ISSN (e): 2250 – 3005 || Volume, 13 || Issue, 1|| Jan.-Feb. – 2023 ||
International Journal of Computational Engineering Research (IJCER)

Detection and Classification of Skin Cancer Lesions Using Machine Learning Techniques

[1]Rakesh Kumar Lenka, [2]Pravat Kumar Jena, [2]Dr. Rajesh Kumar Sahoo
[1,2,3]Department of Computer Sc. & Engg.
[1,2,3]Ajay Binay Institute Of Technology

ABSTRACT:
Skin disorders are a common health concern, making early and accurate diagnosis crucial for effective treatment. Among these, skin cancer is one of the most prevalent and life-threatening diseases, where early detection plays a vital role in improving patient outcomes. This paper proposes a machine learning (ML) based approach utilizing convolutional neural networks (CNN), random forest (RF), support vector machine (SVM), and K-nearest neighbors (KNN) to predict and diagnose skin cancer with high reliability. These techniques are employed to differentiate between benign and malignant skin lesions. The model's effectiveness is enhanced through image preprocessing, including image enhancement, feature extraction, and segmentation. By leveraging ML, the proposed approach improves early detection and reduces the risk of misdiagnosis in skin cancer diagnosis.
Keywords: Skin Cancer, Machine Learning, Classification, SVM, CNN, Random Forest(RF).

I. INTRODUCTION

Skin cancer is a highly prevalent and life-threatening disease worldwide, with millions of new cases diagnosed each year. Early detection and accurate classification of skin cancer lesions are essential for improving patient outcomes, reducing mortality rates, and ensuring timely treatment. Traditional diagnostic methods, such as clinical examinations by dermatologists, dermoscopic analysis, and biopsy-based histopathological evaluation, are often time-consuming, subjective, and dependent on medical expertise. However, CNNs and machine learning methods, especially deep learning, have demonstrated remarkable effectiveness in identifying different kinds of skin cancer lesions and evaluating medical pictures.These ML-based approaches leverage large datasets of dermoscopic and clinical images to train models capable of recognizing patterns, textures, and features indicative of malignant or benign lesions. Beyond supervised learning, unsupervised and reinforcement learning techniques have also been explored to improve model performance and adaptability. Various machine learning algorithms, including SVM, RF, DT and KNN have been widely used to enhance classification accuracy. Leveraging machine learning and artificial intelligence can significantly improve the early detection and classification of skin cancer lesions, leading to better patient outcomes, reduced healthcare costs, and increased diagnostic accuracy in dermatology. This paper is structured as follows: Section 2 reviews related work on skin cancer prediction and detection, Section 3 presents the proposed methodology, Section 4 discusses the results, and Section 5 concludes with key findings and future research directions.

II. RELATED WORK

Deep neural networks were utilized by Esteva et al. [1] to classify skin cancer at a dermatologist level, achieving remarkable accuracy in detecting melanoma. It also demonstrated how deep learning techniques can reliably identify cancerous skin lesions. In order to diagnose dermoscopic melanoma, Haenssle et al. [2] compared the diagnostic performance of a deep learning-based convolutional neural network (CNN) to dermatologists. The CNN demonstrated the potential of deep learning systems to improve skin cancer prediction by outperforming a sizable group of dermatologists. Similarly, Brinker et al. [3] compared deep learning algorithms with dermatologists in the classification of dermoscopic melanoma images. Their findings confirmed that deep learning models performed exceptionally and validating their effectiveness in predicting skin cancer. Z. Ma and J. M. R. S. Tavares [4] explored a robust and flexible segmentation method for dermoscopic images, leveraging color space transformation and deformable models to effectively delineate skin lesions. Goswami et al. [5] emphasized the importance of using photographs for skin condition classification, as accurate diagnosis often relies on subjective characteristics that utilize deep learning based approaches for disease classification.

Al-Mani et al. [6] examined automated skin lesion diagnosis using images and proposed an integrated diagnostic system that comprising lesion border segmentation, lesion classification, and the identification of distinct skin conditions.

III. METHODOLOGY

In order to increase diagnosis accuracy, a machine learning-based method that uses CNN, RF, SVM, and KNN is used for skin cancer identification. Data acquisition, the first step in the procedure, involves gathering a variety of skin lesion photos from dermatological databases that are openly accessible. Preprocessing techniques, including image enhancement and contrast adjustment are applied to improve image quality. To ensure reliable analysis, segmentation techniques are used to separate the lesion from the surrounding skin. To differentiate between benign and malignant tumors, the collected characteristics are subsequently input into a number of machine-learning classifiers, including RF, SVM, and KNN. Each model's performance is assessed using common measures including F1-score, recall, accuracy, and precision. These machine learning models suggested methodology is applied to increase diagnostic accuracy, which eventually leads to more successful skin cancer therapy.

IV. RESULTS AND DISCUSSIONS

Several machine-learning models are assessed for accuracy, precision, recall, and F1 score, among other important performance metrics, to identify and classify skin cancer. With a maximum accuracy of 94.3%, CNN demonstrated better performance in differentiating between benign and malignant skin lesions, as shown in Table 1. CNN was likewise the best with 92.8% accuracy, 94.6% recall, and 93.7% F1-score. An excellent substitute, the RF classifier obtained an accuracy of 87.5%, a precision of 85.9%, a recall of 86.2%, and an F1-score of 86.0%. But it performs somewhat worse than CNN. With an F1-score of 81.0%, recall of 81.7%, accuracy of 82.6%, and precision of 80.3%, SVM shown limits in complicated pattern recognition but reasonable efficacy in classification. Because of its sensitivity, the KNN model may not be as good at detecting skin cancer. Its accuracy was the lowest at 79.8%, while its precision, recall, and F1-score values were 77.5%, 78.9%, and 78.2%, respectively. Overall, CNN outperforms the other models due to its deep learning capabilities, making it the most reliable choice for early and accurate skin cancer detection.

Table 1**:** Performance Metrics of Skin Cancer Detection Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	94.3	92.8	94.6	93.7
Random Forest (RF)	87.5	85.9	86.2	86.0
Support Vector Machine (SVM)	82.6	80.3	81.7	81.0
K-Nearest Neighbors (KNN)	79.8	77.5	78.9	78.2

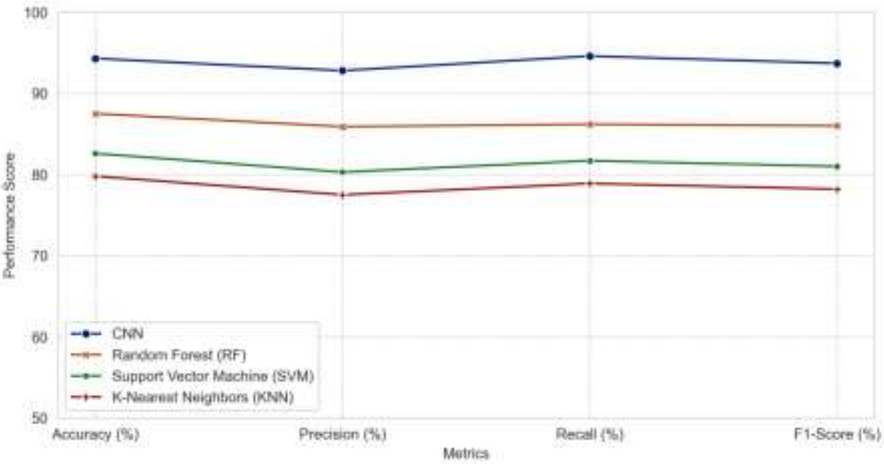

Fig.1: Graphical representation of performance comparison of skin cancer

Accuracy, precision, recall, and F1-score are the four main metrics used to evaluate the efficacy of the CNN, RF, SVM, and KNN machine learning models, as shown in Fig. 1. CNN continuously performs better than the other models on every parameter, obtaining the greatest performance ratings, especially in accuracy and recall, demonstrating its greater capacity to identify skin cancer. RF follows as the second best model, demonstrating stable performance but slightly lower scores compared to CNN. SVM and KNN show lower scores, with KNN performing the worst across all metrics, suggesting that it is less effective for detecting the skin cancer.

V. CONCLUSIONS AND FUTURE SCOPE

This skin cancer detection and categorization approach shows great promise for enhancing early diagnosis and lowering the rate of misdiagnosis. Using sophisticated models like CNN, RF, SVM, and KNN, the system can accurately differentiate between benign and malignant skin lesions. Among these models, CNN exhibits the highest accuracy, making it the most reliable method for skin cancer prediction. In the future, it can be extended by incorporating more diverse and larger datasets to enhance model generalization across various skin types and ethnic backgrounds. Hybrid deep learning models combining CNN with other deep learning architectures, such as transformer-based models, can further optimize accuracy and computational efficiency.

REFERENCES

[1] R. Sumithra, M. Suhilb, and D. S. Guruc, “Segmentation and classification of skin lesions for disease diagnosis,” Procedia Computer Science, vol. 45, pp. 76–85, 2015.

[2] R. Yasir, M. S. I. Nibir, and N. Ahmed, “A skin disease detection system for financially unstable people in developing countries,” Global Science and Technology Journal, vol. 3, no. 1, pp. 77–93, 2015.

[3] International P.Anitha, G.Krithka, Mani Deepak Choudhry, "Machine Learning Techniques for learning features of any kind of data: A Case Study," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Vol.3,No.12,2014.

[4] Z. Ma and J. M. R. S. Tavares,” A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model,” IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 2, pp. 615-623, 2016.

[5] T. Goswami, V. K. Dabhi, and H. B. Prajapati, “Skin Disease Classification from Image - A Survey,” International Conference Advancement Computer. Communication System ICACCS, pp. 599–605, 2020.

[6] M. A. Al-many, D. Kim, and T. Kim, "Computer Methods and Programs in Biomedicine Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification,” Computional Methods Programs Biomedicine, vol. 190, p. 105351, 2020.