International Journal of Advances in
Soft Computing and Intelligent Systems (IJASCIS)
2026, Vol 05, Issue 01, 01-14
ISSN: 3048-4987

Original Paper

A Comparative Analysis of Supervised Machine Learning Algorithm for Automatic Skin Disease Identification

Hiral S. Padhariya [a,*], Shailesh Chaudhari[b]

[a]Veer Narmad South Gujarat University, Surat, Gujarat, India.

[b]Department of Information and Communication Technology, Veer Narmad South Gujarat University, Surat, Gujarat, India.

[a] hiralpadhariya91@gmail.com, [b]sachaudhari@vnsgu.ac.in

ABSTRACT

Dermatologist disease comprises a broad spectrum of conditions that impair the integrity and physiological function of skin. Diagnosing skin conditions can be both time consuming and subjective, when conducted manually by medical professionals. This highlights the need for automatic prediction of skin disorder. In our prior research, we focused on classifying skin lesion using KNN model, conducted using the HAM10000 dataset. The KNN model demonstrated accuracy rate of 97%, indicating an efficient approach to support preliminary recognition and classification of skin lesion. Further, extending the earlier work, in this study we conduct a comparative evaluation of multiple supervised learning techniques for skin disorder identification tested on HAM10000 and ISIC datasets. Our analysis focuses on metrics to gauge the effectiveness of each method in classifying skin lesion. Our findings reveal that SVM, achieve better results than conventional machine learning approaches in terms of accuracy and robustness.

KEYWORDS

Medical Imaging, Automated Learning Algorithms, Digital Image Analysis, Supervised Learning Techniques.

1. INTRODUCTION

Skin Disease are among the leading health concerns affecting millions of individuals and manifesting in various forms, from benign conditions like infections, acne, eczema and many more [1]. The International public health under the United Nations estimates that skin disorders account for a substantial share of global health issues, contributing to both physical discomfort and psychological distress. Early and accurate diagnosis is vital for effective treatment and management; however, the reliance of human dermatologists introduces challenges such as variability in expertise, accessibility and time constraints [2].

Machine Learning is transforming the healthcare industry through enhanced diagnosis and detection of epidermis. Further, advancements in digital imaging and advances in machine learning have enabled the development of novel methods to improve diagnostic processes in dermatology [3]. Methods under the Supervised algorithmic techniques have gained prominence as a transformation solution, enabling machine-driven recognition and dermatological classification for clinical image analysis. By training algorithms on extensive dataset of labelled skin images, these models can learn to recognize patterns and features indicative of various conditions, thereby providing rapid and reliable diagnostic support [4].

The Landscape of Guided learning methods encompasses a diverse array of algorithms, including Linear Classifier that maximizes the margin between data classes, Probabilistic classifier based on Bayes' theorem, K-Nearest Neighbor (KNN), Random Forest and Decision tree. Each method comes with possesses unique traits that influence its performance, robustness and applicability in the scope of skin condition identification [5]. In our field of study, we distinguished and assessed several techniques, including Kernel-based classifier employing a Radial Bases Function (RBF), K-Nearest Neighbor (KNN), Linear Kernel-based Support Vector Model, Naïve Bayes, Random Forest, Polynomial Support Vector Classifier and Decision Tree. Besides that, HAM10000 and ISIC raw data, which are forthrightly accessible assets for lesion classification, have contributed to the current model.

The key objective of the findings is: First, to rigorously evaluate the predictive performance of selected algorithms using key metrics such as accuracy, F1 score, recall and precision; second to access their real-world computational effectiveness. By systematically comparing these algorithms, this study intends to elucidate which methods offer most promise for enhancing skin identification and supporting professionals in their diagnostic endeavors. By leveraging the capabilities of these algorithms, we can develop sophisticated automated systems that empower dermatologists with timely and informed decision-making tools, ultimately leading to improved patient outcomes and advancing in field of dermatology.

Research Study have identified a wide range of impactful applications among broad range of algorithms describes that Support Vector Machine was extensively applied in skin cancer detection [6]. For instance, the study conducted by author in [7] finds that SVM are profoundly primarily adopted method for classification. The research outcomes reveal will assist healthcare professionals in diagnosing skin disease at early stage and preventing further deterioration. The research in [8] utilized KNN for classifying multiple forms of skin disorder. Decision Tree and their ensemble variant, Random Forest strategy was carried out to predict skin disorder based on image features.

This paper offers an in-depth description of the relative effectiveness among the diverse methods in supervised learning framework for skin disease prediction. Researcher can choose the best supervised model for their research, guided by essential comparative performance comparison metric

2. RESEARCH BACKGROUND

The following section comprises an outline of recent data-driven models aimed at classifying skin illness.

In [9], the author introduced a new approach for classifying and segmenting dermatological lesion employing SVM model and K-Nearest Neighbor (KNN) classifier for feature extraction. The study achieved an F1 score of 34% with KNN, 46.71% with SVM, and 61%, when combining both algorithms. In [10], the researcher designed a system to discerning among skin disease based on skin color by analyzing texture features. The study also compared the outcome of K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), achieving an impressive accuracy of 98.2%. In [11], the author explores the Gray Level Cooccurrence Matrix (GLCM) and K-Nearest Neighbor (KNN) implementation in classifying two types of skin disorder using ISIC dataset. Various KNN parameters were evaluated, including a single neighbor and distance metrics. The results yielded maximum accuracy for the various dataset size.

In [12], author suggested approach that classifies skin ailment using computer vision and computer vision techniques. A Gaussian filter is applied to reduce noise, and the Grab Cut technique is then used for segmentation. Margin-based classifier was used for classification, with an emphasis on textural characteristics such as eccentricity, area and perimeter. The m-Skin Doctor device exhibits a 75% specificity and 80% sensitivity. In [13], the author introduces a classification system that employs the K-Nearest Neighbor (KNN) algorithm to effectively differentiate between normal and malignant skin lesion. KNN reached a notable accuracy of 98% in this classification task. In [14], author developed dynamic graph cut technique for segmenting skin lesion, complemented by naïve bayes classifier with the objective to classify illness of skin. The level of excellence for keratosis is 92.9%, for benign its 94.3% and for melanoma it is 91.2%.

In [15], author examine information gathered from tertiary medical facilities in Alappuzha and Kottayam, India, Kerela. Employing the naïve bayes technique, author determine the prevalence rate of different dermatological conditions and assess their likelihood. In [16], author make an attempt to compare random forest and decision tree method using various data set. Feature representations with superior spatial precision are generated by this method while preserving spatial information. The method outperformed current algorithm with respect to accuracy and showed resilience to skin image aberrations like hair filaments. In [17], the author uses both over-sampling and under-sampling strategies by using Random Forest model, demonstrating high performance in predicting cases of lumpy infection.

In [18], the study addresses the problems with unbalanced dataset and the shortcoming of decision tree technique as discussed by the author. The findings emphasize the obligation of sophisticated classification strategies and data balancing approaches. Moreover, significance of acquiring features and pre-processing. In [19], the author investigates how well Random Forest and Linear Regression method perform, seeking to ascertain which variation leads the others using the Mean squared loss function and R-squared measures. Important parameters were identified using the Boruta feature selection approach, revealing that Random Forest performs noticeably better in terms of predictive power than Linear Regression. In [20], the study examines the efficacy of Support Vector Classifier approach and provides an epidemiological framework for categorizing common skin disorder. The author advocate for SVM as a reliable predictive tool, achieving a 95.39% success rate with a diversity of kernels.

In [21], the author incorporates support vector system and radial basis kernel technique exhibit 86% sensitivity, 84% specificity, 88% precision and 85% accuracy rate to distinguish skin tumors in dermoscopic visuals. In [22], the author leveraged Support Vector technique along with linear kernel to detect and recognize erythemato-squamous illness. Their methodology yields greater accuracy contrary to other kernel approaches. Three-fold cross-validation was used thereby maximizing the model parameters, highlighting their significance in the process of classification.

The literature on machine learning for skin condition diagnosing hinders boundaries such as lack of longitudinal studies to gauge model efficacy over time, inadequate research on uncommon skin condition, and the requirement for efficient machine learning tool integration into clinical workflows constitute a few underneath the foremost gaps.

The overarching target of the research was to explore accessible, efficient and transparent machine learning solutions that could be developed to facilitate early identification of epidermal condition using public dataset. The aim was not to build accurate yet practical models, enabling access to machine learning for real-world healthcare, especially in resource-limited settings.

3. METHODOLOGY

This segment provides a thorough examination of the approach deployed to assess different machine learning techniques applied to automatically recognize skin condition using HAM10000 and ISIC research data.

3.1 Block Diagram of Methodology

Figure 1 illustrates the structured workflow intended for identifying skin ailments, covering key stages such as preprocessing, feature identification, and model evaluation. Each phase is integral to ensuring precise and legitimate skin classification disorders [23]. The key aspects of the procedure are elaborated in the below sub-section.

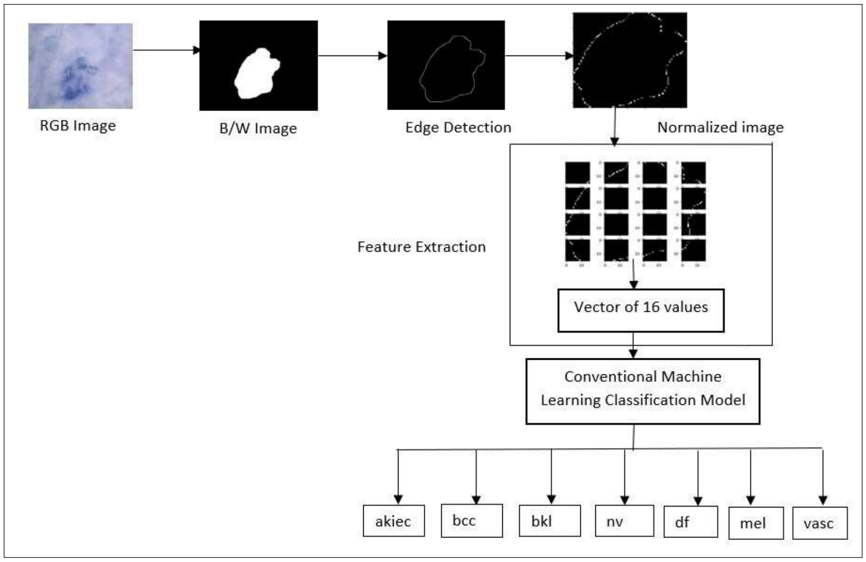

Fig. 1. Schematic representation of proposed diagnostic model.

3.2 Dataset Description

In this study, we employed the HAM10000 and ISIC dataset, recognized among top tier, large scale collection available to public for skin disease classification. This dataset encompasses real-world variations in lesion appearance, skin pigmentation and imaging conditions, providing a practical and challenging benchmark for machine learning classification models. The detailed description of the dataset used in proposed work is described below.

HAM10000 Dataset: HAM10000 Dataset includes 10015 dermatoscopic visuals which are segmented into seven types of skin abnormalities: vascular lesions, benign keratosis, actinic keratosis, melanoma, squamous cell carcinoma, seborrheic keratosis and basal cell carcinoma [24].

ISIC Dataset: The ISIC Dataset purpose is to make image diagnosis belonging to eight skin condition with 25,331 images: actinic keratosis, vascular lesion, squamous cell carcinoma, melanoma, melanocytic nevus, dermatofibroma, benign keratosis and basal cell carcinoma [25].

For this present research, we established a dataset that comprised of 100 images across seven distinct categories of HAM10000 dataset and similarly, 100 images over seven different categories of ISIC dataset. For maintaining uniform representation across all classes throughout the study, we extracted 100 images per category to ensure a balanced dataset.

3.3 Image Preprocessing Pipeline

Effective preprocessing is critical step in any machine learning pipelines, notably when addressing with medical data such as dermatoscopic cutaneous lesion images. So, before feeding the data from the

dermal in relation to machine learning model, preprocessing steps were applied to strengthen image quality, ensure consistency, and extract relevant features.
The preprocessing pipeline used in this study includes the following stages [26]:

Image Acquisition: During this stage, the dataset was imported and prepared for preliminary inspection.

Image Resizing: All images were resized to 64×64 pixels to standardize input dimensions and reduce computational overhead.

Grayscale Conversion: The input image is read in its original color format (Red, Green, Blue channels). To reduce computational complexity and focus on intensity-based features, the RGB images were converted to grayscale.

Image Normalization: To ensure numerical stability and improve model convergence, pixel brightness levels were scaled to a uniform range. This was achieved by normalizing each pixel to the [0,1] interval through division by 255, standardizing the input across all images.

Label Encoding: Each lesion type was converted into a unique integer label to support multiclass classification. This paradigm changes feasible machine learning models to process categorical diagnoses as numerical values.

Segmentation: In this step, thresholding segments the lesion by conversion of grayscale images to binary form, separating the foreground (lesion) from the background. This simplifies the image and isolates the region of interest for further analysis.

Edge Detection and Contour Extraction: In our present work we adopted Canny edge detection algorithm to precisely delineate the lesion edges. Contours are extracted for outline lesion edges and support shape-based feature analysis for classification.

Bounding Box Computation: To localize the lesion area, a bounding rectangle is generated around each detected. The resulting region of interest enables targeted feature extraction while reducing unnecessary computational overhead.

Feature Extraction: As convention models for machine learning cannot effectively process raw image pixels; it is necessary to extract handcrafted attributes that encapsulate relevant visual information. At this phase each image is segmented into a 4×4 grid equal size of (16 x 16 pixel), for 16 non-overlapping regions to enable localized analysis of color, texture, and shape characteristics. Further, for each region, the pixel density is determined by dividing the count of white pixels by that of black pixels, capturing lesion intensity.

This procedure is applied to each of the 16 regions, generating a feature vector of 16 values that captures spatial distribution characteristics.

Feature Scaling: Standardization was applied using a Standard Scaler to normalize features scaled to zero mean and unit variance, boosting model effectiveness and convergence.

Data Splitting: The raw data was split to create distinct evaluation and training purpose to ensure robust model evaluation. A method of stratified splitting was put forward to maintain the original class distribution across all subsets. The data was partitioned using a common ratio of 80% for training, and 20% each for testing.

3.4 Analysis and Comparative Study of Computational Learning Models for Skin Inflammation Detection

In this element of the research, we evaluate and contrast multiple standard machine learning procedures aimed at detecting inflammation in visual data of skin lesion. In this study, we employ five distinct classifiers: K-Nearest Neighbor (KNN), Naïve Bayes, Decision Tree, Random Forest and Support Vector Machine (SVM) to evaluate and compare their performance for classification of skin disorders.

K-Nearest Neighbors (KNN): During this research, we harnessed the k-Nearest Neighbors (k-NN) algorithm as a classification approach, chosen for its straightforward implementation, robustness, and non-parametric characteristics [27].

During our examination of skin lesion detection, the k-Nearest Neighbors (k-NN) algorithm facilitates the grouping of lesions through the metric-based assessment of extracted from dermoscopic visual data. Each lesion is mapped to a point within a high-dimensional feature space, where decisions making stage are made by evaluating the predominant class among the 'k' nearest neighbors, determined using an appropriate distance metric as the Euclidean distance. Within this analysis, the k-Nearest Neighbors classifier was configured with value k=1 to optimize classification performance.

Naïve Bayes: Naive Bayes represents a family of probabilistic classifiers founded on Bayes' Theorem, with the key assumption that given the class, the features are conditionally independent. The Gaussian Naive Bayes variant is particularly well-suited for handling continuous data, under the assumption that the feature values are normally distributed [28].

For present research, Naive Bayes is applied in skin lesion identification by estimating the probability of a lesion belonging to a particular class based on extracted features. The Gaussian variant is effective for continuous data, making it suitable for classifying skin conditions using features derived from medical images. The present study utilizes the Gaussian Naive Bayes classifier to perform classification tasks within the model.

Decision Tree: Decision Tree is a non-parametric, supervised learning approach trained to execute classification by deriving decision criteria from data analysis. For skin ailment classification, Decision Trees are utilized to differentiate between conditions by learning a series of feature-based rules to construct a tree-like model that supports accurate and interpretable diagnosis of various skin lesions [29].

For study, decision tree was configured using the Gini impurity criterion for node splitting, with no restriction on maximum tree depth. Besides that, the model required one at minimum sample per leaf node and a we used minimum of five samples to split an internal node.

Random Forest: Random Forest aggregates the forecasting from a group of decision tree trained on different data samples, using majority voting to ascertain the final classification and improving model generalization by reducing overfitting and increasing accuracy [30].

In this analysis, the Random Forest algorithm was utilized owing to its proficiency to effectively model non-linear relationships and its resilience to noisy features. Additionally, its interpretability and capability to assess feature importance make it particularly valuable for clinical decision-making. Random Forest classifier was tuned using minimum of 2 samples per leaf node and a minimum of 5 samples required to split an internal node.

Support Vector Machine (SVM): SVM is a robust supervised learning algorithm widely applied in binary classification. It operates by finding the optimal hyperplane that maximizes the margin between data points of different classes. It employs kernel methods to transform the data segmented to a higher-dimensional space, enabling the uncovering of an optimal decision boundary between classes [31].

Support vector-based classifier is widely used in skin disease detection by classifying skin lesion images based on extracted features. After feature extraction, SVM identifies an optimal hyperplane that separates different lesion classes. Through kernel transformations, SVM is capable of handling non-linear separations in data, boosting its effectiveness in distinguishing between complex skin lesion categories [32]. A regularization constant of 10 was chosen for the SVM, and class weights were not specified, implying equal importance across classes.

3.5 Hyperparameter Optimization

Hyperparameter tuning was conducted in our research towards better model superior classification results using grid search combined with cross-validation, ensuring the selection of ideal model parameters [33]. Table 1 presents the tuned hypermeters settings optimized for the proposed research.

Table 1. Hyperparameter optimization in dermatologist diagnosis

Algorithm	Tuned Hyperparameter
KNN	n_neighbors = 1, weight = ‘uniform’, metric = ‘euclidean’
Naïve Bayes	var_smoothing = 1e-9
Decision Tree	max_depth = 5, min_samples_split = 2, min_samples_leaf=1, criterion = ‘entropy’
Random Forest	n_estimators=50, max_depth=None, min_samples_split=2, max_features = ‘sqrt’, min_samples_leaf = 1
SVM (Kernel = Linear)	C = 10, class_weight = None, max_iter = 500
SVM (Kernel = Polynomial)	C = 100, degree = 1, coef0 = 0, class_weight = None
SVM (Kernel = Radial Base Function)	C = 10, gamma = 0.1, class_weight = None

4. MODEL TRAINING AND EVALUATION

In this study, five traditional machine learning algorithms: SVM, Random Forest, Decision Tree, k-NN and Naive Bayes were trained on handcrafted feature vectors derived from dermatoscopic skin lesion images. We incorporated HAM10000 dataset which contains 10015 images and ISIC dataset which comprises 25,331 visuals. To prepare dataset we acquired 100 data samples cross all category. The dataset was segmented into training and evaluation phase with an 80:20 ratio prior to model development and 5-fold cross-validation was applied for model generalization and hyperparameter tuning. Feature normalization was performed using StandardScaler, and optimal parameters were identified via GridSearchCV. The best configuration pertaining to indivudial model was selected for final evaluation.

After model tuning, effectiveness was analyzed on the unseen test set to test their classification proficiency. Several standard evaluation metrics were used, including accuracy, precision, recall, and F1-score, each providing insights into different aspects of model performance. A confusion matrix was also analyzed to assess true and false classification rates, which is crucial due to the clinical impact of misdiagnoses. Results showed that SVM exhibited the most favorable results across selected metrics [34].

5. RESULTS AND DISCUSSION

The current section examines the outcomes of applying several machine learning models to autonomous detection of skin disease.

5.1 Performance Comparison

The Figures 2 and 3 compares the performance assessment of multiple supervised machine learning model. Each algorithm follows a 5-fold cross validation on the HAM10000 dataset and ISIC dataset. This reduces the possibilities of overfitting or underfitting to any particular subset of data and assures that the evaluation is reliable.

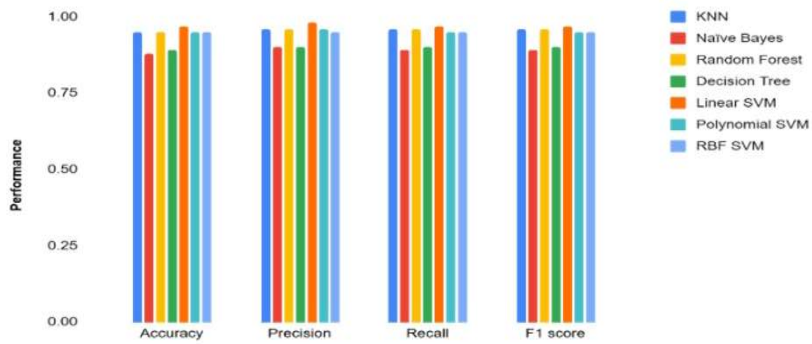

Fig. 2. Performance comparison Graph on the HAM10000 dataset.

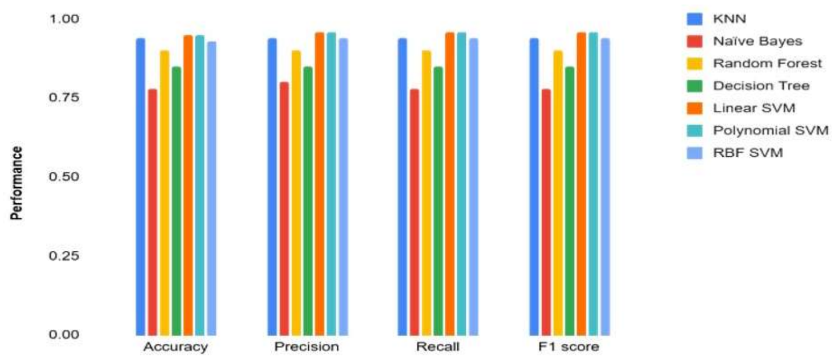

Fig. 3. Performance comparison Graph on the ISIC dataset

The outcomes presented in Figure 2 and 3 reveals that the KNN algorithm succeeds well on ISIC and HAM10000 publicly accessible dataset. The Naïve Bayes algorithm exhibits superior precision on the HAM10000 dataset in comparison to the ISIC dataset. The Random Forest and Decision Tree algorithms yield comparable and favorable results on HAM10000 and ISIC dataset. Support Vector Machine achieves superior outcome with respect to performance on the two datasets in comparison with other algorithms.

The findings indicate that the potency of Random Forest, K-Nearest Neighbor and Support Vector Classifier was equivalent, ranging from 95-97% on ISIC and HAM10000 datasets. KNN and Random Forest performs well but has a lesser precision, while SVM has a high precision for HAM1000 and ISIC datasets. When contrast against SVM, KNN and Random Forest; Decision Tree and Naïve Bayes performed more dire. Analogously, in contrast to Bayesian classifier and Decision Tree; KNN, Random Forest and SVM outperforms them on both the datasets, according to insights conducted from distinct skin categories. Overall, the Support Vector Method outperformed the others in corresponds to ISIC and HAM10000 datasets in relation to widely used machine learning framework, SVM continuously produced the best accuracy, recall, F1 score and precision [35].

5.2 Comparative Analysis Conducted on Several Skin Category

The Figures 4 and 5 illustrates analysis of performance across distinct skin categories using various supervised machine learning techniques applied to the ISIC and HAM10000 datasets that are openly accessible.

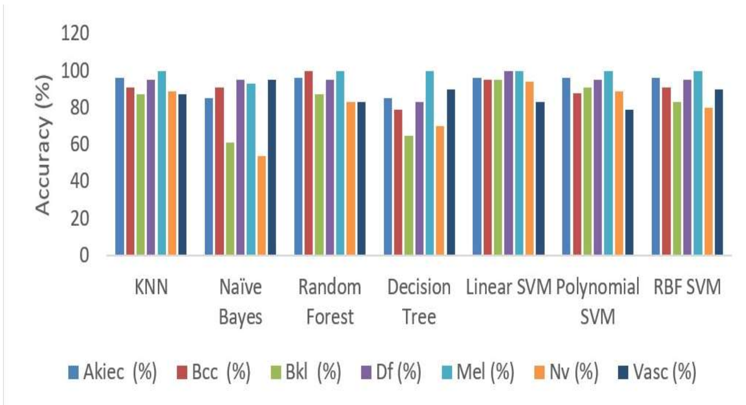

Fig. 4. Comparative research across different skin types on HAM10000 dataset.

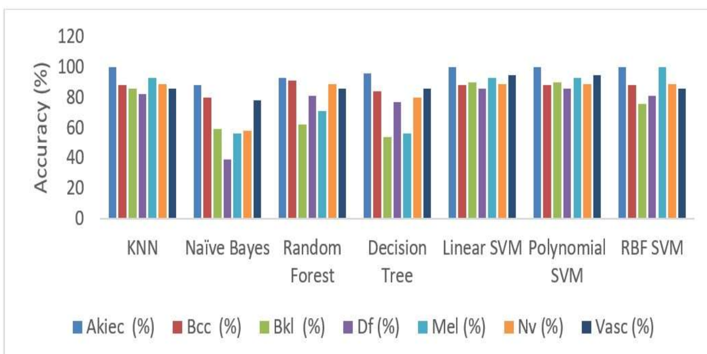

Fig. 5. Comparative research across different skin types on ISIC dataset.

After analyzing the findings produced by multiple machine learning models across different skin segments, as depicted in the image below, we found that KNN classifier demonstrates the most favorable performance on the Actinic Keratoses and Melanoma pigmented lesions, while also delivering strong results across all seven skin lesions for HAM10000 and ISIC datasets. The Naïve Bayes model yields unfavorable results on the Melanocytic nevi for both the datasets. Additionally, it shows suboptimal performance on the Benign Keratosis, Dermatofibroma and Melanoma skin categories in the ISIC dataset. The Random Forest approach show high effectiveness results when applied to the Basal Cell Carcinoma and Melanocytic nevi skin categories. Also, it delivers favorable outcomes for all the seven skin categories analyzing samples from HAM10000 collection, outperforming its results using the ISIC archive of dermatoscopic dataset.

In divergence from other computational models discussed above, the SVM delivers more impactful results and achieve overall outstanding results on all seven skin categories in both datasets. In particular, the linear SVM achieves the best accuracy, ranging from 95-100% across all seven skin categories based on data from the HAM10000 collection. It also performs well on the ISIC dataset. differing from other skin categories in both datasets, polynomial and radial basis kernel SVM provides superior results for the Melanoma and Actinic Keratoses skin categories.

In summary, the insights interpreted from our study showed that the margin-based classifier works most effective on both the HAM10000 and ISIC image repositories for all seven skin classifications, whereas KNN and Tree-based ensemble learning model delivers strong predictive performance for a minimal variety of skin disorder. When compared to other conventional machine learning models, SVM consistently scored the leading precision.

6. CONCLUSIONS

The prevalence rate of various cutaneous problems has been escalating regularly. Traditional diagnostic procedures are expensive and also laborious practices. The current research explored the implementation of AI-based predictive classification of skin health conditions via machine learning, utilizing features extracted from dermatoscopic images. A range of supervised learning algorithms for autonomous skin disease recognition are highlighted by comparison study over HAM10000 and ISIC datasets. Specifically, the study evaluated and compared the performance of five commonly used classification algorithms: Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), k-Nearest Neighbors (k-NN) and Naive Bayes (NB). Each model underwent trained with 16 values of feature vector. The training pipeline incorporated standardized preprocessing and feature normalization, followed by robust model validation through 5-fold cross-validation. Hyperparameter was fine-tuned using GridSearchCV to identify the best configurations for each algorithm. The ultimate model performance was analyzed on a distinct test set applying a set of performance indicators, notably preciseness, granularity, consistency, F1-score. Findings indicate that KNN, SVM and Random Forest performs enhance performance in comparison to Naïve Bayes and Random Forest. The proposed model is capable of being applicable to other domains as well. Additionally, this model may be used for various disease along with diverse medical diagnose, incorporating MRIs and X-rays. Also, it improves diagnostic accuracy by precisely identifying skin lesions, reducing human error, and ensuring consistent evaluations. It also enables non-invasive monitoring and remote assessments through telemedicine, streamlining diagnosis and supporting personalized treatment, thereby enhancing patient care. This research is subject to certain constraints, including reliance on handcrafted features that may miss subtle image patterns. Model results might fluctuate across diverse datasets with different image qualities, patient demographics, or disease types. Moreover, clinical application requires integration with realtime imaging systems and further validation in practical settings. For future advancements the current work can be extended by applying diverse machine learning strategies like deep learning and also improvement can also be done by increasing the dataset. Also, we may focus on leveraging deep learning approaches to derive relevant feature or hybrid models combining traditional ML with CNN representations. The ultimate aim is to develop interpretable, intelligent diagnostic tools to support early detection and management of skin diseases.

Authors Contributions:

All authors have equally contributed.

REFERENCES

1. Bhadula, S., Sharma, S., Juyal, P., & Kulshrestha, C. (2019). Machine learning algorithms-based skin disease detection. International Journal of Innovative Technology and Exploring Engineering, 9(2), 4044-4049.
2. Goswami, T., Dabhi, V. K., & Prajapati, H. B. (2020, March). Skin disease classification from image-a survey. In 2020 6th international conference on advanced computing and communication systems (ICACCS) (pp. 599-605). IEEE.
3. Raut, R., Borole, Y., Patil, S., Khan, V. N., & Takale, D. G. (2022). Skin disease classification using machine learning algorithms. NeuroQuantology, 20(10), 9624-9629.
4. Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised machine learning: a brief primer. Behavior therapy, 51(5), 675-687.
5. Muhammad, I. (2015). Supervised machine learning approaches: A survey. ICTACT Journal on Soft Computing.
6. Ahammed, M., Al Mamun, M., & Uddin, M. S. (2022). A machine learning approach for skin disease detection and classification using image segmentation. Healthcare Analytics, 2, 100122.
7. Banasode, P., Patil, M., & Ammanagi, N. (2021, February). A melanoma skin cancer detection using machine learning technique: support vector machine. In IOP conference series: materials science and engineering (Vol. 1065, No. 1, p. 012039). IOP Publishing.
8. Sandhya, N., Bharadwaj, R. S. R., Reddy, R., Reddy, N. A., Maniteja, R., & Reddy, R. E. skin disease classification.
9. Sumithra, R., Suhil, M., & Guru, D. S. (2015). Segmentation and classification of skin lesions for disease diagnosis. Procedia computer science, 45, 76-85.
10. Nosseir, A., & Shawky, M. A. (2019, April). Automatic classifier for skin disease using k-NN and SVM. In Proceedings of the 8th International Conference on Software and Information Engineering (pp. 259-262).
11. Araaf, M. A., Nugroho, K., & Setiadi, D. R. I. M. (2023). Comprehensive analysis and classification of skin diseases based on image texture features using K-nearest neighbors algorithm. Journal of Computing Theories and Applications, 1(1), 31-40.
12. Taufiq, M. A., Hameed, N., Anjum, A., & Hameed, F. (2016). m-Skin Doctor: a mobile enabled system for early melanoma skin cancer detection using support vector machine. In eHealth 360°: International Summit on eHealth, Budapest, Hungary, June 14-16, 2016, Revised Selected Papers (pp. 468-475). Cham: Springer International Publishing.
13. Hatem, M. Q. (2022). Skin lesion classification system using a K-nearest neighbor algorithm. Visual Computing for Industry, Biomedicine, and Art, 5(1), 7.
14. Balaji, V. R., Suganthi, S. T., Rajadevi, R., Kumar, V. K., Balaji, B. S., & Pandiyan, S. (2020). Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier. Measurement, 163, 107922.
15. Manjusha, K. K., Sankaranarayanan, K., & Seena, P. (2014). Prediction of different dermatological conditions using naïve Bayesian classification. International Journal of Advanced Research in Computer Science and Software Engineering, 4(1).

16. Dhivyaa, C. R., Sangeetha, K., Balamurugan, M., Amaran, S., Vetriselvi, T., & Johnpaul, P. (2024). retracted article: Skin lesion classification using decision trees and random forest algorithms. Journal of Ambient Intelligence and Humanized Computing, 15(Suppl 1), 157-157.
17. Suparyati, S., Utami, E., & Muhammad, A. H. (2022). Applying different resampling strategies in random forest algorithm to predict lumpy skin disease. Journal RESTI, 6(4), 555-562.
18. Rismayanti, N., & Azzahrah, S. F. (2024). Classification of Skin Diseases using Decision Tree Algorithm on an Imbalanced Dataset. International Journal of Artificial Intelligence in Medical Issues, 2(2), 83-94.
19. Srivastava, N., Lamba, T., & Agarwal, M. (2019, November). Comparative analysis of different machine learning techniques. In International Conference on Futuristic Trends in Networks and Computing Technologies (pp. 245-255). Singapore: Springer Singapore.
20. Parikh, K. S., & Shah, T. P. (2016). Support vector machine–a large margin classifier to diagnose skin illnesses. Procedia Technology, *23*, 369-375.
21. Rahajeng, A. C. D., Nuh, M., & Hikmah, N. F. (2020, November). An Evaluation Performance of Kernel on Support Vector Machine to Classify the Skin Tumors in Dermoscopy Image. In 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM) (pp. 76-81). IEEE.
22. Basu, A., Roy, S. S., & Abraham, A. (2015, February). A novel diagnostic approach based on support vector machine with linear kernel for classifying the erythemato-squamous disease. In 2015 international conference on computing communication control and automation (pp. 343-347). IEEE.
23. Padhariya, H. S., & Chaudhari, S. (2023, February). Medical Image Processing and Machine Learning: A Study. In International Conference on Intelligent Computing and Networking (pp.107-117). Singapore: Springer Nature Singapore.
24. Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-sources dermatoscopic images of common pigmented skin lesions. Scientific data, 5(1), 1-9.
25. Lins Joy, Y., & Jerine, S. (2025). Efficient Skin Disease Diagnosis Using Optimized nnU-Net Segmentation and Hybrid E-Cap Net with UFO-Net. Optical Memory and Neural Networks, 34(3), 402-417.
26. Pugazhenthi, V., Naik, S. K., Joshi, A. D., Manerkar, S. S., Nagvekar, V. U., Naik, K. P., ... & Sagar, K. (2019). Skin disease detection and classification. International Journal of Advanced Engineering Research and Science (IJAERS), 6(5), 396-400.
27. Sindayigaya, L., & Dey, A. (2022). Machine learning algorithms: a review. Inf Syst J, 11(8), 1127-1133.
28. Reddy, E. M. K., Gurrala, A., Hasitha, V. B., & Kumar, K. V. R. (2022). Introduction to Naive Bayes and a review on its subtypes with applications. Bayesian reasoning and gaussian processes for machine learning applications, 1-14.
29. Navada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011, June). Overview of use of decision tree algorithms in machine learning. In 2011 IEEE control and system graduate research colloquium (pp. 37-42). IEEE.
30. Kulkarni, V. Y., & Sinha, P. K. (2012, July). Pruning of random forest classifiers: A survey and future directions. In 2012 International Conference on Data Science & Engineering (ICDSE) (pp. 64-68). IEEE.
31. Chandra, M. A., & Bedi, S. S. (2021). Survey on SVM and their application in image classification. International Journal of Information Technology, 13(5), 1-11.
32. Patle, A., & Chouhan, D. S. (2013, January). SVM kernel functions for classification. In 2013 International conference on advances in technology and engineering (ICATE) (pp. 1-9). IEEE.
33. Bechelli, S., & Delhommelle, J. (2022). Machine learning and deep learning algorithms for skin cancer classification from dermoscopic images. Bioengineering, 9(3), 97.
34. Bhavsar, H., & Ganatra, A. (2012). A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4), 2231-2307.
35. Sheth, V., Tripathi, U., & Sharma, A. (2022). A comparative analysis of machine learning algorithms for classification purpose. Procedia Computer Science, 215, 422-431.