ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

Graptolite Image Classification Using Two-stage Transfer Learning and EfficientNet

Yuantong Liang[1, a], Chaofeng Li[1, b], Honglin Liu[2, 3, 4, c], Jiehui Zhang[5, d], Yubing Ji[5, e], and Xiaobo Li[2, 3, 4, f]

[1] Institute of Logistics Science and Engineering, Shanghai Maritime University, Shanghai 201306, China

[2] China Petroleum Exploration and Development Research Institute, Beijing 100083, China

[3] PetroChina unconventional oil and gas Key Laboratory, Beijing 100083, China

[4] National energy shale gas research and development (Experiment) center, Langfang 065007, China

[5] Zhejiang Oilfield, Petrochina, Hangzhou 211762, China

[a] yuantong_liang@163.com, [b] wxlichaofeng@126.com, [c] liuhonglin69@petrochina.com.cn,

[d] zhangjh85@petrochina.com.cn, [e] jiyb85@petrochina.com.cn, [f] lixiaobo69@petrochina.com.cn

Abstract

Graptolites provide an important basis for the fine division of shale strata. Due to the large variety of graptolites and small differences between species, it is difficult to identify graptolite images manually. It is necessary to study the automatic classification method of graptolite images. Nowadays, deep learning has made great progress in image classification, and it is feasible to apply it to graptolite automatic classification. In this paper, by collecting the original graptolite images and data expansion, we constructed the graptolite image dataset, including 22 genera and 51 species, with a total of 125375 images. Then, we proposed a graptolite image classification method using two-stage transfer learning based on taxonomic category and EfficientNet-b5 convolutional neural network (CNN). Our experiments show the average classification accuracy reached 94.3% for 22 genera, and 92.73% for 51 species, and using two-stage transfer learning efficiently improve the classification accuracy of CNN.

Keywords

Graptolite Fossils; Image Classification; Convolutional Neural Network; Taxonomic Category; Transfer Learning.

1. Introduction

It is very reliable to judge the stratigraphic age according to graptolites [1]. In the exploration and development of shale gas, experts can accurately determine the shale stratum by distinguishing the types of graptolites. However, it is very difficult to identify graptolite images manually because of the large variety of graptolites and the small differences between species. With the development of artificial intelligence technology, it is of great significance to use artificial intelligence algorithm to automatically identify the types of graptolite fossils [2, 3].

In the early stage, many scholars have studied the approaches to the automatic classification of fossil images. In 1996, Culverhouse et al. [4] combined the morphological features obtained from image processing with statistics to classify field-collected dinoflagellates. Some scholars directly input

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

images into models, such as COGNIS proposed by Nederbragt et al. [5] in 2005. In 2013, Schulze et al. [6] combined morphological features with artificial neural network (ANN) classifier for the identification of phytoplankton. Some scholars also combine images with morphological features, such as SYRACO for automatic classification of calcareous nannofossils proposed by Barbarin [7] in 2014. However, the training of shallow convolutional neural network (CNN) tried in the early stage consumes a lot of time. For instance, it takes 30 hours to train COGNIS on the dataset of 2000 images, and the robustness of these algorithms can not be deeply analyzed.

With the improvement of computing power, some scholars began to study the methods for fossil image classification based on deep CNN. For example, in 2017, Zhong et al. [8] used pre-trained VGGNet [9] and ResNet [10] for transfer learning to classify foraminifera images.

At present, there is no special method for graptolite image classification. Aiming at this vacancy, this paper proposes a method for graptolite image classification using two-stage transfer learning and EfficientNet. The main contributions are summarized as follows.

- We collected 1015 original images of graptolite fossils. Using the data expansion methods of size scaling, center clipping, random translation, random flip, boundary expansion and rotation, a graptolite fossil image dataset is constructed, including 22 genera and 51 species, with a total of 125375 images. Some typical graptolite fossil images are shown in Figure 1.
- By combining the characteristics of biological taxonomic category with deep transfer learning, a graptolite image classification algorithm using two-stage transfer learning based on graptolite taxonomic category and EfficientNet is proposed.

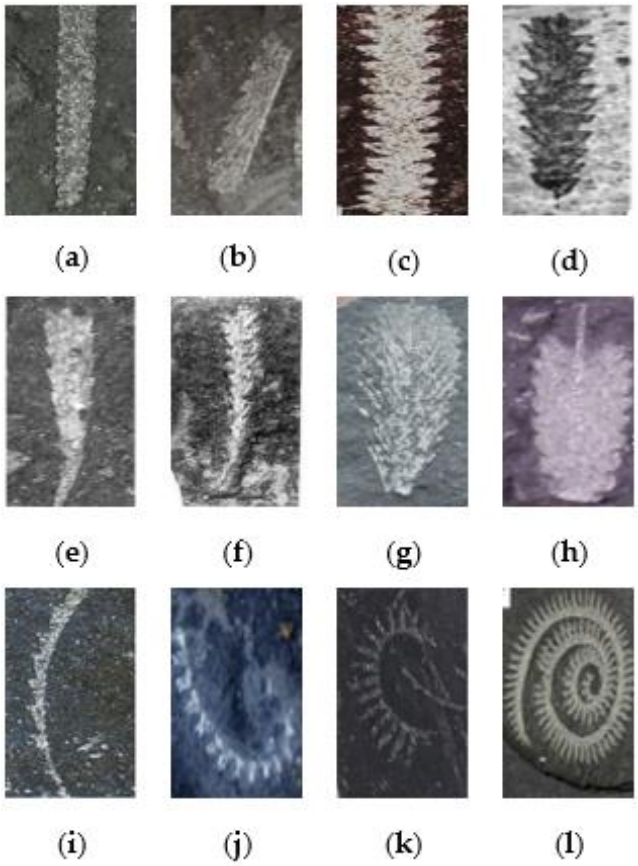

Figure 1. Some typical graptolite fossil images. (a) Climacograptus Tubuliferus; (b) Pristiograptus Regularis; (c) Anticostia Tenuissima; (d) Normalograptus Persculptus; (e) Akidograptus Ascensus; (f) Parakidograptus Acuminatus; (g) Petalolithus Folium; (h) Cystograptus Vesiculosus; (i) Huttagraptus Typicus; (j) Demirastrites Triangulates; (k) Rastrites Orbitus; (l) Oktavites Spiralis.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

The rest of this article is organized as follows. Section 2 introduces the establishment process of graptolite image dataset, including original image acquisition and data expansion. Section 3 describes the graptolite classification algorithm using two-stage transfer learning and EfficientNet. Section 4 reports the experiment and analysis. Finally, the conclusions are given in Section 5.

2. Graptolite Image Dataset

2.1 Original Images

As shown in Figure 2, we obtained the original images by photographing the graptolites in the wild and core samples.

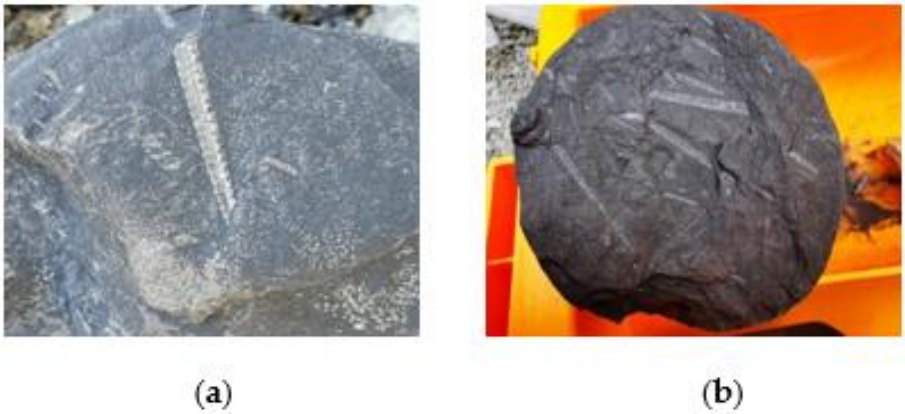

Figure 2. Collected graptolite fossil images. (a) The field outcrop sample; (b) The core sample.

2.2 Data Expansion

In order to enrich the graptolite image dataset and improve the generalization ability of the model [11, 12], combined with the characteristics of graptolite image, we adopts data expansion methods including boundary expansion, rotation, size scaling, center clipping, random flip and random translation to increase the number of graptolite images.

Specifically, we use the vision toolkit torchvision for data expansion. The parameter settings are as follows. The expanded width is 50, 100, 150 and 200 pixels respectively. The rotation angle is between 0 and 360 degrees (15 degrees apart). Size scaling sets the shorter edge of the image to 256 pixels, and the other edge is adjusted according to the original scale. The side length of the center clipping is 224 pixels. The probability of per-forming a random flip is 50%. The amplitude of random translation is limited to 5% of the edge length of the image. The effect of data expansion is shown in Figure 3.

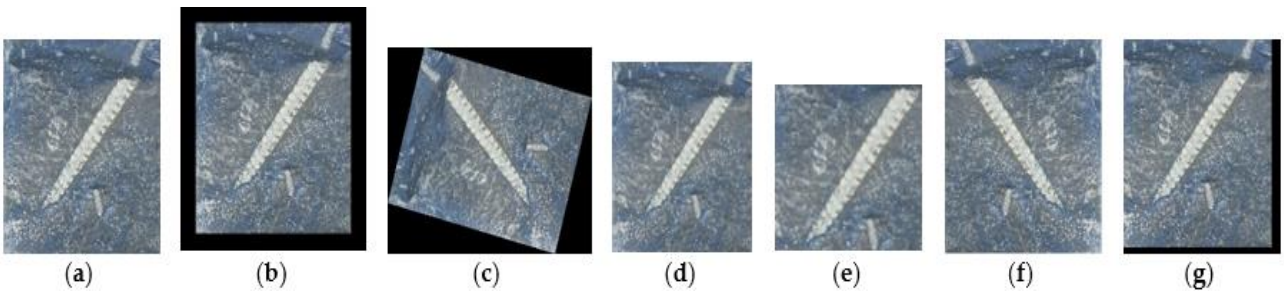

Figure 3. Data expansion. (a) The original image; (b) Boundary expansion; (c) Rotation; (d) Size scaling; (e) Center clipping; (f) Random flip; (g) Random translation.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

3. Proposed Methods

3.1 EfficientNet Convolutional Neural Network

In 1998, Lecun [13] proposed the first CNN which is called LeNet. It includes convolution layer, pooling layer and full connection layer, which lays a foundation for the development of subsequent CNN. With the emergence of large-scale datasets such as ImageNet and the continuous improvement of computing power, in 2012, Krizhevsky et al. [14] used GPU to train AlexNet on large-scale datasets for the first time. AlexNet contains local response normalization (LRN) layer, which creates a competition mechanism for the activities of local neurons and enhances the generalization ability of the model. In 2014, Simonyan et al. believed that the role of LRN layer is limited and the number of network layers needs to be increased to improve network performance. Therefore, VGGNet proposed by them has larger network parameters and higher algorithm complexity. At the same time, it also takes up more memory. Excessively increasing the number of network layers will cause vanishing gradient. Therefore, ResNet proposed by He et al. in 2015 contains residual structure, which strengthens the information exchange between the front and rear layers through the "shortcuts". Further, in 2017, DenseNet proposed by Huang et al. [15] maximized the information exchange and dense connection between the front and back layers enables each layer to be directly connected with the input and the loss values of all previous layers, so that it can achieve better performance than ResNet with less parameters and calculation.

However, the above models are generally lack of research on the joint debugging of network parameters [16, 17] (i.e. resolution, network depth and network width). EfficientNet [18] solved this problem, and its core idea is composite scaling, which uses a composite coefficient ϕ to uniformly scale the network width, depth and resolution:

$$\begin{aligned} &Depth: d = \alpha^{\phi} \\ &Width: w = \beta^{\phi} \\ &Resolution: r = \gamma^{\phi} \\ &s.t. \quad \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\ &\qquad \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned} \tag{1}$$

α, β, γ are constants that can be determined by a small grid search. ϕ is a user-specified coefficient that controls how many more resources are available for model scaling, while α, β, γ specify how to assign these extra resources to network depth d, width w, and resolution r respectively. The structure of EfficientNet is shown in Figure 4.

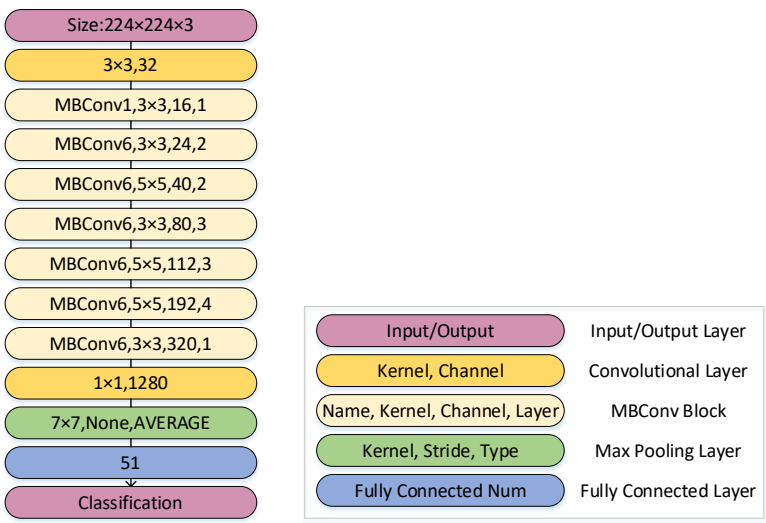

Figure 4. Schematic structure diagram of EfficientNet with different detail structure consisting of convolution layer, maxpooling layer, fully connected layer and MBConv block.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

3.2 Transfer Learning

Human beings have the inherent ability to transfer knowledge across tasks. The knowledge gained in the process of learning a task can be used to solve related tasks. The higher the degree of task relevance, the easier it is for us to Transfer or cross utilize knowledge [19]. Therefore, referring to human learning style, transfer learning is not only a learning paradigm to overcome isolation, but also an idea to use the knowledge obtained from a task to solve related tasks [20]. Figure 5 shows the principle of transferring existing knowledge to new related tasks.

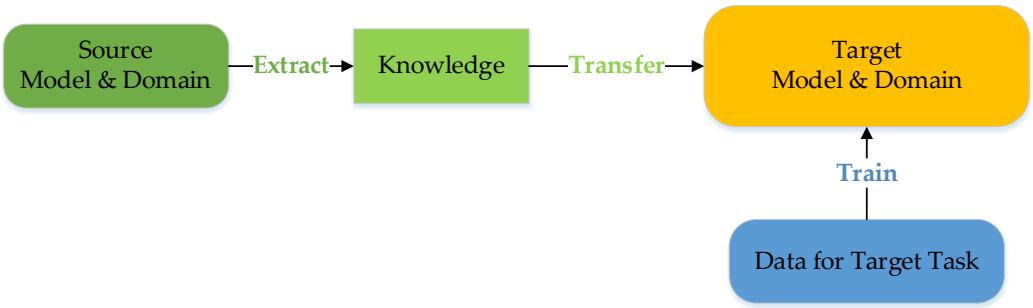

Figure 5. The principle of transfer learning.

We can use domain, task and marginal probability to describe a framework of transfer learning. A domain D can be defined as a tuple containing two elements, one element is the feature space x, and the other element is the marginal probability P(X), where X represents a sample point. X ={x1, x2, ... , xn}, where xi represents a specific vector. X∈x. Thus, the domain D is defined as:

$$D = \{x, P(X)\} \tag{2}$$

On the other hand, a task T can be defined as a tuple containing two elements, one of which is the feature space γ, The other element is the objective function f, which can be expressed as P(γ|X). Therefore, the task T is defined as:

$$T = \{\gamma, P(\gamma \mid X)\} \tag{3}$$

Using this framework, we can define transfer learning as a process. The goal is to use the knowledge of the source task TS in the source domain DS to improve the target task TT in the target domain DT [21].

The main advantages of transfer learning are as follows.

- Improve the baseline performance. When we use the knowledge of isolated learners (also known as ignorant learners) in the source model, the baseline performance may be improved due to this knowledge transfer.
- Shorten the training time of models. Using the knowledge of the source model is conducive to learning the target task in an allround way and shortening the training time of models.
- Improve final performance.

3.3 Two- stage Transfer Learning based on EfficientNet

The animal classification system mainly consists of kingdom, phylum, class, order, family, genus and species. Organisms are classified level by level according to the degree of similarity and difference between organisms and the distance of genetic relationship.

As for the graptolite species classification, this paper proposes a method using two-stage transfer learning based on graptolite taxonomic category and EfficientNet. Considering that it is difficult to directly use EfficientNet to distinguish the characteristics between graptolite species, we divided the training process into two stages. Firstly, we use graptolite genus dataset to train the pre-trained

Volume 8 Issue 4, 2022
DOI: 10.6919/ICJE.202204_8(4).0049

EfficientNet to get the model for graptolite genus classification. On this basis, we input the graptolite species dataset, and finally get the model for graptolite species classification through training. The training process is shown in Figure 6.

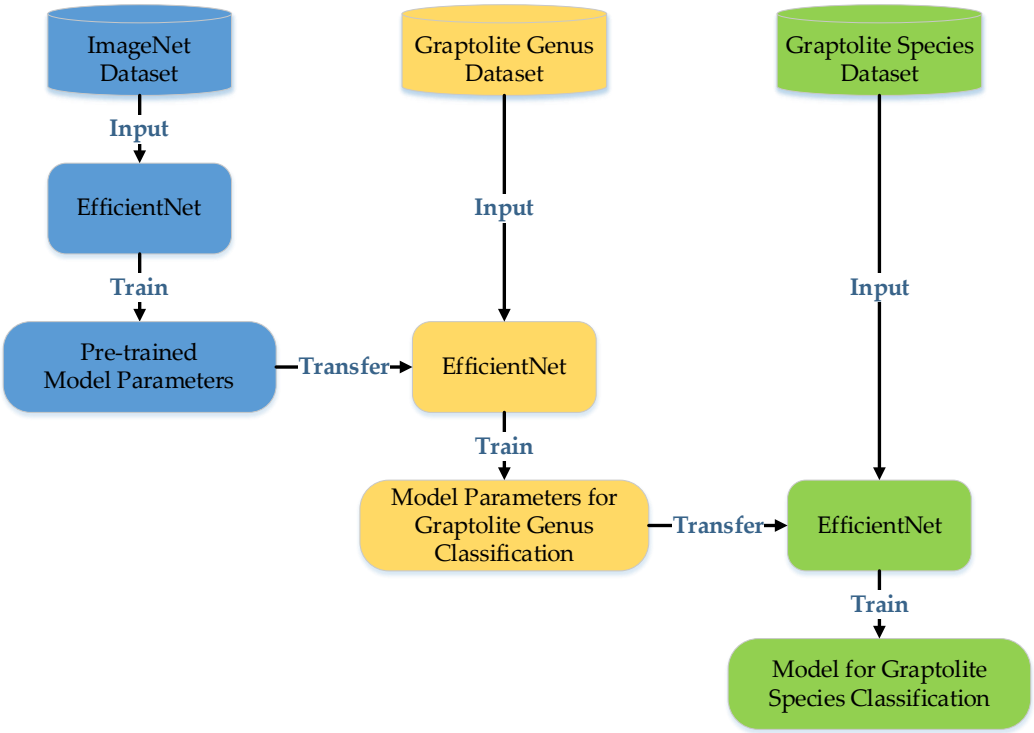

Figure 6. The method of graptolite image classification using two-stage transfer learning and CNN.

In the first stage, the ImageNet dataset is input into the initialized EfficientNet. After training, we get the pre-trained model Ghs1(Gp1(Gt1)). Ghs1 is the first source domain classifier. Gp1 is the feature extractor for ImageNet dataset, and Gt1 is part of the network to be transferred. Then, the knowledge learned from the ImageNet dataset is transferred to the model for graptolite genus classification, which is recorded as Ghs2(Gp2(Gt1) at this time. Ghs2 is the second source domain classifier, and Gp2 is the feature extractor for graptolite genus dataset. After sending graptolite genus dataset into the model Ghs2(Gp2(Gt1) for training, we obtaine the final model for graptolite genus classification, which is recorded as Ghs2(Gp2(Gt2). Gt2 is part of the network to be transferred in the second stage. The loss of the trained feature extractor Gp2 is defined as:

$$L_{t1}=\frac{1}{N_n}\sum_{i=1}^{Nn} L\,(G_{hs2}(G_{p2}(G_{t1}(x_n))),y_n) \tag{4}$$

In the second stage, the knowledge learned from graptolite genus dataset is transferred to the model for graptolite species classification, which is recorded as Ght(Gp3(Gt2) at this time. Ght is the target domain classifier, and Gp3 is the feature extractor for graptolite species dataset. After sending graptolite species dataset into the model Ght(Gp3(Gt2) for training, we obtaine the final model for graptolite species classification. The loss of the trained feature extractor Gp3 is defined as:

$$L_{t2}=\frac{1}{N_n}\sum_{i=1}^{Nn} L\,(G_{ht}(G_{p3}(G_{t2}(x_n))),y_n) \tag{5}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

4. Experiment and Analysis

4.1 Evaluating Indicator

1) Average classification accuracy (ACC). We count the classification accuracy of graptolite images of each type, and take the average.

2) Network parameters (Params). It contains the parameters of convolution layer and full connection layer:

$$Param_{conv} = (kernel \times kernel) \times channel_{input} \times channel_{output} \tag{6}$$

$$Param_{fc} = weight_{in} \times weight_{out} \tag{7}$$

3) Floating-point operations (FLOPs). It is used to measure the computational complexity of the model. The number of FLOPs is actually the number of multiplication and addition operations in the model [22].

4.2 Comparison of Backbone Networks

We randomly divided the graptolite species dataset into training set and test set according to the ratio of 9 : 1. The backbone networks, including AlexNet, VGGNet (VGG-Net-11, VGGNet-16), ResNet (ResNet-18, ResNet-101), DenseNet (DenseNet-121, Dense-Net-201) and EfficientNet (EfficientNet-b0, EfficientNet-b3, EfficientNet-b5), are measured by accuracy among 100 epochs to obtain the maximum accuracy every 25 epochs. The backbone networks based on two-stage transfer learning are trained in graptolite dataset by species. We use Adam optimizer with 0.001 learning rate, and crossentropy loss function, and 224 × 224 input size is chosen to train the backbone networks. The results are shown in Figure 7, Table 1, and Figure 8.

As shown in Figure 7, the bar in the box diagram shows maximum, upper quartile, median, lower quartile, and minimum from top to bottom. The maximums represent the accuracy that the backbone networks may achieve in the final training, and the minimums mean the minimum accuracy of the backbone networks during the training. The data between the maximums and the minimums show the degree of change in the accuracy of the model during the training. Intuitively, EfficientNet reach better performance than the other backbone networks, and the accuracy of EfficientNet-b5 is the highest. These results confirm EfficientNet transfer well and achieve state-of-the-art accuracy.

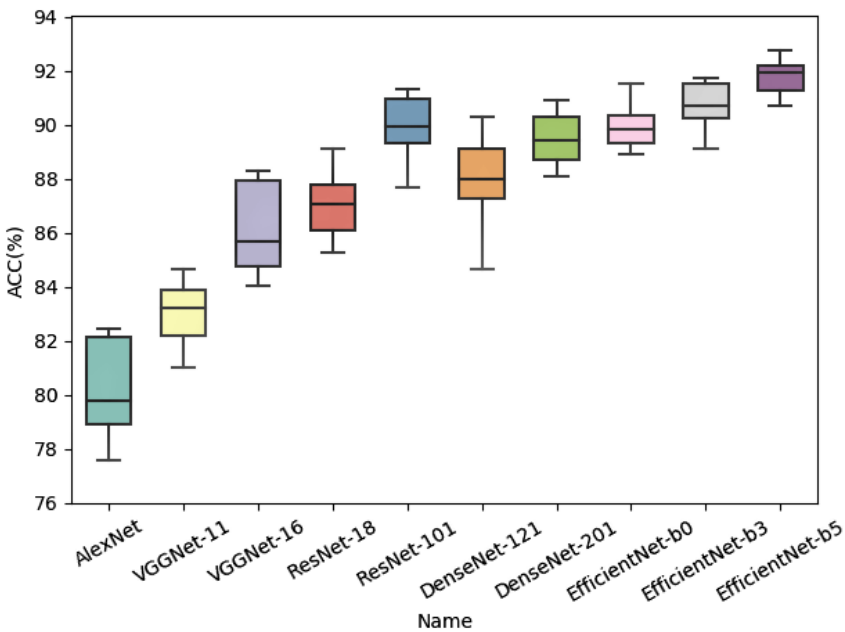

Figure 7. Accuracy of the proposed backbone networks to show data dispersion during the training. The maximum, upper uartile, median,lower quartile, and minimum is shown from top to bottom.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

Table 1 indicates that the maximum accuracy of the backbone networks on the test set of graptolite species dataset with respect to the parameters of models and the computational complexity of the considered architectures for a single forward pass, namely floating-point operations (FLOPs). Figure8 reports the accuracy versus the parameters and FLOPs of the backbone networks. The ball size corresponds to the FLOPs of models. Intuitively, with the increase of parameters and FLOPs, the accuracy of the same series backbone networks is constantly improved, such as VGGNet, ResNet, DenseNet and EfficientNet.

Table 1. Accuracy measures obtained for the proposed CNN approaches using the graptolite dataset by species. The parameters (Params) and FLOPs of the backbone networks are shown.

Name	ACC (%)	Params (M)	FLOPs (B)
AlexNet	82.42	57	0.7
VGGNet-11	84.65	129	7.6
VGGNet-16	88.28	134	15.5
ResNet-18	89.09	11	1.8
ResNet-101	91.31	43	7.9
DenseNet-121	90.30	7	0.6
DenseNet-201	90.91	20	2.8
EfficientNet-b0	91.52	4	0.4
EfficientNet-b3	91.72	12	1.8
EfficientNet-b5	92.73	30	9.9

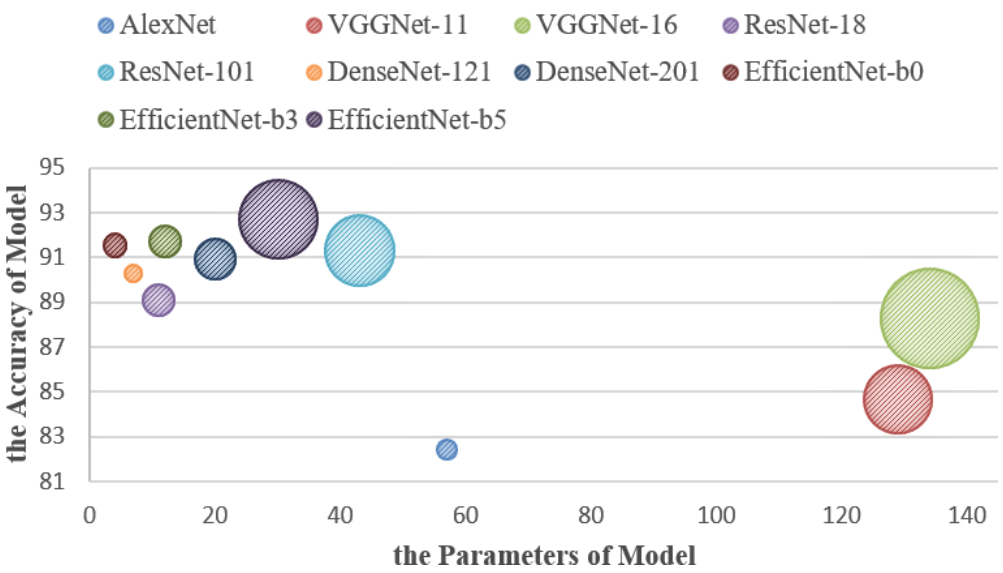

Figure 8. Ball chart reporting the accuracy versus the parameters and FLOPs of the backbone networks. The size of each ball corresponds to the FLOPs of the model.

4.3 Ablation Experiment

EfficientNet-b5, EfficientNet-b5 based on transfer learning and EfficientNet-b5 based on two-stage transfer learning are respectively used for classification of graptolites by species. The results are shown in Table 2.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202204_8(4).0049

Table 2. Accuracy measures obtained for the proposed three methods using the graptolite dataset by species. The training time and iterations of the methods are shown.

Methods	ACC (%)	Time (s)	Epochs
EfficientNet-b5	85.25	6750	37
EfficientNet-b5 (transfer learning)	87.47	3910	23
EfficientNet-b5 (two-stage transfer learning)	92.73	3230	19

As shown in Table 2, the classification accuracy obtained by two-stage transfer learning is the highest, which is 5.26% and 7.48% higher than that by transfer learning and backbone network, respectively. At the same time, the training time of two-stage transfer learning is the shortest, which is approximately 17% and 49% shorter than that of transfer learning and backbone network, respectively. Therefore, the use of two-stage transfer learning not only improves the classification accuracy, but also shortens the training time, and these results are completely consistent with the advantages of transfer learning.

5. Conclusion

At present, graptolite image classification mainly depends on sample comparison and expert experience, which is time-consuming and laborious. This paper proposes a method for graptolite image classification using two-stage transfer learning based on taxonomic category and EfficientNet. By comparing the representative CNNs in image classification, we select EfficientNet-b5 as the backbone network. Based on EfficientNet-b5, deep transfer learning is used in graptolite genus classification, and the accuracy is 94.3%. Furthermore, using two-stage transfer learning based on graptolite taxonomic category to deal with graptolite species classification, we obtained an accuracy of 92.73%. The experimental results show that the proposed method combines the characteristics of graptolite taxonomic category with deep transfer learning, which not only improves the average accuracy of classification, but also shortens the time for training deep CNN. Since it is difficult to collect graptolite images, only two taxonomic categories of genus and species are considered in this paper. In the future, we can continue to expand the graptolite image dataset, enrich the graptolite taxonomic categories, and analyze the effect of the combination of multiple graptolite taxonomic categories and deep transfer learning.

Acknowledgments

Thank the editor-in-chief and other experts for their suggestions.

References

[1] H. Wang, Z. Shi, Q. Zhao, et al. Stratigraphic framework of the Wufeng-Longmaxi shale in and around the Sichuan Basin, Energy Geoscience, vol. 1 (2020), 124-133.

[2] H.S. Munawar, A.W.A. Hammad, S.T.A. Waller. Review on flood management technologies related to image processing and machine learning, Automation in Construction, vol. 132 (2021), 103916.

[3] R.J.S. Raj, S.J. Shobana, I.V. Pustokhina, et al. Optimal feature selection-based medical image classification using deep learning model in internet of medical things, IEEE Access, vol. 8 (2022), 58006-58017.

[4] P.F. Culverhouse, R.G. Simpson, R. Ellis, et al. Automatic classification of field-collected dinoflagellates by artificial neural network, Marine Ecology Progress Series, vol. 139 (1996), 281-287.

[5] J. Bollmann, P.S. Quinn, M. Vela, et al. Image calibration, filtering, and processing, Image Analysis, Sediments and Paleoenvironments, vol. 1 (2005), 35-58.

[6] K. Schulze, U.M. Tillich, T. Dandekar, et al. PlanktoVisionan automated analysis system for the identification of phytoplankton, BMC bioinformatics, vol. 14 (2013), 1-10.

[7] N. Barbarin. La reconnaissance automatisée des nannofossiles calcaires du cénozoïque. PhD thesis, Aix-Marseille Université, Aixen-Provence, France, 2014.

[8] B. Zhong, Q. Ge, B. Kanakiya, et al. A comparative study of image classification algorithms for Foraminifera identification, Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence, vol. 1 (2017), 1-8.

[9] K. Simonyan, A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv: 1409.1556.

[10] K. He, X. Zhang, S. Ren, et al. Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, vol. 1 (2016), 770-778.

[11] P.L. Bartlett, P.M. Long, G. Lugosi, et al. Benign overfitting in linear regression, Proceedings of the National Academy of Sciences, vol. 117 (2020), 30063-30070.

[12] M. Gour, S. Jain, T. Sunil. Residual learning based CNN for breast cancer histopathological image classification, International Journal of Imaging Systems and Technology, vol. 30 (2020), 621-635.

[13] Y. Lecun, L. Bottou, Y. Bengio, et al. Gradient-based learning applied to document recognition, Proceedings of the IEEE, vol. 86 (1998), 2278-2324.

[14] A. Krizhevsky, I. Sutskever, G.E. Hinton. Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, vol. 25 (2012), 1097-1105.

[15] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. arXiv 2018, arXiv: 1608.06993v5.

[16] E. Weinan, M. Chao, W. Lei. A comparative analysis of optimization and generalization properties oftwo-layer neural network and random feature models under gradient descent dynamics, Science China Mathematics, vol. 63 (2020), 1235.

[17] D.X. Zhou. Theory of deep convolutional neural networks: Downsampling, Neural Networks, vol. 124 (2020), 319-327.

[18] X. Yin, L. Liu, H. Liu, et al. Heterogeneous cross-project defect prediction with multiple source projects based on transfer learning, Mathematical Biosciences and Engineering, vol. 17 (2020), 1020-1040.

[19] B.M. Lake, T.D. Ullman, J.B. Tenenbaum, et al. Building machines that learn and think like people, Behavioral and brain sciences, vol. 40 (2017), 253.

[20] F. Zhuang, Z. Qi, K. Duan, et al. A comprehensive survey on transfer learning, Proceedings of the IEEE, vol. 109 (2020), 43-76.

[21] M. Tan, Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv 2020, arXiv: 1905.11946v5.

[22] C. Zhang, Y. Chen, X. Yang, et al. Improved remote sensing image classification based on multi-scale feature fusion, Remote Sensing, vol. 12 (2020), 213.