© June 2024 | IJIRT | Volume 11 Issue 1 | ISSN: 2349-6002

Skin Cancer Detection System Using Deep Learning

KANAMARLAPUDI VENKATA NAGA LAKSHMI SRUJANA[1], VEMULA RAMALAKSHMI[2]
[1, 2] Potti Sri Ramulu Chalavadi Mallikarjuna Rao College of Engineering and Technology

Abstract— My work presents a cutting-edge Convolutional Neural Network (CNN) based Skin Cancer Identification system that is improved by Batch Normalization methods. With CNNs being well known for image analysis, the model's adoption guarantees a strong basis, which is intended to solve the worldwide health concern of skin cancer[12].Batch Normalization mitigates issues associated with internal covariate shifts in the CNN architecture by normalizing inputs and enhancing stability. An important breakthrough in the identification of skin cancer has been made by the model, which consistently adapts to a variety of datasets.The imbalances in datasets connected to skin cancer are also subjected to modern imbalance learning algorithms. The model greatly increases sensitivity and accuracy, especially for minority classes, by taking into consideration class imbalances through the use of SMOTE, under sampling, and oversampling. The CNN- based model performs better in recognizing lesions associated with skin illnesses than earlier benchmarks, according to experiments. Intentionally using imbalanced learning techniques improves generalization, whereas group normalization [9] guarantees stability. Because of its versatility, the model may be used in real-world scenarios and effectively solves the complicated issues faced by healthcare practitioners. The study [15] raises the bar for technological integration in the ongoing fight against skin disorders by highlighting the wider impact of state-of-the-art deep learning algorithms in medical picture interpretation, beyond local applications . with in the medical field.

Index Terms- Skin cancer, technology, healthcare, deep learning, CNN (convolutional neural network), and imbalance learning; batch normalization.

I. INTRODUCTION

Skin cancer is a common and occasionally lethal type of cancer that poses a serious danger to global health. Better patient outcomes and a successful course of therapy depend on an accurate and timely diagnosis.

Convolutional Neural Networks are one specific kind of deep learning technology (CNNs). [17], have transformed medical image processing in the last few years, opening up fascinating new possibilities for better and more accurate diagnosis. The study intends to progress dermatology by offering a thorough examination of CNN-based skin cancer classification. The research utilized a dataset [3] called "hmnist_28_28_RGB.csv," which includes high-resolution RGB images of skin lesions [11] labeled with seven distinct types of skin cancer: melanoma, actinic keratoses, melanocytic nevi, benign keratosis-like lesions, dermatofibroma, melanoma, and basal cell cancer [13].

We use sophisticated approaches like oversampling to address the problem of unequal class distribution in medical datasets and guarantee that every class of skin cancer is equally represented in the training set [21].

This strategy is essential for keeping the model from behaving in a way that benefits the majority class and for improving its capacity to recognize less common but no less important circumstances.

The foundation of our study is the development and assessment of a CNN model using Keras and TensorFlow. Convolutional and pooling layers, batch normalization, dropout for regularization, and a final SoftMax layer for multiclass classification are all part of the thoughtfully designed model architecture. The model is optimized using the Adam optimizer and sparse categorical cross-entropy loss over a number of training epochs. [5]

By evaluating the model's performance on a different test set, we determine how useful it is and offer insights via measures like accuracy and a thorough confusion matrix. The outcomes show that the model is capable of correctly classifying skin cancer lesions and identifying areas for improvement. Long-term patient care could be improved by deep learning in dermatological diagnostics through the promotion of early therapeutics and improved healthcare outcomes in skin cancer detection.

II. LITERATURE SURVEY

This study compares previous investigations on CNN-Based Deep Learning for Skin Cancer Identification and Categorization.

Numerous deep learning and machine learning techniques have been studied recently in relation to the identification and categorization of skin cancer.

Dr. Maya V. Karki and Vidya M. wrote a paper on the use of neural networks, KNN, SVM, and Naive Bayes in the diagnosis of skin cancer. Although it demonstrates greater accuracy than manual methods [15], it has a number of shortcomings, including a limited dataset and feature-based evaluation. In a different study, Wassim El Hajj Chehade, Ali El-Zaart, and Riham Abdel Kader employed Otsu's Algorithm and iterative threshold computing to segment skin pictures[1].

Their investigation demonstrates that skin cancer images with a lognormal distribution do not fit Otsu's Gaussian assumption.

On the other hand, Sugiyanto, I. Slamet, and A. A. Nugroho use CNN in conjunction with stochastic gradient descent to diagnose skin cancer; they achieve an 80% accuracy rate on the Skin Cancer Dataset HAM10000 training set [17]. The following shortcomings have been identified: inconsistent data, large-scale dataset management, and variation in skin lesions [5]. Turning to deep learning architectures, François Chollet's Xception (using depthwise separable convolutions) outperforms rival systems like Inception in terms of processing efficiency while achieving competitive performance [6] in picture categorization tasks.

Depthwise separable convolutions are used to illustrate how important well-designed models are.

The detection of skin cancer is studied by Tausif Diwan, Jitendra V. Tembhurne, Hemprasad Y. Patil, and Nachiketa Hebbar [14] using a combination of machine learning and deep learning techniques, such as ResNet, Logistic Regression, Contourlet Transform, VGG19, InceptionV3, and InceptionV3. They acknowledge that their dataset is small and that generalization and explainability are challenging, but they nevertheless employ CNNs with transfer learning to accomplish automatic feature extraction. Furthermore, dermatologists currently offer deep neural network classification of skin cancer at the dermatologist level, as demonstrated by [7]; this highlights the efficacy of convolutional neural networks (CNNs) in image classification tasks and aims to provide an objective, scalable, and effective approach to skin cancer classification.

Studies on the transferability of features in deep neural networks, the use of deep convolutional networks for better skin lesion classification, the application of Inception and ResNet architectures for deep learning for skin lesion diagnosis,[3] and the multimodal fusion of clinical and imaging data for skin cancer risk stratification are some additional contributions. Together, these researchers make progress in the field by tackling difficulties related to dataset size, imbalanced datasets, interpretability, and the need for larger and more representative datasets for comprehensive skin cancer detection and diagnosis.

III. METHODOLOGY

The RGB images of seven different types of skin cancer based on the distribution of skin lesions were taken from "hmnist_28_28_RGB.csv," which provided the dataset for this comprehensive approach to skin cancer identification. To ensure randomization in the dataset, rows are shuffled when imported into a Pandas Data Frame [2], readying it for exploratory analysis. To strengthen the model's robustness, more images are added, especially ones that mimic blur and dim lighting.

The dataset is then split 80/20 between the training and testing sets, and meaningful class names are assigned using a dictionary. The imbalanced-learn library's Random Over Sampler function is used to oversample minority classes in order to correct the class imbalance. This guarantees a training set that is more balanced, which is crucial for related to medicine [21].

Label	Full name	Image count
akiec	Actinic Keratoses	327
bcc	Basal cell carcinoma	514
bkl	Benign keratosis	1099
df	Dermatofibroma	115
mel	Melanoma	1113
nv	Melanocytic nevi	6705
vasc	Vascular skin lesions	142

Fig:1 Seven different typesof Skin Lesions with their count [17]

After the data has been prepared, the Convolutional Neural Network (CNN) model is architecturally developed using

TensorFlow and Keras Sequential API. Dropout layers are used for regularization, batch normalization is used for regularization, and maxpooling layers are used to mitigate overfitting. Dense layers with softmax activation and convolutional layers triggered by rectified linear units (ReLUs) both meet the multi-class classification objective. The next steps involve compiling and optimizing the model. With a learning rate of 0.001 and a loss function of sparse categorical cross-entropy, the Adam optimizer is used [16]. Based on validation accuracy, the Model Checkpoint callback saves the model that performs the best, and it tracks the training process. The CNN model is trained on the balanced dataset with a 20% validation split, using a preset batch size and number of epochs.

The provided neural network architecture is designed exclusively for image classification and is meant to find patterns in grayscale photos with 28 by 28 pixels. It is made up of many Conv2D layers that extract hierarchical features from the input data using different filter sizes.

[20] MaxPooling2D layers systematically reduce spatial dimensions after each convolutional layer in order to retrieve important information. After a few convolutional and dense layers, batch normalization layers are methodically added for stable activations throughout training. In order to facilitate the shift to densely connected layers, A flattened layer is added in order to convert the 3D tensor output from the convolutional layers into a 1D tensor. Dropout layers, which have a 50% dropout rate, are placed after the flattened layer and the two Dense layers in order to avoid overfitting. They achieve this by turning off some of the units at random while they are being trained.

The model concludes with a sequence of Dense layers that result in a seven-unit output layer that reflects a multi-class categorization task. There are a total of 504,103 parameters linked to the Batch Normalization [9] layers, of which 502,983 are trainable and 1,120 are not. The architecture is described in depth, however certain crucial components— such as the activation functions, loss function selection, optimizer selection, and training dataset characteristics—are left out. These unidentified components have a major impact on the model's performance and its applicability to real- world occurrences [10].

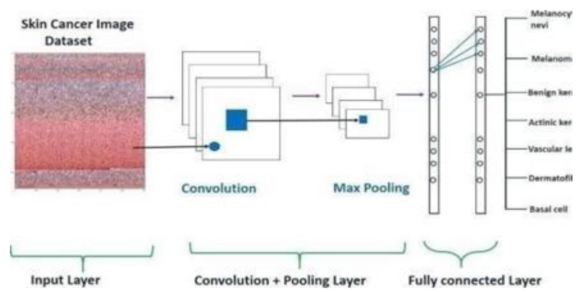


Fig2: Representing the CNN model complete architecture.[19]

The convolutional neural network (CNN) model continues to improve after 50 epochs; accuracy reaches an incredible 99.74%, and the training loss drops to 0.0079. Even with a marginally lower validation accuracy of 96.35%, the model exhibits strong generalization. With about eleven minutes and forty-nine seconds of training, the CNN shows good computational performance [18]. The model is positioned as a powerful tool for identifying skin cancer and may be successfully utilized in real-world circumstances, especially for skin cancer diagnosis, because of its high accuracy and low loss [1].

After training is complete, a new test set is used to evaluate the model's generalization capacity. A confusion matrix is produced as part of performance analysis to show how well the model categorizes different types of skin cancer. The trained model [7], which is kept in a file named "skin_cancer.h5", is the result of this work. The predictive capabilities of the model are demonstrated in a clinical context, underscoring its potential utility. Additional testing using external images, located in the "Test Images" subfolder, demonstrates how well the model predicts various cases of skin cancer.

IV. RESULT

The model's suggested outcome is the most accurate and efficient prediction of skin cancer. Of the seven types of skin cancer, the strongly matched type will have the highest percentage. It should be 50% or higher for at least one to confirm the type of cancer.

Among the numerous visualizations created to demonstrate the performance of the model are a heatmap of the confusion matrix and accuracy plots for training and validation.

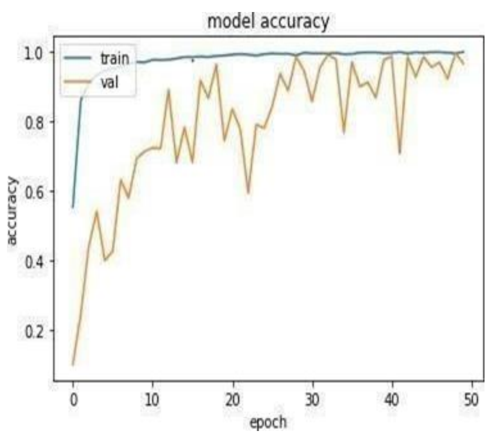


Fig3: Model Accuracy of data [19]

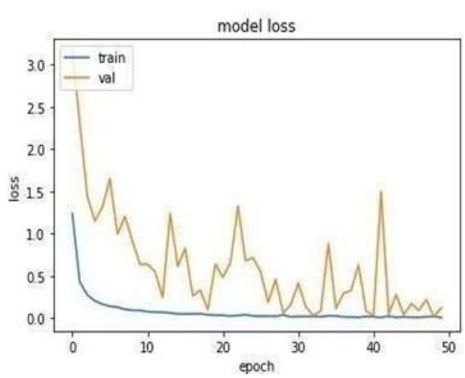


Fig4: Model Loss ofdata [19]

Fig5: Representing in the form of a confusion matrix.[19]

In order to provide a qualitative assessment of various types of skin cancer, as opposed to merely focusing on categories [4], and to assess the model's ability to learn and generalize, random images from the training and testing sets were used.

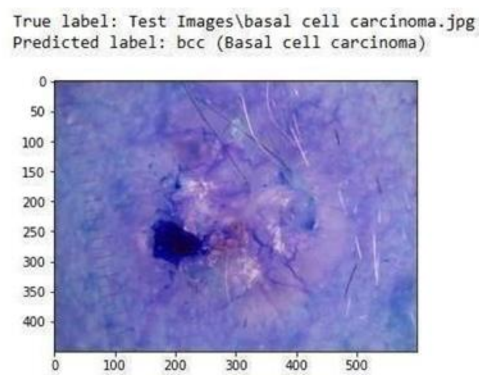


Fig6: Basal cell carcinoma [19]

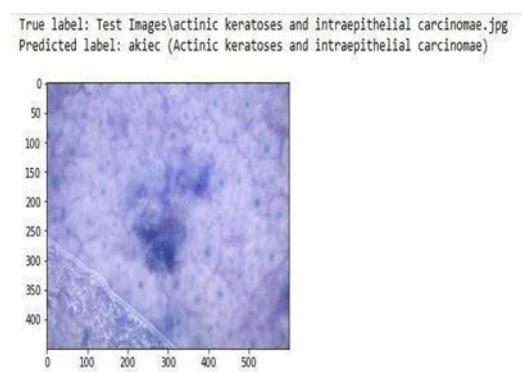


Fig7: Actinic keratoses and intraepithelialcarcinomas [19]

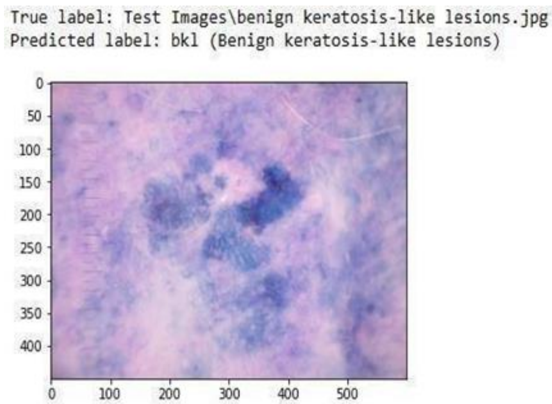


Fig8: Benign keratosis-like lesions [19]

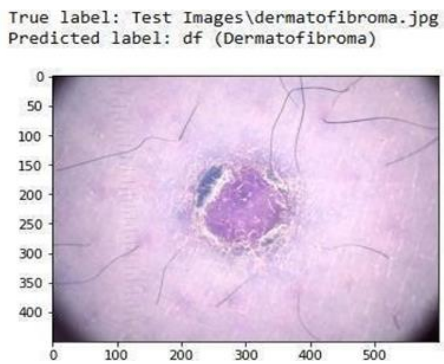


Fig 9: Dermatofibroma [19]

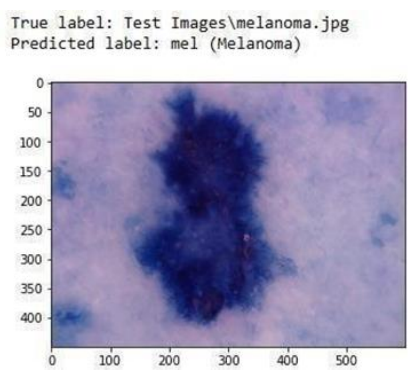


Fig10: Melanoma [19]

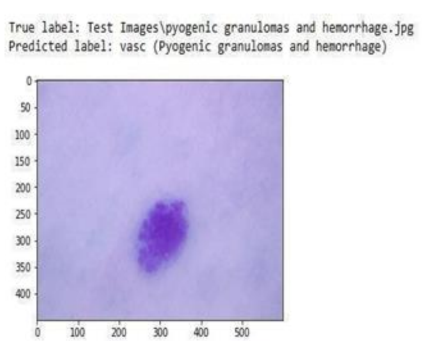
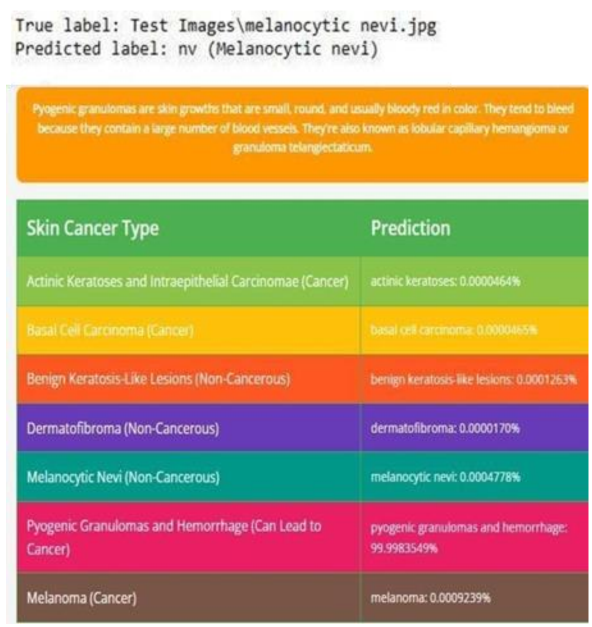


Fig11: Pyogenic granulomas and hemorrhage [19]

With the use of the CNN architecture and specialized techniques like imbalance learning and batch normalization [9], we were able to create a skin cancer prediction model with 98% accuracy for seven different types of lesions. This is an a percentage representation of the results, with the highest percentage going to the skin cancer kind that the algorithm identified.

Fig:12 Upload the image for prediction [19]



CONCLUSION

Convolutional Neural Networks [8] are used in this work to address classification problems related to skin cancer. By utilizing a varied RGB dataset and strategic approaches, such as oversampling, the model's generalization to different skin cancer types is improved. The results provide precise lesion classification, bolstered by succinct visual aids. The deployed model, which is saved in'skin_cancer.h5' for easy integration into healthcare systems, shows promise in practical situations. Subsequent endeavors entail enhancing the model's performance and investigating a variety of datasets to enhance its generalizability, so augmenting the revolutionary utilization of deep learning in the healthcare domain [12].

REFERENCES

[1] Alquran, Hiam, et al. "The melanoma skin cancer detection and classification using support vector machine." 2017 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT). IEEE, 2017.

[2] M. Hossain, K. Sadik, M. M. Rahman, F. Ahmed, M. N. Hossain Bhuiyan and M. M. Khan, "Convolutional Neural Network Based Skin Cancer Detection (Malignant vs Benign)," 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2021, pp. 0141-0147.

[3] P. Dubal, S. Bhatt, C. Joglekar and S. Patil, "Skin cancer detection and classification," 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), Langkawi, Malaysia, 2017, pp. 1- 6.

[4] O. Abuzaghleh, B. D. Barkana and M. Faezipour, "Noninvasive Real- Time Automated Skin Lesion Analysis System for Melanoma Early Detection and Prevention," in IEEE Journal of Translational Engineering in Health and Medicine, vol. 3, pp. 1-12, 2015.

[5] Dhatri Raval, Jaimin N. Undavia, A Comprehensive assessment of Convolutional Neural Networks for skin and oral cancer detection using medical images, Healthcare Analytics, Volume 3,2023,100199, ISSN 2772-4425.

[6] Nidhi Bansal, S. Sridhar, HEXA-GAN: Skin lesion image inpainting via hexagonal sampling based generative adversarial network, Biomedical Signal Processing and Control, Volume 89,2024

[7] R. Ashraf et al., "Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection," in IEEE Access, vol. 8, pp. 147858-147871, 2020.

[8] L. Riaz et al., "A Comprehensive Joint Learning System to Detect Skin Cancer," in IEEE Access, vol. 11, pp. 79434- 79444, 2023.

[9] Ragab, Mahmoud, et al. "Early and accurate detection of melanoma skin cancer using hybrid level set approach." Frontiers in Physiology 13 (2022):965630.

[10] Monika, M. Krishna, et al. "Skin cancer detection and classification using machine learning." Materials Today: Proceedings 33 (2020): 4266- 4270.

[11] V. Nwogu and I. Singh, "Analyzing skin lesions in dermoscopy images using convolutional neural networks," in Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 4035–4040, Miyazaki, Japan,2018.

[12] Wang H, Chen X, Moss RH, Stanley RJ, Stoecker WV, Celebi ME, Szalapski TM, Malters JM, Grichnik JM, Marghoob AA, Rabinovitz HS, Menzies SW. Watershed segmentation of dermoscopy images using a watershed technique. Skin Res Technol. 2010 Aug;16(3):378-84.

[13] Salih, Shakar, Al-Raheym, Shereen. (2018). Comparison of skin lesion image between segmentation algorithms. Journal of Theoretical and Applied Information Technology. 96. 6085.

[14] Skin cancer detection using ensemble of machine learning and deep learning techniques Jitendra V. Tembhurne, Nachiketa Hebbar, Hemprasad Y. Patil, Tausif Diwan.

[15] Diab, Amal G., Nehal Fayez, and Mervat Mohamed El-Seddek. "Accurate skin cancer diagnosis based on convolutional neural networks." Indonesian Journal of Electrical Engineering and Computer Science 25.3 (2022): 1429-1441

[16] N. Nigar, M. Umar, M. K. Shahzad, S. Islam, and D. Abalo, “A deep learning approach based on explainable artificial intelligence for skin lesion classification,” IEEE Access, vol. 10, pp. 113715–113725, 2022

[17] Adla, Devakishan, et al. “Deep learning-based computer aided diagnosis model for skin cancer detection and classification.” Distributed and Parallel Databases 40.4 (2022):717-736.

[18] Skin Cancer Detection using Machine LearningTechniques Vidya M, Dr. Maya V Karki