2023 Global Conference on Information Technologies and Communications (GCITC)
Karnataka, India. Dec 1-3, 2023

2023 Global Conference on Information Technologies and Communications (GCITC) | 979-8-3503-0816-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/GCITC60406.2023.10425889

Skin Cancer Detection Using the HAM10000 Dataset: A Comparative Study of Machine Learning Models

NandaKiran Velaga
Department of Computer Science and Engineering
National Institute of Technology
Raipur, India
nandakiranvelaga09@gmail.com

Venkata Revanth Vardineni
Department of Computer Science and Engineering
National Institute of Technology
Jamshedpur, India
vvrevanth@gmail.com

Phanindra Tupakula
Department of Computer Science and Engineering
National Institute of Technology
Jamshedpur, India
phanindra197@gmail.com

Jaanaki Swaroop Pamidimukkala
Department of Computer Science and Engineering
SRM University
Andhra Pradesh, India
jaanakiswaroop@gmail.com

***Abstract*—Skin cancer is a prevalent and potentially life-threatening disease that requires early and accurate detection for effective treatment. Existing skin cancer detection procedures face limitations in terms of accuracy and efficiency. To overcome these limitations, we put forward a novel solution that leverages machine learning models to classify seven types of skin cancers using the HAM10000 dataset. This dataset consists of diverse clinical images, providing a valuable resource for developing robust classification models. Our approach involves preprocessing the HAM10000 dataset to ensure data consistency and enhance model performance. We apply various machine learning models, including KNN, Decision Tree, Random Forest, Ridge Classifier, and SVM, to classify the different types of skin cancers. Through rigorous evaluation, our models demonstrate promising results. Among the models tested, the Random Forest algorithm exhibits superior performance, achieving a validation set accuracy of 95.5% and a test set accuracy of 95.6%. These results highlight the effectiveness of our proposed methodology in accurately classifying skin cancers and surpassing the performance of traditional detection procedures. By utilizing machine learning techniques, our solution offers an improved approach to skin cancer detection, addressing the limitations of existing methods. The integration of Random Forest further enhances accuracy, enabling early and accurate diagnosis, which is crucial for successful treatment outcomes and improved patient care.**

***Index Terms*—Skin cancer, HAM10000 dataset, classification, machine learning, imbalanced classes, synthetic minority over-sampling technique (SMOTE).**

I. Introduction

Skin cancer is a significant global health concern, affecting millions of people worldwide. It is characterized by the abnormal growth of skin cells, primarily caused by overexposure to ultraviolet (UV) radiation from the sun or artificial sources such as tanning beds. The prevalence of skin cancer continues to rise, making it a pressing public health issue that requires effective detection and treatment strategies. [1] [2]

According to the World Health Organization (WHO), skin cancer ranks among the most prevalent forms of cancer. commonly diagnosed cancers worldwide [1]. Here are some details about skin cancer:

- Incidence: Skin cancer accounts for approximately one-third of all diagnosed cancer cases globally. Each year, there are over 2 million new cases of non-melanoma skin cancer and around 132,000 cases of melanoma.
- Mortality: While non-melanoma skin cancers, such as basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), have low mortality rates, melanoma is a highly aggressive form of skin cancer that can metastasize and lead to death. In 2020, an estimated 66,000 deaths were attributed to melanoma.
- Regional Variations: The incidence and mortality rates of skin cancer vary across different regions. Countries with fair-skinned populations and high levels of UV radiation exposure, such as Australia, New Zealand, and parts of Europe and North America, have higher skin cancer rates compared to regions with lower UV exposure.
- Risk Factors: The primary risk factor for skin cancer is UV radiation exposure, including both excessive sun exposure and the use of artificial tanning devices. Other risk factors include fair skin, a history of sunburns, a family history of skin cancer, a weakened immune system, and exposure to certain chemicals or substances.

A. Challenges in Skin Cancer Detection

Detecting skin cancer poses several challenges for individuals and healthcare professionals alike:

1) Visual Assessment: Visual examination of skin lesions can be subjective, leading to potential errors and misdiagnosis.

Distinguishing between benign and malignant lesions can be challenging, even for experienced dermatologists.

2) Limited Access to Dermatologists : The availability of dermatologists may be limited in certain regions, leading to long waiting times for appointments and delays in diagnosis and treatment. This can adversely affect patient outcomes and survival rates.

3) Misdiagnosis: Skin cancer can be misdiagnosed or overlooked, resulting in delayed treatment and potentially worse patient outcomes. The precision of diagnosis may differ based on the level of expertise of the healthcare provider and the availability of diagnostic tools.

Current Procedures for Skin Cancer Detection: Currently, skin cancer detection primarily relies on visual examination and biopsy. Dermatologists visually inspect suspicious skin lesions, often using dermoscopy, a technique that magnifies the skin's surface for detailed examination [5] [23]. If necessary, a biopsy is performed to acquire a tissue sample for further analysis by a pathologist.

B. Role of Machine Learning in Skin Cancer Detection

Machine learning (ML) techniques hold promise in improving skin cancer detection and diagnosis. ML algorithms have the ability to examine extensive datasets of images and clinical data, identifying patterns and indicators of skin cancer. Here's how ML can contribute to skin cancer detection:

1) Automated Image Analysis: ML algorithms can analyze images of skin lesions, comparing them to a database of known cases to determine the likelihood of malignancy. This can aid in identifying suspicious lesions that may require further examination.

2) Decision Support Systems: ML algorithms can provide decision support to dermatologists, suggesting diagnoses or ranking the likelihood of malignancy based on input data. This aids healthcare professionals in making well-informed decisions about patient care treatment pathways.

3) Telemedicine and Remote Consultations: ML-based image analysis can be integrated into telemedicine platforms, enabling remote consultations between patients and dermatologists. This is particularly beneficial in underserved areas where access to dermatological

To address the challenges of early detection, we have utilized the HAM10000 dataset, a publicly available collection of dermatoscopic images [3] [22]. The dataset consists of 10,015 images, encompassing seven different types of skin lesions: actinic keratoses and intraepithelial carcinoma, basal cell carcinoma (bcc), benign keratosis-like lesions, dermatofibroma, melanocytic nevi, pyogenic granulomas and hemorrhage, and melanoma.

Among the choices, the main challenges in constructing a reliable classification model for skin cancer is the presence of imbalanced classes within the dataset. To overcome this challenge, we have employed the Synthetic Minority Oversampling Technique (SMOTE), a popular approach for addressing class imbalance, to augment the minority classes.

We have conducted experiments using various machine learning models to classify skin lesions accurately. The models include k-Nearest Neighbors (KNN), Decision Tree Classifier, Random Forest, Ridge Classifier, and Support Vector Machines (SVM). Our evaluation metrics include validation set accuracy, the validation dataset used for testing accuracy, providing insights into the effectiveness of every model.

This research paper aims to contribute to the existing body of knowledge Within the domain of skin cancer identification. By leveraging the HAM10000 dataset and employing state-of-the-art machine learning algorithms, we have achieved promising results in accurately classifying various different varieties of skin cancers. Our findings have the potential to enhance early diagnosis and assist dermatologists in order to make well-informed decisions about patient care treatment pathways.

The remaining sections of this paper are structured in the following manner: Section 2 provides an overview of related work in skin cancer detection and the utilization of machine learning techniques.

Section 3 details the methodology employed in our study, including data preprocessing and model selection. The experimental outcomes and findings are presented in Section 4 performance evaluation. Lastly, Section 5 concludes the paper, summarizing the contributions and outlining potential future research directions.

II. RELATED WORK

In recent research endeavors, skin cancer detection and categorization have gotten a lot of attention. Various studies have explored the application of machine learning algorithms [26] and deep learning techniques to improve the accuracy and efficiency of skin cancer detection.

Esteva et al. [6] conducted a study where they trained a deep learning convolutional neural network (CNN) on a large dataset of dermatology images to detect skin cancer. Their model achieved impressive accuracy in distinguishing between malignant melanoma and benign skin lesions. The effectiveness of the CNN model was improved by using transfer learning methods.

In another study, Winkler et al. [7] proposed a skin cancer classification system that combined multiple machine learning algorithms. They used random forest, support vector machines (SVM), and CNN models to achieve accurate classification of skin lesions. The ensemble approach demonstrated improved performance in identifying various types of skin cancer.

Tschandl et al. [3] focused on developing a deep learning algorithm for skin cancer classification. In order to accomplish this, they used a DenseNet architecture and a sizable dataset of clinical photos.The authors' proposed model for identifying melanoma showed great sensitivity and specificity, demonstrating the promise of deep learning for this application.

Additionally, researchers have explored the integration of advanced image processing techniques for skin cancer detection. Kaur et al. [11] proposed a computer-aided diagnosis system that combined image segmentation, feature extraction,

and classification algorithms to accurately identify different skin cancer types. Their system showed promising results in distinguishing between malignant and benign lesions.

In summary, the existing literature underscores the effectiveness of machine learning algorithms, particularly CNN and ensemble models, in accurately detecting and classifying various types of skin cancer [8] [9] [10] [16] [17] [18] [19] [21]. Transfer learning and advanced architectures, such as DenseNet, have shown improvements in model performance. Furthermore, there is potential for improved accuracy and efficiency of skin cancer detection systems by the combination of image processing techniques with machine learning methodologies [24].

III. PROPOSED METHOD

Proposed method for skin cancer detection consists of the following steps: dataset selection and description, pre-processing, and methodology(Fig.1).

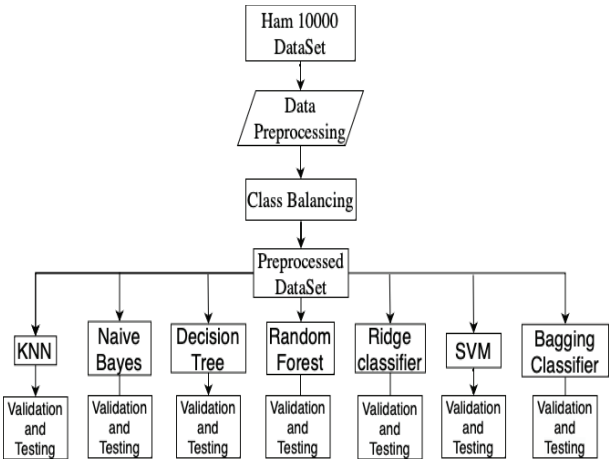

Fig. 1. FlowChart of Proposed Methodology

A. DATASET

1) HAM 10000: The Ham 10000 dataset serves as the foundation for our research. It comprises a collection of clinical images representing different types of skin lesions, including Actinic Keratoses, Basal Cell Carcinoma, Benign Keratosis-like Lesions, Dermatofibroma, Melanocytic Nevi , Vascular Lesions, and Melanoma(fig.2). Each class contains a varying number of samples, resulting in an unbalanced dataset as shown in (Fig.3). [3] [20]

B. PRE-PROCESSING

To ensure the dataset's suitability for training machine learning models, we employ pre-processing techniques. This includes resizing the images to a standardized resolution, normalizing the pixel values, and converting the images to convert it into a format that is appropriate for further analysis. Additionally, we address the issue of class imbalance by applying class balancing techniques to augment the samples of each class,Specifically, we use the Synthetic Minority

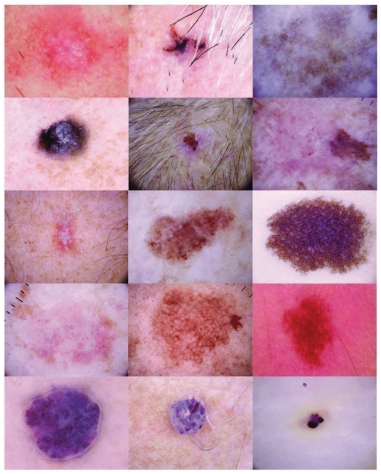

Fig. 2. Sample from the HAM 10000 dataset

Over-sampling Technique (SMOTE) [4] to augment the samples of each class and achieve a more uniform distribution. This ensures that the models are trained on a balanced dataset(fig.3).The balanced dataset contains 6,705 samples per class.

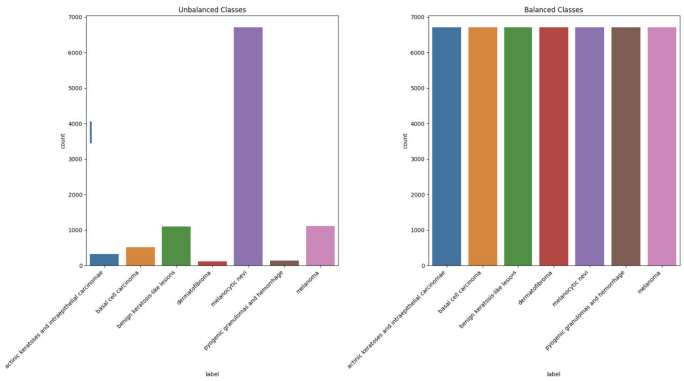

Fig. 3. Bar Chart Representation of Unbalanced and Balanced Datasets

C. METHODOLOGY

In our methodology, we employ various machine learning models for skin cancer classification. The models used include K-Nearest Neighbors (KNN), Decision Tree Classifier, Random Forest, Ridge Classifier, and Support Vector Machines (SVM). These models are selected based on their effectiveness in classification tasks and compatibility with the characteristics of the dataset.

First, we train each model on the pre-processed and balanced dataset. During the training process, hyperparameters are optimized using a validation set. This helps fine-tune the models for improved performance. After training, we proceed to evaluate the performance of each machine learning model using a dedicated test dataset. This evaluation phase is crucial as it helps us assess the model's ability to generalize its learned patterns to unseen data accurately.

Through this methodology, we aim to develop an effective skin cancer detection mechanism that can accurately classify new, unseen skin lesion images into their respective categories, enabling early detection and timely treatment.

1) K-Nearest Neighbours Classifier: Our methodology for training and evaluating a K-Nearest Neighbors (KNN) classifier involves a structured approach. We initially split the dataset into three sets: training, validation, and test, using the train-test-split function from scikit-learn. This method ensures training on a significant portion of the data and evaluation on separate, unseen datasets. We then create a KNN classifier with k=5, train it on the training set, and assess its precision on the validation set, achieving an accuracy of approximately 0.930. To enhance performance further, we conduct hyperparameter tuning using the validation set, optimizing the number of neighbors (n-neighbors) through the GridSearchCV function, resulting in the best hyperparameters and an optimal n-neighbors value. Using these best hyperparameters, we train a final KNN model on the combined training and validation sets. This model is applied to the test set, yielding an accuracy of approximately 0.942.

2) Decision Tree Classifier: Our methodology for training and evaluating a Random Forest classifier using scikit-learn follows a systematic approach. We begin by dividing the data into three sets: a 70% training set and a combined 30% validation-test set, further divided equally into a validation set and a test set. This division allows for robust training on a significant portion of the data and thorough evaluation on entirely separate, unseen datasets.

Next, we create a Decision Tree classifier using the DecisionTreeClassifier class from scikit-learn, configuring it with a random state of 42. The model is then trained on the training set using the fit method. Subsequently, the trained Decision Tree model is employed to predict class labels for both the validation and test sets. Precision is evaluated using the accuracy-score function, resulting in a validation set accuracy of approximately 0.955 and a test set accuracy of approximately 0.956 (rounded to three decimal places).These accuracy values demonstrate the model's effectiveness in accurately classifying new and unseen data during the validation and testing phases. The process maintains a record of these accuracy measurements for both sets to comprehensively assess the model's performance on unseen data.

3) Random Forest Classifier: Our methodology for training and evaluating a Random Forest classifier with scikit-learn involves a systematic approach. We begin by splitting the data into three distinct sets: a 70% training set and a combined 30% validation-test set, which is further divided equally into a validation set and a test set. This division ensures comprehensive training on a substantial portion of the data and rigorous evaluation on entirely separate, unseen datasets.

Next, we create a Random Forest classifier using the RandomForestClassifier class from scikit-learn, configuring it with 100 decision trees (n-estimators=100) and a random state of 42. The model is then trained on the training set using the fit method. The trained Random Forest model is utilized to predict class labels for both the validation and test sets. Precision is evaluated using the accuracy-score function, resulting in a validation set accuracy of approximately 0.955 and a test set accuracy of approximately 0.956 (rounded to three decimal places).These accuracy values demonstrate the model's effectiveness in accurately classifying new and unseen data during the validation and testing phases. The process maintains a record of these accuracy measurements for both sets to comprehensively assess the model's performance on unseen data.

4) Ridge Classifier: In our methodology for training and evaluating a Ridge Classifier with scikit-learn, we initiate the process by splitting the dataset into three distinct sets: training, validation, and test, using the train-test-split function. This division entails a 70% training dataset and a combined 30% validation-test set, further divided equally into a validation dataset and a testing dataset. This meticulous separation ensures that the model is trained on a substantial portion of the data while being rigorously evaluated on entirely separate, unseen datasets.

Following data splitting, we standardize the dataset's features using the StandardScaler to ensure a mean of 0 and a standard deviation of 1 for each feature. This standardization is meticulously applied to the training, validation, and test sets independently to prevent data leakage. Subsequently, we construct a Ridge Classifier model, initialize it with an alpha value of 1.0, and a random state of 42, and then train it on the standardized training dataset. The trained model is leveraged to predict class labels for both the standardized validation and test sets. Precision is calculated using the accuracy-score function, resulting in a validation set accuracy of approximately 0.902 and a test set accuracy of approximately 0.905 (rounded to three decimal places). These values signify the model's effectiveness in classifying new, unseen data during the validation and testing phases. The accuracy values are recorded throughout the process for a comprehensive assessment of the model's performance on unseen data.

5) Support Vector Machine: In our methodology for training and evaluating an SVM Classifier using scikit-learn, we adopt a systematic approach. Initially, the dataset is divided into three sets: a 70% training set and a combined 30% validation-test set, which is further evenly split into a validation set and a test set. This division ensures robust training on a significant portion of the data while rigorously evaluating the model on entirely separate, unseen datasets.

Next, we create an SVM model utilizing the SVC class from scikit-learn, configuring it with a radial basis function (RBF) kernel, a regularization parameter (C) of 1.0, gamma set to 'scale', and a random state of 42. The RBF kernel allows for modeling complex, non-linear decision boundaries, making it a suitable choice for the SVM.

The SVM model is then trained on the training set using the fit method. Subsequently, the trained SVM model is applied to predict class labels for both the validation and test sets. We gauge its performance using the accuracy-score function in scikit-learn, resulting in a validation set accuracy of

approximately 0.909 and a test set accuracy of approximately 0.913 (rounded to three decimal places). These accuracy values reflect the model's effectiveness in classifying new and unseen data during the validation and testing phases. Throughout this process, we continuously monitor and record accuracy measurements for both validation and test sets, enabling a comprehensive assessment of the model's performance on unseen data.

IV. Result and Discussion

This skin cancer detection model utilized machine learning algorithms, namely K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Ridge Classifier, and Support Vector Machine (SVM), to identify skin cancer using the HAM10000 dataset.

The experimental results reveal varying levels of performance among the different machine learning algorithms used in this model. The K-Nearest Neighbors algorithm achieved a test accuracy of 94.2% and a validation accuracy of 93%, indicating its effectiveness in accurately classifying instances of skin cancer based on the features extracted from the balanced HAM10000 dataset. However, it should be stressed that choice of the value for K in KNN can significantly impact the algorithm's performance. Fig. 4 displays the confusion matrix and heatmap obtained through the K-Nearest Neighbors (KNN) methodology.

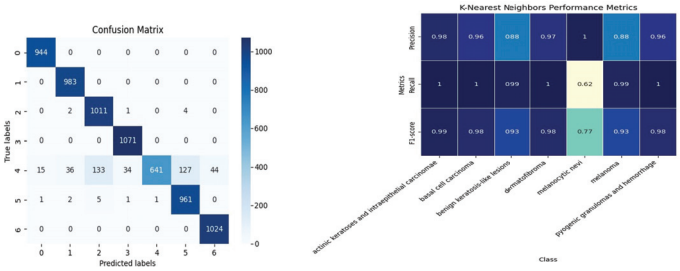

Fig. 4. Confusion Matrix and Heatmap of KNN classifier

The Decision Tree classifier attained a test accuracy of 84.3% and a validation accuracy of 84.8%. The relatively lower performance may be attributed to the inherent limitations of decision trees, such as overfitting to the training data and sensitivity to minor variations in the dataset. Fig.5 displays the confusion matrix and heatmap obtained through the Decision-Tree Methodology.

Conversely, the Random Forest model exhibited outstanding performance with a test accuracy of 95.6% and a validation accuracy of 95.5%. The ensemble approach of combining multiple decision trees in Random Forest mitigates the limitations of individual decision trees, resulting in improved accuracy and generalization. Fig. 6 displays the confusion matrix and heatmap obtained through the Random Forest methodology.

The Ridge Classifier, with a test accuracy of 90.5% and a validation accuracy of 90.2%, showcased consistent performance in the skin cancer detection project using the balanced dataset. This suggests that the Ridge Classifier is a reliable

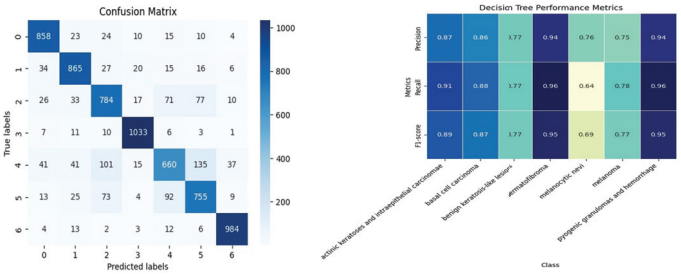

Fig. 5. Confusion Matrix and Heatmap of Decision Tree Classifier

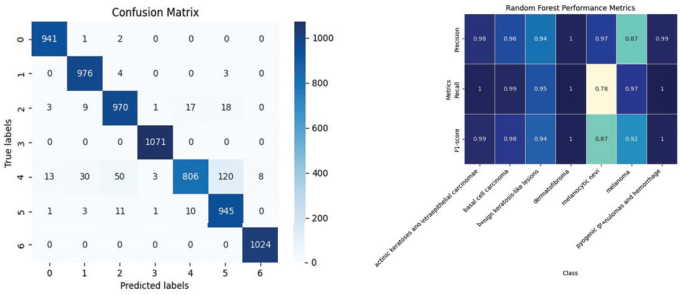

Fig. 6. Confusion Matrix and Heatmap of Random Forest Classifier

choice for classification tasks, particularly when dealing with datasets with high dimensionality. Fig.7 displays the confusion matrix and heatmap obtained through the Ridge Classifier.

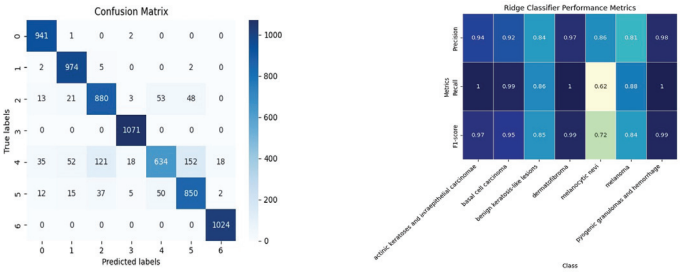

Fig. 7. Confusion matrix and Heatmap of Ridge classifier

Similarly, the Support Vector Machine (SVM) algorithm achieved a test accuracy of 91.2% and a validation accuracy of 90.9%. Fig. 8 displays the confusion matrix and heatmap obtained through the SVM.

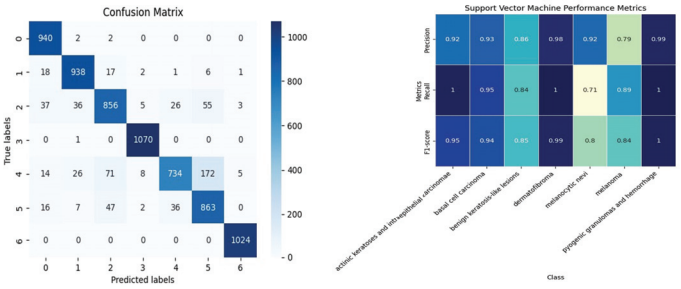

Fig. 8. Confusion Matrix and Heatmap of SVM

Table. IV provides the information of key metrics of each model and the fig.9 shows the visual representation of Testing and Validation Graphs . Overall, The Random Forest method performed best in terms of accuracy, with a test accuracy of 95.6% and a validation accuracy of 95.5%. These findings support the use of the Random Forest ensemble method on the well-balanced HAM10000 dataset for the skin cancer diagnosis job.

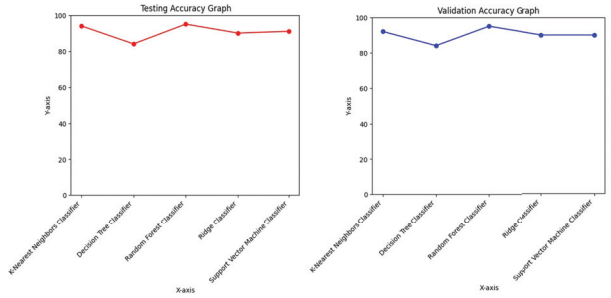

Fig. 9. Accuracy graphs of Testing and Validation

TABLE I
MODEL PERFORMANCE METRICS

Model Name	Accuracy	Recall	Precision	F2-Score
K-Nearest Neighbor Classifier	0.942	0.943	0.946	0.937
Decision Tree Classifier	0.843	0.843	0.840	0.840
Random Forest Classifier	**0.956**	**0.956**	**0.957**	**0.955**
Ridge Classifier	0.905	0.906	0.903	0.901
Support Vector Machine Classifier	0.912	0.913	0.913	0.910

V. CONCLUSION

In conclusion, our skin cancer detection model employed machine learning algorithms, including K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Ridge Classifier, and Support Vector Machine (SVM), on the balanced HAM10000 dataset. The effectiveness of these algorithms in making reliable skin cancer diagnoses was measured.

The experimental results demonstrated varying levels of effectiveness among the different algorithms. The Random Forest algorithm showcased superior performance with a test accuracy of 95.6% and a validation accuracy of 95.5%, being the best-performing algorithm tested here. The ensemble-based approach of Random Forest, combining multiple decision trees, proved to perform exceptionally well at accurately classifying instances of skin cancer.

Moreover, the KNN algorithm achieved a test accuracy of 94.2% and a validation accuracy of 93%, showcasing its potential in accurately classifying skin cancer cases, considering the features extracted from the balanced dataset. Alternatively, the Decision Tree classifier achieved relatively lower test and validation accuracy of 84.3% and 84.8%, respectively. This suggests that decision trees alone may struggle to capture the complexity of skin cancer patterns in the dataset.

The Ridge Classifier and SVM algorithms exhibited consistent performance with test accuracy of 90.5% and 91.2%, and validation accuracy of 90.2% and 90.9%, respectively. Both algorithms demonstrated their capability to handle high-dimensional data and make accurate predictions in the skin cancer detection task.

Overall, our findings highlight the potential of machine learning algorithms in automating the detection of skin cancer, surpassing the accuracy achieved in previous studies [5] [6] [7] [11] [12] [13] [14] [15] [25]. The Random Forest algorithm, in particular, holds promise for developing accurate and reliable skin cancer detection systems.

By advancing the field of automated skin cancer detection, we aim to contribute to improved early diagnosis, better patient outcomes, and reduced burden on healthcare professionals. Future research should explore additional algorithms, feature engineering techniques, and ensemble methods to further enhance the performance of skin cancer detection systems.

REFERENCES

[1] Fijałkowska M, Koziej M, Antoszewski B. Detailed head localization and incidence of skin cancers. Sci Rep. 2021 Jun 11;11(1):12391. doi: 10.1038/s41598-021-91942-5. PMID: 34117314; PMCID: PMC8196200.

[2] Gruber P, Zito PM. Skin Cancer. [Updated 2023 May 14]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK441949/

[3] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data, 5(1), 1-9.

[4] G. A. Pradipta, R. Wardoyo, A. Musdholifah, I. N. H. Sanjaya and M. Ismail, "SMOTE for Handling Imbalanced Data Problem : A Review," 2021 Sixth International Conference on Informatics and Computing (ICIC), Jakarta, Indonesia, 2021, pp. 1-8, doi: 10.1109/ICIC54025.2021.9632912.

[5] Dildar M, Akram S, Irfan M, Khan HU, Ramzan M, Mahmood AR, Alsaiari SA, Saeed AHM, Alraddadi MO, Mahnashi MH. Skin Cancer Detection: A Review Using Deep Learning Techniques. Int J Environ Res Public Health. 2021 May 20;18(10):5479. doi: 10.3390/ijerph18105479. PMID: 34065430; PMCID: PMC8160886.

[6] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 Feb 2;542(7639):115-118. doi: 10.1038/nature21056. Epub 2017 Jan 25. Erratum in: Nature. 2017 Jun 28;546(7660):686. PMID: 28117445; PMCID: PMC8382232.

[7] Winkler JK, Sies K, Fink C, Toberer F, Enk A, Deinlein T, Hofmann-Wellenhof R, Thomas L, Lallas A, Blum A, Stolz W, Abassi MS, Fuchs T, Rosenberger A, Haenssle HA. Melanoma recognition by a deep learning convolutional neural network-Performance in different melanoma subtypes and localisations. Eur J Cancer. 2020 Mar;127:21-29. doi: 10.1016/j.ejca.2019.11.020. Epub 2020 Jan 20. PMID: 31972395.

[8] Nugroho, A. A., Slamet, I., & Sugiyanto, S. (2019, December). Skins cancer identification system of HAMl0000 skin cancer dataset using convolutional neural network. In AIP conference proceedings (Vol. 2202, No. 1). AIP Publishing.

[9] Fu'adah, Y. N., Pratiwi, N. C., Pramudito, M. A., & Ibrahim, N. (2020, December). Convolutional neural network (cnn) for automatic skin cancer classification system. In IOP conference series: materials science and engineering (Vol. 982, No. 1, p. 012005). IOP Publishing.

[10] Garg, R., Maheshwari, S., & Shukla, A. (2021). Decision support system for detection and classification of skin cancer using CNN. In Innovations in Computational Intelligence and Computer Vision: Proceedings of ICICV 2020 (pp. 578-586). Springer Singapore.

[11] Hasan, M. R., Fatemi, M. I., Monirujjaman Khan, M., Kaur, M., & Zaguia, A. (2021). Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks. Journal of Healthcare Engineering, 2021.
[12] Thurnhofer-Hemsi, K., & Domínguez, E. (2021). A convolutional neural network framework for accurate skin cancer detection. Neural Processing Letters, 53(5), 3073-3093.
[13] Moldovan, D. (2019, November). Transfer learning based method for two-step skin cancer images classification. In 2019 E-Health and Bioengineering Conference (EHB) (pp. 1-4). IEEE.
[14] Ali, M. S., Miah, M. S., Haque, J., Rahman, M. M., & Islam, M. K. (2021). An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. Machine Learning with Applications, 5, 100036.
[15] Iqbal, I., Younus, M., Walayat, K., Kakar, M. U., & Ma, J. (2021). Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. Computerized medical imaging and graphics, 88, 101843.
[16] Younis, H., Bhatti, M. H., & Azeem, M. (2019, December). Classification of skin cancer dermoscopy images using transfer learning. In 2019 15th International Conference on Emerging Technologies (ICET) (pp. 1-4). IEEE.
[17] Rezaoana, N., Hossain, M. S., & Andersson, K. (2020, December). Detection and classification of skin cancer by using a parallel CNN model. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 380-386). IEEE.
[18] Agarwal, K., & Singh, T. (2022). Classification of skin cancer images using convolutional neural networks. arXiv preprint arXiv:2202.00678.
[19] Janoria, H., Minj, J., & Patre, P. (2020, November). Classification of skin disease from skin images using transfer learning technique. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 888-895). IEEE.
[20] Chang, C. H., Wang, W. E., Hsu, F. Y., Chen, R. J., & Chang, H. C. (2022, February). AI HAM 10000 Database to Assist Residents in Learning Differential Diagnosis of Skin Cancer. In 2022 IEEE 5th Eurasian Conference on Educational Innovation (ECEI) (pp. 1-3). IEEE.
[21] Chaturvedi, S. S., Gupta, K., & Prasad, P. S. (2021). Skin lesion analyser: an efficient seven-way multi-class skin cancer classification using MobileNet. In Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2020 (pp. 165-176). Springer Singapore.
[22] Jerant, A. F., Johnson, J. T., Sheridan, C. D., & Caffrey, T. J. (2000). Early detection and treatment of skin cancer. American family physician, 62(2), 357-368.
[23] Dorrell, D. N., & Strowd, L. C. (2019). Skin cancer detection technology. Dermatologic clinics, 37(4), 527-536.
[24] Jain, S., & Pise, N. (2015). Computer aided melanoma skin cancer detection using image processing. Procedia Computer Science, 48, 735-740.
[25] Nahata, H., & Singh, S. P. (2020). Deep learning solutions for skin cancer detection and diagnosis. Machine Learning with Health Care Perspective: Machine Learning and Healthcare, 159-182.
[26] Vidya, M., & Karki, M. V. (2020, July). Skin cancer detection using machine learning techniques. In 2020 IEEE international conference on electronics, computing and communication technologies (CONECCT) (pp. 1-5). IEEE.