# Image Classification of Algal Species Applied Deep Learning Algorithms

Springer

May 2023 · Wireless Personal Communications

DOI: 10.1007/s11277-023-10488-z

Roojwan Scddeek Hawezi



## Abstract and figures

Classifying images is one of the most common methods adopted in the current processes and the current development of scientific experiments that provide for the topics of change and development in the field of classification and make them into groups, perhaps small or large, and according to the size of the data used and stored in memory, as is the case in our work. In this study, 2749 microscopic images of a type of algae were used to observe the nature of its shape, accuracy, and quality of similarity between them; it is considered one of the essential classifications and may need some accuracy in its classification, so it needs experts. The program Mat-lab was used to obtain good results in the analysis and collection of the data used in the data and the nature of the work. We have used a developed program that is currently popular in classification processes to get results quickly, which is Convolution Neural Network (CNN), which is part of machine learning through which can collect data and the images are easily and quickly the user, which is a supported language in network analyzer and using quantitative algorithms for neural network and alexnet. Finally, using CNN's deep learning method to show that there are changes in accuracy, up and down, in the data used may be a challenging area to know the causes and to delve into the results that may lead to a good analysis.

# Image Classification of Algal Species Applied Deep Learning Algorithms

Roojwan Scddeek Hawezi[1]

Accepted: 24 April 2023


**Abstract**
Classifying images is one of the most common methods adopted in the current processes and the current development of scientific experiments that provide for the topics of change and development in the field of classification and make them into groups, perhaps small or large, and according to the size of the data used and stored in memory, as is the case in our work. In this study, 2749 microscopic images of a type of algae were used to observe the nature of its shape, accuracy, and quality of similarity between them; it is considered one of the essential classifications and may need some accuracy in its classification, so it needs experts. The program Mat-lab was used to obtain good results in the analysis and collection of the data used in the data and the nature of the work. We have used a developed program that is currently popular in classification processes to get results quickly, which is Convolution Neural Network (CNN), which is part of machine learning through which can collect data and the images are easily and quickly the user, which is a supported language in network analyzer and using quantitative algorithms for neural network and alexnet. Finally, using CNN's deep learning method to show that there are changes in accuracy, up and down, in the data used may be a challenging area to know the causes and to delve into the results that may lead to a good analysis.

**Keywords** Deep learning · Image classification · Algae image · Convolution neural network

## 1 Introduction

Image classification is categorizing and classifying an image of pixels or vectors on specific rules. It is based on using computer vision to do multiple takes, such as segmentation of semantics and detection of semantics classification in other terms means how to make a classification between objects, that is to be the task for anyone; image classification has been a needed mission with computer and within the field of the vision of the computer, and it represents to the labeling of images into one of several predefined classes. It refers to

✉ Roojwan Scddeek Hawezi
roojwan.hawezi@epu.edu.iq

[1] Department of Information System Engineering, Erbil Technical Engineering College, Erbil Polytechnic University, Erbil, Iraq

Published online: 11 May 2023

the task of the information in the image classification classes [1]. Image classification has become one of the essential fields at present, and due to its usefulness in modifying images and obtaining the great result, and is used in practical experiments to illustrate the machine learning process; in addition to that, the role of image classification has become an essential part in digital images due to the importance in the digital processing, and the digital images are one of the essential parts of the technologies in the image and the classifying, it is an advanced technology that enables us to process the digital images using computers, and it has mainly based on images. Classifying an image is among the complex operations that require accuracy and are considered difficult due to differences, fragmentation, and points of similarity, and experts have been working on it. In addition, the classification process depends on the features provided by the user entering the image he needs [2, 3].

Digital processing allows the users to take the digital images as input and perform it to the different algorithms to generate an output, the types of images, greyscale images, and RGB images. AI (Artificial Intelligence) is one of the basic machine learning methods used for the classification of images because AI is one of the branches of computer science that pursues making computers or machines like human beings; it is accomplished by studying the human brain, to know how the human thinks, and acts like the human, and to solve the problem like in algorithms because the AI is one of the essential parts in the machine learning, and has many features for user interface and computer [4, 5].

Artificial intelligence is an innovative machine used in intelligent computing systems for algorithms. It works on its various tasks independently without needing intervention or giving it directly from the user. On the other hand, all machines automatically need human guidance to work steadily [6]. In addition, AI is one of the successful techniques of deep learning, which has successfully developed information processing, analysis, and classifications. Deep learning is one of the essential and developed fields in machine learning, and many successful results and experiences have been achieved in scientific experiments [7].

The machine learning algorithm is used for improving during solving a specific task by using experience [8]. A number of the tasks can be solved by using machine learning algorithms that fall under the tasks category of regression or category of classification. For example, the algorithm predicts several values given some input for regression problems. Moreover, the time and temperature of the day inside a city will predict the consumption of electric energy for the city for the next hour. In addition, machine learning algorithms can solve other tasks, including machine translation, transcription, missing values imputation, clustering, denoising, and ranking [9, 10]. Different categories can be found under the logarithm of machine learning. These kinds depend on the availability of labeling or not labeling the data and the methods of learning patterns used in the data.

Supervised and unsupervised image classification are the two main methods of classification [11]. Supervised classification involves "the classification of pixels of unknown identity by means of a classification algorithm using the spectral characteristics of pixels of known in informational class (referred to as training data) identified by the analyst" [12]. For example, an unsupervised classification can be used during the training of unlabeled data. In contrast, the goal is unlabeled data for finding good patterns by grouping similar instances together in a process called clustering [13].

Unsupervised classification involves "the separation of image pixels into natural groupings based upon similar spectral characteristics means of a classification algorithm and the result an assignment of those groupings to informational classes by the analyst" [14]. During supervised classification, the analyst has complete control of the informational classes or categories to be assigned in the final classification. This allows for easier comparison with other classifications by using identical classes for both. Moreover,

in Unsupervised classification, there is a lower chance of errors from humans because it is no need to make many decisions in the classification process [15].

Another kind of machine learning is reinforcement learning methods learned through interactions with the environment. Based on these interactions, the learner receives awards, aiming to maximize the total rewards [16]. Moreover, Semi-supervised learning uses a mixture of labeled and unlabeled data in training, typically in situations where unlabeled data are abundant, but labels are expensive to obtain [17].

Due to the various factors such as inter-class similarities and significant intra-class variations, Image classification is a challenging task. Therefore, deep learning is essential to deal with this kind of data. Deep learning could be defined as a subset of machine learning, essentially a neural network with three or more layers. These neural networks simulate the human brain's behavior—albeit far from matching its ability—which allows one to "learn" from large amounts of data. In contrast, deep learning can exploit Convolutional Neural Network (CNN) to automatically learn multi-level representations from raw data using multiple linear and non-linear layers. Deep CNN can discover complex data structures without handcrafting features, thus applicable to many domains[18]. (Fig. 1).

Our work goal, Classifying the microscopic alga photos according to their (shape and color) into four categories:

A- Green vegetative cells.
B- Red vegetative cells.
C- Green cyst.
D- Red cyst.

The total number of images are 2749 and separated into four classes. Figure 2 shows the dataset represented.

The paper's organization is as follows the second part is the literature review; the method and material are in section three, section four presents the result and discussion, and finally presents the conclusion and future work.

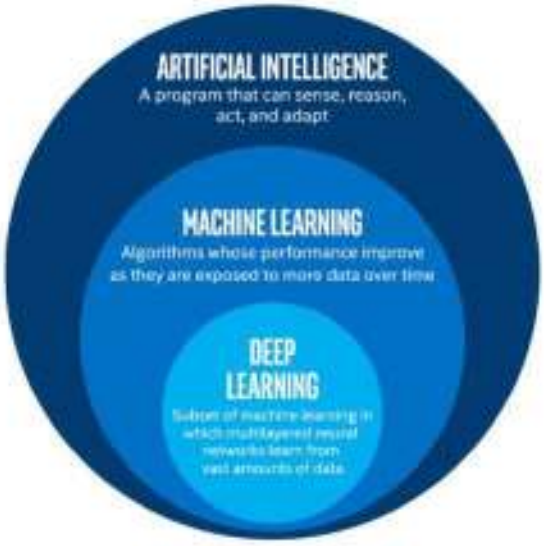



Fig. 1 Artificial intelligence, machine learning, deep learning

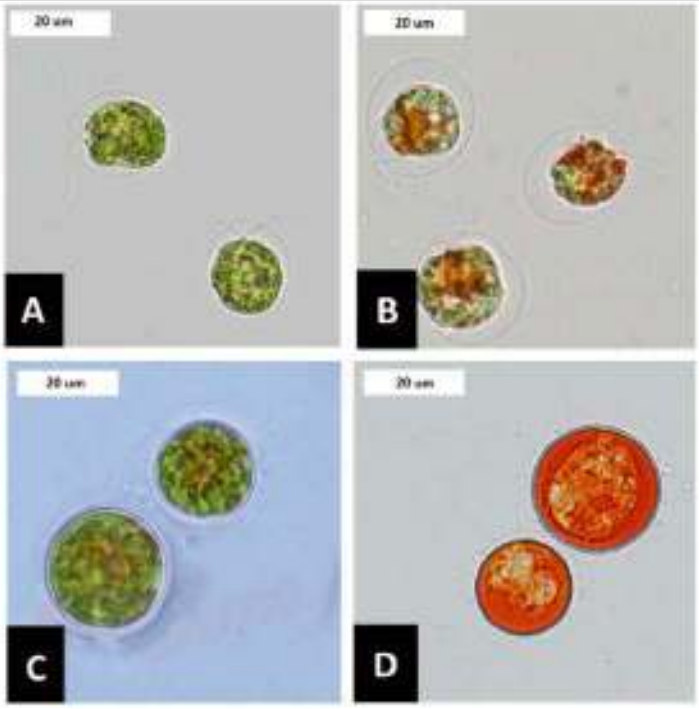



Fig. 2 Algae dataset categories

## 2 Literature Review

Research on freshwater species recognition is an integral part of the actions for the protection of the freshwater environment. It is also an under-exploited application area in the computer vision community. The water quality affects some freshwater species, such as phytoplankton species (Algae).

The Algae that used in this research is called Haematococcus Pluvialis is a type of green aquatic algae; it is a single green alga that grows in special water conditions and has many unique properties and these algae form a type of green mobile cell types under the influence of conditions and pressures in terms of growth and nutrition. For example, when the conditions become the water is terrible regarding growth and nutrition. Under the influence of high temperature, the cells become without flagella [8]; simultaneously, the algae turn from green to red and form sacks with a thick and heavy cell wall. The reason for changing this pigment (of algae) to red is the increased continuous absorption of the green pigment (chlorophyll) and the accumulation of the (astaxanthin) red pigment [8].

Algae is a basic form of aquatic life that depends on identifying and classifying the different forms to identify the quality of aquatic life. Each depends on the form of nutrition and heat you need. Algae are a form of various microscopic aquatic life. They are essential in many fields and are considered an actual product in many cases in an essential aquatic food chain and oxygen in an aquatic ecosystem. In managing water resources, we can use algae as a biological indicator, for example, to identify the types

**Table 1** Initializing input data normalization for (300 Epoch)

| Epoch | Iteration | Time Elapsed (hh:mm:ss) | Mini-batch Accuracy | Mini-batch Loss | Base Learning rate |
|---|---|---|---|---|---|
| 1 | 1 | 0:00:20 | 30.77% | 1.4116 | 1.00E-04 |
| 50 | 50 | 0:11:18 | 41.35% | 1.2861 | 1.00E-04 |
| 100 | 100 | 0:20:36 | 46.15% | 1.2406 | 1.00E-04 |
| 150 | 150 | 0:31:15 | 45.19% | 1.2205 | 1.00E-04 |
| 200 | 200 | 0:40:11 | 50.00% | 1.1884 | 1.00E-04 |
| 250 | 250 | 0:49:01 | 48.08% | 1.1508 | 1.00E-04 |
| 300 | 300 | 0:57:51 | 75.88% | 0.4425 | 1.00E-04 |

**Table 2** Initializing input data normalization for (20 Epoch)

| Epoch | Iteration | Time Elapsed (hh:mm:ss) | Mini-batch Accuracy | Mini-batch Loss | Base Learning rate |
|---|---|---|---|---|---|
| 1 | 1 | 00:00:00 | 26.56% | 8.1102 | 0.1000 |
| 20 | 20 | 00:00:02 | 100.00% | -0.0000e+00 | 0.1000 |

of changes that occur in the aquatic environment since algae are very sensitive to identifying changes that occur [19].

Therefore, algae are considered among the essential and delicate issues within the classification and identification of their types in the management of water resources, and thus require effort and time to obtain them and require more explanations and specialized experts for the process of achieving and obtaining them. Therefore, in this field, we can classify them and divide them using automated processes through which we can learn about the method of division and how to identify the species and the nature that live on them, so we will use an essential process in terms of images in order to obtain a classification of images [19].

Ohnuki [8], during his research about Haematococcus Pluvialis image classification, developed a monitoring System (HaematoCalMorph) for Green Algal. Five steps represent the process of the program. First, color information was pre-processed from bright-field microscopic images (Step 1). Second, detect the edges to recognize the boundary of segments (Step 2). distinguish the foreground from the background (Step 3). Third, foreground regions were further classified into the algal cell, colorless cell, and non-cell regions (Step 4), and finally, calculate the morphological parameters (Step 5).

In another search for microalgae identification [10], researchers used 3,423 pictures and processed them, and these pictures contain examples of 24 micro and major microalgae types. Of the processes, which are the spectrum of distance, absorption, color, and various measurements, these groups were processed by a standard for measurement in the classification of algae, and the results depended on the selection of features and accuracy of color adopted in the process of classification of algae.

During other research about microalgae [12], for the classification of microalgae research, the technique of deep learning was postponed to solve the problem of the classification to solve the problem and then obtain correct classification results for the data. Tables 1, 2, 3

Table 3 Deep learning network analyzer

| | Name | Type | Activations | Learnable |
|---|---|---|---|---|
| 1 | image input | Image input | 227 *227*1 | – |
| 2 | FC | Fully Connected | 1*1*4 | Weights 4*51,529 |
| 3 | softmax | softmax | 1*1*4 | – |
| 4 | Class output | Classification output | 1*1*4 | – |

## 3 Material and Methods

### 3.1 Background

Plankton [20] are different kinds of aquatic micro-organisms, they and be of different sizes and can not go against the current of water. They can be of different kinds, such as archaea, bacteria, protozoa, algae, and microscopic animals. Plankton communities could be divided according to the trophic into three groups: producers, decomposers, and consumers. Phytoplankton contains the producers who can do photosynthesis, while zooplankton has the consumers who eat other plankton. The phytoplankton and bacterioplankton contained the decomposers, which can break down the nutrients [21]. Many species of plankton can work in different and multiple trophic environments, which they called as mixotrophy [22].

### 3.2 Algae Dataset

In this research, we will use a processing technique to classify the algae in terms of type, color, accuracy, and clarity, algae are often very accurate, which makes them difficult to classify, so it requires an automated classification technique to identify the types of algae. In this research, the microscopic images were used to identify the type of green haematococcus algae, so (2749 images) of a unicellular flagellated green alga were selected for this study [17].

This algae forms flagellated green motile cells under favorable growth conditions. However, when the conditions become unfavorable, for example, under nutrient depletion, continuous light, or high temperature, the cells become spherical with no flagella and then turn from green to red to form cysts (resting cells) with a thick and heavy resistant cell wall [15].

Figure 3 shows the data set used in the training set. We classify our data into 4 classes, green vegetative cells (547 images), red vegetative cells (896 images), pictures of green cysts (652 images), and the red cyst (654 pictures).

### 3.3 Proposed Method

Deep learning (CNN) is one of the methods adopted in analyzing the data that will be used in this research to identify the classification methods for the approved image, which helps in understanding the advantages and can go into the depths to obtain a good classification process. The default architecture of CNN and two method of transfer learning (Mobile net, Alex net). The proposed method for CNN is shown in Fig. 4, which starts by collecting

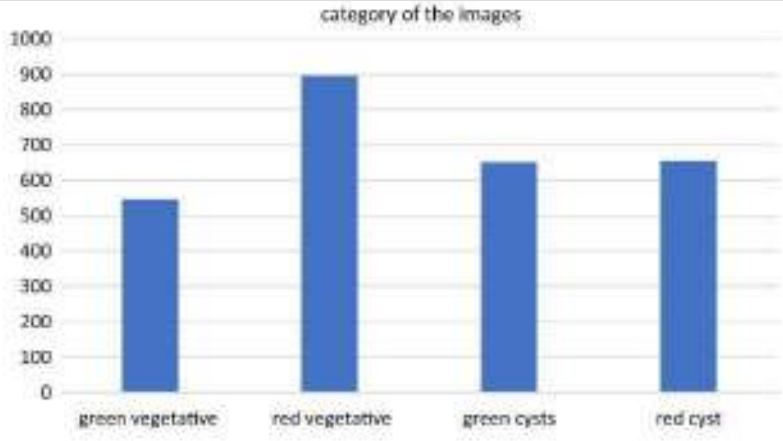



Fig. 3 Classes of the data

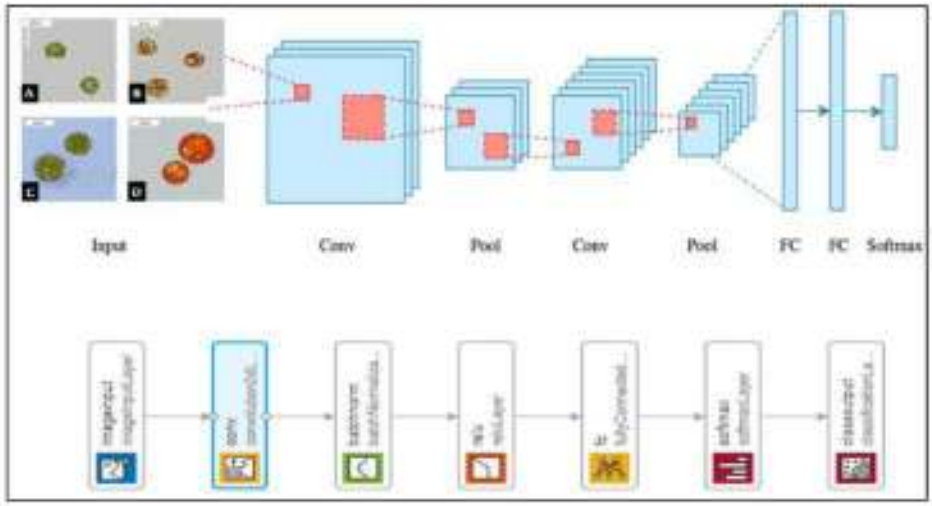



Fig. 4 CNN proposed methods

the datasets. A convolutional neural network (CNN) comprises multiple building blocks, such as convolution layers (image input layer), pooling layers, and fully connected layers. It is designed to automatically and adaptively learn spatial hierarchies of features through a backpropagation algorithm.

The proposed was applied with epochs with different numbers (100, 200, 500). It gains different results at each Iteration, Time Elapsed, and Mini-batch Accuracy in deep learning (network Analyzer). The method was applying the training of the initializing input data normalization and classifying all the training of the four classes of images ( red v. cells, green v. cells, red cyst, green cyst) and subplot to(subplot(5,5, i), and the result is shown in the (Fig. 5) the classification training: Five rows, five columns, and for subplot (5,2, i): 2 columns, 5 row. (Fig. 6)

The overall procedure shown in the Fig. 7, the procedure of the training and testing.

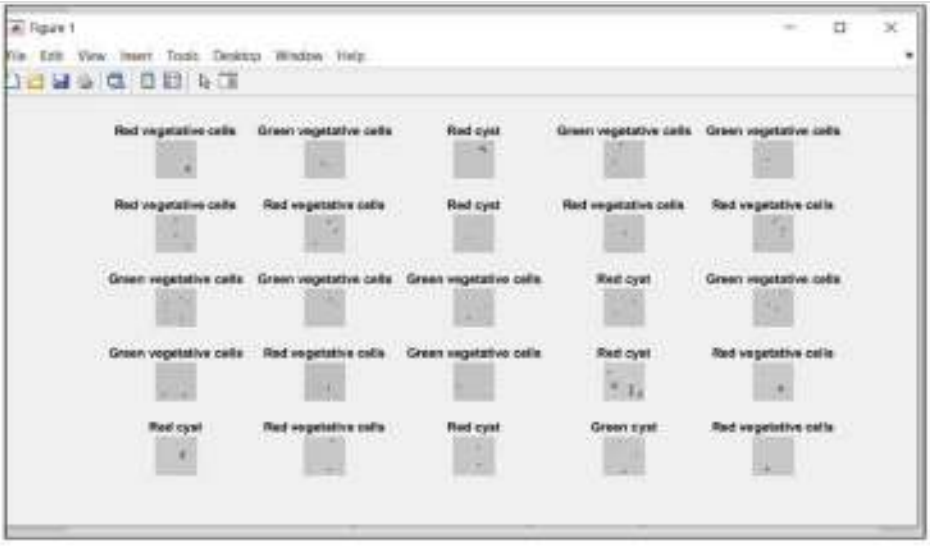



Fig. 5 The classification training(subplot (5,5,i))


Fig. 6 The classification training (subplot (5,2,i))

## 4 Results and Discussion

### 4.1 Results of CNN

The dataset was divided into two subgroups in this method: training and testing. The data were randomly distributed, with the training set consisting of 2199 images (80%) and the testing set consisting of 550 images (20%).

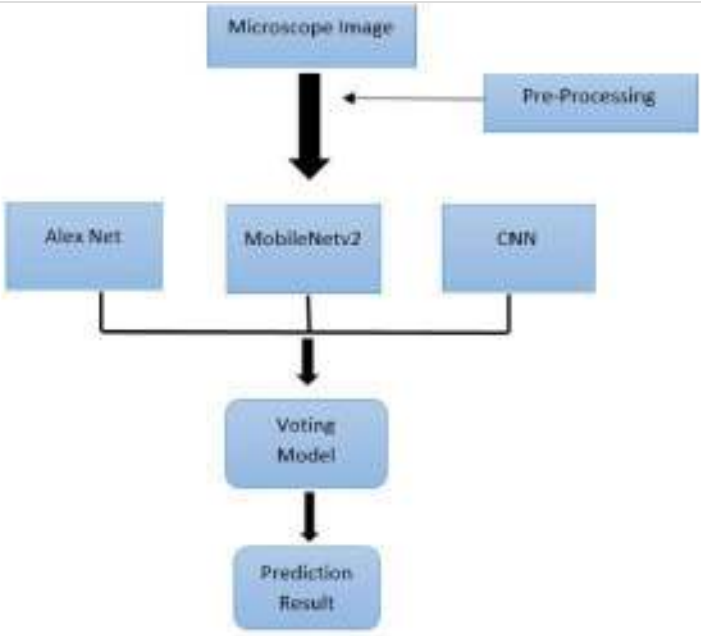



Fig. 7 Overall procedure of the proposed method

The result of implementation (Initializing input data normalization) and the training option is shown below in the table:

With properties of Series Network: Layers: [25×1 net.CNN.layer.Layer] InputNames: {'data'}.

OutputNames: {'output'}.

With properties of SeriesNetwork: Layers: [25×1 net.CNN.layer.Layer].

With properties of the layers: 4×1 Layer array and with the other 4 layers And some properties of Series Network.

### 4.2 Result of Transfer Learning

The authors also implement some of the transfer learning separated each algorithm like ( Mobile Net, and Alex Net). After that, the authors applied the ensemble for the mentioned algorithms. The result of Mobile Net is as below: Fig. 8 show the Model loss and model accuracy. (Fig. 9)

## 5 Conclusion & Future Work

A total of 2749 images of microalgae were classified by a processing technique using a deep learning method (CNN) which is an improved method technique. It is a developing method that can organize the data and gain results well because the images contain more than 2700 images with an input layer [227 227 3]. Hence, we have to classify them

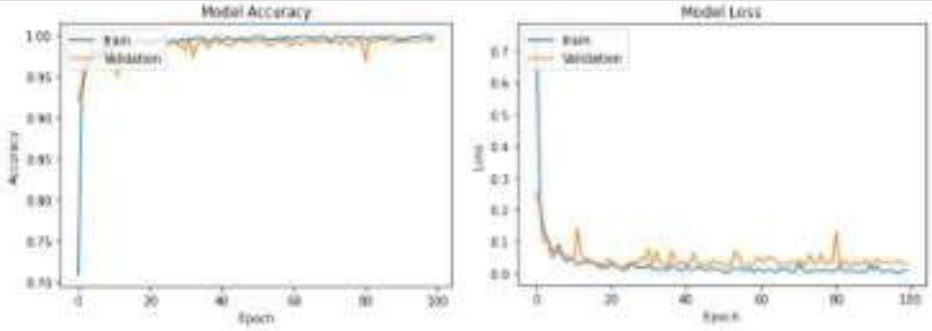



Fig. 8 Model Loss and Model Accuracy of Mobile Net

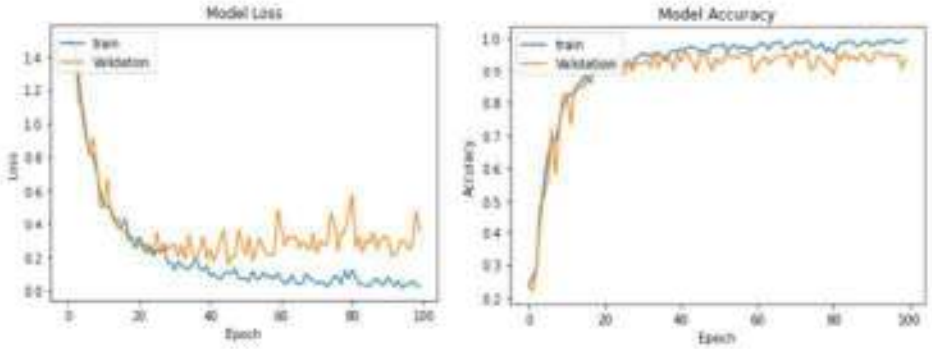



Fig. 9 Model Accuracy and loss model of Alex net

into four classes; our algae dataset is green vegetative cells with 547 images, red vegetative cells were 896 pictures, pictures of the green cyst (652 images), and the red cyst was 654 pictures. At first, the proposed method applied the classification on the algae spaces and was used at different times with the plots and training progress. As a result, it gained different results regarding the accuracy and loss of precision that changes at different times (10, 20, 5). The method adopted in this approach contains eight layers. Still, we used at least 4–5 of them, which are the image input layer, fully connected, softmax, and classification output layers which are used in the application of the diagram, and then the results were visible in the training of network analyzer to show the analysis results and the training progress, and for The design had a feature that was GUI. Its graphic user interface was used in the field of CNN, and the user could choose any image he wanted and display the first text area for the path of the image.

**Acknowledgements** Not applicable

**Author Contributions** Corresponding author contributed significantly to the research and this paper, and the first author is the main contributor.

**Funding** The author received no specific funding for this study.

**Availability of Data and Material** Not applicable.

## Declarations

**Competing interests** The authors declare that they have no competing interests.

## References


1. Hawezi, R. S., Khoshaba, F. S., & Kareem, S. W. (2022). A comparison of automated classification techniques for image processing in video internet of things. *Computers and Electrical Engineering, 101*, 108074.
2. Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation, 29*(9), 2352–2449.
3. Bajaj, R., & Sharma, V. (2018). Smart Education with artificial intelligence-based determination of learning styles. *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2018.05.095
4. M. Amin Ali, S. Wahhab Kareem and A. S. Mohammed, "Evaluation of Electrocardiogram Signals Classification Using CNN, SVM, and LSTM Algorithm: A review," In: *2022 8th International Engineering Conference on Sustainable Technology and Development (IEC)*, 2022, pp. 185–191.
5. Deepa, S. N. (2011). A survey on artificial intelligence approaches for medical image classification. *Indian Journal of Science and Technology*. https://doi.org/10.17485/ijst/2011/v4i11.35
6. Al-Jumur, S. M. R. K., Kareem, S. W., & Yousif, R. Z. (2021). Predicting temperature of Erbil City applying deep learning and neural network. *Indonesian Journal of Electrical Engineering and Computer Science, 22*(2), 944–952.
7. Enderle, D., & Weih, R. C., Jr. (2005). Integrating supervised and unsupervised classification methods to develop a more accurate land cover classification. *Journal of the Arkansas Academy of Science, 59*, 65–73.
8. Ohnuki, S., Nogami, S., Ota, S., Watanabe, K., Kawano, S., & Ohya, Y. (2013). Image-based monitoring system for green algal haematococcus pluvialis (chlorophyceae) cells during culture. *Plant and Cell Physiology, 54*(11), 1917–1929. https://doi.org/10.1093/pcp/pct126
9. Ali, O. M. A., Kareem, S. W., & Mohammed, A. S. (2022). Comparative evaluation for two and five classes ECG signal classification: Applied deep learning. *Journal of Algebraic Statistics, 13*(3), 580–596.
10. Corrêa, L., Drews, P., Botelho, S., de Souza, M.S., Tavano, V.M., Deep learning for microalgae classification. 2017.
11. Krizhevsky, A., Sutskever, I., Hinton, G.E., Image net classification with deep convolutional neural networks. In *Proceedings of Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
12. Stoecker, D. K. (1998). Conceptual models of mixotrophy in planktonic protists and some ecological and evolutionary implications. *European Journal of Protistology, 34*(3), 281–290.
13. Mohammed, A. S., Mohammed, A. S., & Kareem, S. W. (2022). Deep learning and neural network-based wind speed prediction model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 30*(3), 403–425.
14. Ward, B.A., Follows, M.J. (2016) Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. In *Proceedings of the national academy of sciences*, 113, pp. 2958–2963
15. Do, T.-T., Ong, B.-N., Le, T.-L., Nguyen, T.-C., Tran-Thi, B.-H., Thu Hien, B. T., Melkonian, M., & Tran, H.-D. (2021). Growth of haematococcus pluvialis on a small-scale angled porous substrate photobioreactor for green stage biomass. *Applied Science, 11*, 1788. https://doi.org/10.3390/app11041788
16. Basiroh, B., Priyatno, P., Kareem, S. W., & Nurdiyanto, H. (2021). Analysis of expert system for early diagnosis of disorders during pregnancy using the forward chaining method. *International Journal of Artificial Intelligence Research, 5*(1), 44–52.
17. Mostafa, N., Omar, H., Tan, S. G., & Napis, S. (2011). Studies on the genetic variation of the green unicellular alga haematococcus pluvialis (chlorophyceae) obtained from different geographical locations using ISSR and RAPD molecular marker. *Molecules, 16*, 2599–2608. https://doi.org/10.3390/molecules16032599
18. LeCun, Y., Haffner, P., Bottou, L., Bengio, Y. (1999) Object recognition with gradient-based learning. In *lecture notes in computer science: shape, contour and grouping in computer vision*, pp. 319–345. Springer.

19. Promdaen, S., Wattuya, P., & Sanevas, N. (2014). Automated microalgae image classification. *Procedia Computer Science*, *29*, 1981–1992.
20. Rodrigues, F.C.M., Hirata, N.S., Abello, A.A., Leandro, T., La Cruz, D., Lopes, R.M. and Hirata Jr, R., Evaluation of transfer learning scenarios in plankton image classification. In *Proceedings of international joint conference on computer vision, imaging and computer graphics theory and applications (5: international conference on computer vision theory and applications)*, pp. 359–366, 2018.
21. Orenstein, E.C., Beijbom, O., Transfer learning and deep feature extraction for planktonic image data sets. In *Proceedings of 2017 IEEE winter conference on applications of computer vision*, pp. 1082–1088, 2017.
22. Muhamad, H. A., Kareem, S. W., & Mohammed, A. S. (2022, February). A comparative evaluation of deep learning methods in automated classification of white blood cell images. In *2022 8th International engineering conference on sustainable technology and development (IEC)* (pp. 205–211). IEEE.



**Roojwan Scddeek Hawezi** I got the BSc degree from Baghdad University in Physical science, then I got a higher Diploma in Computer science. Back in 2003, I got my MSc degree then finally the Ph.D. in 2007 in Computer science from Baghdad too. I'm a specialist in Digital Image Processing. I'm the former head of the Information System Eng. (ISE) Department

of Algae is very difficult and needs high level expertise. In this case, it becomes error prone and tidy task for biologist. In this paper , we presented high end, fast, accurate and efficient system based on image processing and machine learning algorithms for classification of...



**Deep Learning and Transfer Learning Approaches for Image Classification**

Article Full-text available

June 2019

Sajja Tulasi Krishna · Hemantha kumar Kalluri

Women Deep Learning is-one of the machine learning areas, applied in recent areas. Various techniques have been proposed depends on varieties of learning, including un-supervised, semi-supervised, and supervised-learning. Some of the experimental results proved that the deep learning systems are performed well compared to conventional machine learning systems in image processing, computer vision and...



**Iris Recognition using Convolutional Neural Network**

Conference Paper Full-text available

November 2020

Yuan Zhuang · Joon Huang Chuah · Chee Onn Chow · Marcus Guozong Lim





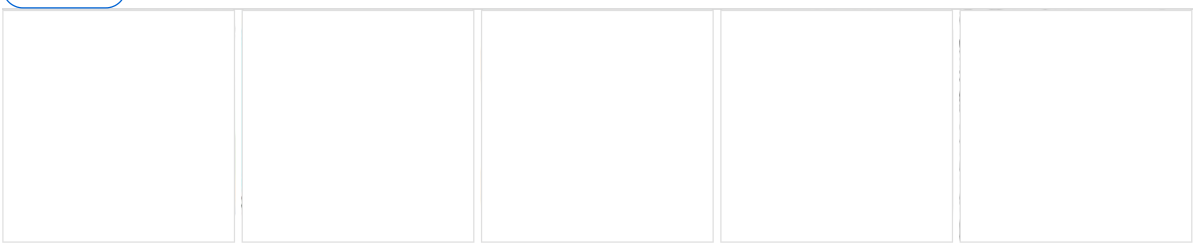
**Fixed-Point Implementation of Convolutional Neural Networks for Image Classification**

Conference Paper Full-text available

October 2018

Chun Yan Lo · Francis C. M. Lau · Chiu-Wing Sham





**Family Coat of Arms and Armorial Achievement Classification**

Chapter Full-text available

January 2020

Martin Sustek · Frantisek Vidensky · Frantisek Zboril · Frantisek Zboril

This paper presents an approach to classification of family coats of arms and armorial achievement. It is difficult to obtain images with coats of arms because not many of them are publicly available. To the best of our knowledge, there is no dataset. Therefore, we artificially extend our dataset using Neural Style Transfer technique and simple image transformations. We describe our dataset and the division int...