-9_54.
[25] R. K. Mishra, G. Y. S. Reddy, and H. Pathak, “The Understanding of Deep Learning: A Comprehensive Review,” *Mathematical Problems in Engineering*, vol. 2021. Hindawi Limited, 2021. doi: 10.1155/2021/5548884.
[26] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, and D. De, “Fundamental concepts of convolutional neural network,” in *Intelligent Systems Reference Library*, Springer, 2019, pp. 519–567. doi: 10.1007/978-3-030-32644-9_36.
[27] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00444-8.
[28] A. Mosavi, S. Ardabili, and A. R. Varkonyi-Koczy, “List of Deep Learning Models,” 2019, doi: 10.20944/preprints201908.0152.v1.
[29] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, “Recent Advances in Recurrent Neural Networks,” Dec. 2017, [Online]. Available: http://arxiv.org/abs/1801.01078
[30] A. Biswal, “Recurrent Neural Network (RNN) Tutorial: Types and Examples,” Feb. 14, 2023. https://www.simplilearn.com/tutorials/deep-learning-tutorial/rnn (accessed Apr. 03, 2023).
[31] R. Deepa, E. Tamilselvan, E. S. Abrar, and S. Sampath, “Comparison of Yolo, SSD, Faster RCNN for Real Time Tennis Ball Tracking for Action Decision Networks,” in *2019 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, IEEE Computer Society, Dec. 2019, pp. 1–4. doi: 10.1109/ICACCE46606.2019.9079965.
[32] F. Joiya, “OBJECT DETECTION: YOLO VS FASTER R-CNN,” *International Research Journal of Modernization in Engineering Technology and Science*, Sep. 2022, doi: 10.56726/irjmets30226.
[33] J. A. Kim, J. Y. Sung, and S. H. Park, “Comparison of Faster-RCNN, YOLO, and SSD for Real-Time Vehicle Type Recognition,” in *2020 IEEE International Conference on Consumer Electronics - Asia, ICCE-Asia 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/ICCE-Asia49877.2020.9277040.

[34] Kukil and S. Rath, "YOLOv7 Object Detection Paper Explanation & Inference," 2022. https://learnopencv.com/yolov7-object-detection-paper-explanation-and-inference/ (accessed Apr. 03, 2023).

[35] G. Boesch, "YOLOv7: The Most Powerful Object Detection Algorithm (2023 Guide)," 2023. https://viso.ai/deep-learning/yolov7-guide/ (accessed Apr. 03, 2023).

[36] A. Ayele, S. Benor, and A. Suresh, "Isolation and Morphological Identification of Some Indigenous Microalgae from Ethiopia for Phycoprospecting," *Ethiopian Journal of Science and Sustainable Development*, vol. 6, no. 2, p. 2019, 2019, doi: 10.20372/ejssdastu:v6.i2.2019.102.

[37] F. Liu, C. Zhang, Y. Wang, and G. Chen, "A review of the current and emerging detection methods of marine harmful microalgae," *Science of the Total Environment*, vol. 815, p. 152913, 2022, doi: 10.1016/j.scitotenv.2022.152913.

[38] J. J. Chin Chwan Chuong, M. Rahman, N. Ibrahim, L. Y. Heng, L. L. Tan, and A. Ahmad, "Harmful Microalgae Detection: Biosensors versus Some Conventional Methods," *Sensors*, vol. 22, no. 9, 2022, doi: 10.3390/s22093144.

[39] R. Tambaru, A. I. Burhanuddin, A. Massinai, and M. A. Amran, "Detection of marine microalgae (Phytoplankton) quality to support seafood health: A case study on the west coast of south sulawesi, indonesia," *Biodiversitas*, vol. 22, no. 11, pp. 5179–5186, 2021, doi: 10.13057/biodiv/d221156.

[40] Z. Xu, Y. Jiang, J. Ji, E. Forsberg, Y. Li, and S. He, "Classification, identification, and growth stage estimation of microalgae based on transmission hyperspectral microscopic imaging and machine learning," *Opt Express*, vol. 28, no. 21, p. 30686, 2020, doi: 10.1364/oe.406036.

[41] R. Pardeshi and P. D. Deshmukh, "Classification of Microscopic Algae: An Observational Study with AlexNet," *Advances in Intelligent Systems and Computing*, vol. 1118, no. June, pp. 309–316, 2019, doi: 10.1007/978-981-15-2475-2_29.

[42] P. Coltelli, L. Barsanti, V. Evangelista, A. M. Frassanito, and P. Gualtieri, "Water monitoring: Automated and real time identification and classification of algae using digital microscopy," *Environ Sci Process Impacts*, vol. 16, no. 11, pp. 2656–2665, 2014, doi: 10.1039/c4em00451e.

[43] S. Promdaen, P. Wattuya, and N. Sanevas, "Automated microalgae image classification," *Procedia Comput Sci*, vol. 29, pp. 1981–1992, 2014, doi: 10.1016/j.procs.2014.05.182.

[44] A. Dhindsa, S. Bhatia, S. Agrawal, and B. S. Sohi, "An Improvised Machine Learning Model Based on Mutual Information Feature Selection Approach for Microbes Classification," *Entropy*, vol. 23, no. 2, p. 257, 2021, doi: 10.3390/e23020257.

[45] M. A. A. Mosleh, H. Manssor, S. Malek, P. Milow, and A. Salleh, "A preliminary study on automated freshwater algae recognition and classification system.," *BMC Bioinformatics*, vol. 13 Suppl 17, 2012, doi: 10.1186/1471-2105-13-s17-s25.

[46] D. J. Armaghani, P. G. Asteris, B. Askarian, M. Hasanipanah, R. Tarinejad, and V. Van Huynh, "Examining hybrid and single SVM models with different kernels to predict rock brittleness," *Sustainability (Switzerland)*, vol. 12, no. 6, pp. 1–17, Mar. 2020, doi: 10.3390/su12062229.

[47] Z. Zhuo, H. Wang, R. Liao, and H. Ma, "Machine Learning Powered Microalgae Classification by Use of Polarized Light Scattering Data," *Applied Sciences (Switzerland)*, vol. 12, no. 7, Apr. 2022, doi: 10.3390/app12073422.

[48] H. Huang *et al.*, "The Identification of Spherical Engineered Microplastics and Microalgae by Micro-hyperspectral Imaging," *Bull Environ Contam Toxicol*, vol. 107, no. 4, pp. 764–769, 2021, doi: 10.1007/s00128-021-03131-9.

[49] K. Thurnhofer-Hemsi, E. López-Rubio, M. A. Molina-Cabello, and K. Najarian, "Radial basis function kernel optimization for Support Vector Machine classifiers," Jul. 2020, [Online]. Available: http://arxiv.org/abs/2007.08233

[50] S. Haochen, X. Haipeng, Z. Jianjiang, Ning Li, and Z. Huiyu, "Radial Basis Function Kernel Parameter Optimization Algorithm in Support Vector Machine Based on Segmented Dichotomy," in *The 2018 5th International Conference on Systems and Informatics (ICSAI 2018),* 2018.

[51] I. Correa, P. Drews, S. Botelho, M. S. De Souza, and V. M. Tavano, "Deep learning for microalgae classification," in *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, Institute of Electrical and Electronics Engineers Inc., 2017, pp. 20–25. doi: 10.1109/ICMLA.2017.0-183.

[52] L. Barsanti, L. Birindelli, and P. Gualtieri, "Water monitoring by means of digital microscopy identification and classification of microalgae," *Environ Sci Process Impacts*, vol. 23, no. 10, pp. 1443–1457, Oct. 2021, doi: 10.1039/d1em00258a.

[53] P. Otálora, J. L. Guzmán, F. G. Acién, M. Berenguel, and A. Reul, "Microalgae classification based on machine learning techniques," *Algal Res*, vol. 55, no. March 2021, doi: 10.1016/j.algal.2021.102256.

[54] M. E. Sonmez, N. Eczacıoglu, N. E. Gumuş, M. F. Aslan, K. Sabanci, and B. Aşikkutlu, "Convolutional neural network - Support vector machine based approach for classification of cyanobacteria and chlorophyta microalgae groups," *Algal Res*, vol. 61, Jan. 2022, doi: 10.1016/j.algal.2021.102568.

[55] P. Qian *et al.*, "Multi-Target Deep Learning for Algal Detection and Classification," May 2020, doi: 10.1109/EMBC44109.2020.9176204.

[56] S. Abdullah, Z. Ali, A. Khan, A. Hussain, A. Athar, and H. C. Kim, "Computer Vision Based Deep Learning Approach for the Detection and Classification of Algae Species Using Microscopic Images," Water, vol. 14, no. 14, pp. 2219, Jul. 2022. [Online].Available: https://doi.org/10.3390/w14142219.

[57] M. Cao, J. Wang, Y. Chen, and Y. Wang, "Detection of microalgae objects based on the Improved YOLOv3 model," *Environ Sci Process Impacts*, vol. 23, no. 10, pp. 1516–1530, Oct. 2021, doi: 10.1039/d1em00159k.

[58] J. Salido, C. Sánchez, J. Ruiz-Santaquiteria, G. Cristóbal, S. Blanco, and G. Bueno, “A low-cost automated digital microscopy platform for automatic identification of diatoms,” *Applied Sciences (Switzerland)*, vol. 10, no. 17, Sep. 2020, doi: 10.3390/app10176033.
[59] J. L. Deglint, C. Jin, and A. Wong, “Investigating the Automatic Classification of Algae Using Fusion of Spectral and Morphological Characteristics of Algae via Deep Residual Learning,” Oct. 2018, [Online]. Available: http://arxiv.org/abs/1810.10889.
[60] 1. C. Safi, B. Zebib, O. Merah, P. Y. Pontalier, and C. Vaca-Garcia, "Morphology, composition, production, processing and applications of Chlorella vulgaris: A review," Renewable and Sustainable Energy Reviews, vol. 121, p. 109676, Jan. 2020. [Online]. Available: https://doi.org/10.1016/j.rser.2019.109676.
[61] S. Hena, N. Fatihah, and N. H. M. Yasin, "Recent advances in microalgal biofuels production," Biotechnology for Biofuels, vol. 13, p. 144, Aug. 2020. [Online]. Available: https://doi.org/10.1186/s13068-020-01785-4.
[62] P. L. Show et al., "Microalgae biomolecules: Chemistry and bioactivity," Microorganisms, vol. 9, no. 3, p. 485, Mar. 2021. [Online]. Available: https://doi.org/10.3390/microorganisms9030485.
[63] Y. Wang et al., "Microalgal biorefineries: From concept towards reality," Green Chemistry, vol. 22, no. 9, pp. 2752-2770, May 2020. [Online]. Available: https://doi.org/10.1039/D0GC00182A.
[64] P. Geada et al., "Microalgae cultivation and biofuels production: A Portuguese perspective," Energies, vol. 14, no. 13, p. 4065, Jun. 2021. [Online]. Available: https://doi.org/10.3390/en14134065.