Classification of Radiolarian Fossil Images Using Generalized Feed Forward Neural Network

Ms. [1]Trushali S. Ruikar
Student of HVPM'S College of Engineering and Technology Amravati (India)

Mr.Vijay L. Agrawal[2]
Associate Professor in Dept. (Electronic and Telecommunication) of HVPM'S
College of Engineering and Technology (India)

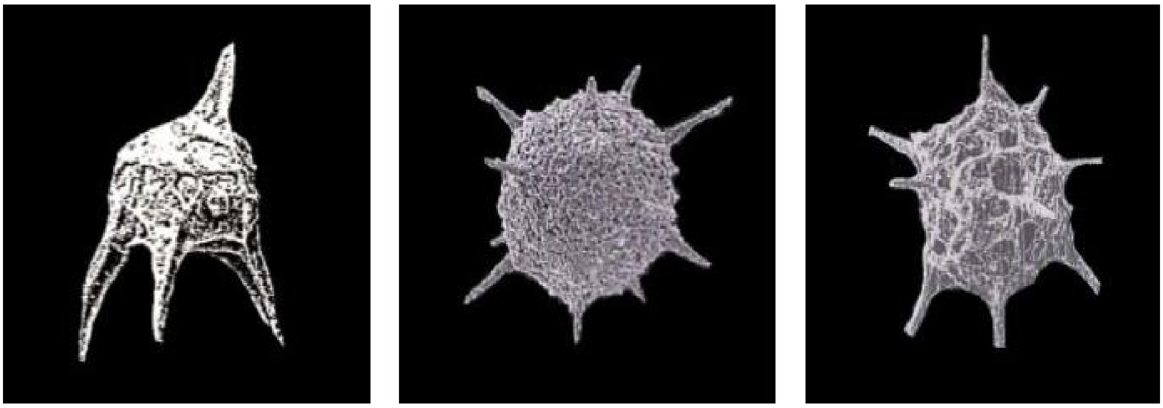
ABSTRACT: Radiolarians are some kind of planktonic protozoa and are important biostratigraphic and paleoenvironmental indicators for palaeogeographic reconstructions. Radiolarian paleontology is still considered to be the most affordable way to date deep ocean sediments. Conventional methods for identifying radiolarians are time consuming and cannot be scaled by the detail or scope required for large-scale studies. Automatic image classification allows these analyzes to be done quickly. In this study, a method for automatic classification of fosilized radiolarian images obtained by Scanning Electron Microscope (SEM) using neural network. High classification performances were obtained with the generated models. The Efficient classifiers based on Multilayer Perceptron (MLP) Neural Network. A separate Cross-Validation dataset is used for proper evaluation of the proposed classification algorithm with respect to important performance measures, such as MSE and classification accuracy. The Average Classification Accuracy of MLP Neural Network comprising of one hidden layers with 14 PE's organized in a typical topology is found to be superior (86.74 %) for Training and cross-validation. Finally, optimal algorithm has been developed on the basis of the best classifier performance.
KEYWORDS: Neural solution, MatLab, Microsoft excel,SEM images

I. INTRODUCTION

Radiolarians are a type of planktonic protozoa suspended along the water column in the oceans. The skeleton of the radiators consists of pure amorphous silica and is the most characteristic morphological feature of the organism. Its skeletons are complex and very diverse in terms of architecture. They are considered to be important biostratigraphic and paleoenvironmental indicators for palaeogeography reconstruction. Radiolarians have an increasing value as depth, palaeoclimate, and palaeoenvironmental indicators. Radiolaria paleontology is the cheapest and relatively quick way to date deep ocean sediments.

Dumitrica [1] and Pessagno Jr et al. Has become important and facilitated by the rock removal methods defined by [2]. These methods basically have the same procedure, but the acid type used differs according to the type of rock. The rock samples which are found to contain radiospheric fossils are treated with a mixture of concentrated acids (10%) (hydrofluoric, hydrochloric, nitric and acetic) and water (90%). It is subjected to a 24 hour wash followed by washing and sieving. Finally, under a binocular microscope, the help of a fine brush is used to collect the radiolar shells from the part remaining on the sieve. The resulting radiolar fossils are displayed under the Scanning Electron Microscope (SEM) for detailed taxonomic studies. For the systematic identification of fossils examined, it is necessary to scan the work done up to daylight. The skeletal morphological character as well as the geometry of the shells, including the number of various structures such as spins, feet, spikes and pores on the skeleton, are important to determine the genus and species of radiolarians. As a result of taxonomic research, palaeontologists work on the radiolar assemblages they have obtained and reveal data about when and where the sample collapses

in geological time. After this process, the paleontologist forms biostratigraphy of the study area and can make paleogeographic model interpretations.
Identification of radiolayers with morphological richness difficult to define because they have many generations and species are fossils. Using machine learning methods As with the work done with the plankton image it may facilitate the operation. With this study,microfossil SEM images Classification System which was developed to simulate human experience in the recognition of underwater shapes by using Pattern Averaging and Back Propagation Learning Algorithm, will be presented.The reliability and the success of these systems are depend on the effectiveness of applied data pre-processing techniques and neural networks which can be used for efficient modeling of human's visual system during the recognition or classification of patterns. Neural networks have an important part in the modelling of human experience and decision making process into computers. some microfossil SEM images are shown in below figure 1



(a) (b) (c)

Figure 1:(a) Hozmadia reticulata (b) Triassospongosphaera multispinosa (c) Pentactinocapsa awaensis

II. RELATED WORK

Automatic classification approaches have previously been used in taxonomy have been used in their work. Apostol et al. [3] fossil a method for classifying radiological images He has proposed. A dialect using Fourier transform,rotation and scale insensitive method, Fimbres-Castro and others [4]. Radyoolar images Another method that uses transfer learning on the Felt et al. [5].It also has similar features with radiolarians automatic classification of plankton images and There are various methods for image retrieval. Interest areas (ROIs) can be used for segmentation, feature extraction,
finding similarity or image classification is the main steps are. For example, many visual segmentation algorithms It is absent; threshold base, color base, texture base, model base and so on. [6]. Wavelet identifiers [7, 8] and textural properties [9], morphological properties such as granulometric properties [10] features, contours and bounds, oriented gradients (HOG) histogram [11] and so on. classification of features [12, 13] and image access approaches. In diatom and phytoplankton image classification [14-16] Fourier transformation and use of Fourier masks studies are available in the literature.

III METHOD

Figure3 Methodology of work

It is classification of Radiolarian Fossil Images Using Neural Network Approaches.. Data acquisition for the proposed classifier designed for the classification of Radiolarian Fossil Images. The most important un correlated features as well as coefficient from the images will be extracted .In order to extract features WHT transformed domain will be used.

3.1*Neural Networks*

Following Neural Networks are tested:

- **Feed-Forward Neural Networks**

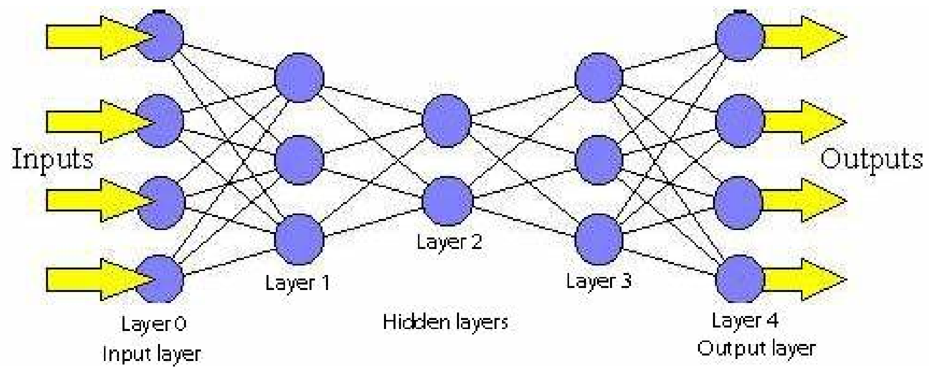


Figure 3.1 A feed-forward network.

- Feed-forward networks have the following characteristics:

1. Perceptrons are arranged in layers, with the first layer taking in inputs and the last layer producing outputs. The middle layers have no connection with the external world, and hence are called hidden layers.
2. Each perceptron in one layer is connected to every perceptron on the next layer. Hence information is constantly "fed forward" from one layer to the next., and this explains why these networks are called feed-forward networks.
3. There is no connection among perceptrons in the same layer.

- A single perceptron can classify points into two regions that are linearly separable. Now let us extend the discussion into the separation of points into two regions that are not linearly separable. Consider the following network:

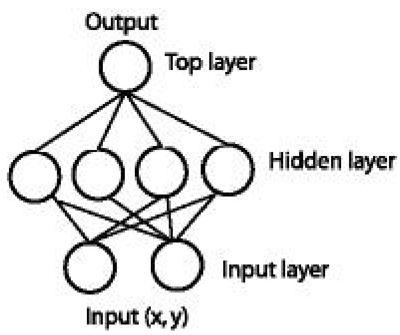


Figure 3.2. A feed-forward network with one hidden layer.

- The same (x, y) is fed into the network through the perceptrons in the input layer. With four perceptrons that are independent of each other in the hidden layer, the point is classified into 4 pairs of linearly separable regions, each of which has a unique line separating the region.

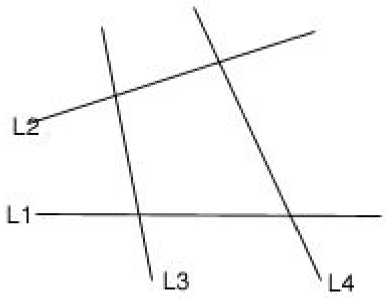


Figure.3.3 lines each dividing the plane into 2 linearly separable regions.

- The top perceptron performs logical operations on the outputs of the hidden layers so that the whole network classifies input points in 2 regions that might not be linearly separable. For instance, using the AND operator on these four outputs, one gets the intersection of the 4 regions that forms the center region.

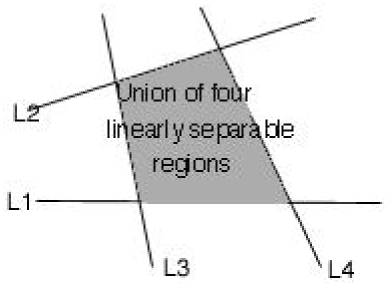


Figure 3.4 Intersection of 4 linearly separable regions forms the center region.

- By varying the number of nodes in the hidden layer, the number of layers, and the number of input and output nodes, one can classification of points in arbitrary dimension into an arbitrary number of groups. Hence feed-forward networks are commonly used for classification.

- ***Learning Rules used***:
  - ***Momentum***

Momentum simply adds a fraction m of the previous weight update to the current one. The momentum parameter is used to prevent the system from converging to a local minimum or saddle point. A high momentum parameter can also help to increase the speed of convergence of the system. However, setting the momentum parameter too high can create a risk of overshooting the minimum, which can cause the system to become unstable. A momentum coefficient that is too low cannot reliably avoid local minima, and can also slow down the training of the system.

- ***Conjugate Gradient***

CG is the most popular iterative method for solving large systems of linear equations. CG is effective for systems of the form A=xb-A (1) where x _is an unknown vector, b is a known vector, and A _is a known, square, symmetric, positive-definite (or positive-indefinite) matrix. (Don't worry if you've forgotten what "positive-definite" means; we shall review it.) These systems arise in many important settings, such as finite difference and finite element methods for solving partial differential equations, structural analysis, circuit analysis, and math homework.

Developed by Widrow and Hoff, the delta rule, also called the Least Mean Square (LMS) method, is one of the most commonly used learning rules. For a given input vector, the output vector is compared to the correct answer. If the difference is zero, no learning takes place; otherwise, the weights are adjusted to reduce this difference. The change in weight from ui to uj is given by: dwij = r* ai * ej, where r is the learning rate, ai represents the activation of ui and ej is the difference between the expected output and the actual output of uj. If the set of input patterns form a linearly independent set then arbitrary associations can be learned using the delta rule.

It has been shown that for networks with linear activation functions and with no hidden units (hidden units are found in networks with more than two layers), the error squared vs. the weight graph is a paraboloid in n-space. Since the proportionality constant is negative, the graph of such a function is concave upward and has a minimum value. The vertex of this paraboloid represents the point where the error is minimized. The weight vector corresponding to this point is then the ideal weight vector.

- ***Quick propagation***

Quick propagation (Quickprop) [1] is one of the most effective and widely used adaptive learning rules. There is only one global parameter making a significant contribution to the result, the e-parameter. Quick-propagation uses a set of heuristics to optimise Back-propagation, the condition where e is used is when the sign for the current slope and previous slope for the weight is the same.

- ***Delta by Delta***

Developed by Widrow and Hoff, the delta rule, also called the Least Mean Square (LMS) method, is one of the most commonly used learning rules. For a given input vector, the output vector is compared to the correct answer. If the difference is zero, no learning takes place; otherwise, the weights are adjusted to reduce this difference. The change in weight from ui to uj is given by: dwij = r* ai * ej, where r is the learning rate, ai represents the activation of ui and ej is the difference between the expected output and the actual output of uj. If the set of input patterns form a linearly independent set then arbitrary associations can be learned using the delta rule.

It has been shown that for networks with linear activation functions and with no hidden units (hidden units are found in networks with more than two layers), the error squared vs. the weight graph is a paraboloid in n-space. Since the proportionality constant is negative, the graph of such a function is concave upward and has a minimum value. The vertex of this paraboloid represents the point where the error is minimized. The weight vector corresponding to this point is then the ideal weight vector. [10]

IV.Simulation Results

1) Computer Simulation

The GFF neural network has been simulated for 76 different images of Radiolarian Fossil images out of which 60 were used for training purpose and 16 were used for cross validation.

The simulation of best classifier along with the confusion matrix is shown below :

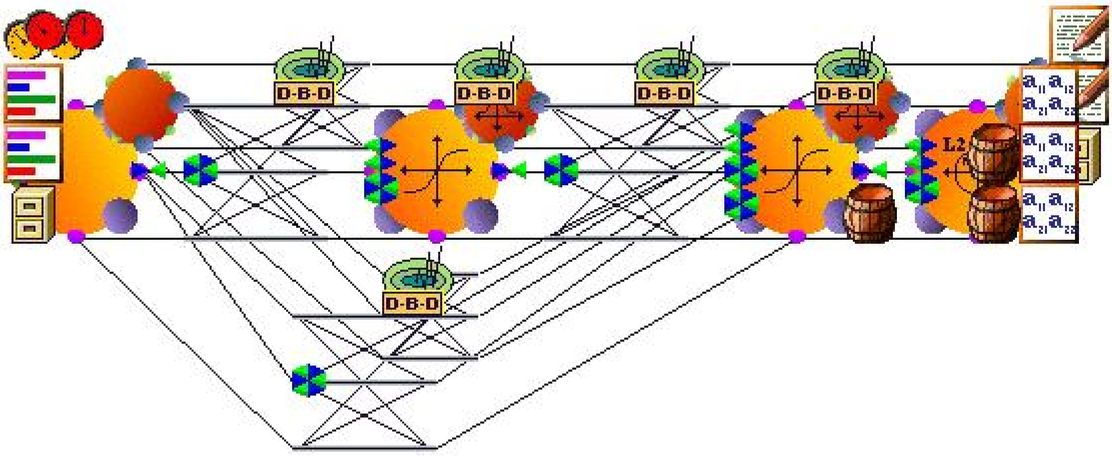


Figure.3.1 GFF neural network trained with DBD learning rule

2) Results

Output / Desired	*NAME(HOZMADIA RETICULATA)*	*NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)*	*NAME(PENTACTINOCAPSA AWAENSIS)*
NAME(HOZMADIA RETICULATA)	5	1	0
NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)	0	4	2
NAME(PENTACTINOCAPSA AWAENSIS)	0	0	3

Table I. Confusion matrix on CV data set

Output / Desired	*NAME(HOZMADIA RETICULATA)*	*NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)*	*NAME(PENTACTINOCAPSA AWAENSIS)*
NAME(HOZMADIA RETICULATA)	18	0	0
NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)	0	20	2
NAME(PENTACTINOCAPSA AWAENSIS)	2	0	19

TABLE II. Confusion matrix on Training data set

Here Table I and Table II Contend the C.V as well as Training data set.

Performance	*NAME(HOZMADIA RETICULATA)*	*NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)*	*NAME(PENTACTINOCAPSA AWAENSIS)*
MSE	0.036882344	0.010472041	0.022051877
NMSE	0.373433734	0.106029415	0.127585859
MAE	0.122096281	0.059051699	0.090455911
Min Abs Error	0.002606556	0.00761174	0.003839355
Max Abs Error	0.519284323	0.358497372	0.431063713
R	0.827740306	0.979398548	0.941459492
Percent Correct	100	80	60

TABLE III. Accuracy of the network on CV data set

Performance	*NAME(HOZMADIA RETICULATA)*	*NAME(TRIASSOSPONGOSPHAERA MULTISPINOSA)*	*NAME(PENTACTINOCAPSA AWAENSIS)*
MSE	0.0007593	0.000788512	0.000264437
NMSE	0.013147708	0.009410913	0.001425058
MAE	0.022338146	0.021875387	0.012289394
Min Abs Error	3.35649E-05	0.000250013	0.000157662
Max Abs Error	0.051531923	0.052780109	0.045616751
R	0.995032434	0.996426622	0.999359714
Percent Correct	90	100	90.47

TABLE IV. Accuracy of the network on training data set

Here Table III and Table IV Contain the C.V and Training result and show the 86.74% percent accuracy.

V. Conclusion and Future Work

From the results obtained it concludes that the GFF Neural Network with DBD (delta by delta) and hidden layer 1 with processing element 14 gives best results of 93.49% in Training while in Cross Validation it gives 80% so overall result is 86.74%.

VI. ACKNOWLEDGMENT

We are very grateful to our HVPM College of Engineering and Technology to support and other faculty and associates of ENTC department who are directly & indirectly helped me for these paper

References

[1] P. Dumitrica, "Family Eptingiidae n. fam., extinct Nassellaria (Radiolaria) with sagittal ring," Dari de seama ale sedintelor, Institutul de Geologie si Geofizica, Bucharest, vol. 64, pp. 27-38, 1978.
[2] E. A. Pessagno Jr and R. L. Newport, "A technique for extracting Radiolaria from radiolarian cherts," Micropaleontology, pp. 231-234, 1972.
[3] L. A. Apostol, E. Márquez, P. Gasmen, and G. Solano, "RadSS: A radiolarian classifier using support vector machines," in 2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA), 2016, pp. 1-6.
[4] C. Fimbres-Castro, J. Alvarez-Borrego, I. Vazquez-Martinez, T. L. Espinoza-Carreon, A. E. Ulloa-Perez, and M. A. Bueno-Ibarra, "Nonlinear correlation by using invariant identity vectors signatures to identify plankton," Gayana, vol. 77, pp. 105-124, 2013.

[5] A. S. Keceli, A. Kaya, and S. U. Keceli, "Classification of radiolarian images with hand-crafted and deep features," Computers & Geosciences, vol. 109, pp. 67-74, Dec 2017.
[6] L. G. Shapiro and G. C. Stockman, Computer Vision. New Jersey: Prentice-Hall, 2001.
[7] K. Arai, "Image Retrieval and Classification Method Based on Euclidian Distance Between Normalized Features Including Wavelet Descriptor," Image, vol. 2, 2013.
[8] J. Landre and F. Truchetet, "Hierarchical architecture for content-based image retrieval of paleontology images," in Electronic Imaging 2002, 2001, pp. 138-147.
[9] Q. Hu and C. Davis, "Automatic plankton image recognition with co-occurrence matrices and support vector machine," Marine Ecology Progress Series, vol. 295, pp. 21-31, 2005.
[10] F. Zhao, F. Lin, and H. S. Seah, "Binary SIPPER plankton image classification using random subspace," Neurocomputing, vol. 73, pp. 1853-1860, Jun 2010.
[11] H. Bi, Z. Guo, M. C. Benfield, C. Fan, M. Ford, S. Shahrestani, et al., "A semi-automated image analysis procedure for in situ plankton imaging systems," PloS one, vol. 10, p. e0127121, 2015.
[12] K. Arai and C. Rahmad, "Content Based Image Retrieval by using Multi Layer Centroid Contour Distance," International journal of advanced research in artificial intelligence, vol. 2, 2013.
[13] C. Rahmad and K. Arai, "Comparison Contour Extraction Based on Layered Structure and Fourier Descriptor on Image Retrieval," International Journal of Advanced Computer Science and Applications, vol. 6, pp. 71-74, Dec 2015.
[14] S. Solorza and J. Alvarez-Borrego, "Position and rotation-invariant pattern recognition system by binary rings masks," Journal of Modern Optics, vol. 62, pp. 851-864, 2015.
[15] A. S. Ventura, J. A. Borrego, and S. Solorza, "Adaptive nonlinear correlation with a binary mask invariant to rotation and scale," Optics Communications, vol. 339, pp. 185-193, Mar 15 2015.
[16] C. Barajas-Garcia, S. Solorza-Calderon, and J. Alvarez-Borrego, "Classification of fragments of objects by the Fourier masks pattern recognition system," Optics Communications, vol. 367, pp. 335-345, May 15 2016.
[17] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, pp. 211-252, Dec 2015.
[18] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and F. F. Li, "ImageNet: A Large-Scale Hierarchical Image Database," Cvpr: 2009 Ieee Conference on Computer Vision and Pattern Recognition, Vols 1-4, pp. 248-255, 2009.
[19] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," Statistics surveys, vol. 4, pp. 40-79, 2010.
[20] P. E. Hart, D. G. Stork, and R. O. Duda, "Pattern classification," John Willey & Sons, 2001.