**1. Introduction**

Melanoma is an aggressive type of skin cancer that can spread to other organs. Automatic and early diagnosis of melanoma is essential for administering effective treatment and increasing survival chances. Since medical skin cancer diagnosis employs the Asymmetry, Border, Colour, Diameter and Enlargement (ABCDE) guideline, the extraction and identification of such discriminative and significant morphological characteristics play a crucial role in attaining accurate diagnosis rates. However, it is still a challenging task for the retrieval of such distinguishing attributes, owing to the fine-grained variability in the appearance of benign and cancerous skin lesions.



Fig 1.1: Skin Cancer

Evolutionary algorithms[2] possess powerful search capabilities, and have been widely used for solving various feature selection challenges. Owing to the comparatively simple underlying concepts and relatively few user-defined parameters, WOA[3] has been widely studied for feature selection tasks.

**2. Literature and Methodology**

**2.1 Whale Optimisation Algorithm**

WOA is gaining its interest in many application domains and it fits in to the stochastic population-based algorithm[3]. WOA imitates the bubble-net feeding characteristic that is present in the foraging behaviour of the whales. The whale chases near to the region with entrapping the victim in a net of bubbles. Formerly this has been created while swimming in a ‘6’-shaped manner[3].

WOA involves two stages namely

1. Exploitation stage (encircling a victim and coiling bubble-net offensive manner).
2. Exploration stage (probing randomly for the victim).

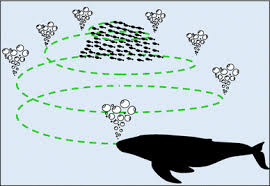


Fig 2.1: Bubble net whale hunting behaviour

WOA is employed for feature selection by which each feature subset is considered to be a position of a whale. Every subset possibly has N number of features, where Nis the number of features in the original set. The less the number of features in the solution and the higher the classification accuracy, the better is the solution. WOA begins with a set of randomly generated solutions (subsets) called population. After that, the available solution will be assessed by the employed fitness function

**2.2 Fitness Function**

The fitness function employed in this research work is modeled by poising between the number of selected features in each solution (minimum) and the classification accuracy (maximum) obtained by using these selected features as depicted in Eq. (1)

------1

Where *A* is accuracy and *n* is no of features selected, are constants in [0,1] and =1.

* 1. **Proposed Skin Cancer Detection System**

We propose an intelligent system for benign and malignant skin lesion classification. The proposed system consists of five key stages.

1. Pre-processing
2. Feature extraction
3. BWAO-based feature selection
4. Classification

**Flow Chart**

Pre-processing

Feature Extraction

GLRLM

Color

Shape

Feature Selection

Binary Whale Optimization Algorithm

Deep Learning

Classification

GLCM

Noise Removal

Median Filtering

Gaussian Filtering

Segmentation

OTSU Thresholding

Skin Cancer Images

**3. Work done**

**3.1 Pre-Processing**

The pre-processing techniques[1] are applied to lesion images for noise filtering. Specifically, to remove ‟salt and pepper‟ noise, median filtering is performed by diminishing the effects of thin hairs and air bubbles.

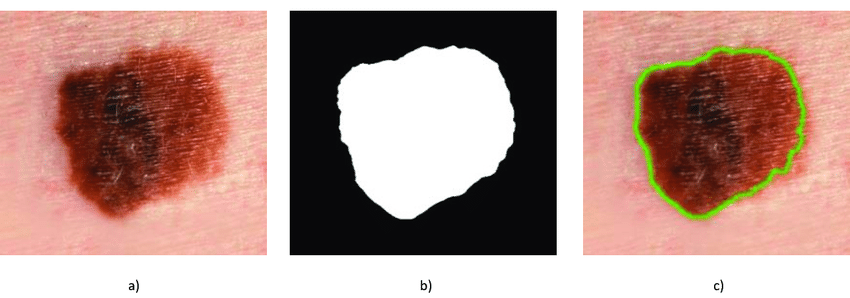


Fig 3.1: Lesion area segmentation

The grayscale conversion transforms the original RGB images into grayscale. OTSU threshold method[1] is used to obtain the lesion area.

**3.2 Feature Extraction**

Several feature extraction methods are used to extract shape, colour, and texture features from the separated lesion regions[1]. According to the ABCDE guideline for clinical skin cancer diagnosis, three categories of features play very important roles in distinguishing between benign and malignant lesions, namely

1. Shape features such as asymmetry, border irregularity, and compactness.
2. Color features such as relative chromaticity, and differences in

lightness and color.

1. Gray Level Run Length Matrix (GLRLM) based texture features. As such, these shape, color, and texture features are extracted in this research.

The RGB colour space is employed for feature extraction. Color features such as relative chromaticity and ratio of red, green and blue, and factors exhibited with respect to the lesion’s tone are extracted. The following colour features are also retrieved, i.e. variance, entropy, skewness, correlation, Principal Component Analysis (PCA) variance, mean of image darkness, variance of image darkness, mean & standard for both lesion and skin, and average colour of red, green and blue. Moreover, four orientations (0, 45, 90, and 135) of the GLRLM-based texture features are retrieved, whereby each level embeds 11 different emphases. These 11 statistics include Short Run Emphasis, Long Run Emphasis, Gray-Level Nonuniformity, Run Length Nonuniformity, Run Percentage, Low Gray-Level Run Emphasis, High Gray-Level Run Emphasis, Short Run Low Gray-Level Emphasis, Short Run High Gray-Level Emphasis, Long Run Low Gray-Level Emphasis, and Long Run High Gray-Level Emphasis. Pyradiomics was used to extract shape features, GLCM, GLRLM based texture features. Overall, we have extracted 102 features from the images.

**3.3 Feature Selection using BWOA**

Since the features are not equally important for the identification of lesion types, we propose a BWOA[4] model for feature optimization and dimensionality reduction. The aim is to identify the most discriminative features and remove redundant ones.

The original WOA[3] algorithm has the disadvantage of slow convergence and easy to fall into the local optimal in the search process, so BWOA[4] algorithm is used in this feature selection[6].

In this section the mathematical model of encircling prey, spiral bubble-net feeding maneuver, and search for prey is first provided.

**3.3.1. Encircling prey**

Humpback whales can recognize the location of prey and encircle them. Since the position of the optimal design in the search space is not known a priori, the WOA[4] algorithm assumes that the current best candidate solution is the target prey or is close to the optimum. After the best search agent is defined, the other search agents will hence try to update their positions towards the best search agent. This behaviour is represented by the following equations:

-------2

-------3

where indicates the current iteration, and are coefficient vectors, is the position vector of the best solution obtained so far, is the position vector, | | is the absolute value. It is worth mentioning here that should be updated in each iteration if there is a better solution.

-------4

-------5

where a is linearly decreased from 2 to 0 over the course of iterations (in both exploration and exploitation phases) and is a random vector in [0,1].

**3.3.2 Binary Transfer Function for Updating Positions**

Due to the difference in the position update method of the WOA algorithm in the binary space, an association rule needs to be established to achieve conversion of the whale position between 0 and 1. The binary transfer function for update is as follows.

-------6

-------7

Where denotes the feature of the j-th dimension of the i-th whale, and is a random number within [0, 1].

**3.3.3 Bubble-net attacking method (exploitation phase)**

In order to mathematically model the bubble-net behaviour of humpback whales, two approaches are designed as follows:

**1.Shrinking encircling mechanism:** This behaviour is achieved by decreasing the value of in the Eq.(4). Note that the fluctuation range of is also decreased by . In other words is a random value in the interval [−,] where is decreased from 2 to 0 over the course of iterations. Setting random values for in [−1,1], the new position of a search agent can be defined anywhere in between the original position of the agent and the position of the current best agent.

**2.Spiral updating position:** This approach first calculates the distance between the whale located at (*X,Y*) and prey located at (*X\*,Y\*).* A spiral equation is then created between the position of whale and prey to mimic the helix-shaped movement of humpback whales as follows:

-------8

where and indicates the distance of the i-th whale to the prey (best solution obtained so far), *b* is a constant for defining the shape of the logarithmic spiral, is a random number in [−1,1].

Note that humpback whales swim around the prey within a shrinking circle and along a spiral-shaped path simultaneously. To model this simultaneous behaviour, we assume that there is a probability of 50% to choose between either the shrinking encircling mechanism or the spiral model to update the position of whales during optimization. The mathematical model is as follows:

-------9

where is a random number in [0,1]. In addition to the bubble-net method, the humpback whales search for prey randomly. The mathematical model of the search is as follows.

**3.3.4 Search for prey (exploration phase)**

The same approach based on the variation of the vector can be utilized to search for prey (exploration). In fact, humpback whales search randomly according to the position of each other. Therefore, we use with the random values greater than 1 or less than −1 to force search agent to move far away from a reference whale. In contrast to the exploitation phase, we update the position of a search agent in the exploration phase according to a randomly chosen search agent instead of the best search agent found so far. This mechanism and || > 1 emphasize exploration and allow the WOA algorithm to perform a global search. The mathematical model is as follows:

-------10

-------11

where is a random position vector (a random whale) chosen from the current population.

**3.4. Classification**

Deep learning methods[5] aim at learning feature hierarchies, where features at higher levels of the hierarchy are formed using the features at lower levels. Much better results could be achieved in deeper architectures when each layer is pretrained with an unsupervised learning algorithm. Then the network is trained in a supervised mode using back-propagation algorithm to adjust weights.

Deep neural network is used for classification[5] by giving selected features as input for every whale in each iteration. 10 fold cross validation is employed in this deep neural network.

**4. Results**

The proposed model is implemented with various fitness functions. To evaluate the efficiency of each strategy in the proposed BWOA model on skin cancer detection, we conducted evaluation dedicated to the HAM10000 data set. And we got best accuracy of 85.889 with fitness function in Eq(1) and it has selected 38 features.

**5. Conclusion**

In this project, we proposed an approach for skin cancer detection using bio inspired algorithm for feature selection. We used median filtering to remove ‟salt and pepper‟ noise and OTSU threshold method to obtain the lesion area. Features are extracted from the segmented images. We have applied BWOA algorithm on the extracted features to get more relevant features[6] and these features are fed to a deep neural network[5] for classification.

We are planning to obtain better results by finding novel fitness function.

**6. References**

[1] J. Glaister, A. Wong, D.A. Clausi, “Segmentation of Skin Lesions from Digital Images Using Joint Statistical Texture Distinctiveness”, *IEEE Transactions on Biomedical Engineering*, 61(4), pp.1220–1230, 2014.

[2] Teck Yan Tan, Li Zhang, Siew Chin Neoh, Chee Peng Lim, “Intelligent Skin Cancer Detection Using Enhanced Particle Swarm Optimization”, *Knowledge-Based Systems,* pp.118-135, October 2018.

[3] Gehad Ismail Sayed, Ashraf Darwish, Ella Hassanien, “A New Chaotic Whale Optimization Algorithm for Features Selection”, *Journal of Classification*, pp.300-344, 2018.

[4] Seyedali Mirjalili, Andrew Lewis, “The Whale Optimization Algorithm”, *Advances in Engineering Software*, pp.238-253, 2016.

[5] F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng, A. Bovik, "Melanoma classification on dermoscopy images using a neural network ensemble mode", IEEE Trans. Med. Imag., vol. 36, no. 3, pp. 849-858, Mar. 2017.

[6]  J. Zhang, Y. Xia, Y. Xie, M. Fulham, D. D. Feng, "Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features", IEEE J. Biomed. Health Infom., vol. 22, no. 5, pp. 1521-1530, Sep. 2018.