Detection of Skin disease using Metaheuristic supported Artificial Neural Networks

Shouvik Chakraborty, Kalyani Mali
Department of Computer Science & Engineering
University of Kalyani
Kalyani, Nadia, West Bengal, India
shouvikchakraborty51@gmail.com
kalyanimali1992@gmail.com

Sankhadeep Chatterjee
Student Member, IEEE
Department of Computer Science & Engineering
University of Calcutta
Kolkata, West Bengal, India
chatterjeesankhadeep.cu@gmail.com

Soumen Banerjee, Kaustav Guha Mazumdar, Mainak Debnath, Pikorab Basu, Soumyadip Bose
Student Member, IEEE
Department of Electronics & Communication Engineering
University of Engineering & Management
New Town, Kolkata, West Bengal, India
prof.sbanerjee@gmail.com

Kyamelia Roy
Department of Electronics & Communication Engineering
Future Institute of Engineering & Management
Sonarpur, Kolkata, West Bengal, India
kyamelia2015@gmail.com

***Abstract*—Automated, efficient and accurate classification of skin diseases using digital images of skin is very important for bio-medical image analysis. Various techniques have already been developed by many researchers. In this work, a technique based on meta-heuristic supported artificial neural network has been proposed to classify images. Here 3 common skin diseases have been considered namely angioma, basal cell carcinoma and lentigo simplex. Images have been obtained from International Skin Imaging Collaboration (ISIC) dataset. A popular multi objective optimization method called Non-dominated Sorting Genetic Algorithm - II is employed to train the ANN (NNNSGA-II). Different feature have been extracted to train the classifier. A comparison has been made with the proposed model and two other popular meta-heuristic based classifier namely NN-PSO (ANN trained with Particle Swarm Optimization) and NN-GA (ANN trained with Genetic algorithm). The results have been evaluated using various performances measuring metrics such as accuracy, precision, recall and F-measure. Experimental results clearly show the superiority of the proposed NN-NSGA-II model with different features.**

Keywords—Skin disease detection, neural network, hybrid methods, meta-heuristic

I. INTRODUCTION

Skin related diseases including skin cancers are the most frequent diseases in the whole world. It has been found that one in every five Americans suffers from skin cancer in their whole lifetime. Among various categories of skin related diseases, malignant melanoma is one of the major reasons for more than 10000 deaths annually in the USA [1]. However, in most of the cases, early diagnosis and treatment can completely cure the disease [2]. Early stage detection of the disease type is important since appropriate treatments can be applied based on the type of the disease [3]. Dermoscopy is one of the most popular and important non-invasive method for identification of pigmented skin lesions. Examinations performed by naked eye have some drawbacks like accuracy related issues, limitations of human observer etc. Computer aided techniques can efficiently analyze and reveal subsurface structures of the skin lesions. This method gives better differentiation among different types of lesion their appearance and features. However, visual experiments and observations of dermoscopic images by experts depend on the experience of the expert. To reduce the diagnostic errors and overcome the difficulty and subjectivity of human understanding, computerized tools for dermoscopy images has become an attractive research domain. Various computer assisted methods for lesion segmentation and classification are dependent on lesion edge computation and some general features. That is why some of the methods cannot provide generalization power due to variations in dermoscopic image samples. The major reasons behind these variations are different angles, irregular zooming, lightening conditions etc. Moreover, images may contain some artifacts which can affect the classification results. Different methods have been developed to analyze and segment biomedical images of different modalities [4, 5, 6]. Different contour based methods can be used in the analysis of biomedical images [7].

In this work, a computer aided automatic method is proposed, which applies NSGA II trained neural networks for skin lesions classification. The network has been trained by different images obtained from the ISIC 2017 challenge. The trained NSGA II supported classifier was then employed on

978-1-5386-2215-5/17/$31.00 ©2017 IEEE

different unlabeled images. The obtained results have been compared with other methods like NN-PSO, NN-GA.

II. Feature Extraction

Feature extraction is one of the major step in any classification oriented problems [8]. Features are the key for both training and testing purposes. To extract features, at the beginning, Single-level discrete 2-D wavelet transform has been applied 3 times using Daubechies DB4 filter on the binary image. It is illustrated in Fig. 1.The Principle Component Analysis technique has been employed on the approximation coefficients matrix cA and details coefficients matrices cH, cV, and cD. Create gray-level co-occurrence matrix from the matrix obtained by applying PCA. Then using the matrix obtained by PCA, 13 distinct features have been selected and computed as described below.

(a) Mean: It is simply the mean of the matrix elements.

(b) Standard Deviation: We have to compute the standard deviation of matrix elements.

(c) Entropy: Entropy is a statistical metric of randomness that can be used to find characteristic of the texture of the input skin image. Entropy is defined as given in equation 1.

$$Entropy = -sum(p * \log_2 p) \tag{1}$$

where p contains the count of the histogram.

(d) RMS: The root-mean-square level of a vector X, is defined as equation 2.

$$X_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |X_n|^2} \tag{2}$$

(e) Variance: The variance is the square of the standard deviation.

(f) Smoothness: It is defined as equation 3.

$$Smoothness = 1 - \frac{1}{1-a} \tag{3}$$

where a is sum of the matrix obtained from PCA.

(g) Kurtosis: It is a measures how a distribution is prone to outlier of a distribution. The kurtosis of the normal distribution is 3. If a distribution is more prone to outlier than the normal distributions, then the value of the kurtosis will greater than 3 for that distribution; similarly distributions that are less prone to outlier have kurtosis value less than 3. The kurtosis of a distribution is defined as equation 4.

$$k = \frac{E(x-\mu)^4}{\sigma^4} \tag{4}$$

where μ is the mean of x, σ is the SD of x, and $E(t)$ represents the expected value of the term t. kurtosis computes a sample version of this population value.

(h) Skewness: Skewness is a metric that measures the asymmetry of the data around the mean. If skewness value is less than zero, the data are distributed more to the left side of the mean compared to the right. Similarly, if the value of the skewness is greater than 0, then we can conclude that the data are spread out more to the right side of the sample mean than left. The skewness of the normal distribution (or any other perfectly symmetric distribution) is zero. The skewness of a distribution is given in equation 5.

$$s = \frac{E(x-\mu)^3}{\sigma^3} \tag{5}$$

where μ is the mean of x, σ is the SD of x, and E(t) denotes the expected value of the parameter t. Skewness generally computes a sample version of this population value.

(i) Inverse Difference Movement: It is defined as equation 6.

$$IDM = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{G(i,j)}{1+(i-j)^2} \tag{6}$$

(j) Contrast: It is a metric to compute the intensity contrast between a pixel and its neighbor over the whole image. It is defined as equation 7.

$$contrast = \sum_{i,j} |i-j|^2 \; p(i,j) \tag{7}$$

For a constant image, the value of the contrast is 0.

(k)Correlation: It is a metric to measure how a pixel is correlated to its neighbor over the total image. It is defined as equation 8.

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j} \tag{8}$$

The value of the correlation is undefined for a constant image.

(l) Energy: It is the sum of squared elements in the gray-level co-occurrence matrix. It is defined as equation 9.

$$Energy = \sum_{i,j} p(i,j)^2 \tag{9}$$

For the constant image, the value of the energy is 1.

(m) Homogeneity: It is metric to measure the closeness of the distribution of elements in the gray-level co-occurrence matrix to the gray-level co-occurrence matrix diagonal. It is defined as equation 10.

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{10}$$

III. Proposed Work

The NN [9] is neuron interaction inspired model that can be efficiently used to learn complex patterns. Understanding patterns lead to future prediction of certain value. The basic structure of a typical NN consists of several artificial neurons that are arranged into multiple layers [10 - 14]. Neurons of a layer connect with the neurons of its preceding and succeeding layers. During the training phase, the data instances are fed to the ANN and the response is compared with the expected outcome. The difference between the expected outcome and response given by the NN is used as an error measure. This error quantity is utilized to adjust the weights associated with the NN to improve its performance [15 - 20]. The same process is carried out for each data instance. At the last step of training phase, it is expected to have the optimal set of weights that used to minimize the misclassification probability. Throughout the testing stage, the NN is tested with a set of unknown data and its responses are checked to measure its performance. Fig.2 depicted a typical NN.

The traditional training algorithms do not ensure the optimal training as they may trap in local optima while searching for the best solution. For ensuring that the NN obtains the optimal set of weights during the training phase ensures, the error is to be minimized, which can be considered as an optimization problem. In the current work, minimizing the error is used as the objective function to attain the optimal weights set. Thus, meta-heuristic algorithms can be employed to ensure convergence to the global optima. Consequently, meta-heuristic optimization methods, such as SA based methods [21, 22, 23], GA [24, 25], PSO [26, 27], and NSGA-II [28, 29] can be employed to train the NN.

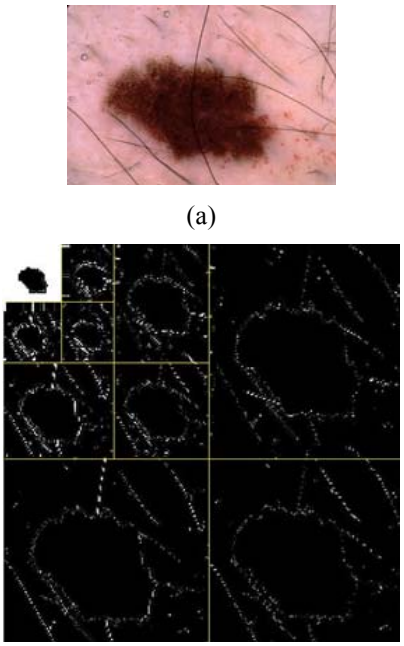

Fig. 1. (a) Original Image (b) After Applying Daubechies DB4 filter 3 times on the binary image

Generally, several problems involve multiple objectives optimization at the same time to achieve a potent solution The multi-objective optimization [30] can formally be defined by finding the vector $\overrightarrow{x_p} = [x_1, x_2, \ldots, x_n]^T$ of n decision variables such that

$$\overrightarrow{f(\overrightarrow{x_p})} = \left[f_1(\vec{x}), f_2(\vec{x}), \ldots, f_n(\vec{x}) \right]^T$$

satisfies some constraints. The vector $\overrightarrow{x_p}$ is considered Pareto optimal if only there is no $\vec{x}$ exist such that $\forall i \in \{1, 2, \ldots, n\}$, $f_i(\vec{x}) \leq f_i\left(\overrightarrow{x_p}\right)$ and there exist at least one i such that $f_i(\vec{x}) < f_i\left(\overrightarrow{x_p}\right)$ .The set of solutions which is generated due to the Pareto optimality are generally addressed as non-dominated solutions. The working principle of multi-objective GA (NSGA-II) is analogous to the GA. However, the optimization process in the NSGA-II tries to optimize multiple objectives, which are sometimes contradictory to each other. A general NSGA-II algorithm has been depicted in Fig. 3.

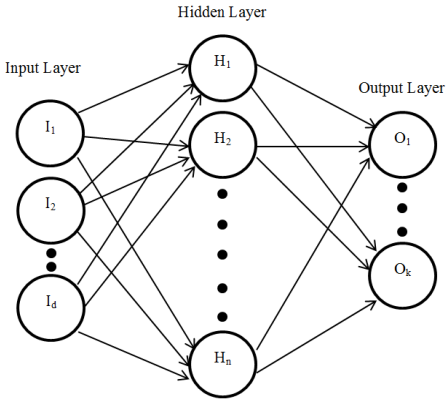

Fig. 2. One hidden layer NN structure

The selection process comes after a ranking of the current population, which is achieved in the current work by a non-dominated sorting [31] that used in the NSGA-II algorithm. For selection purpose a binary tournament method has been utilized. In NSGA-II, the non-dominated sorting process is faster than the previous version. In addition, it also assures elitism of next generation, which is a key factor of successful and efficient convergence of the algorithm.

IV. Dataset Description

The International Skin Imaging Collaboration (ISIC) is one of the popular efforts to improve and enhance the skin disease diagnosis [32]. It has recently started efforts to aggregate an open access dataset of dermoscopy images. The dataset contains more than 25000 images of different types of diseases. The images are screened for both privacy and quality

standards. The clinical metadata is associated with each image which has been provided by the renowned skin disease experts. Clinical relevance of the test samples can be ensured by the various national and international participation in image contribution. One of the major goals of this challenge was to give a snapshot from the ISIC Archive to enhance development of automated skin disease diagnosis procedures from dermoscopic images. The images are of different resolutions (from 1022 X 767 to 6748 X 4499), angles of photography and lightening conditions. Some images of this dataset contain some artifacts. Three types of images have been considered for this work namely Angioma, Basal Cell Carcinoma, Lentigo Simplex. Sample test images have been shown in Fig. 4.

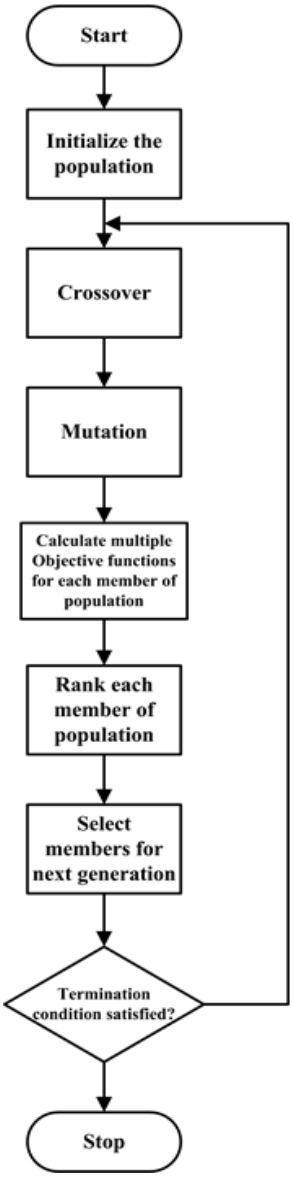

Fig. 3. Flow chart of NSGA-II

V. Results and Discussions

The simulations are carried out by using the proposed NN-NSGA-II model as explained in section II. The flow of the training phase is as in [33]. For single objective optimization RMSE [34] and for multi objective optimization RMSE along with Maximum error has been used as objective function [33]. The experimental setup of NN-GA and NN-PSO is as in [34, 35]. The performances of the classifiers are measures in terms of confusion matrix based performance measuring methods.

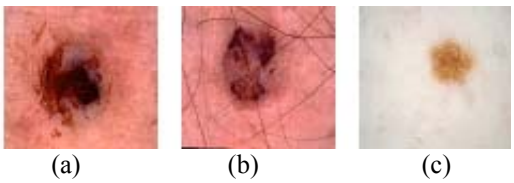

Fig. 4. Different types of skin diseases under test: (a) Angioma (b) Basal Cell Carcinoma (c) Lentigo Simplex

Table I reports the comparative study of the proposed method with NN-GA and NN-PSO. The performance of NN-GA is poor with 72.02% accuracy. However, the precision of the same is 93.6%, the recall and F-measure reveals that the classifier is actually poor at detecting the skin diseases under study. NN-PSO reports a significant improvement with an accuracy of 82.42% and 93.8% precision, 81.54% recall and 87.24% F-measure. The performance of NN-NSGA-II is superior to other classifiers under study. It achieved 87.92% accuracy with 94.2% precision, 87.5% recall, and 90.73% F-measure. Graphical representation of the results is given in Fig. 5.

TABLE I. Comparison of Performance Measures of NN-NSGA-II with NN-GA and NN-PSO

Parameters	NN-GA	NN-PSO	NN-NSGA-II
Accuracy	72.02	82.42	87.92
Precision	93.6	93.8	94.2
Recall	61.91	81.54	87.5
F-Measure	74.53	87.24	90.73

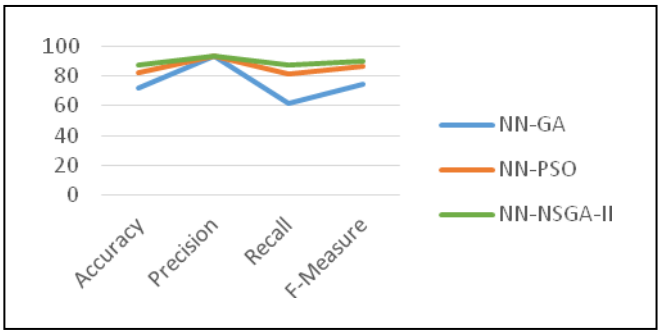

Fig. 5. Graphical Representation of the results

VI. Conclusion

In this work, the application of metaheuristic supported artificial neural network based architecture has been investigated, which were employed and trained for dermoscopic image classification. Results of 3 different methods have been compared in table 1 and Fig.5. From the obtained results, it can be concluded that NSGA II trained ANN gives better results than the other methods with 87.92% of accuracy. Same approach on a larger training data set could potentially lead to a better predictive model.

References

[1] M. F. Møller. "A scaled conjugate gradient algorithm for fast supervised learning",. Neural networks, vol. 6, no. 4, pp. 525-533, 1993.

[2] S. Chatterjee, S. Sarkar, S. Hore, N. Dey, A. S. Ashour, and V. E. Balas, "Particle swarm optimization trained neural network for structural failure prediction of multistoried RC buildings". Neural Computing and Applications, pp. 1-12. 2016.

[3] S. Chatterjee, S. Ghosh, S. Dawn, S. Hore, and N. Dey, "Forest Type Classification: A hybrid NN-GA model based approach" Information Systems Design and Intelligent Applications, pp. 227-236, Springer India. 2016

[4] S. Chakraborty, S. Chatterjee, N. Dey, A. S. Ashour, A. S. Ashour, F. Shi and Kalyani Mali, "Modified Cuckoo Search Algorithm in Microscopic Image Segmentation of Hippocampus", Microscopy Research and Technique, 00; pp. 1-22, 2017

[5] S. Hore, S. Chakraborty, A. S. Ashour, N. Dey, A. S. Ashour, D. S. Pistolla, T. Bhattacharya, and S. R. Chaudhuri, "Finding Contours of Hippocampus Brain Cell Using Microscopic Image Analysis", Journal of Advanced Microscopy Research vol. 10, no. 2, pp. 93-103, 2015.

[6] S. Hore, S. Chakraborty, S. Chatterjee, N. Dey, A. S. Ashour, L. V. Chung, and D. N. Le. "An Integrated Interactive Technique for Image Segmentation using Stack based Seeded Region Growing and Thresholding" International Journal of Electrical and Computer Engineering (IJECE) vol. 6, no. 6. 2016.

[7] S. Chakraborty, M. Roy and S. Hore, "A Study on Different Edge Detection Techniques in Digital Image Processing", Feature Detectors and Motion Detection in Video Processing, pp. 100-122, 2016.

[8] S. Hore, S. Chatterjee, S. Chakraborty and R. K. Shaw, "Analysis of Different Feature Description Algorithm in object Recognition", Feature Detectors and Motion Detection in Video Processing, pp. 66-99, 2016.

[9] S. Hore, S. Chatterjee, S. Sarkar, N. Dey, A. S. Ashour, D. B. Timar and V. E. Balas , "Neural-based Prediction of Structural Failure of Multi Storied RC Buildings", Structural Engineering and Mechanics, Vol. 58, No. 3, pp.459-473, May, 2016.

[10] S. Banerjee, S. Chatterjee, and R. Chakraborty. "Studies on the performance characteristics of a typical Collision-free Transportation System."

[11] S. Hore, S. Chatterjee, S. Chakraborty and R. K. Shaw, "Analysis of Different Feature Description Algorithm in object Recognition", Feature Detectors and Motion Detection in Video Processing, pp. 66-99, 2016.

[12] S. Hore, S. Chatterjee, N. Dey, A. S. Ashour, and V. E. Balas. "A Vulnerability Analysis Mechanism Utilizing Avalanche Attack Model for Dependency-Based Systems." In Advances in Computing Applications, pp. 243-262. Springer Singapore, 2016.

[13] S. Chatterjee, R. Chakraborty, N. Dey, and S. Hore. "A quality prediction method for weight lifting activity." pp. 95-8. 2015.

[14] S. Chatterjee, and D. Banerjee. "A Novel Boolean Expression based Algorithm to find all possible Simple Paths between two nodes of a Graph." International Journal of Advanced Research in Computer Science 5, no. 7, 2014.

[15] S. Chatterjee, Sankhadeep, S. Paladhi, S. Hore, and N. Dey. "Counting all possible simple paths using Artificial Cell division mechanism for directed acyclic graphs." In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on, pp. 1874-1879. IEEE, 2015.

[16] S. Chatterjee, S. Paladhi, and R. Chakraborty. "A Comparative Study On The Performance CharacteristicsOf Sudoku Solving Algorithms." IOSR Journals (IOSR Journal of Computer Engineering) vol. 1, no. 16: pp 69-77.

[17] S. Hore, S. Chakraborty, S. Chatterjee, N. Dey, A. S. Ashour, L. V. Chung, and D. N. Le. "An Integrated Interactive Technique for Image Segmentation using Stack based Seeded Region Growing and Thresholding." International Journal of Electrical and Computer Engineering (IJECE) vol 6, no. 6, pp. 2773-2780, 2016.

[18] S. Hore, S. Chatterjee, V. Santhi, N. Dey, A. S. Ashour, V. E. Balas, and F. Shi. "Indian Sign Language Recognition Using Optimized Neural Networks." In Information Technology and Intelligent Transportation Systems, pp. 553-563. Springer International Publishing, 2017.

[19] S. Chatterjee, R. Chakraborty, S. Hore: "A Quality Prediction method of Weight Lifting Activity". Proceedings of the Michael Faraday IET International Summit - 2015 (MFIIS-2015), An IET International Conference On September, pp. 12 - 13, 2015 in Kolkata, India.

[20] S. Chatterjee, S. Sarkar, N. Dey, A. S. Ashour, S. Sen: "Hybrid Non-dominated Sorting Genetic Algorithm-II-Neural Network approach toward Water Quality Classification", Advancements in Applied Metaheuristic Computing, IGI Global, June 2017.

[21] S. Chakraborty and S. Bhowmik, "Job Shop Scheduling using Simulated Annealing", In Proc. 1st International Conference IC3A 2013, JIS College of Engineering, Kalyani, India, pp. 69-73, McGrawHill Publication, January 2013.

[22] S. Chakraborty and S. Bhowmik, "Blending Roulette Wheel Selection with Simulated Annealing for Job Shop Scheduling Problem", In Proc. Michael Farady IET International Summit MFIIS 2015, Kolkata, India, vol. 2, IET, September 2015.

[23] S. Chakraborty, S. Bhowmik, "An Efficient Approach to Job Shop Scheduling Problem using Simulated Annealing", International Journal of Hybrid Information Technology, Vol. 8 No. 11, pp. 273-284, 2015.

[24] S. Chatterjee, R. Nag, N. Dey and A. S. Ashour, "Efficient Economic Profit Maximization: Genetic Algorithm based Approach". World Conference on Smart Trends in Systems, Security and Sustainability (WS4 2017) At London, UK, Volume: Springer LNNS series, February 2017. (In press)

[25] S. Chakraborty, A. Seal and M. Roy, "An Elitist Model for Obtaining Alignment of Multiple Sequences using Genetic Algorithm", In Proc. 2nd National Conference NCETAS 2015, MIET, Bandel, India, pp. 61-67, International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), 2015.

[26] S. Chatterjee, S. Banerjee, P. Basu, M. Debnath and S. Sen, "Cuckoo Search coupled Artificial Neural Network in Detection of Chronic Kidney Disease", International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, 2017. (in press)

[27] S. Chatterjee, S. Hore, N. Dey, S. Chakraborty and A. S. Ashour, "Dengue Fever Classification using Gene Expression Data: A PSO based Artificial Neural Network Approach". 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications; Springer, June 2016.

[28] S. Chatterjee, S. Banerjee, K. G. Mazumdar, S. Bose and S. Sen, "Non-dominated Sorting Genetic Algorithm II supported Neural Network in Classifying Forest Types", International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, 2017. (in press)

[29] S. Chatterjee, S. Sarkar, S. Hore, N. Dey and A. S. Ashour, F. Shi, D. N. Le, "Structural Failure Classification for Reinforced Concrete Buildings Using Trained Neural Network based Multi-Objective Genetic Algorithm", Structural Engineering and Mechanics, Techno Press 2017. (in press)

[30] C. A. C. Coello, "A comprehensive survey of evolutionary-based multiobjective optimization techniques", Knowledge and Information systems, vol. 1, no. 3, pp. 129-156. 1999.

[31] K. Deb, “Nonlinear goal programming using multi-objective genetic algorithms”. Journal of the Operational Research Society, pp. 291-302. 2001.

[32] International Skin Imaging Collaboration Website. Available: http://www.isdis.net/index.php/isic-project

[33] S. Chatterjee, S. Sarkar, N. Dey, S. Sen, T. Goto and N. C Debnath, “Water Quality Prediction: Multi Objective Genetic Algorithm coupled Artificial Neural Network based approach", IEEE 15th International Conference of Industrial Informatics INDIN'2017, Emden, Germany, July 2017. (in press)

[34] S. Chatterjee, N. Dey, A. S. Ashour and C. V. A. Drugarin, “Electrical Energy Output prediction using Cuckoo Search supported Artificial Neural Network". World Conference on Smart Trends in Systems, Security and Sustainability (WS4 2017) At London, Volume: Springer LNNS series, February 2017. (In press)

[35] S. Chatterjee, R. Nag, S. Sen and A. Sarkar, “Towards Golden Rule of Capital Accumulation: A Genetic Algorithm Approach". 16th International Conference on Computer Information Systems and Industrial Management Applications (CISIM 2017), Bialystok, Poland. (in press).