International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Electronic Health Data Analysis for Skin Cancer Prediction using BAT Algorithm

Abhishek Kumar Sahu[1], Manish Rohilla[2]

[1]M.Tech Scholar Technocrats Institute of Technology Excellence.
[2]Asst. Professor, CSE, Technocrats Institute of Technology Excellence.

ABSTRACT

Skin lesions refer to specific areas on the skin that display unusual growth patterns or appearances, which may be harmless (benign) or cancerous (malignant). Accurate identification and classification of these lesions are vital for the early detection and successful management of skin cancer. Nonetheless, current methods for distinguishing between malignant and non-malignant lesions using dermoscopic images often fall short in effectively extracting relevant features and are prone to overfitting, especially when dealing with imbalanced datasets. This study focuses on enhancing the classification accuracy of skin cancer by improving the quality and representation of features within dermoscopic images. This paper has proposed a model that filter noise from the input data and extract learning features. For feature extraction model has uses BAT genetic algorithm. BAT based selected pixels were used for the learning of convolutional model. Experiment was done real dataset and result shows that proposed Bat Algorithm Based Skin Cancer Detection (BASCD) has improved the classification accuracy of work.

Keywords: Digital Image Processing, Skin Cancer Diagnosis, Machine learning, Medical Image Diagnosis, Feature Extraction, Segmentation.

INTRODUCTION

Skin cancer has emerged as one of the most prevalent forms of cancer in recent years [1]. Given that the skin is the body's largest organ, it is not surprising that skin cancer is the most commonly diagnosed cancer in humans [2]. Skin cancer is generally categorized into two main types: melanoma and nonmelanoma skin cancers [3]. Among these, melanoma is considered the most aggressive, though it is relatively rare and highly lethal.

The skin is composed of two primary layers. The outermost layer, known as the epidermis, consists of flat cells called squamous cells. Beneath these lie the basal cells, which are round in shape. Interspersed among the basal cells are melanocytes — the cells responsible for producing melanin, the pigment that gives skin its color. Beneath the epidermis lies the dermis, which houses blood vessels and glands such as sweat glands [4]. Skin cancer occurs when these skin cells grow uncontrollably, forming either benign or malignant tumors. The classification of skin cancer is based on the type of skin cell where abnormal growth begins. The three most common forms are melanoma, basal cell carcinoma, and squamous cell carcinoma. Melanoma arises from melanocytes and is the deadliest form. Basal cell carcinoma originates from basal cells, while squamous cell carcinoma develops from squamous cells [5].

There are various techniques available for skin cancer detection. Traditional diagnostic methods include visual inspection by dermatologists and biopsies. However, these approaches come with certain

drawbacks. A biopsy, though commonly used, is invasive, painful, time-consuming, and in some cases may increase the risk of spreading the disease [6].

As a result, early detection remains a critical aspect of successful skin cancer treatment [7]. The biopsy procedure typically involves removing a portion of the suspicious lesion for microscopic examination to determine if it is cancerous. This method is not only uncomfortable but also slow and costly. In contrast, advancements in computer-aided technologies offer non-invasive, faster, and more cost-effective alternatives for identifying skin cancer. These automated systems aim to distinguish between melanoma and nonmelanoma cases using advanced imaging and classification techniques.

Rest of paper was organized into few sections. Next section brief other researcher work of medical image classification. Further proposed model was brief with block diagram and each block was detailed. Next section shows the experimental values of the model. Finally whole work outcome was concluded with future map.

Related Work

Venugopal et al. [8], introduced a deep learning method utilizing the Efficient Net architecture to identify skin cancer. Work applied transfer learning with fine-tuning on the EfficientNetV2-M model, demonstrating that even a limited dataset could perform well when enhanced with generalization techniques like data augmentation. The model achieved up to 95% accuracy in results, with a 0.49 score in multiclass tasks and an AUC of 0.99 in binary classification—surpassing many previous methods. This version of Efficient Net was chosen due to its strong performance on complex datasets and reliable output during testing.

Dildar et al. [9], analyzed several AI techniques applied to skin lesion detection, including CNNs, ANNs, KNNs, and GANs. Their comparison, based on the ISIC dataset, highlighted the superior performance of convolutional networks in medical image analysis.

Gouda et al. [10], conducted research on skin lesion images. They employed ESRGAN for image enhancement, alongside a custom convolutional model and pre-trained networks like ResNet50, InceptionV3, and Inception ResNet. Among these, InceptionV3 achieved the highest classification accuracy at 85.7%. Their work emphasized improving early diagnosis through advanced deep learning and visual enhancement techniques.

H. L. Gururaj et al. [11] introduced a model that incorporates several data pre-processing steps, including sampling, the dull razor method for artifact removal, and image segmentation using an autoencoder-decoder architecture. To enhance performance, the model utilizes transfer learning with pre-trained networks such as DenseNet169 and ResNet50 for training and evaluation.

In [12], Hajiarbabi M. et al. proposed a melanoma detection framework based on transfer learning. The system begins with preprocessing steps to reduce noise and correct illumination issues in dermoscopic images. A critical part of this process involves image augmentation, which applies various transformations to expand the dataset size, thereby improving training efficiency. Subsequently, a convolutional neural network (CNN) is trained using ImageNet pre-trained weights. In the final step, the network undergoes fine-tuning to tailor it specifically for distinguishing melanoma from benign lesions. The system analyzes the lesion in three sequential stages, each focusing on the central region of the image suspected to contain cancer. Outputs from these stages are then combined and processed through a fully connected neural network for classification.

Entesar Hamed I. et al. [13] developed a two-phase approach using the HAM10000 dataset, which includes a diverse collection of skin lesion images. The initial phase involves transfer learning with frozen layers, while the second phase involves unfreezing and fine-tuning all layers to adapt the model more closely to the dataset's specific characteristics.

F. P. Carranco-Avila et al. [14] investigated the relationship between genetic mutations commonly found in several cancer types and their potential link to depression. The study analyzed genomic data extracted from cancer patients via cBioPortal and compared it with genetic variants associated with depressive disorders. Advanced clustering techniques such as HJ biplot K-means and DBSCAN were applied in a two-dimensional space to identify meaningful patterns. This analysis produced a dataset used to train and evaluate machine learning and deep learning classifiers. The clustering results revealed genes related to depression and mutation trends across multiple cancers, including skin, uterine, cervical, stomach, and prostate cancers.

Proposed Skin Cancer Prediction Model

This section of paper details multiclass image segmentation for skin lesion. Proposed model Bat Algorithm Based Skin Cancer Detection (BASCD) is shown in fig. 1. Explanation of each block of figure like pre-processing, clustering, etc. was detailed in this section.

Pre-Processing The input dermoscopic image is resized to a fixed dimension p × q and then converted into grayscale to minimize computational cost. This standardized format is used for further processing in the BAT-based detection model [15].

PSI ← Grayscale(Resize(Input_Image, p, q))

Initialize Bat Population Each bat in the population represents a potential segmentation solution using c cluster centers (similar to antibodies in AIA). The population of n bats is initialized using random Gaussian-distributed cluster centers based on the grayscale image.

BP ← Gaussian(PSI, c) # Bat Population Matrix of n bats
fi ← Frequency of each bat
vi ← Velocity of each bat
xi ← Position (cluster centers)
Ai ← Loudness
ri ← Pulse rate

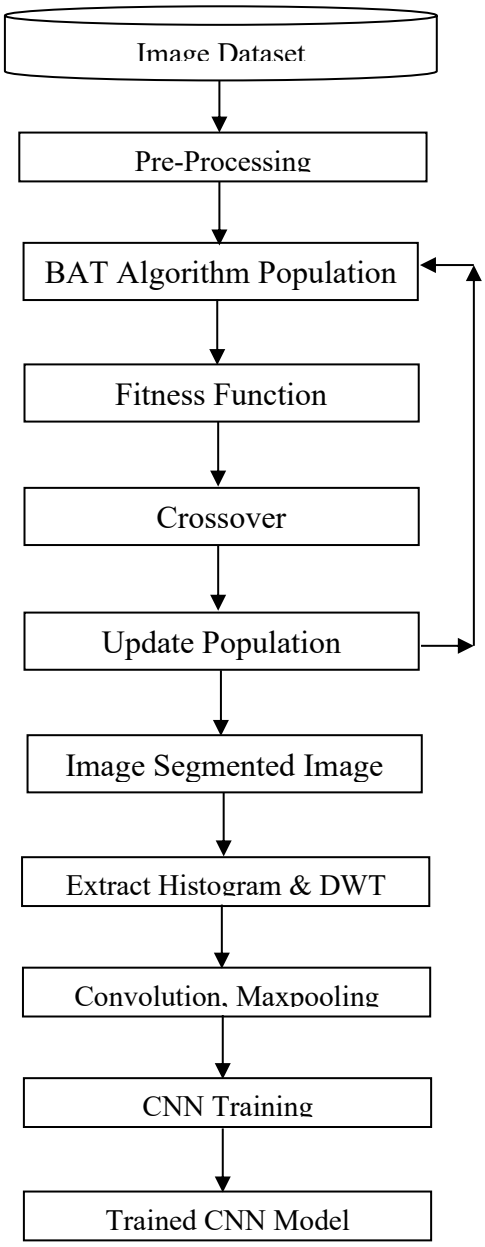


Fig. 1 Proposed model BASCD training model.

Fitness Evaluation (Affinity Analogy) The fitness (similar to affinity in AIA) is calculated by measuring how well the cluster centers represent the pixel groups. [15] For each bat, the distance between its cluster centers and image pixels is computed, grouping pixels by nearest cluster.

Fit(x_i) = ∑_(x=1)^q ∑_(y=1)^p Min[(x_i_c - PSI(x, y))^2]

Where x_i_c is the cluster center for the i-th bat.

Update Bat Population (Movement & Mutation) For each bat in the population, update its velocity, frequency, and position based on the global best solution found so far. This mimics the echolocation mechanism of bats [16].

fi ← fmin + (fmax - fmin) * rand()

v_i ← vi + (xi - xbest) * f_i

xi ← xi + vi

If a random number > pulse rate r_i, generate a local solution near the best bat using a random walk:

xi ← xbest + ε * Ai
Where ε is a small random number.
Loudness and Pulse Rate Adjustment (Exploitation & Exploration) Loudness and pulse rate are updated dynamically to balance exploration and exploitation:
Ai ← α * Ai
ri ← ri * [1 - exp(-γ * t)]
Where α and γ are constants and t is the current iteration.
Selection of Best Bat (Cloning Analogy) After evaluating all bats, the one with the best fitness is chosen. Other bats may copy parts of the best solution (similar to cloning and mutation in AIA) to enhance convergence.
x_best ← Best(BP, Fit)
BP ← Local_Exploration(BP, x_best)
Image Segmentation
After T iterations, the final cluster centers from the best bat are used to segment the grayscale image into regions (e.g., infected and non-infected areas). Segmentation is performed by assigning each pixel to the nearest cluster center.
SLDI ← Segment(PSI, x_best)
The segmented lesion detection image (SLDI) highlights infected regions (e.g., in black) and healthy skin (in white).
Extrat Histogram & DWT Feature:
This work extract histogram from the segmented image and DWT feature from same [16, 17]. This features were used for the training of neural network. Sixteen values of histogram and 16 values from the DWT low frequency region were used.

Convolution

The features extracted from the image using the BAT algorithm are further refined using a convolutional operation. In Convolutional Neural Networks (CNNs), this operation involves sliding a small kernel (also known as a filter) over the two-dimensional input to detect local patterns such as edges, textures, and shapes. This process enables CNNs to learn spatial hierarchies effectively, making them simpler to train and less prone to overfitting compared to traditional models [18].
The convolution process can be mathematically expressed as:
C←Convolution(EF, s, p, F_c)
Here, EF is the input feature block, s (stride) determines the step size of the filter movement, p (padding) refers to the addition of zero rows or columns around the input to maintain dimensionality, and Fc is the convolutional filter applied to extract features.

Max-pooling

The concept behind max-pooling is to reduce the spatial dimensions of feature maps while preserving the most significant information. This is achieved by selecting the maximum value within each region of the input, thereby expanding the receptive field and reducing computational complexity. In CNNs, this type of subsampling can be performed either through maximum pooling or average pooling. By down sampling the feature maps by a factor of a, the effectiveness of the convolutional layers in capturing global patterns

is significantly enhanced. Typically, pooling is performed with a stride of 2 (s = 2), and it is used in conjunction with convolutional layers to build more abstract, higher-level representations.
The max-pooling operation is represented as:
C←Max pooling (C, s, p, F_m)
Where C is the input feature map, s is the stride, p is padding, and Fm refers to the pooling filter.

Training of Deep Neural Networks

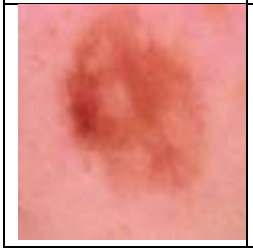
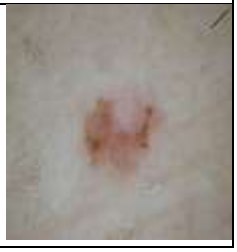
Rather than classifying each pixel individually, Deep Neural Networks (DNNs) are designed to output a segmented version of the entire input image in one forward pass. In this architecture, convolution and pooling operations retain the structural layout of the image across successive hidden layers, ensuring that each pixel's contextual information is preserved. The resulting feature maps contain complete spatial information necessary for accurate pixel-wise classification.
The input feature block EF, after being processed by convolutional operations, is used as the training input for the DNN. The corresponding label, whether the region contains a tumor or not, serves as the target output. Once trained, the DNN can directly accept the processed image block and classify it as either tumor or non-tumor, providing an efficient solution for medical image segmentation and classification.

Experiment and Results

The proposed model was implemented using MATLAB, a widely-used software platform for numerical computing, algorithm development, and data visualization. The experiments were conducted on a computer system equipped with 4 GB of RAM, an Intel Core i3 processor from the 6th generation series, and running the Windows 10 operating system. Despite the system's modest hardware configuration, the model was successfully trained and tested, demonstrating its efficiency and suitability for environments with limited computational resources. For performance evaluation and comparison, the results of the proposed model were benchmarked against existing state-of-the-art models, particularly DenseNet, as referenced in [11].

Table 1 Precision value of Skin cancer detection models

Testing Image Dataset	DenseNet [11]	BASCD
30	0.7	0.6875
60	0.875	0.7391
90	0.76	0.7568
120	0.7222	0.8039
150	0.6964	0.8023

Benign	Maligant
	

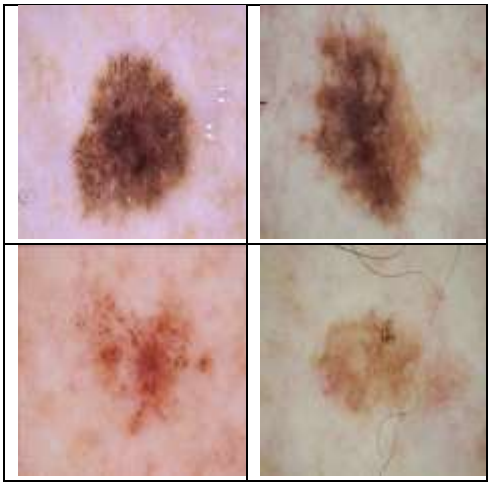


Fig. 1 Skin Cancer Dataset [19].

Table 2 Precision value of Skin cancer detection models.

Testing Image Dataset	DenseNet [11]	BASCD
30	0.4667	0.7857
60	0.6364	0.8095
90	0.5278	0.8485
120	0.4561	0.8039
150	0.4588	0.8118

Table 2 shows the skin cancer precision values on different set of testing images. It was found that with increase in number of testing images precision value increases. Further it was found that use of BAT algorithm has increases the feature quality by removing unwanted information from the training data. Precision value was enhanced by 37.28% as compared to DenseNet model.

Table 3 Recall value of Skin cancer detection models.

Testing Image Dataset	DenseNet []	BASCD
30	60.7143	73.33
60	76.19	77.27
90	67.1429	80.56
120	63.3929	82.46
150	62.5	81.36

Table 2 shows skin cancer models recall values. It was found that use of Dep Learning model for learning of histogram and DWT feature has increases the work performance. Proposed BASCD model has increase the recall value 0.95%.

Table 4 F-Measure value of Skin cancer detection models.

Testing Image Dataset	DenseNet [11]	BASCD
30	0.56	0.7333
60	0.7368	0.7727
90	0.623	0.8
120	0.5591	0.8039
150	0.5532	0.807

Inverse Average of precision and recall values shown in table 4. It was shown use of color histogram feature and frequency DWT features has enhanced the learning of Deep learning model. BAT base image segmentation has increases the work performance.

Table 5 Accuracy value of Skin cancer detection models.

Accuracy of Input skin cancer image correct class identification by comparing models were shown in table 5. Use of BAT model has increases the work performance by **16.46%** as compared to existing model.

CONCLUSION

Medical image diagnosis by computer algorithms provides great support and reduces the work load of medical officers. This paper has work for proposed a model that increases the correct class prediction accuracy of skin cancer detection. Paper has extract histogram feature and DWT feature from the segmented image. BAT algorithm was used for the image segmentation. Extracted features are used for the training of neural network. Experiment was done on real dataset of skin cancer. Result shows that Precision value was enhanced by 37.28% as compared to Dense Net model , it was found that use of BAT model has increases the work performance by 16.46% as compared to existing model. In future researcher uses other format of images for predicting the image class.

REFERENCE

1. Ashraf R., Afzal S., Rehman A.U., Gul S., Baber J., Bakhtyar M., Mehmood I., Song O.Y., Maqsood M. Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection. IEEE Access. 2020;8:147858–147871.
2. Byrd A.L., Belkaid Y., Segre J.A. The Human Skin Microbiome. Nat. Rev. Microbiol. 2018;16:143–155.
3. Elgamal M. Automatic Skin Cancer Images Classification. IJACSA. 2013.
4. Jones, O. T., Ranmuthu, C. K. I., Hall, P. N., Funston, G. & Walter, F. M. Recognising skin cancer in primary care. Adv. Ther. 37(1), 603–616 (2020).
5. Leiter, U., Eigentler, T. & Garbe, C. Epidemiology of skin cancer. Adv. Exp. Med. Biol. 810, 120–140 (2014).
6. Andre, G. C. & Renato, A. The impact of patient clinical information on automated skin cancer detection. Comput. Biol. Med. 116, 103545 (2020).
7. Khan M.Q., Hussain A., Rehman S.U., Khan U., Maqsood M., Mehmood K., Khan M.A. Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer. IEEE Access. 2019;7:90132–90144.
8. Venugopal, V., Raj, N. I., Nath, M. K. & Stephen, N. A deep neural network using modified efficientnet for skin cancer detection in dermoscopic images. Decision Analytics Journal 8, 100278 (2023).

9. Dildar, M. et al. Skin cancer detection: a review using deep learning techniques. International journal of environmental research and public health 18, 5479 (2021).
10. Gouda, W., Sama, N.U., Al-Waakid, G., Humayun, M., Jhanjhi, N.Z. Detection of skin cancer based on skin lesion images using deep learning. In Healthcare, vol. 10, 1183 (MDPI, 2022).
11. H. L. Gururaj, N. Manju, A. Nagarjun, V. N. M. Aradhya and F. Flammini, "DeepSkin: A Deep Learning Approach for Skin Cancer Classification," in IEEE Access, vol. 11, pp. 50205-50214, 2023.
12. Hajiarbabi M. Skin cancer detection using multi-scale deep learning and transfer learning. J Med Artif Intell 2023;6:23.
13. Entesar Hamed I. Eliwa. "Enhancing Skin Cancer Diagnosis Through Fine-Tuning of Pretrained Models: A Two-Phase Transfer Learning Approach". International Journal of Breast Cancer, 17 February 2025.
14. F. P. Carranco-Avila et al., "Correlation Between Depression-Associated Genes and Cancer Types: Predicting Cancer Based on Mutation Frequencies," in IEEE Access, vol. 13, pp. 60930-60943, 2025.
15. H. Chen et al., "Distributed Text Feature Selection Based On Bat Algorithm Optimization," 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Metz, France, 2019, pp. 75-80.
16. Aydin Y. A Comparative Analysis of Skin Cancer Detection Applications Using Histogram-Based Local Descriptors. Diagnostics (Basel). 2023 Oct 6;13(19):3142.
17. Md. Mahbubur Rahman, Mostofa Kamal Nasir, Md. Nur-A-Alam, Md. Saikat Islam Khan. "Proposing a hybrid technique of feature fusion and convolutional neural network for melanoma skin cancer detection, Journal of Pathology Informatics". Volume 14, 2023, 100341.
18. Neha Sharma, Vibhor Jain, Anju Mishra. "An Analysis Of Convolutional Neural Networks For Image Classification". Procedia Computer Science, Volume 132, 2018, Pages 377-384.
19. http://www.dermnet.com/dermatology-pictures-skin-disease-pictures