International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0212
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1304 - 1306

Skin Cancer Prediction Using KNN

A Navaneetha Krishnan[1], A Punitha Bridjith[2], M Nandha Kumar[3], K Priyadharshini[4]
[1,2,3]UG Scholar, Dept. of CSE, Kamaraj college of Engg. & Tech., Virudhunagar, Tamilnadu, India.
[4]Assistant Professor, Dept. of CSE, Kamaraj college of Engg. & Tech., Virudhunagar, Tamilnadu, India.
Emails ID: *21ucs043@kamarajengg.edu.in[1], 21ucs044@kamarajengg.edu.in2, 21ucs124@kamarajengg.edu.in[3], priyadharshinicse@kamarajengg.edu.in[4]*

Abstract
This Skin cancer is one of the most prevalent and life-threatening diseases worldwide, making early detection essential for effective treatment. This project aims to develop a skin cancer prediction system using the K-Nearest Neighbors (KNN) algorithm, a simple yet powerful machine learning technique. The dataset includes skin lesion images or structured data containing features like color, texture, and shape. Preprocessing techniques such as image resizing, normalization, and feature extraction are applied to enhance model accuracy. The KNN model is trained and tested to classify lesions as benign or malignant, with evaluation metrics including accuracy, precision, recall, and F1-score. This system provides a cost-effective, non-invasive, and accessible solution for preliminary skin cancer diagnosis, assisting dermatologists in early detection. Future enhancements may involve optimizing the K-value, improving dataset diversity for better generalization, and integrating real-time processing for immediate predictions. Additionally, developing a user-friendly web or mobile application can increase accessibility for both patients and healthcare professionals. The project emphasizes the potential of machine learning in medical diagnostics, reducing dependency on expensive and time-consuming traditional testing methods. By improving early detection rates, this system can significantly contribute to better patient outcomes and lower mortality rates associated with skin cancer.

Keywords: *Image processing, KNN algorithm, Dermatoscopic image dataset, Report generation, Real-world visualization.*

1. Introduction
Skin cancer is one of the most common forms of cancer worldwide, affecting millions of people every year. Early detection and accurate diagnosis are crucial to improving treatment outcomes and increasing survival rates. Traditional methods of skin cancer diagnosis rely heavily on visual examination by dermatologists, followed by biopsy and laboratory tests, which can be time-consuming and sometimes subjective. With the rapid advancement of technology, machine learning techniques have emerged as powerful tools to assist in medical diagnostics. Among these, the K-Nearest Neighbors (KNN) algorithm offers a simple yet effective approach for classification problems such as skin cancer prediction. KNN operates by finding the closest data points, or "neighbors," in the feature space and classifying new cases based on the majority class of these neighbors. In this project, we aim to develop a skin cancer prediction model using the KNN algorithm by leveraging a dataset containing relevant patient information and skin lesion characteristics. By preprocessing the data, normalizing features, and optimizing model parameters, the KNN model can provide reliable predictions that support early diagnosis. This not only helps healthcare professionals make informed decisions but also contributes to raising awareness about the potential of AI-driven solutions in the fight against skin cancer. [1]

1.1.Methods of Skin cancer prediction
- **Data Collection:** Gather a suitable dataset containing patient details and skin lesion characteristics. Dataset can include tabular data (age, gender, lesion location, etc.) or images of skin lesions.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0212
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1304 - 1306

- **Data Preprocessing:** Handle missing values and clean the data. Split the dataset into training and testing sets (70% training and 30% testing).
- **Feature Extraction**: For image data, extract features using image processing techniques or deep learning models.
- **Model Building**: Choose appropriate value for k. Use knn for to train the model on training data. Model **Evaluation:** Test the trained models on testing dataset. Evaluate model using metrics like accuracy, precision and confusion matrix.
- **Hyperparameter Tuning:** Experiment with different value of k and distance metrics
- **Result Interpretation** Analyze the output of the model.

2. Methodology Used

The Methods that we followed for the implementation of skin cancer prediction using knn are mentioned here. It includes tables, figures and flowchart.

2.1.Tables

The table which is presented below holds the value of cancer cell incubation values with IC 50. These are the main parameters for cancer cell prediction. Also we can know its severeness and diagnostic level. (Table 1) [2]

2.2.Figures

In this Figures column the flowchart and input images are provided. These figures is provide for easier understanding and our working method. (Figure 1)

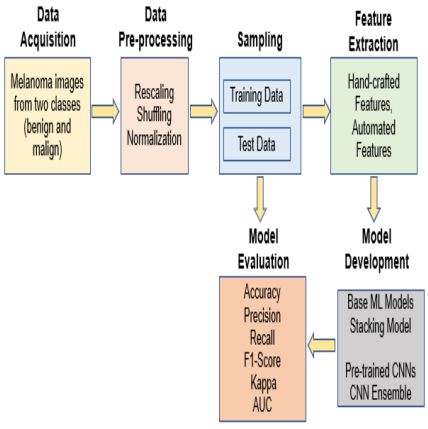

Figure 1 Flowchart of Our Model

Table 1 Input Parameters for Skin Cancer Cell Incubation

NE	48	MCF-7	380 ± 19
		HTC 116	~1200
SLN	24	MCF-7	525 ± 21
		HTC 116	655 ± 33
	48	MCF-7	157 ± 7
		HTC 116	590 ± 37
Doxorubicin	24	MCF-7	1.30 ± 0.06
		HTC 116	0.60 ± 0.03
	48	MCF-7	0.17 ± 0.01
		HTC 116	0.14 ± 0.01
NE + doxorubicin	24	MCF-7	0.52 ± 0.03
		HTC 116	0.39 ± 0.02
	48	MCF-7	0.16 ± 0.01
		HTC 116	0.10 ± 0.01
NE + thymoquinone	24	MCF-7	7.00 ± 0.50
		HTC 116	7.00 ± 0.50
	48	MCF-7	6.00 ± 0.50
		HTC 116	5.00 ± 0.40
SLN + doxorubicin	24	MCF-7	0.35 ± 0.02
		HTC 116	0.34 ± 0.02
	48	MCF-7	0.14 ± 0.01
		HTC 116	0.08 ± 0.01
SLN + thymoquinone	24	MCF-7	5.00 ± 0.40
		HTC 116	5.00 ± 0.40

3. Results and Discussion

3.1.Results

Here the result of this project is displayed.

3.2.Discussion

This skin cancer prediction project using the K-Nearest Neighbors (KNN) algorithm demonstrate the potential of machine learning in supporting early and accurate diagnosis. The model achieved promising accuracy, indicating that KNN is effective in classifying skin lesions based on extracted features. The confusion matrix and evaluation metrics such as

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0212
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1304 - 1306

precision, recall, and F1-score further confirm the model's ability to distinguish between cancerous and non-cancerous cases. By tuning the value of 'k' and applying proper data normalization, the performance improved, highlighting the importance of preprocessing in KNN-based models. Although KNN is simple and interpretable, it can be sensitive to noisy data and the choice of features. Overall, the project successfully shows that integrating KNN with healthcare data can aid dermatologists in early detection and potentially improve patient outcomes. (Figure 2) [3]

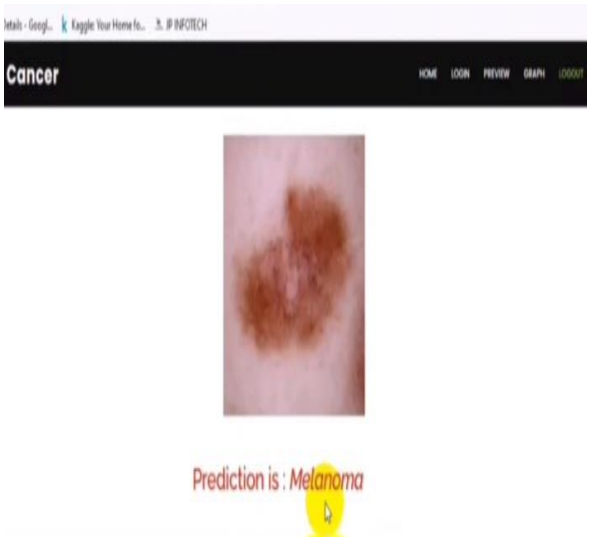

Figure 2 Result of this Project

Conclusion

In conclusion, the skin cancer prediction system using the K-Nearest Neighbors (KNN) algorithm effectively demonstrates the capability of machine learning in assisting early diagnosis. By leveraging patient data and lesion characteristics, the model provides reliable predictions that can support clinical decision-making. While KNN offers simplicity and interpretability, further improvements with advanced feature extraction and larger datasets can enhance accuracy. This project highlights the potential of AI-driven tools in healthcare, contributing to faster detection and better outcomes for skin cancer patients.

Acknowledgements

I would like to thank my project mentor Mrs.K.Priyadharshini mam for guiding us to complete this project. Also I would like to thank my teammates for completing the project and also helps me in publishing this work.

References

This is our reference(base) paper for this project.

Title: A convolutional neural network framework for accurate skin cancer detection.

Abstract: This paper uses the technique of Convolutional Neural Network (CNN), which make use of deep learning techniques for skin cancer detection. Through evaluation of research on skin cancer detection using CNN, the effectiveness and performance of these techniques in early and efficient diagnosis of skin cancer were established. Year – 2021. For image dataset - ISIC archive (International Skin Imaging Collaboration).

[1]. Y Jusman, "Performance of Multi-Layer Perceptron and Deep Neural Networks in Skin Cancer Classification", IEEE Transactions, vol. 10, pp. 118198-118212, 2022.

[2]. R Raja Subramanian, "Skin Cancer Classification using CNN", IEEE Conference 2021 11th International Conference on Cloud Computing Data Science & Engineering 10.1109/Confluence51648.2055, 2021.

[3]. "Research on Skin Cancer Cell Detection using Image Processing", 2020 IEEE International Conference on Computational Intelligence and Computing Research, pp. 1-8, 2020.