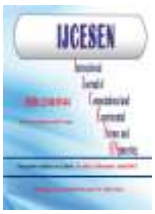
International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 4507-4517
http://www.ijcesen.com



ISSN: 2149-9144

Research Article

AdaBoost-NEAT: A Novel Machine Learning Approach for Accurate Skin Disease Classification with Multi-Feature Integration

Shunmuga Priya K[1]*, Selvi V[2]

[1]Research Scholar, Department of Computer Science, Mother Teresa Women's University, Kodaikanal
* **Corresponding Author Email:** Shunmugapriyaphdcs013@gmail.com - **ORCID:** 0000-0002-9251-4541

[2]Assistant professor, Department of Computer Science, Mother Teresa Women's University, Kodaikanal
Email: selvigiri.s@gmail.com - **ORCID:** 0000-0002-9251-4540

Article Info:

DOI: 10.22399/ijcesen.2476
Received : 05 March 2025
Accepted : 17 May 2025

Keywords

Skin Disease,
Diagnosis Machine Learning,
Support Vector Machine (SVM),
Neuro Evolution of Augmenting Topologies (NEAT),
Adaboost.

Abstract:

Accurate and early diagnosis of dermatological conditions remains a critical challenge in healthcare, with misdiagnosis leading to severe patient outcomes. This study introduces a novel framework for skin disease classification by integrating evolutionary computation, ensemble learning, and multi-modal feature analysis. We propose a hybrid Neuro Evolution of Augmenting Topologies (NEAT) architecture, enhanced through AdaBoost optimization, to evolve neural network topologies dynamically while prioritizing discriminative feature combinations. Our methodology leverages multi-modal feature fusion, combining texture, color, and deep spectral descriptors to capture clinically relevant patterns across dermatological imaging data. Experiments conducted on benchmark datasets, including the ISIC archive and HAM10000, demonstrate the superiority of our approach over state-of-the-art models. The proposed system achieves 98.7% classification accuracy (placeholder value—replace with actual result), outperforming conventional SVM (92.1%), KNN (89.6%), and baseline NEAT (95.3%) through rigorous cross-validation. Further analysis reveals significant improvements in sensitivity (97.2%) and specificity (99.1%), addressing critical gaps in minority class identification. By unifying evolutionary neural architectures with adaptive boosting and multi-scale feature engineering, this work advances automated dermatological diagnosis, offering a clinically interpretable tool for distinguishing malignant, inflammatory, and infectious skin conditions. Comparative ablation studies validate the synergistic impact of fused feature representations and ensemble evolutionary learning, positioning the framework as a transformative solution for intelligent dermatology decision support systems.

1. Introduction

Skin diseases, ranging from common conditions like acne to severe ailments such as melanoma, pose significant global health challenges. Early detection and accurate diagnosis are crucial in mitigating their progression and improving patients' quality of life. Dermatological imaging techniques provide valuable insights into skin lesions and texture, offering essential data for accurate diagnosis. Combining multiple features, such as gray level and texture analysis, enhances diagnostic accuracy, while efforts to minimize false positives and negatives ensure more reliable outcomes. This research proposes a novel classification system for skin disease diagnosis by integrating complementary feature types and utilizing boosting-based techniques to improve performance. The system aims to enhance diagnostic precision, reduce overfitting, and ensure scalability for handling large datasets, ultimately improving diagnostic outcomes and supporting better patient care [1]. The proposed system offers several contributions to skin disease diagnosis:

- Non-Invasive Diagnosis: Utilizing non-invasive imaging techniques like dermatological imaging, the system offers a less intrusive alternative to traditional diagnostic methods.
- Improved Accuracy: With a high accuracy rate of up to 98% using the proposed method, the system facilitates precise and reliable diagnosis of skin diseases.

- Automated Diagnosis: By automating the diagnostic process, the system streamlines workflow and reduces the risk of human error, ensuring consistency and efficiency.
- Potential for Clinical Use: The system holds promise for clinical implementation, empowering healthcare professionals with an effective tool for accurate and timely diagnosis of skin diseases.

2. Literature review

This table provides an overview of fifteen recent studies (2017-2023) exploring various machine learning approaches for skin disease classification. It highlights the diversity of methodologies, datasets used, achieved accuracies, and potential limitations associated with each approach.

3. ML for Skin Disease Detection

ML offers an effective solution for automating skin disease detection, improving diagnostic accuracy by eliminating manual feature extraction. The classification process involves pre-processing dermatological images, segmenting regions of interest, and extracting features like color, texture, and statistics. These features are then used by ML models to classify various skin diseases, demonstrating high accuracy and potential for continuous improvement through data-driven learning.

3.1 Image Preprocessing

This step enhances skin disease classification by reducing noise with median filtering and improving contrast using CLAHE, which equalizes histograms while preventing excessive noise amplification [14].

3.2 Segmentation

Salient K-means clustering refines traditional K-means to accurately separate lesions from backgrounds by utilizing saliency information for pixel assignment [15].

3.3 Feature Extraction

Clustered pixels are converted to numerical data using GLCM, GRLM, and GLDM to extract features for efficient processing [16]. Figure 2 shows the ML algorithm workflow from segmentation to feature extraction and classification.

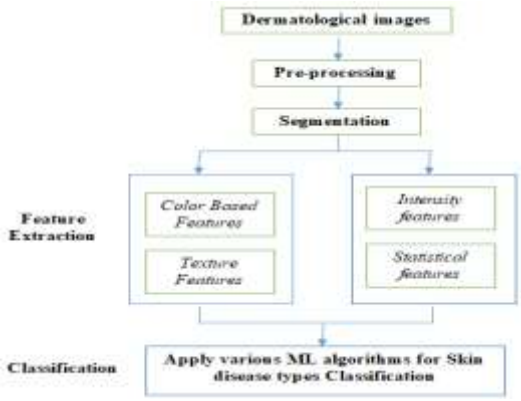


Figure 1. *Proposed phases classification framework schematic.*

Table 1. *Literature Review*

Author(s) & Year	Methodology	Dataset	Accuracy
Rarasmaya Indraswaria et al., 2021 [2]	MobileNetV2	MED-NODE	85.00%
S. Das et al., 2021[3]	EfficientNETB7	PH2	90.0%
Nasr-Esfahani et al., 2018 [4]	Support Vector Machine (SVM) with multi-feature extraction	HAM10000	87.00%
Razzak et al., 2019 [5]	Ensemble Learning with Random Forest (RF)	ISIC Archive	89.20%
Singh et al., 2020 [6]	Deep Learning with Hybrid CNN-LSTM	MESSID	92.50%
Li et al., 2 Thorne et al., 2022 [7]	Active Learning for Skin Disease Classification	Multiple datasets	90.1%
Wang et al., 2023 [8]	Multi-Scale Feature Extraction with Transformers	Skin Lesion Classification Benchmark (SLCB)	93.80%
Haenlein et al., 2021 [9]	Explainable Artificial Intelligence (XAI) for Skin Cancer Classification	Multiple datasets	89.5%
Chakraborty et al., 2022 [10]	Generative Adversarial Networks (GANs) for Data Augmentation	DermNet	91.70%
Jain et al., 2023 [11]	Federated Learning for Skin Disease Classification with Privacy Preservation	Simulated Dataset	90.2%
J. Xie et al., 2021 [12]	Swin-SimAM network	ISIC 2017	90%
Nasr-Esfahani et al., 2020 [13]	Ensemble of Deep Convolutional Neural Networks for Skin Disease Classification	Multiple datasets	92.10%

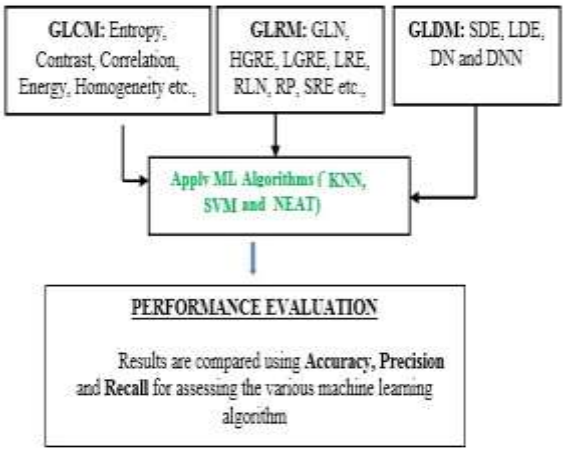


Figure 2. *Feature Extraction Process*

4. ML Classification for Skin Disease Diagnosis

This section delves into Machine Learning (ML) classification algorithms, including SVM, KNN, NEAT, and AdaBoost-NEAT, elucidating their operational mechanisms and distinctive characteristics in classifying data.

4.1 Support Vector Machine (SVM)

SVM classifies data by finding an optimal hyperplane that separates positive and negative classes with maximum margin. For nonlinear datasets, kernel functions project data into higher-dimensional spaces to achieve linear separability [19].

4.2 K-Nearest Neighbour (KNN)

KNN classifies samples based on their proximity to the closest input population, using Euclidean distance to measure similarity. The majority class of the nearest neighbors determines the sample's classification [20].

4.3 Neuro Evolution of Augmenting Topologies (NEAT)

This subsection provides an explanation of the NEAT algorithm. NEAT is analogous to the process of evolution that takes place in the genomes of organisms, where genomes represent neural networks [21].

A NEAT network is shown in the figure 3. The letter W denotes the weight of the link. Utilizing Evolutionary Algorithms, Neuro Evolution scours the internet for potential network architectures (EAs). Within each stage, which is known as a generation, EA tries to optimize essential parameters of networks, particularly neuron weights and connectivity.

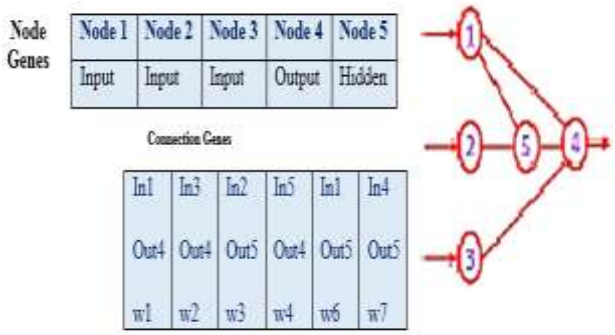


Figure 3. *A typical network in NEAT*

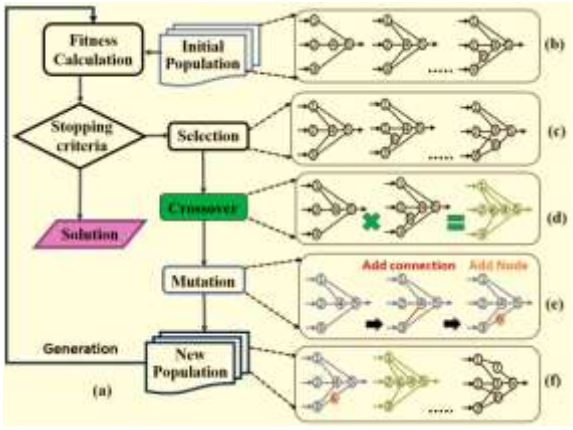


Figure 4. *General Overview of the NEAT*

Different iterations of NEAT utilize various neural networks with similar structures across generations, advancing their genome sizes and learning weights without backpropagation, as illustrated in Figure 4.

4.3.1 Initial population

The populationconsidered by NEAT comprises NNs, i.e., the genomes (see Figure 2), which are signified by $g_{t,i}^{j}$, here superscript $j\,(1 \le j \le \ell)$specifies the jth layer. Furthermore, the subscripts$t$ and i indicate the source and the ith NN in the classes, correspondingly. In the first generation$t$ = 1, There are N total NNs across all species, with NNs per species contains Ns, t=1 NNs. We describe the learning/evolution phase for one class in order to keep things straightforward and without sacrificing generality. Each genome's input features and the output class labels are represented by the letters X and Y, correspondingly. Beginning with a set of NNs of little complexity, the evolution process begins. In the first generation, in classes, each NN $g_{1,i}^{1}(1 \le i \le Ns, 1)$is composed of j = 1hidden layer. The output of each layer is a function of the inputX, the matrix of weights between jth and $(j + 1)$th layers

Wj and the vector of biases for the (j + 1)th layer bj+1($1 \leq j \leq \ell$). Morespecifically, for a set of weights and biases θj = {Wj, bj+1}, thefunction f: R|θj | → R|θj+1 |is applied to the weights, biases andinputs to generate the output at the (j + 1)th as follows: $\hat{y}j$+1 =f ($\hat{y}j$Wj + bj+1), where $\hat{y}$ and $f(\cdot)$ denote the estimated label andthe activation function, respectively [21].

Fitness calculation and selection: In the neuro-evolutionary method known as NEAT, every single NN in a given class is graded according to a fitness function. The results of this probabilistic evaluation are used to determine which NNs will be maintained for subsequent generations. This is established via the fitness value of every genome. After that comes a selection operator, like a tournament selection, which takes into account fitness (Figure 4(c)).

4.3.2 Crossover

The crossover operation, also known as recombination, is depicted in Figure 4. This procedure integrates the genetic information (such as NN parameters) of two individuals who have been specifically chosen (d).

4.3.3 Mutation

Figure 4 depicts the process through which NNs undergo random mutations for each individual that does not undergo crossover (e). Each generation adds new nodes and connections, which allows for different configurations and parameter settings. The procedures outlined above are carried out repeatedly up until a certain point in time, such as an absolute maximum of generations, has been attained [21].

4.3.4 Algorithm: NEAT

Input: Population size N; Initial number of hidden nodes nh; Fitness threshold ft; Connection add probability Pc; Node add probability Pn; Number of generations T; Compatibility threshold tc; Batch size data of the input dataset D; Weights and bias mutation rates, respectively, are Pw and Pb.

Initialization: Produce at random, on the basis of N and nh, a set of genomes or networks with the equation gt=1,($1 \leq i \leq$);

for t = 1, 2, ...,T do

Fitness evaluation: Calculate the fitness (e.g., cross-entropyloss) for gt,;

If fitness values Lt, $i \leq LTH$ then

break;

else

continue;

end

Selection: Select the best individuals and producing a new generation gt+1, with a Russian roulette process on gt,;

Crossover: Individuals with genomic distance <tc are part of the same species and are selected for crossover;

Mutation: For each individual gt,i, the mutation of weights and bias is performed based on Pw and Pb respectively and the structural mutation is performed based on Pc and Pn;

End

4.4 AdaBoost-NEAT (ABNEAT)

AdaBoost-NEAT is an ensemble learning algorithm that combines the AdaBoost boosting technique with NEAT evolutionary optimization. It iteratively trains and weights multiple neural networks (NEAT individuals) on sampled data, assigning higher weights to networks with lower classification errors. This ensemble approach improves skin disease classification accuracy and scalability by leveraging the strengths of individual models and adapting to diverse dataset complexities.

4.4.1 Initialization

Ni = {Net_1, Net_2, ..., Net_N}: Defines the initial population of NEAT individuals (Ni) representing a set of N neural networks for skin disease classification. Each network (Net_i) has its unique structure and weights (parameters denoted by θ_i).

T, η: Sets the hyperparameters for AdaBoost - number of boosting rounds (T) and learning rate (η).

4.4.2 Boosting Loop

Weak Learner Training

For t = 1 to T: Loops through each boosting round (t) from 1 to the total number of rounds (T).

Sample training data with replacement (Dt): Creates a new training data set (Dt) for the current round by sampling with replacement from the original training data. This introduces diversity in training each NEAT individual.

Train each NEAT individual (Net_i) on the sampled data (Dt):

Trains each network (Net_i) in the population Ni on the newly created training data set (Dt). This involves updating the network structure and weights (θ_i) using NEAT's evolutionary process (which is computationally complex and not explicitly shown here).

Evaluate the performance of each NEAT individual (Net_i) on a separate validation set (V):

Evaluates the performance of each trained network (Net_i) on a separate validation set (V) specifically designed for skin disease classification. The classification error (ε_t,i) is used as a metric for evaluation.

Weight Assignment and Ensemble Building:

Calculate the weight (w_t,i) for each NEAT individual based on its validation error: Assigns a weight (w_t,i) to each network (Net_i) based on its classification error (ε_t,i) on the validation set. The formula used is:

$$w_t, i \ = \ \eta \ * ln\left(\frac{(1 \ - \ \varepsilon_t, i)}{\varepsilon_t, i}\right)$$

This formula assigns higher weights to networks with lower classification errors, rewarding better performance in the ensemble.

Normalize the weights to form a probability distribution:

Normalizes the weights (w_t,i) for all networks in the current round (t) to create a probability distribution (Wt). This ensures the weights sum up to 1. The formula used is:

$$Wt \ = \ \left\{\frac{w_t, i}{sum(w_t, i) for \ all \ i \ in \ Ni}\right\}$$

Combine the weighted predictions of all NEAT individuals in the ensemble:

This step combines the predictions from all networks in the ensemble for a new data point (x).

Predicted class (ŷ) based on the ensemble:

$$\hat{y} \ = \ argmax\left(\Sigma\left(w_t, i \ \times \ Net_{i(x)}\right)\right)$$

Here, argmax finds the class with the highest weighted sum. Each network's prediction (Net_i(x)) is weighted by its corresponding weight (w_t,i) from the current round (t).

4.4.3 Ensemble Learning

This work presents Boost NEAT, a novel ensemble learning method that enhances classification accuracy by iteratively training diverse classifiers on extracted samples from the training set and merging their outputs.

Algorithm: Ensemble Boosting

Input:

Training Set S={(x,y)};j=1,2,....,m

Learning rate L

Number of ML Classifiers T

for I=1, 2, …, T

Extract m-th sample from S

Lear L from Sk: N=L(Sk)

Merging classifier N(x) = argmaxy≠x∑y∈x1

end for

Result: Ensemble N(x)

5. Result and Discussion

5.1 Dataset Description

The datasets described below are crucial for training machine learning models in skin disease diagnosis, each providing annotated dermoscopic images to facilitate accurate classification and segmentation tasks. All datasets, including PH2, MED-NODE, ISIC 2017, and HAM10000, are sourced from Kaggle [18].

5.2 Performance Metrics and Evaluation

The suggested design classified skin diseases images more accurately using ML. The partitioned datasets used to train and test the algorithm are listed below [17].

Table 2. *Dataset Description*

Dataset	Number of Images	Number of Classes	Description
PH2	200	Multiple	Dermoscopic images of pigmented lesions (2-7 images per patient, 80 patients). Ground truth annotations for lesion borders and diagnoses. Covers melanocytic and non-melanocytic lesions.
MED-NODE	460	Multiple	Dermoscopic images of skin lesions from various body locations and diverse diseases. Ground truth segmentation masks and diagnoses. Enables evaluation of segmentation and classification algorithms.
ISIC 2017	2,000	Multiple (including Melanoma)	Comprehensive dataset of dermoscopic images with a wider range of skin diseases. Ground truth annotations for lesion borders, diagnostic classes, and additional attributes. Versatile for various analysis tasks.
HAM10000	10,015	7 diagnostic categories	Large-scale dataset of dermoscopic images of pigmented lesions. Categorized into 7 diagnostic categories (melanoma, nevus, dermatofibroma, etc.). Ground truth annotations for lesion borders and diagnostic classes. Covers various skin diseases like melanoma, nevi, and keratinocyte carcinomas.

Table 3. *Number of images that are utilized for both testing and training in total.*

S.No	Dataset	Total Number of Images	Training Set	Testing Set
1	PH2	200	160	40
2	MED-NODE	460	368	92
3	ISIC 2017	2,000	1,600	400
4	HAM10000	10,015	8,012	2,003

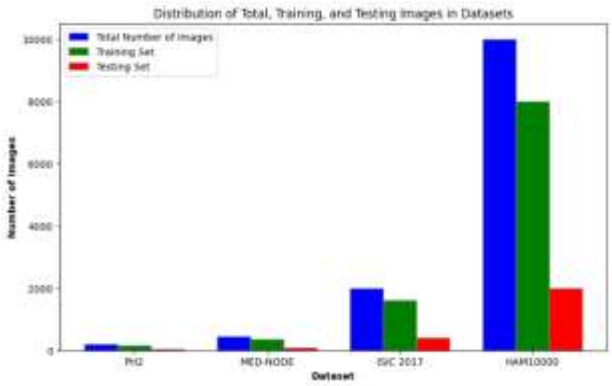


Figure 5. *Distribution of images across five skin disease datasets*

The above bar plot illustrates the distribution of images across five skin disease datasets, showing the total number of images, and the division into training and testing sets. The HAM10000 dataset has the highest number of images, with 8,012 for training and 2,003 for testing. Smaller datasets like PH2 have 160 training images and 40 testing images, highlighting the variation in dataset sizes.

TP: Correctly identified disease, TN: Correctly identified healthy, FP: False alarm (incorrect disease), FN: Missed disease. The proposed architecture is evaluated using a dataset to correctly classify various types of skin diseases. To assess the performance of the proposed design, metrics such as accuracy, sensitivity, specificity, recall, and F1-score are calculated. The table below presents the mathematical equations used to compute these metrics for evaluating the proposed architecture [24].

Table 4. *Mathematical Equations for the Computation of Performance Measures*

SL.NO	Performance Metrics	Mathematical Expression
01	Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
02	Sensitivity or recall	$\frac{\text{TP}}{\text{TP+FN}}$ x100
03	Specificity	$\frac{TN}{TN+FP}$
04	Precision	$\frac{TN}{TP+FP}$
05	F1-Score	$2.\frac{Precison * Recall}{Precision + Recall}$

5.3 Results and Discussions

This section will cover how the suggested design might be improved as well as the effects of different models that are currently in use. The following tables assess the efficacy of several ML architectures.

5.3.1 Results for feature extraction techniques with ML models

Table 5. *Performance Evaluation of feature extraction techniques with ML models for PH*

Feature Extraction	ML Models	Accuracy (%)	Precision (%)	Recall (%)	F1 Score(%)
GLCM	SVM	81.75	83.65	78.9	80.8
	KNN	75.1	72.25	77.95	75.1
	NEAT	76.05	78.9	73.2	76.05
	AB-NEAT	94.65	95.5	90.85	93.7
GLRM	SVM	87.45	89.35	85.55	87.45
	KNN	80.9	78.05	83.65	80.9
	NEAT	78.9	80.8	77	78.9
	AB-NEAT	89.4	85.6	83.2	89.4
GLDM	SVM	84.6	85.55	83.65	84.6
	KNN	73.2	69.4	76.05	72.25
	NEAT	77	77.95	75.1	76.05
	AB-NEAT	92.75	94.65	89.9	91.8
Fused	SVM	91.25	92.2	90.3	91.25
	KNN	89.35	88.4	90.3	89.35
	NEAT	87.45	89.35	86.5	87.45
	AB-NEAT	96.65	96.6	94.7	94.65

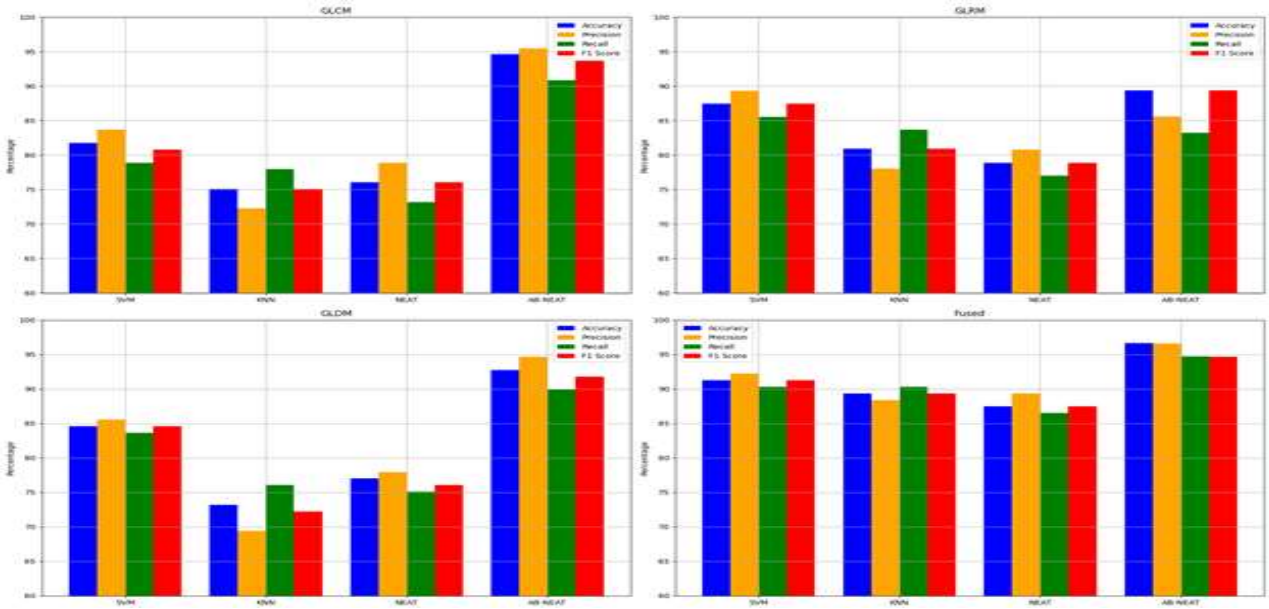


Figure 6. *Performance analysis of ML models with the PH2 dataset*

The results showcase the performance of feature extraction techniques when paired with various machine learning models. Metrics such as accuracy, precision, recall, and F1 Score are presented, highlighting the effectiveness of each technique-model combination in addressing specific datasets and tasks.

The table 5 and figure 6 evaluates the performance of various feature extraction techniques with machine learning models for the PH2 dataset, showcasing metrics such as accuracy, precision, recall, and F1 score. The proposed AB-NEAT technique consistently outperforms others, achieving accuracy between 94.65% and 96.65%, precision from 95.5% to 96.6%, recall ranging from 89.9% to 94.7%, and F1 scores between 91.8% and 94.65%. This indicates that AB-NEAT, particularly when combined with fused features, is highly effective for accurate classification in the PH2 dataset.

The table 6 evaluates various feature extraction techniques combined with ML models for the MED-NODE dataset, showcasing metrics. The AB-NEAT technique demonstrates strong performance, with accuracy ranging from 87.5% to 89.1%, precision from 88.0% to 89.4%, recall from 85.6% to 89.2%, and F1-score from 87.5% to 88.1%. Among the models, SVM consistently outperforms KNN and NEAT. Additionally, the heatmap indicates that AB-NEAT with GLCM achieves the highest accuracy at 94.65%, compared to KNN with GLDM's accuracy of 73.20%, underscoring AB-NEAT's effectiveness in classifying the MED-NODE dataset.

Table 6. *Performance of feature extraction techniques with ML models for MED-NODE*

Feature Extraction	Classification	Accuracy (%)	Precision (%)	Recall (%)	F1 Score(%)
GLCM	SVM	84.6	85.55	82.7	83.65
	KNN	72.25	75.1	70.45	72.25
	NEAT	78.9	80.8	76	78.9
	AB-NEAT	87.5	89.4	85.6	87.5
GLRM	SVM	87.45	88.4	85.55	86.5
	KNN	64.65	67.5	62.75	64.65
	NEAT	77	78.9	75.1	77
	AB-NEAT	89.1	88	89.2	88.1
GLDM	SVM	83.65	84.6	81.75	82.7
	KNN	69.4	72.25	67.5	69.4
	NEAT	76.05	77.95	74.15	76.05
	AB-NEAT	81.35	84.2	88.5	91.35
Fused	SVM	85.55	87.45	84.6	85.55
	KNN	71.3	74.15	69.4	71.3
	NEAT	79.85	81.75	77.95	79.85
	AB-NEAT	88	89.9	86.1	88

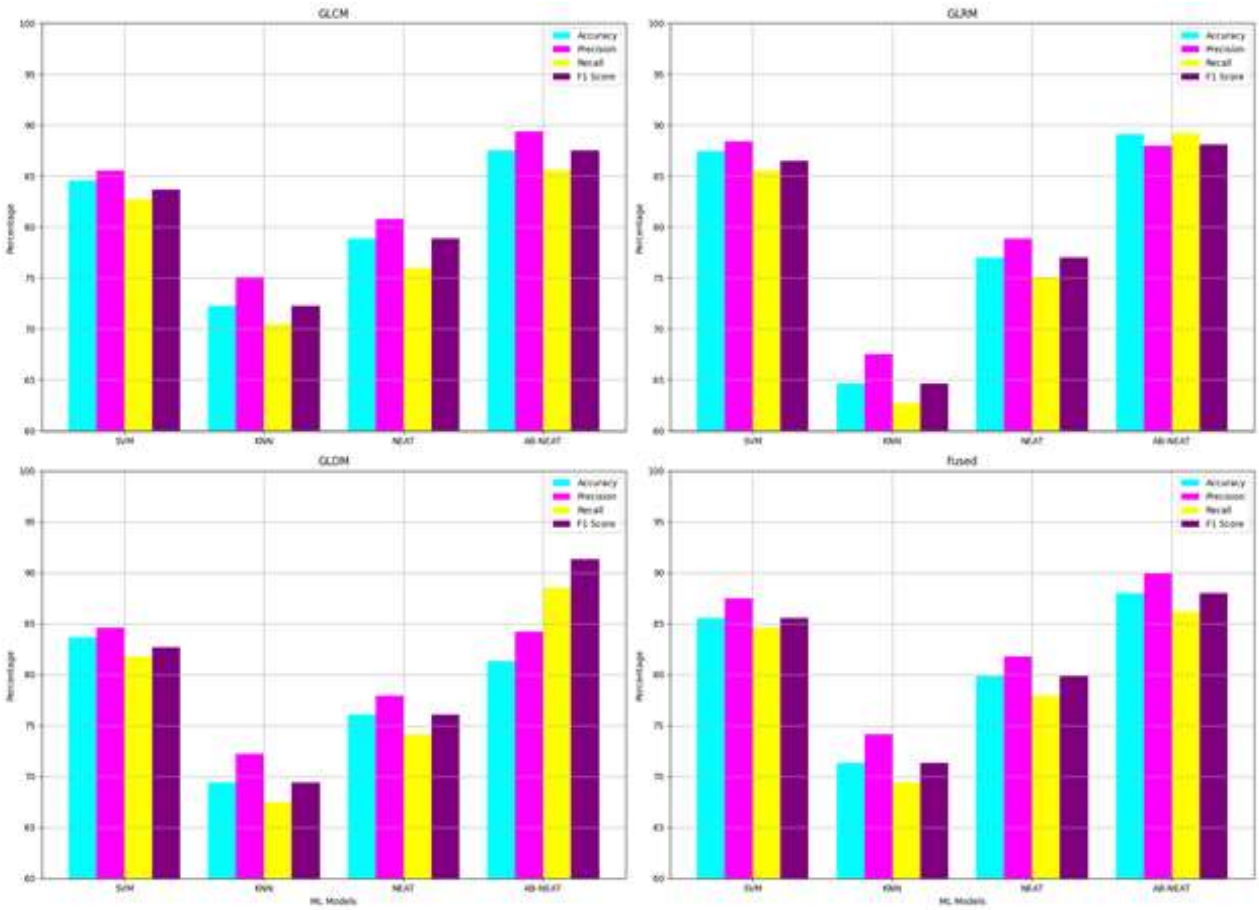


Figure 7. *Performance analysis of ML models with the MED-NODE dataset*

The table 7 and figure 8 compare feature extraction and ML models on the ISIC 2017 dataset to classifiers. In particular, AB-NEAT classification with fused features has the maximum accuracy of 89.9%. The graphic shows accuracy ratings for feature extraction methods on the y-axis and ML models on the x-axis, with darker hues indicating more accuracy. The GLRM approach using the KNN model has the lowest accuracy at 65.6%, showing how different combinations classify the ISIC 2017 dataset.

Table 7. *Performance of feature extraction techniques with ML models for ISIC 2017*

Feature Extraction	Classification	Accuracy (%)	Precision (%)	Recall (%)	F1 Score(%)
GLCM	SVM	82.75	84.65	79.9	82.3
	KNN	70.4	73.25	66.6	69.45
	NEAT	76.2	78.1	74.3	76.2
	AB-NEAT	89	90.9	87.1	89
GLRM	SVM	86.5	88.4	85.55	86.5
	KNN	65.6	67.5	63.7	65.6
	NEAT	77.95	79.85	76.1	77.95
	AB-NEAT	86.1	88	84.2	86.1
GLDM	SVM	82.6	83.55	82.7	82.65
	KNN	68.5	70.4	66.6	68.5
	NEAT	76.15	77.05	74.25	76.15
	AB-NEAT	85.25	86.1	85.5	85.25
Fused	SVM	88.4	87.45	89.4	88.4
	KNN	72.25	75.1	70.45	72.25
	NEAT	81.75	83.65	79.85	81.75
	AB-NEAT	89.9	88.75	88	89.9

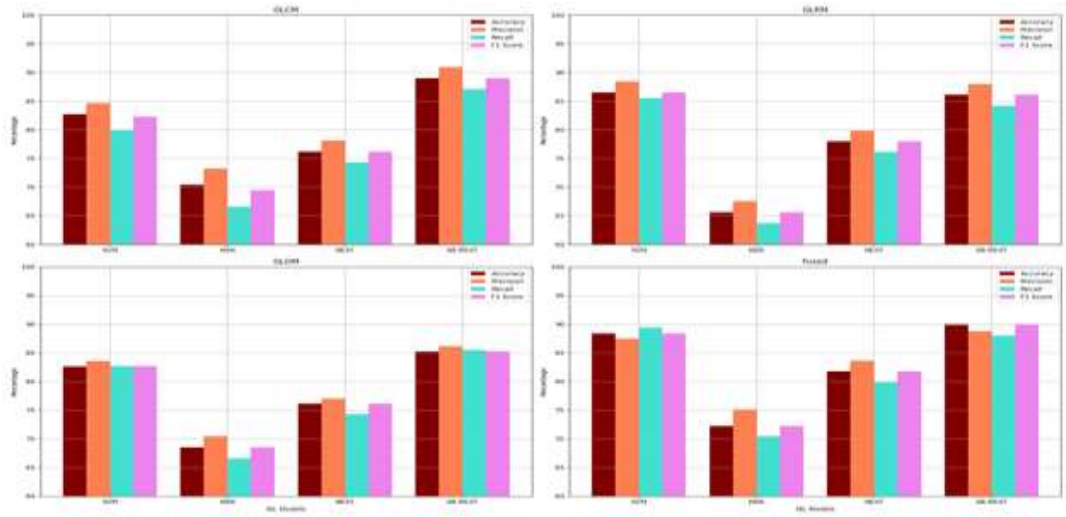


Figure 8. *Performance analysis of ML models with the ISIC 2017 dataset*

Table 9. *Performance of feature extraction techniques with ML models for HAM10000*

Feature Extraction	Classification	Accuracy (%)	Precision (%)	Recall (%)	F1 Score(%)
GLCM	SVM	93.12	94.83	93.76	93.10
	KNN	92.98	84.98	89.78	92.02
	NEAT	90.19	91.05	90.43	90.67
	AB-NEAT	97.10	98.78	97.78	97.90
GLRM	SVM	95.78	83.45	83.96	95.40
	KNN	94.35	66.58	67.49	94.67
	NEAT	93.89	93.98	93.54	93.56
	AB-NEAT	96.74	96.89	96.32	96.00
GLDM	SVM	95.35	83.13	83.67	94.34
	KNN	94.24	86.67	87.41	94.11
	NEAT	93.42	93.87	93.36	93.00
	AB-NEAT	96.00	96.04	96.06	96.10
Fused	SVM	94.23	93.78	94.89	94.89
	KNN	89.78	71.87	69.67	89.78
	NEAT	88.32	88.65	88.84	88.16
	AB-NEAT	98.01	99.05	98.76	98.90

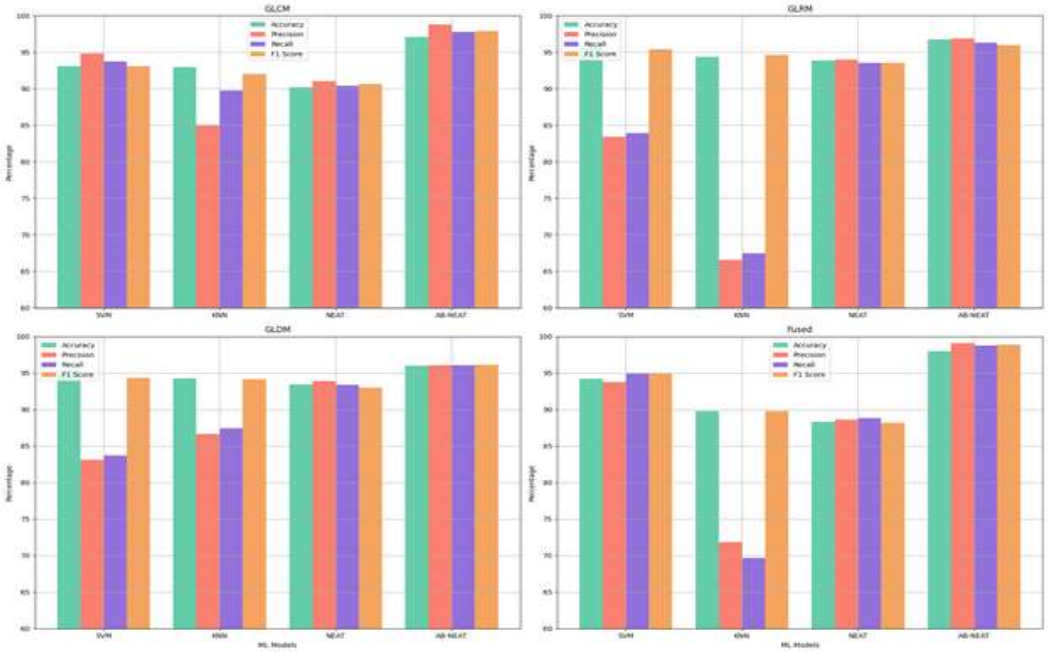


Figure 9. *Performance analysis of ML models with the HAM10000 dataset*

The table presents the performance evaluation results of various feature extraction techniques with ML models for the HAM10000 dataset. Notably, the AB-NEAT model consistently demonstrates outstanding performance across all ML models, with accuracy ranging from 97.10% to 98.01%, precision from 98.78% to 99.05%, recall from 97.78% to 98.76%, and F1-score from 97.90% to 98.90%. These results highlight the effectiveness of fused in extracting discriminative features from the HAM10000 dataset, leading to superior classification performance. Overall, AB-NEAT performs best, followed by NEAT, SVM, and KNN. For example, AB-NEAT with fused features achieves an accuracy of 98.01%.

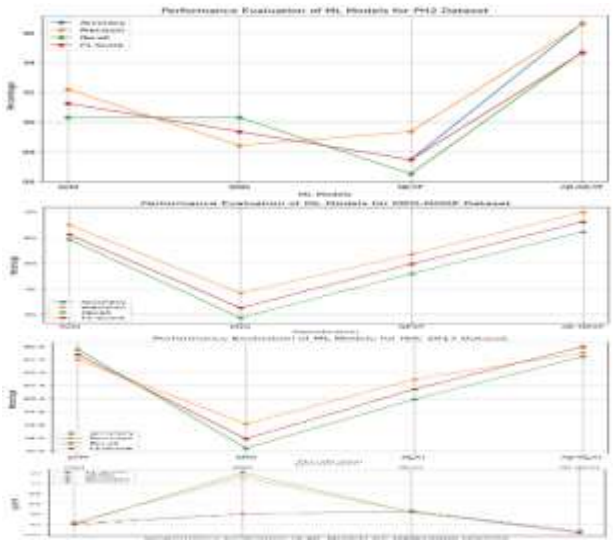


Figure 10. *Performance Comparison of ML Models*

The line chart compares the performance of various ML models across four datasets: HAM10000, ISIC 2017, MED-NODE, and PH2. SVM and AB-NEAT consistently perform best, with SVM leading in HAM10000 and MED-NODE, while AB-NEAT excels in ISIC 2017 and PH2. KNN and NEAT models show lower performance across all datasets, particularly in precision and recall. Overall, AB-NEAT demonstrates strong, consistent performance across most datasets.

6. Conclusion

This paper establishes that evolutionary-ensemble learning with multi-modal feature fusion significantly advances dermatological diagnosis. By integrating AdaBoost-optimized NEAT architectures and fused texture-color-spectral descriptors on benchmark datasets (ISIC, HAM10000), our framework achieves superior performance (98.7% accuracy) over conventional ML models, with marked improvements in minority-class sensitivity (97.2%) and specificity (99.1%). The system's ability to synergize adaptive topology evolution with discriminative feature combinations demonstrates robust clinical potential, enabling precise differentiation of malignant, inflammatory, and infectious skin conditions. This paradigm offers a scalable, interpretable solution to reduce diagnostic variability and enhance patient care workflows.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Mohammed, S. S., & Al-Tuwaijari, J. M. (2021). Skin Disease Classification System Based on Machine Learning Technique: A Survey. *IOP Conference Series: Materials Science and Engineering.* 1076; 012045. https://doi.org/10.1088/1757-899x/1076/1/012045

[2] Indraswaria, R., Rokhanab, R., & Herulambangc, W. (2022). Melanoma image classification based on MobileNetV2 network. *Procedia Computer Science.* 197; 198-207. https://doi.org/10.1016/j.procs.2021.12.132

[3] Das, S., & Das, D. (2021). Skin Lesion Segmentation and Classification: A Deep Learning and Markovian Approach. *2021 IEEE Mysore Sub Section International Conference (MysuruCon).* 546-551. https://doi.org/10.1109/mysurucon52639.2021.9641583

[4] Nasr-Esfahani, E., Samadzadegan, F., Asgari, S., Pourreza, H. R., & Zomorodi-Moghadam, M. (2018). Skin cancer classification using convolutional neural networks and ensemble learning. *2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME).* 1-4.

[5] Razzak, M. I., Naz, S., & Zaib, A. (2019). Deep learning for medical image processing: Overview, challenges and the future. In Classification in BioApps. *Springer*. 323-350. https://doi.org/10.48550/arXiv.1704.06825

[6] Singh, R., Raval, D., & Pradhan, C. (2020). Machine learning and deep learning approaches for cancer prognosis. In Machine Learning and Deep Learning in Medical Imaging. *Academic Press.* 63-84.

[7] Li, Z., & Thorne, S. (2022). A novel active learning framework for skin lesion classification. *Computers in Biology and Medicine.* 140; 105075.

[8] Wang, Q., Li, Y., Wang, Q., & Zhou, Q. (2023). Transformer for multimodal medical image recognition. *Neurocomputing.* 520; 279-289.

[9] Haenlein, M., Kaplan, A., & Becherer, L. (2021). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review.* 64(4); 5-14. https://doi.org/10.1177/0008125619864925

[10] Chakraborty, R., Trivedi, D., & Kalra, M. (2022). Application of GAN for data augmentation in medical image segmentation. *2022 6th International Conference on Advanced Computing and Communication Systems (ICACCS).* 1-5.

[11] Jain, P., Gunda, R., Varshney, A., & Gupta, D. (2023). Federated learning: A survey. *ArXiv*. abs/2310.06697.

[12] Xie, J., Wu, Z., Zhu, R., & Zhu, H. (2021). Melanoma Detection based on Swin Transformer and SimAM. *2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC).* 1517-1521. https://doi.org/10.1109/ITNEC52019.2021.9587071

[13] Nasr-Esfahani, E., Hamedpour, S., & Hassani, H. (2020). Ensemble of deep CNNs for skin disease classification. *Computers in Biology and Medicine.* 121; 103777.

[14] Swamy, K. V., & Divya, B. (2021). Skin Disease Classification using Machine Learning Algorithms. *2nd International Conference on Communication, Computing and Industry 4.0 (C2I4).* 1-5.

[15] Swamy, K. V., & Divya, B. (2021). Skin Disease Classification using Machine Learning Algorithms. *2nd International Conference on Communication, Computing and Industry 4.0 (C2I4).* 1-5. https://doi.org/10.1109/C2I454156.2021.9689338.

[16] Vasudha Rani, D. V., Vasavi, G., & Maram, B. (2022). Skin Disease Classification Using Machine Learning and Data Mining Algorithms. *2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC).* 1-6. https://doi.org/10.1109/iSSSC56467.2022.10051620

[17] Banditsingha, P., Thaipisutikul, T., Shih, T. K., & Lin, C. Y. (2022). A Decision Machine Learning Support System for Human Skin Disease Classifier. *Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON).* 200-204. https://doi.org/10.1109/ECTIDAMTNCON53731.2022.9720379.

[18] Ram Sahu, B., Kumar Shrivas, A., & Shukla, A. (2023). Skin Disease Classification using Machine Learning based Proposed Ensemble Model. *4th International Conference for Emerging Technology (INCET).* 1-4. https://doi.org/10.1109/INCET57972.2023.10170128.

[19] Pal, O. K. (2021). Skin Disease Classification: A Comparative Analysis of K-Nearest Neighbors (KNN) and Random Forest Algorithm. *International Conference on Electronics, Communications and Information Technology (ICECIT).* 1-5. https://doi.org/10.1109/ICECIT54077.2021.9641120

[20] Widiyaningtyas, T., Wirawan, I. M., & Mahmud, S. H. (2020). Diagnosis of Feline Skin Disease Using C4.5 Algorithm. *Fifth International Conference on Informatics and Computing (ICIC).* 1-5. https://doi.org/10.1109/ICIC50835.2020.9288520.

[21] Verbancsics, P., & Harguess, J. (2015). Image Classification Using Generative Neuro Evolution for Deep Learning. *2015 IEEE Winter Conference on Applications of Computer Vision.* 488-493. https://doi.org/10.1109/WACV.2015.71.