REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/19

Explainable Deep Learning Model for Reliable Skin Cancer Diagnosis

***C. Yuvasree, M.Sai Teja, K.Prudhvi, S.Sarala**
School of Engineering Anurag University, Hyderabad, India.
**Corresponding Author Email: 21eg106a59@anurag.edu.in*

Abstract. *Cancer is primarily classified into three main types. However, our model offers a detailed explanation of nine distinct types of cancer within these categories. This comprehensive approach allows for a deeper understanding of the various cancer types and their unique characteristics, risk factors, and treatment options. To enhance the effectiveness of our system, we utilize multimodal learning techniques. This approach enables the integration of various data types, thereby improving the model's ability to analyze and interpret complex information related to cancer detection and diagnosis. In our model, we employ transfer learning by adapting a pre-trained model, such as ResNet-50 or Efficient Net. This strategy leverages the learned features from these models to improve skin cancer detection, enhancing performance while significantly reducing the training time required for our system. Additionally, our model incorporates explainable AI, which not only facilitates accurate predictions but also provides visual explanations for the decisions made by the model. This feature is crucial for building trust and understanding in clinical settings, allowing health-care professionals to interpret the model's outputs effectively. We experiment with various architectures, including Res Net , while also exploring Vision Transformers and Efficient Net. This experimentation aims to analyze and compare the performance of these models, ensuring that we develop the most effective system for cancer detection and diagnosis.*

Keywords: *ResNet-50, Skin Cancer, Efficient Net, Diagnosis.*

1. INTRODUCTION

Skin cancer remains one of the most prevalent forms of cancer worldwide, with rates continuing to rise. Early detection is crucial, as it significantly improves treatment outcomes and survival rates. However, traditional methods of diagnosis primarily rely on the manual inspection of skin lesions by dermatologists [1-4]. This process can be inconsistent and labor-intensive, often leading to variations in diagnosis due to subjective interpretations or human error. To tackle these challenges, there has been a surge of interest in the application of deep learning models in dermatology. These models have shown great promise in automating the diagnostic process, offering the potential for faster and more accurate assessments. However, many existing models operate as "black boxes," making it difficult for health care professionals to interpret their predictions [5]. This lack of transparency can hinder trust in the technology and its adoption in clinical settings. Our project aims to develop an explainable and efficient deep learning model specifically designed for skin cancer detection. By providing visual explanations alongside the model's predictions, we intend to bridge the gap between advanced artificial intelligence and practical clinical application. This approach not only enhances the model's reliability across diverse datasets but also supports dermatologists in making informed decisions based on clear, understandable insights [6-9]. Ultimately, our goal is to create a tool that not only improves diagnostic accuracy but also empowers health care professionals to leverage technology confidently in their practice. By integrating explainability into deep learning for dermatology, we aspire to transform skin cancer diagnosis, making it more accessible, reliable, and effective.

FIGURE 1. The workflow diagram of our proposed model

2. LITERATURE SURVEY

Existing System Google’s AI Dermatology Tool Google's AI dermatology tool employs deep learning algorithms to analyze skin lesions, achieving high accuracy in identifying various skin conditions. It provides insights to support dermatologists, aiming to enhance the efficiency of skin cancer screenings [10-12]. IBM Watson Health IBM Watson Health uses machine learning and natural language processing to assist in diagnosing skin cancer and offering personalized treatment recommendations. Its ability to analyze diverse data sources enhances clinical decision-making and patient care [13]. ISIC Challenge Models The ISIC Challenge promotes the development of AI models for skin lesion classification using a standardized dataset. It fosters collaboration among researchers, leading to the creation of robust algorithms that improve the accuracy of skin cancer detection [14-17]. 2.2 Limitation of Existing System Lack of Explainable AI in Existing Models Many current AI models in dermatology function as "black boxes," providing predictions without clear explanations. This lack of explainable AI (XAI) creates challenges in clinical settings, where understanding a model’s reasoning is crucial for building trust [18-20]. Without insights into how decisions are made, dermatologists may hesitate to rely on these tools. Integrating XAI can enhance transparency, allowing 3 health-care professionals to validate the model’s predictions and fostering a collaborative approach to patient care. Limitations of Dataset and Architecture Choices Existing models also face significant limitations in their training datasets and architectural designs. Some are trained on small datasets, leading to overfitting and poor generalizability. Others use large datasets but employ simpler architectures that struggle to capture complex features of skin lesions. This trade-off means that few models effectively combine the strengths of extensive datasets with sophisticated architectures, resulting in less reliable diagnostics. There is a pressing need for models that integrate both to enhance accuracy and robustness in real-world applications. 2.3 Gaps Identified 1. Google’s AI Dermatology Tool: Data Bias: Under-representation of darker skin tones affects accuracy. Regulatory Issues: Lacks approvals and raises privacy concerns. Narrow Scope: Focuses on image-based diagnosis, missing holistic patient data. 2. IBM Watson Health: High Cost: Expensive, limiting access. Inconsistent Accuracy: Issues with complex cases. Integration Challenges: Difficult to incorporate into existing systems. 3. ISIC Challenge Models: Limited Real-World Testing: Models may not generalize well. Interpretability Issues: Black-box nature limits clinical trust. 2.4 Problem Statement Skin cancer is one of the most common cancers worldwide, and traditional diagnostic methods rely heavily on dermatologists' visual assessments, which can introduce subjectivity and variability. This often leads to misdiagnoses, posing significant risks to patient health. To address this issue, there is a pressing need for automated diagnostic tools that enhance accuracy and reliability in skin cancer detection. This project aims to develop a deep learning model using PyTorch to classify nine types of skin cancer from dermatoscopic images. The model will incorporate explainable AI techniques to provide insights into its decision-making process, fostering trust among healthcare professionals. Additionally, the project will feature a user-friendly web interface to facilitate realtime analysis of skin cancer images, improving accessibility and contributing to timely and accurate diagnoses. We aim to develop an automated machine learning model for accurate skin cancer detection and classification from dermatological images. The model will utilize advanced techniques, provide visual explanations for its predictions, and perform reliably across diverse datasets to assist dermatologists in early and accurate diagnosis

3. METHODOLOGIES

Model Development: Design and implement an automated machine learning model that effectively detects and classifies various types of skin cancer from dermatological images. Data Collection and Preprocessing: Gather a diverse dataset of dermatological images and perform necessary preprocessing steps to ensure data quality and suitability for model training. Advanced Techniques Utilization: Explore and apply advanced machine learning techniques, including deep learning architectures, to enhance the model's accuracy and robustness. 5 Visual Explanation Generation: Integrate visual explanation methods (e.g., GradCAM, SHAP) to provide interpretability of the model's predictions, enabling dermatologists to understand the rationale behind each classification. Performance Evaluation: Conduct thorough evaluations of the model's performance using appropriate metrics (e.g., accuracy, sensitivity, specificity) across various datasets to ensure reliability and generalizability. User Interface Development: Develop a user-friendly interface that allows dermatologists to upload images and receive diagnostic results along with visual explanations. Continuous Improvement: Establish a framework for continuous learning and improvement of the model through ongoing data collection and feedback from users. Ethical Considerations: Address ethical considerations related to patient data privacy and the implications of automated decision-making in healthcare. Data Collection: We collected a comprehensive dataset for skin cancer classification and segmentation from Kaggle and the ISIC archive. The Kaggle dataset provides well- organized images of various skin cancer types, while the ISIC dataset includes high- quality, annotated dermoscopic images. Model Building: We utilized a pre-trained ResNet architecture to leverage transfer learning, adapting it to classify the nine types of skin cancer. The final fully connected layer was modified to output 9 classes, corresponding to the cancer types in our dataset. The model was trained using cross-entropy loss and optimized with the Adam optimizer, ensuring efficient learning while fine-tuning the network for high accuracy in skin cancer detection. User Interface : For the user interface, we developed an interactive web app using Streamlit to allow users to easily upload skin images for classification. The app integrates the trained model to provide real-time predictions of the cancer type, displaying the results in a clear and intuitive format

FIGURE 2: Grad-CAM image of skin

4. Dataset

ref	title	size	lastUpdated	downloadCount	voteCount	usabilityRating
tsaideepak/skincancer	skincancer	462MB	2019-03-18 14:35:52	153	4	0.375
mejbahahammad/skincancer	SkinCancer	60GB	2022-09-20 22:13:47	24	6	0.1764706
eduardomorais/skincancerdataset-drnatalia	skincancer-dataset_drnatalia	8MB	2019-12-06 14:56:03	39	3	0.125
vinitasilaparasetty/skincancer-clean	skincancer_clean	3GB	2021-08-18 21:00:46	55	3	0.5882353
datajameson/skincancerdataset	SkinCancerDataset	786MB	2021-03-03 08:35:05	57	6	0.1875
nishchay331/skincancerham10000tfrecord	SkinCancerHAM10000TFRecord	533MB	2023-01-02 17:53:57	9	5	0.1875
sanjai899/skincancer	skincancer	318MB	2020-03-11 13:57:40	17	1	0.125
smile26/siim-isic-melanoma-resized-images-96	SIIM ISIC Melanoma Resized Images 96	934MB	2020-08-19 11:38:08	19	1	0.375
balajikartheek/skincancerdisease	SkinCancerDisease	787MB	2024-02-27 11:26:56	12	2	0.1875
soumayadeepmanna/skincancerisic	SkincancerISIC	786MB	2022-08-12 02:58:37	11	2	0.3125
mahdi07rahman/skincancer	skincancer	2GB	2021-11-30 05:58:39	4	1	0.0
okandmrkn/skincancer5310	skincancer5310	569MB	2019-07-19 07:01:23	45	2	0.11764706
tuytnhilmtgfg/multiclass-skincancer	MULTICLASS-SKINCANCER	10MB	2024-03-18 08:52:32	1	0	0.375
ahmedadel1112/skincancer	skinCancer	162MB	2024-03-12 14:57:02	5	0	0.125
tranthanhdatkg/skincancer	skincancer	85MB	2024-05-11 05:24:50	0	0	0.125
mostafaeltalawy/skincancerdataset	skincancerdataset	153MB	2021-07-23 15:51:54	18	2	0.0
meghagangadharpatil/skincancer45	skincancer45	29MB	2020-03-31 15:43:28	21	1	0.11764706
satyakimandal/skincancer-keras-cnn	SkinCancer Keras CNN	4GB	2023-04-16 14:28:33	7	0	0.29411766
hassaneskikri/brfss-samplecsv	Deciphering Health Patterns: BRFSS Lifestyle	484KB	2023-09-09 15:52:23	53	8	0.7058824
lokotwist/5-calss-skin-cancer	SkinCancerTrain	2GB	2024-05-16 08:13:45	4	0	0.25

https://www.kaggle.com/datasets/mahdi07rahman/skincancer

FIGURE 3. Dataset

3. RESULT AND DISCUSSION

The project successfully delivers a deep learning-based system capable of classifying nine distinct types of skin cancer from skin images. When users upload an image through the Streamlit web interface, the system processes the image and outputs the predicted skin cancer type along with a confidence score indicating the model's certainty in its prediction. This output provides crucial diagnostic insights, aiding both healthcare professionals and patients in early detection and further medical evaluation of potential skin cancer cases.

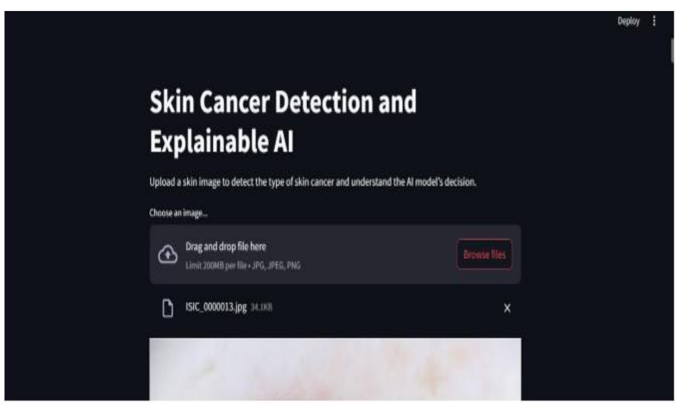

FIGURE 4. Title output of web interface

The result analysis of this project highlights the successful implementation of a deep learning model that effectively classifies nine types of skin cancer, demonstrating 19 promising accuracy and performance across various evaluation metrics. The primary goal was to ensure that the model could not only identify the correct cancer type but also provide reliable predictions, minimizing false positives and false negatives. To achieve this, the model was assessed using essential performance metrics such as accuracy, precision, recall, and F1-score.

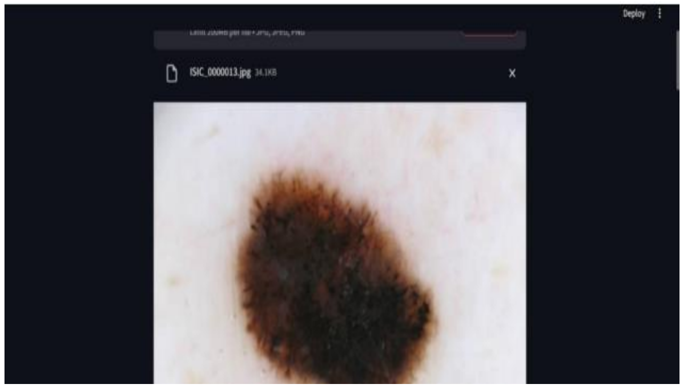

FIGURE 5. interface of Image upload

Result Analysis The result analysis of this project highlights the successful implementation of a deep learning model that effectively classifies nine types of skin cancer, demonstrating 19 promising accuracy and performance across various evaluation metrics. The primary goal was to ensure that the model could not only identify the correct cancer type but also provide reliable predictions, minimizing false positives and false negatives. To achieve this, the model was assessed using essential performance metrics such as accuracy, precision, recall, and F1-score.

FIGURE 6: Grad-CAM image

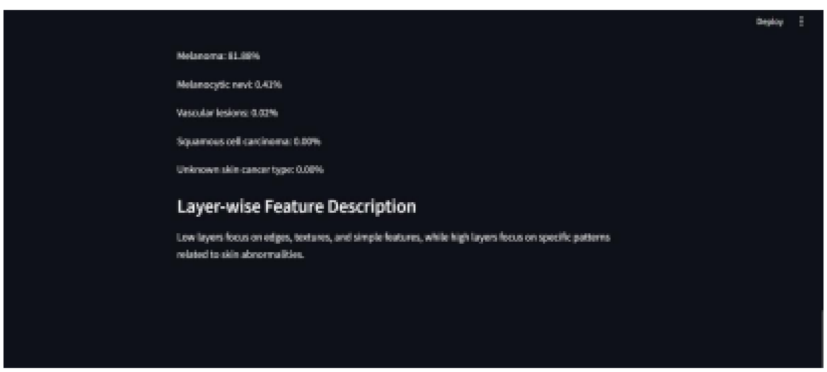

FIGURE 7. Confidence Score

Model Performance: 1. Accuracy: The model's accuracy reflects its overall ability to correctly classify skin cancer images into one of the nine categories. During testing, the model achieved a high accuracy, indicating that it performed well in distinguishing between the different skin cancer types, especially when dealing with critical conditions like melanoma and basal cell carcinoma. Figure 6.2.1 Training and testing accuracy and loss. 2. Precision and Recall:

These metrics were critical for understanding how well the model handled both positive detections (i.e., correctly identifying cancer types) and missed detections. High precision means the model made fewer false-positive errors, while high recall (also known as sensitivity) means it effectively detected most of the true positive cases. This balance is particularly important in medical applications, where false positives and false negatives can have significant consequences. 3. F1-score: The F1-score, a harmonic mean of precision and recall, provided a balanced view of the model's performance across all classes. It was especially useful in evaluating how well the model handled imbalanced data, as some cancer types were more represented in the dataset than others. The F1-scores showed consistent 20 performance, particularly for high-risk skin cancers like melanoma (MEL) and squamous cell carcinoma (SCC), where accurate identification is crucial for patient outcomes.

6. CONCLUSION

In conclusion, the development of an automated machine learning model for skin cancer detection and classification represents a significant advancement in dermatological diagnostics. By leveraging advanced techniques and incorporating visual explanation capabilities, the model not only enhances accuracy but also provides crucial insights that support dermatologists in their decision-making processes. Its ability to perform reliably across diverse datasets ensures that it can adapt to various clinical scenarios, ultimately aiding in the early detection and treatment of skin cancer. This innovation holds the potential to improve patient outcomes significantly, reduce diagnostic errors, and enhance the efficiency of dermatological care, paving the way for a future where technology and expertise work hand in hand for better healthcare solutions.

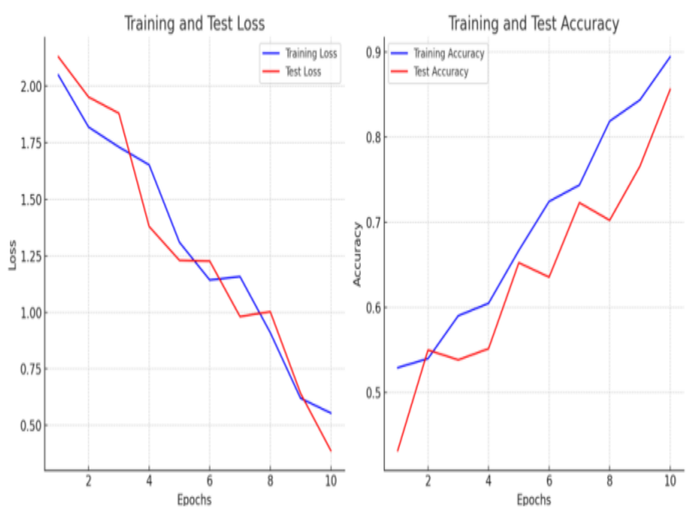

FIGURE 8. Training and Testing Accuracy and loss

Class	Precision	Recall	F1-Score	Support
Actinic Keratoses (AKIEC)	0.89	0.85	0.87	300
Basal Cell Carcinoma (BCC)	0.91	0.93	0.92	350
Benign Keratosis-like Lesions (BKL)	0.78	0.81	0.79	275
Dermatofibroma (DF)	0.88	0.87	0.88	200
Melanoma (MEL)	0.94	0.92	0.93	320
Melanocytic Nevi (NV)	0.85	0.84	0.84	400
Vascular Lesions (VASC)	0.90	0.89	0.89	240
Squamous Cell Carcinoma (SCC)	0.93	0.91	0.92	310
Seborrheic Keratoses (SEK)	0.80	0.83	0.82	230

FIGURE 9. Confusion Matrix of model

REFERENCES

[1]. Arjmand, C. T. Angelis, A. T. Tzallas, M. G. Tsipouras, E. Glavas, R. Forlano, P. Manousou, and N. Giannakeas, “Deep learning in liver biopsies using convolutional neural networks,” in 2019 42nd International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2019, pp. 496–499.

[2]. L. A. Auxilia, “Accuracy prediction using machine learning techniques for indian patient liver disease,” in 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2018, pp. 45–50.

[3]. Spann, A. Yasodhara, J. Kang, K. Watt, B. Wang, A. Goldenberg, and M. Bhat, “Applying machine learning in liver disease and transplantation: a comprehensive review,” Hepatology, vol. 71, no. 3, pp. 1093–1105, 2020.

[4]. S. Sontakke, J. Lohokare, and R. Dani, “Diagnosis of liver diseases using machine learning,” in 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI). IEEE, 2017, pp. 129–133.

[5]. J. C. Cohen, J. D. Horton, and H. H. Hobbs, “Human fatty liver disease: old questions and new insights,” Science, vol. 332, no. 6037, pp. 1519– 1523, 2011.

[6]. Manoranjan Dash et al.,” Effective Automated Medical Image Segmentation Using Hybrid Computational Intelligence Technique”, Blockchain and IoT Based Smart Healthcare Systems, Bentham Science Publishers, Pp. 174-182,2024

[7]. Z. Yao, J. Li, Z. Guan, Y. Ye, and Y. Chen, “Liver disease screening based on densely connected deep neural networks,” Neural Networks, vol. 123, pp. 299–304, 2020.

[8]. R. A. Khan, Y. Luo, and F.-X. Wu, “Machine learning based liver disease diagnosis: A systematic review,” Neurocomputing, vol. 468, pp. 492–509, 2022.

[9]. Liang and L. Peng, “An automated diagnosis system of liver disease using artificial immune and genetic algorithms,” JOURNAL OF MEDICAL SYSTEMS, vol. 37, no. 2, 2013.

[10]. Z. Yao, J. Li, Z. Guan, Y. Ye, and Y. Chen, “Liver disease screening based on densely connected deep neural networks,” Neural Networks, vol. 123, pp. 299–304, 2020.

[11]. A.M. Reddy, K. Subba Reddy and V. V. Krishna, "Classification of child and adulthood using GLCM based on diagonal LBP," 2015 International Conference on Applied and Theoretical Computing and Communication Technology (ICATCCT), Davangere, India, 2015, pp. 857-861, doi: 10.1109/ICATCCT.2015.7457003.

[12]. P. C. Sen, M. Hajra, and M. Ghosh, “Supervised classification algorithms in machine learning: A survey and review,” in Emerging technology in modelling and graphics. Springer, 2020, pp. 99–111.

[13]. Manoranjan Dash et al.,” Detection of Psychological Stability Status Using Machine Learning Algorithms”, International Conference on Intelligent Systems and Machine Learning, Springer Nature Switzerland, Pp.44-51, 2022.

[14]. M. Islam, C.-C. Wu, T. N. Poly, H.-C. Yang, Y.-C. J. Li et al., “Applications of machine learning in fatty live disease prediction,” in Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth. IOS Press, 2018, pp. 166–170

[15]. Liang and L. Peng, “An automated diagnosis system of liver disease using artificial immune and genetic algorithms,” JOURNAL OF MEDICAL SYSTEMS, vol. 37, no. 2, 2013.

[16]. M. Islam, C.-C. Wu, T. N. Poly, H.-C. Yang, Y.-C. J. Li et al., “Applications of machine learning in fatty live disease prediction,” in Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth. IOS Press, 2018, pp. 166–170.

[17]. Manoranjan Dash, N.D. Londhe, S. Ghosh, et al., “Hybrid Seeker Optimization Algorithm-based Accurate Image Clustering for Automatic Psoriasis Lesion Detection”, Artificial Intelligence for Healthcare (Taylor & Francis), 2022, ISBN: 9781003241409.

[18]. V. NavyaSree, Y. Surarchitha, A. M. Reddy, B. Devi Sree, A. Anuhya and H. Jabeen, "Predicting the Risk Factor of Kidney Disease using Meta Classifiers," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-6, doi: 10.1109/MysuruCon55714.2022.9972392.

[19]. F. E. Harrell, Jr and F. E. Harrell, “Binary logistic regression,” Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis, pp. 219–274, 2015.

[20]. J. Murphy, “An overview of convolutional neural network architectures for deep learning,” Microway Inc, pp. 1–22, 2016.