International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Identification and Classification of Animal Kingdom using Image Processing and Artificial Neural Networks

K. Sujatha , V. Srividhya, M. Aruna

Abstract— *The biological kingdom 'Animalia' is composed of multi cellular eukaryotic organisms. Most of the animal species exhibit bilateral symmetry. The hierarchy of biological classification has eight taxonomy ranks. The top position in the hierarchy is occupied by the 'domain' and ending with the lowest position occupied by 'species'. The classification of animal kingdom includes, Porifera, Coelenterata, Platyhelminthes, Aschelminthes, Annelida, Arthropoda, Mollusca, Echinodermata and Chordata. Manual identification of Phylum or class for each and every species, is very tedious, because there exists nearly a millions of species categorized under various classes. Hence an automated system is proposed to be developed using image segmentation and Artificial Neural Networks (ANN) trained with Back Propagation Algorithm (BPA) which is capable of assisting the scientists and researchers for class identification. This system will be useful in Museums and Archeological departments, where a huge variety of species are maintained. The classification efficiency of the proposed system is 89.1%.*

Keywords: *Animal Kingdom, Phylum, Image processing, Artificial Neural Networks, Back Propagation Algorithm.*

I. INTRODUCTION

The Kingdom Animalia, include animals that are eukaryotic, multi-cellular, classes. All the classes of the animal kingdom possess distinctive characteristics. The species belonging to this category extract their energy by consuming plants or animals [1, 2]. A few of the animals share their characteristics; among numerous species which are extremely moveable. The classification of Animal kingdom is based on the presence and absence of backbone or the vertebral column. This classification is based on their characteristics. They are distinguished from algae, flora, and fungi where inflexible cell wall is not present. A few are heterotrophic; where the food digestion takes place in the digestive system present in the interior cavity which again differentiates them from algae and plant life [3, 4]. Those belonging to this category are capable of movement which means that they are motile, apart from certain life stages. The detailed classification is shown in Figure 1.

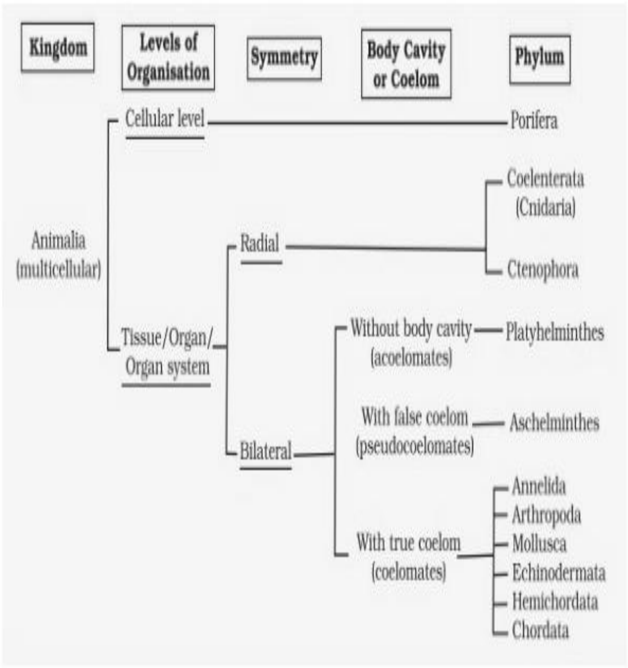


Figure 1. Classification of Animal Kingdom

II. HIGHLIGHTS OF THE WORK

The major objectives are as follows

- To replace the manual mode of Phylum classification
- To automate the Phylum classification for various species as the number is very huge.
- Collection of database with various species belonging to various Phyla.
- Feature Extraction and automatic classification of species for Phyla identification using ANN with BPA.

III. CURRENT STRATEGY FOR PHYLUM IDENTIFICATION OF VARIOUS SPECIES IN ANIMAL KINGDOM

There are more than a million species which are manually classified and described under the animal kingdom. For a newly described species, the manual classification technique proves to be a tedious process in assigning a methodical place in the animal kingdom [5, 6].

Manuscript published on 30 September 2019
* Correspondence Author
K. Sujatha* , Professor, EEE Dept., Dr.MGR Educational and Research Institute, Chennai ,Tamil Nadu, India.
(email: drksujatha23@gmail.com)
V. Srividhya, Asst. Prof., EEE Dept., Meenakshi College of Engineering, India
(email:sripranav2007@gmail.com)
M. Aruna, Asst. Prof., EEE Dept., Meenakshi College of Engineering, Tamil Nadu, India.

Despite the dissimilarities in formation and shape for diversified animals, there are some essential features universal to a variety of each and every species relative to the arrangement of cells, body regularity, and type of coelom, model of digestive, excretory, circulatory and reproductive systems. These are some of the features used for basic classification of species under Animal Kingdom [7, 8].

III. PROPOSED STRATEGY FOR PHYLUM IDENTIFICATION OF VARIOUS SPECIES IN ANIMAL KINGDOM

In the proposed method, the images of the various species corresponding to the different classes of Animal Kingdom are collected. These images are preprocessed for using Gaussian filter for noise removal. Edges are detected using fuzzy logic [9, 10]. Thereafter, features like entropy, correlation and average intensity are extracted from the edge detected images. These features are used as inputs to train the ANN using BPA. The ANN, using its capability to learn, generalizes the characteristics of the species belonging to various phyla, processes them inparallel and finally identifies the class of the species exactly.

IV. MATERIALS AND METHODS

The boundary formed between two homogeneous portions of the image is called as an edge. These boundaries are identified by matching the intensity at the edges for the adjacent pixels. Nevertheless, since identical regions are not well defined, petite variations in intensity between two adjacent pixels doesn't always correspond to an edge. Instead, the intensity variation might represent a shading effect. The image based fuzzy logic approach uses membership functions to describe the extent to which a pixel corresponds to an edge or an identical region through which the features are extracted. The flowchart for fuzzy based image processing approach is indicated in Figure 2.

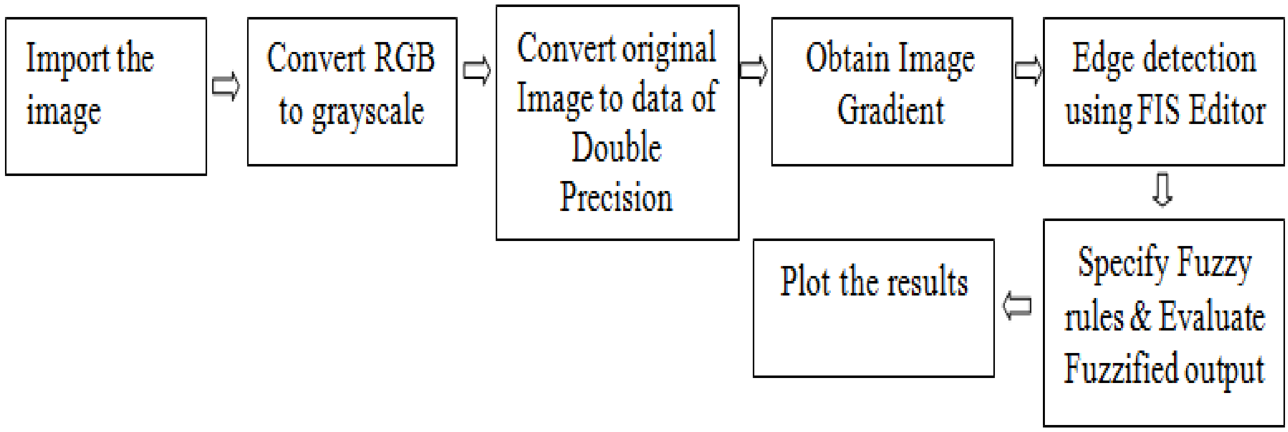


Figure 2. Block Diagram for Fuzzy Logic Image Processing

These features extracted are used for training the ANN trained with BPA. Any complex system can be abstracted in a simple way, or at least dissected to its basic abstract components. Complexity arises by the accumulation of several simple layers. The goal is to explain the working of neural network with a simple example. A supervised neural network can be presented as a black box with two methods to learn and predict the following:

Step 1- ANN Model initialization

Step 2- Propagate the inputs in the forward direction from input to output layer

Step 3- Compute the deviation

Step 4- Apply Generalized Delta Rule

Step 5- Back-propagation

Step 6- Modify the weights

Step 7- Iterate till the ANN converges with a minimum value of Mean Squared Error (MSE)

V. METHODOLOGY

The methodology includes collection of data base consisting of images corresponding to various phyla. These images are then preprocessed for noise removal. Edges are detected to extract useful information from these images. This useful information represents the features. These features are used as inputs to train the ANN with BPA. The schematic is shown in Figure 3.

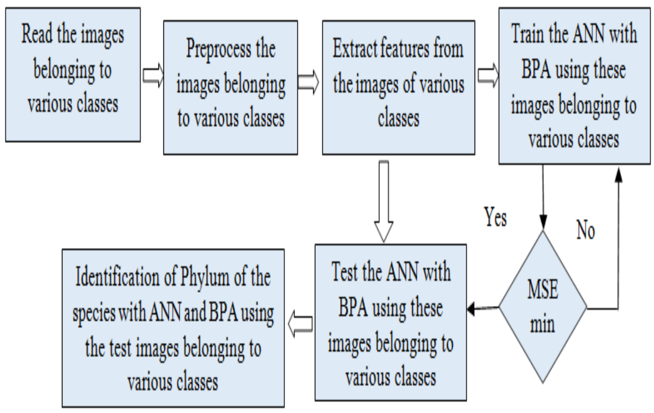
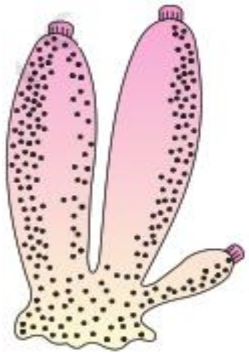
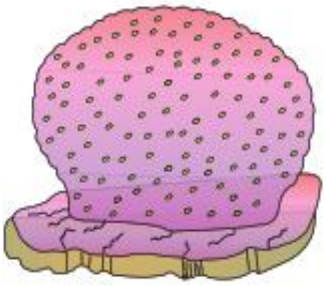
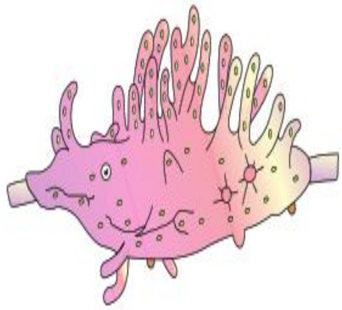
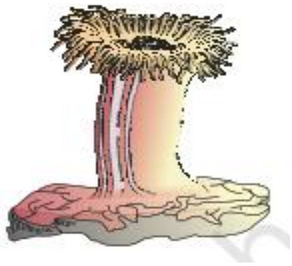
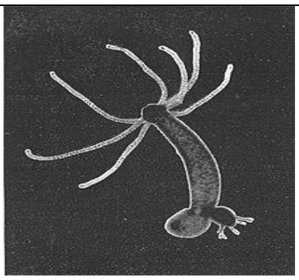
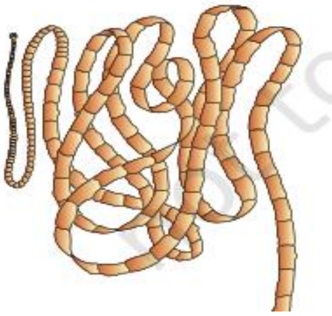
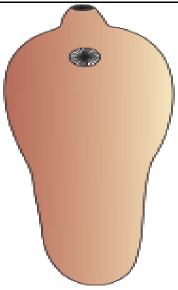
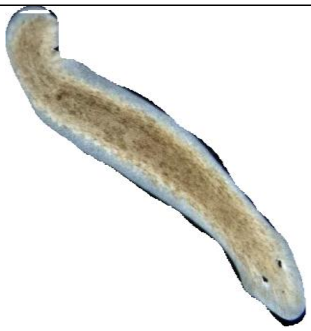
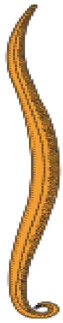
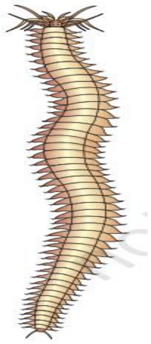
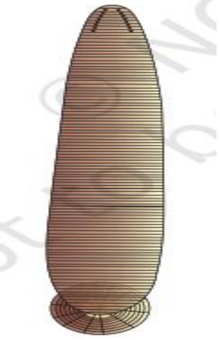
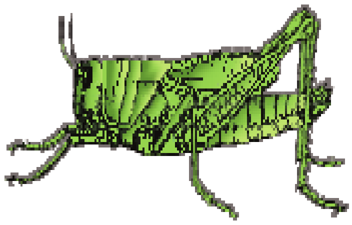
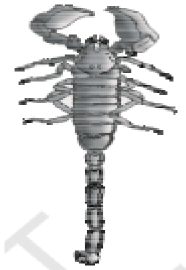
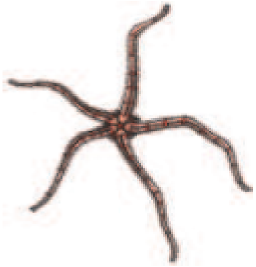
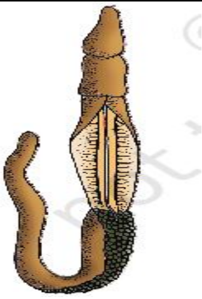


Figure 3. Schematic for Image Processing based Phylum Identification in Animal Kingdom

V. RESULTS AND DISCUSSION

The data base consisting of various images belonging to different classes in Animal Kingdom is gathered from the open source data repository and the samples are displayed in Table 1.

S.No.	Phylum	Name of the Species		
1.	Porifera	 Sycon	 Euspongia	 Spongilla
2.	Coelenterata	 Aurelia	 Adamsia	 Hydra
3.	Platyhelminthes	 Taenia	 Fasciola	 Planaria
4.	Aschelminthes	 Ascaris lumbricoides	 Ancylostoma duodenale	 Wuchereria bancrofti
5.	Annelida	 Nereis	 Hirudinaria	 Polychaetes

6.	**Arthropoda**	 Locust	 Butterfly	 Scorpion
7.	**Mollusca**	 Pila	 Octopus	 Oyster
8.	**Echinodermata**	 Asterias	 Ophiura	 Echinus
9.	**Hemichordata**	 Balanoglossus	 Saccoglossus	
10.	**Chordata**	 Ascidia	 Petromyzon	 Rana Hexadactyla

Nearly 1000 images of the species belonging to various classes are taken from the open source. About 750 images are used for training, 150 images for testing and the remaining 100 images for validation of the ANN architecture trained with BPA. These images are preprocessed for which the edges are detected. Followed by this, the features are extracted from these noise free images which are used as inputs to train the ANN with BPA. The results for training, testing and validation are shown in Figure 4, 5 and 6 respectively.

Retrieval Number C6840098319/2019©BEIESP
DOI: 10.35940/ijrte.C6840.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

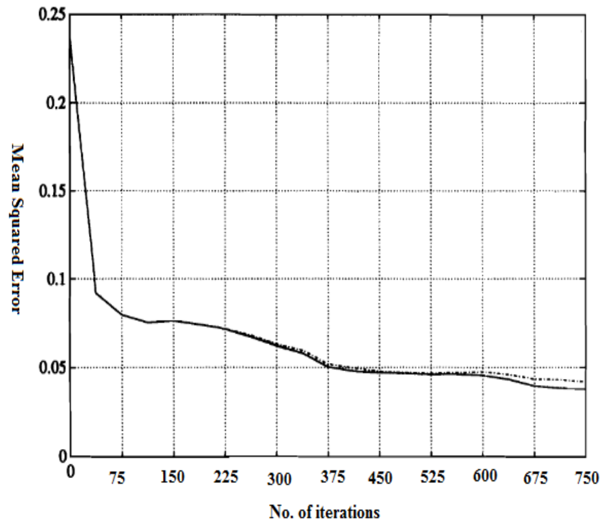


Figure 4. Output for Training the ANN using BPA

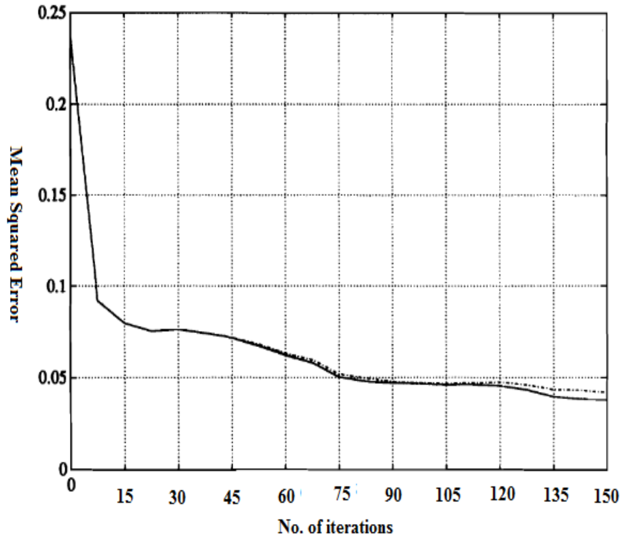


Figure 5. Output for Testing the ANN Architecture using BPA

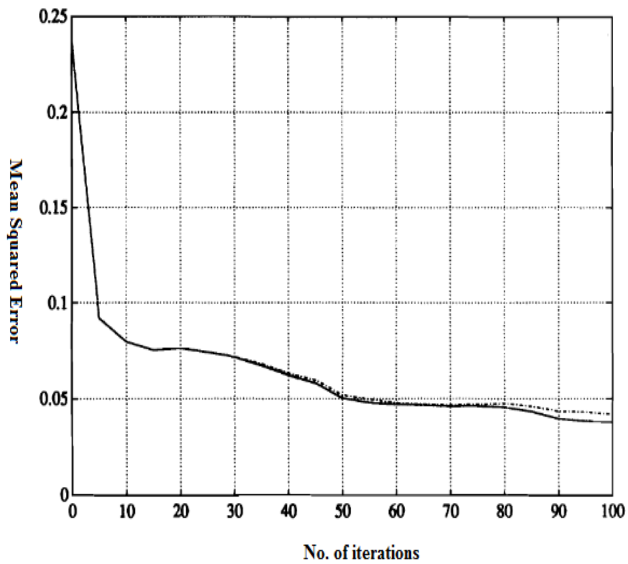


Figure 6. Output for Validation of ANN Architecture

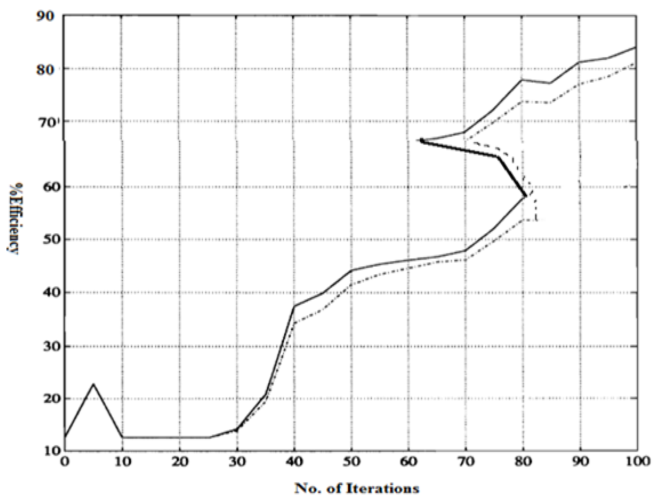


Figure 7. Efficiency for Phylum Identification by the ANN Architecture with BPA

The actual output values from the output layer are matched with the desired output values. The desired output values are used to teach the ANN so that the ANN learns the data pattern corresponding to the various phyla for species identification. For the ANN to converge, the MSE between actual and target output values is found out and propagated back towards the hidden layer and this process continues till the optimal value of MSE is achieved. Each test vector is fed into the input layer and the output is evaluated. The validation is also a similar procedure where the inputs are fed to feed forward network and the outputs are evaluated. The classification efficiency for categorization of phylum for various species is89.1% as inferred from Figure 7.

VI. CONCLUSION

For the ANN trained with BPA, the weights are updated for the input/output vector pair presented to network. The weights are modified after the whole of the training data has undergone one sweep. The final architecture of ANN used, states that, the configuration for the ANN architecture is 103x25x9 which denotes input nodes x hidden nodes x output nodes. Totally 1000 input/output vector pairs are used for training the ANN. The ANN was made to learn the patterns till a minimum value of MSE was achieved with classification efficiency for Phylum identification to be 89.1% with a very short duration of 2.2 seconds for training the data.

REFERENCES

1. McNeill, J., (2012). "International Code of Nomenclature for algae, fungi, and plants (Melbourne Code)", Adopted by the Eighteenth International Botanical Congress, International Association for Plant Taxonomy.
2. "Life sciences". The American Heritage New Dictionary of Cultural Literacy (third ed.). Houghton Mifflin Company. 2005, "Phyla in the plant kingdom are frequently called divisions".

3. Berg, Linda R. (2007). Introductory Botany: Plants, People, and the Environment (2 ed.). Cengage Learning. p. 15. ISBN: 9780534466695.
4. Holt, Jack R.; Iudica, Carlos A. (2016). "Blastocladiomycota". Diversity of Life. Susquehanna University.
5. K. Sujatha, N.P.G. Bhavani, V. Srividhya, "Soft sensor with shape descriptors for flame quality prediction based on LSTM regression", Real time data Analytics for large scale sensor data, Elsevier 2020.
6. Bhavani, NPG.,&Sujatha, K. (2014).Soft Sensor for Temperature Measurement in Gas Turbine Power Plant.International Journal of Applied Engineering Research (pp. 21305-21316).
7. Gang Lu a, Yong Yan a, Steve Cornwell b, Michael Whitehouse b, Gerry Riley b.(2007). Impact of co-firing coal and biomass on flame characteristics and stability, Elsevier (pp. 1133–1140).
8. G Gilabert, G Lu & Y Yan.(2005). Three dimensional visualisation and reconstruction of the luminosity distribution of a flame using digital imaging Techniques, Sensors & their Applications (pp. 167–171).
9. G, Lu.,& Yan.(2009). Advanced Monitoring and Characterization of combustion flames, 20th annual meeting and meeting of the Advanced Power Generation Division, University of Leeds.
10. Sujatha, K., &Pappa, N. (2011). Combustion monitoring of a water tube boiler using a discriminant radial basis network, ISA Transactions (pp. 101–110)..