SKIN CANCER DETECTION USING CONVENTIONAL CNN ARCHITECTURE

PULLAKHANDAM VAMSI KRISHNA MANIKANTA SAI [1], SHAIK MASTAN VALI [2], POTNURU PAVAN KUMAR [3] , M RAVI TEJA[4], Mr. G AMAR TEJ [5]

[1,2,3,4] UG Student, Dept of Electronics and Communication Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, India
[5] Assistant Professor, Dept. of Electronics and Communication Engineering, Vasireddy Venkatadri Institute of Technology, Andhra Pradesh, India

ABSTRACT

Skin cancer is one of the most common forms of cancer worldwide, and its early detection is crucial for effective treatment and patient improvement. Traditionally, the diagnosis of skin lesions has been performed by dermatologists through visual inspection, which can be subjective and may vary in accuracy between the people. Here we proposed the potential of Deep Learning technique to develop a robust detection system of skin cancer by using Convolutional Neural Networks (CNNs).CNN one of a class of Deep Learning models known for their exceptional image recognition capabilities. Our system which uses CNN with different layers with LeakyReLU activation function and Adam optimizer technique to identify skin lesions and to eliminate personal analysis of dermoscopy images by experts, because CNN based algorithm is capable of accurately classifying skin lesions as benign and malignant.

Keywords : *- Adam optimizer, Convolution Neural Networks (CNN), LeakyReLU, Skin Cancer Detection*.

---

1. INTRODUCTION

Skin cancer is a disease characterized by the abnormal growth of skin cells, often caused by exposure to ultraviolet (UV) radiation from the sun or tanning devices. The three primary types of skin cancer are Basal cell carcinoma, Squamous cell carcinoma, and Melanoma. Basal cell carcinoma originates from basal cells in the lower epidermis and is the most common type, usually developing on the head and neck due to sun exposure. Squamous cell carcinoma arises from flat, scale-like squamous cells in the epidermis and is also associated with sun exposure. Melanoma is the most aggressive type, developing from melanocytes and having the potential to spread to other parts of the body if not detected early. Early detection and sun protection are crucial in preventing and managing skin cancer. Diagnosing skin cancer presents challenges for dermatologists due to similarities in the appearance of various pigments. This paper proposes a deep learning approach for multi-class skin cancer classification using the HAM10000 dataset. By leveraging convolutional neural networks, the developed model enhances classification accuracy and reduces loss, thereby improving diagnostic outcomes in skin lesion classification.

2. LITERATURE SURVEY

Vijayalakshmi M M[1] developed "Melanoma Skin Cancer Detection using Image Processing and Machine Learning" explores the use of Back Propagation Algorithm (Neural Networks), Support Vector Machine (SVM), and Convolutional Neural Networks (CNN) for melanoma detection, achieving an accuracy of 85%. The ISIC dataset (International Skin Imaging Collaboration) was utilized for training and evaluation.

"Automated System for Prediction of Skin Disease using Image Processing and Machine Learning" by Lakshay Bajaj, Himanshu Kumar, and Yasha Hasija[2] presents a dual-stage approach combining image processing and machine learning, achieving an accuracy of 90%. The system employs preprocessing techniques and feature extraction in the first stage, followed by the utilization of artificial neural networks in the second stage to identify and predict skin diseases.

The paper titled "Deep CNN and Data Augmentation for Skin Lesion Classification" by Tri-Cong Pham, Chi-Mai Luong, Muriel Visani, and Van-Dung Hoang[3] developed a Deep Convolutional Neural Network (CNN) approach enhanced with data augmentation for skin lesion classification. The study reports improvements in AUC (89.2% vs. 87.4%), AP (73.9% vs. 71.5%), and Accuracy 87.2%. The dataset used for evaluation is a public skin lesion testing dataset.

"Deep Learning for Melanoma Detection in Dermoscopy Images" by Julie Ann A. Salido and Conrado Ruiz Jr[5].The authors employ a deep learning methodology based on Convolutional Neural Networks (CNNs) for the detection of melanoma in dermoscopy images. Dermoscopy is a non-invasive imaging technique widely used by dermatologists for skin cancer diagnosis. The utilization of deep learning techniques, particularly CNNs, allows for automated and accurate classification of dermoscopy images, aiding in the early detection of melanoma. The dataset utilized in this study is the PH2 dataset with an accuracy of 93%.

The paper titled "Skin Cancer Detection Using Ensemble of Machine Learning and Deep Learning Techniques" by Nachiketa Hebbar and Hemprasad Y. Patil[5] presents a novel approach that combines machine learning and deep learning techniques for skin cancer detection. The deep learning model employs state-of-the-art neural networks to extract features from images, while the machine learning model processes image features obtained through techniques such as Contourlet Transform and Local Binary Pattern Histogram. The study achieved an accuracy of 93% using the ISIC Archive dataset, demonstrating the effectiveness of the proposed ensemble approach in accurately detecting skin cancer.

Pratik Kanani and Mamta Padole[6] introduces a deep learning model trained on the HAM10000 dataset to identify seven types of skin cancer. Achieving a test accuracy of 77.98%, validation accuracy of 77.31%, and training accuracy of approximately 82% over 100 epochs, the model outperforms previous studies with an accuracy of 77.03%. Utilizing Google Colab, the study underscores the accessibility of deep learning for skin cancer detection.

3. DATA SET

HAM10000 dataset is obtained from the Kaggle Repository[7].HAM10000 dataset, which includes over 10,000 high-quality dermatoscopic images of different benign and malignant skin diseases such as melanocytic nevi, melanomas, and basal cell carcinomas, is an essential tool for research in dermatology and computer vision. This dataset is widely used by researchers to create and evaluate deep learning and machine learning algorithms for automated diagnosis and classification of skin lesions.

These categories encompass melanocytic nevi (nv), melanomas (mel), benign keratoses (bkl), basal cell carcinomas (bcc), actinic keratoses (akiec), vascular lesions (vasc), and dermatofibromas (df). Table I provides a breakdown of the image count for each class within the dataset.

The dataset is split into the following sections.

TABLE 1: Images in different classes of the dataset

Classes	Testing	Training	Total
Class 0	82	245	327
Class 1	129	385	514
Class 2	275	824	1099
Class 3	29	86	115
Class 4	1677	5028	6705
Class 5	36	106	142
Class 6	279	834	1113

4. PROPOSED WORK

4.1 The flow chat mentioned below explains the work flow of our project

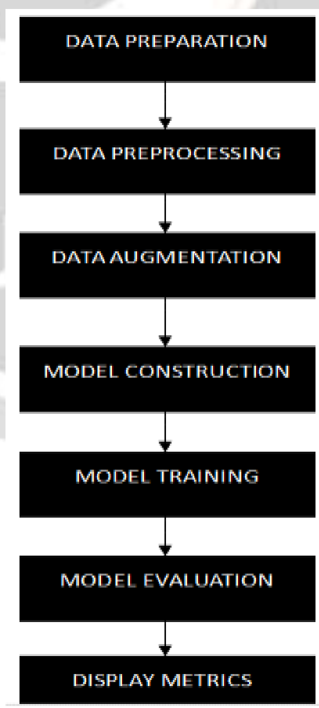

Fig. 1: Flow of the model

4.1.1 DATA PREPARATION :
From Fig 1 step Data Preparation is about importing the skin cancer image data from a CSV file. This file typically contains information about the images, such as pixel values and corresponding label and dataset is divided into two subsets: training and testing sets. It's crucial to ensure that each class is equally represented in both sets to prevent bias during model training and evaluation.

4.1.2 DATA PREPROCESSING :
Class imbalance is a common issue in medical datasets, where some classes may have significantly fewer samples than others. To address this, oversampling techniques are applied to increase the representation of minority classes, ensuring a balanced distribution across all classes. Before feeding the data into the model, pixel values are scaled to a range suitable for processing, typically between 0 and 1. Additionally, the data is reshaped to match the input requirements of the model architecture. Since machine learning models require numerical inputs, categorical labels representing different classes of skin cancer are encoded into numerical format. This ensures compatibility with the model training process.

4.1.3 DATA AUGMENTATION :
Data augmentation techniques are applied to artificially increase the diversity of the training dataset. This involves performing transformations like rotation, flipping, and zooming on the images. By introducing variations in the data, the model becomes more robust and capable of handling different variations of skin lesions.

4.1.4 MODEL CONSTRUCTION:
In designing the Convolutional Neural Network (CNN) architecture for skin cancer classification, we incorporate the Leaky ReLU activation function. Leaky ReLU is chosen for its ability to address the vanishing gradient problem by allowing a small gradient when the input is negative, which can expedite training and improve convergence.

4.1.5 MODEL TRAINING:
For model compilation, we utilize the Adam optimizer, a popular choice for deep learning tasks. Adam optimizer combines the advantages of by maintaining per-parameter learning rates and adapting them over time based on the first and second moments of the gradients. This adaptive learning rate method helps in efficiently updating the model parameters during training and often leads to faster convergence and better performance compared to traditional optimization algorithms.

4.1.6 MODEL EVALUATION:
Once training is complete, the model's performance is evaluated on the testing dataset. This involves measuring metrics such as accuracy and loss to assess how well the model generalizes to unseen data.

4.1.7 DISPLAY METRICS:
After evaluating the trained model on the testing dataset, a comprehensive summary of evaluation metrics is provided to assess the model's performance. This summary includes metrics such as accuracy, loss, and classification reports, which provide insights into the model's predictive capabilities and overall effectiveness. Additionally, the confusion matrix, depicting the actual versus predicted class labels, is presented to visualize the model's performance across different classes. By analyzing these metrics and the confusion matrix, researchers and practitioners can identify the model's strengths and weaknesses, understand its predictive behavior, and make informed decisions to further refine and improve the model's performance if necessary.

4.2 Architecture Of CNN

Convolutional Neural Network (CNN) architecture[8] for the skin cancer detection model using the HAM10000 dataset features a meticulously crafted design with the goal of optimizing accuracy and enhancing robust pattern recognition. Every component within the model, ranging from Conv2D,Polling layer , input layers, Flatten and optimizer and activation function these meticulously configured with specific parameters. The convolutional layers employ diverse filter sizes and activation functions deliberately chosen to encapsulate the various features present in skin lesions image.

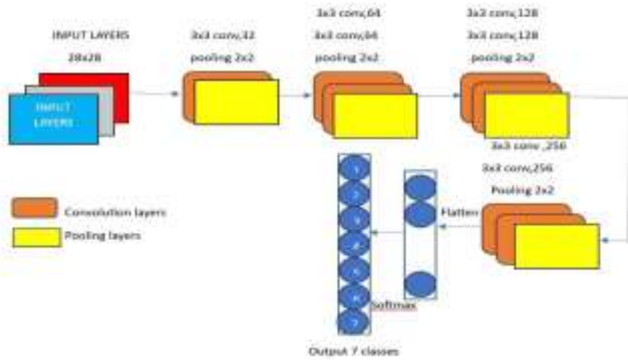

Fig. 2: Architecture of CNN

4.2.1 CONVOLUTIONAL LAYER:

Four convolutional layers are used in the model to extract hierarchical information from photos of skin cancer. The architecture uses multiple convolutional layers, each with a kernel size of 3x3 with 32, 64, and 128 filters. Learnable filters are used by each layer to identify patterns and features, which helps to capture intricate representations. Non-linearity is introduced using leaky ReLU activation functions, which helps the model discover complex relationships in the data. By normalizing activations, batch normalization stabilizes and speeds up training. The model improves classification accuracy by gradually learning to distinguish between various forms of skin cancer through the use of many convolutional layers. All things considered, the convolutional layers are essential to the model's ability to successfully extract discriminative features and produce precise predictions.

4.2.2 POOLING LAYER:

In this model, four max-pooling layers are utilized to reduce the feature map size by selecting the maximum value within 2x2 windows. This strategy enhances computational efficiency while retaining crucial features. Max-pooling aids in promoting generalization, robust feature extraction, and mitigating overfitting, thereby maximizing the effectiveness of the convolutional neural network (CNN) in accurately categorizing images.

4.2.3 FLATTEN LAYER:

A one-dimensional array is created by the flatten layer in the Architecture, which converts the multidimensional output of the previous convolutional layers. the multi-dimensional feature maps from the previous convolutional layers into a one-dimensional vector with 256 elements. Transforming the spatial hierarchical characteristics that the convolutional layers extracted into a manner appropriate for classification requires this reshaping process. The model is able to process the extracted features and feed them into the next fully linked layers by using the feature mappings to their flattening. Because of this, the network can understand intricate correlations and patterns in the data. The CNN model's overall efficacy in skin cancer classification tasks is mostly attributed to the flatten layer, which is crucial in streamlining feature aggregation and category creation.

4.2.4 Adam optimizer:

Adam optimizer is an adaptive learning rate optimization algorithm widely employed in training neural networks. It combines the advantages of both momentum optimization and RMSProp. Adam stands for Adaptive Moment Estimation, and it dynamically adjusts the learning rate for each parameter during training. This adaptation is based

on exponentially decaying average of past gradients and squared gradients. The key benefits of Adam include robustness to noisy gradients, efficient convergence, and automatic adjustment of learning rates based on the specific characteristics of each parameter. Additionally, Adam typically requires less tuning of hyperparameters compared to traditional optimization algorithms, making it a popular choice for optimizing deep learning models.
Formula for Adam Optimizer is below

$$(\theta t+1= \theta t - \text{learning rate x } m\hat{}t / \text{rootover}(v\hat{}t)+\epsilon) \quad (1)$$

θt is is the parameter at time step t.

learning rate is the learning rate hyperparameter.

m^t is the bias-corrected first moment estimate.

v^t is the bias-corrected second moment estimate.

ϵ is a small constant for numerical stability.

4.2.4 Leaky ReLU:

Rectified Linear Units, or leaky ReLUs[9], are activation functions that are frequently employed in neural networks to add non-linearity. It solves the problem of "dying neurons," in which units may fall dormant during training, and is an improvement over the traditional ReLU function. When given positive inputs, the Leaky ReLU function maintains the positive output; but, when given negative inputs, it introduces a modest, non-zero slope. This little change keeps neurons from going dormant by ensuring that they contribute to learning even in the face of unfavorable inputs. Leaky ReLU is a popular activation function in deep learning architectures because it properly addresses the issue of disappearing gradients while retaining the implementation simplicity associated with ReLU.

$$F(x)=\max(a*x ,x) \quad (2)$$

Here a,(alpha) is a small constant value, typically between 0.01 and 0.1. This slope ensures that even negative inputs contribute slightly to the output, preventing "dying ReLU" and allowing the network to learn from a wider range of inputs.

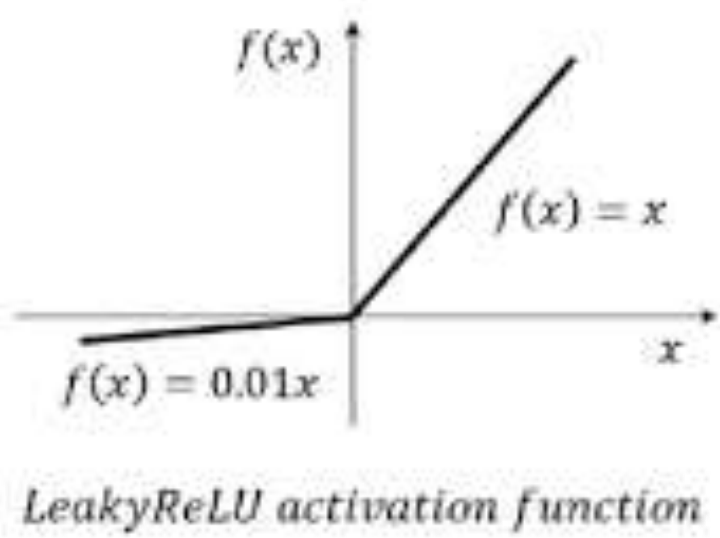

Fig 3 : Graph for Leaky ReLU Activation[10]

5. EXPERIMENT AND RESULTS

The model was trained and validated for 30 consecutive epochs as part of the experimental work.According to the findings, the proposed approach has a accuracy of 96.30 percentage. The below table 2 follows the Precision,Recall and F1 score.

TABLE 2: Evaluation Metrics

classes	precision	recall	f1-score	Support
0	0.80	1.00	0.89	1667
1	1.00	1.00	1.00	1689
2	0.99	1.00	1.00	1651
3	1.00	1.00	1.00	1629
4	1.00	0.76	0.86	1663
5	1.00	1.00	1.00	1680
6	0.97	0.99	0.99	1755
accuracy			0.96	11734
macro avg	0.97	0.96	0.96	11734
weighted avg	0.97	0.96	0.96	11734

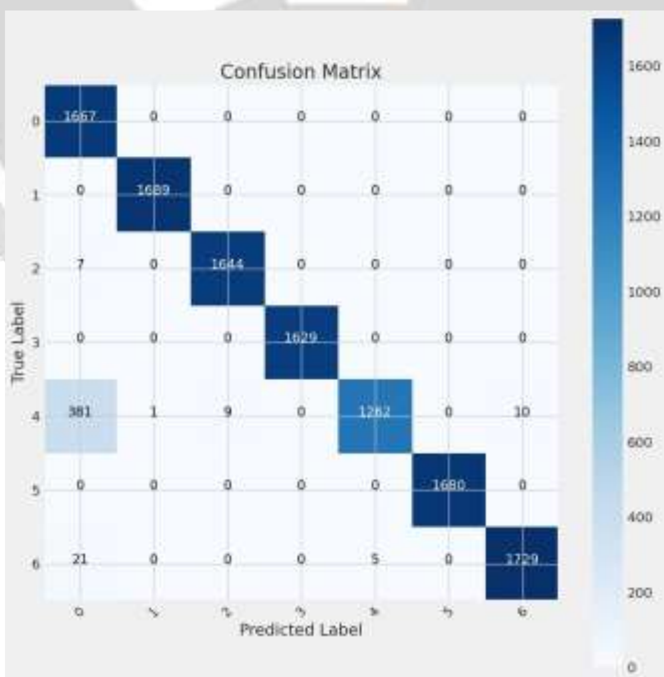

Fig 3: Confusion Matrix

$$\text{Accuracy} = \frac{TP + TN}{TP+TN+FP+FN} = 96 \qquad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} = 80 \qquad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} = 100 \qquad (5)$$

$$\text{F1-score} = 2 * \frac{precision * Recall}{Precision + Recall} = 89 \qquad (6)\ [11]$$

TP is the Instances that are actually positive and are correctly predicted as positive.
TN is Instances that are actually negative and are correctly predicted as negative.
FP is the Instances that are actually negative but are incorrectly predicted as positive.
FN is the Instances that are actually positive but are incorrectly predicted as negative.

The classification report and accuracy score show the performance metrics of a machine learning model on a dataset with seven classes. The precision, recall, and F1-score are reported for each class, indicating the model's ability to correctly identify instances of each class. The support column denotes the number of instances for each class in the dataset. Overall accuracy, as well as macro and weighted averages of precision, recall, and F1-score, provide a comprehensive assessment of the model's overall performance. With an accuracy of 96.37%, the model demonstrates strong performance in accurately classifying instances across multiple classes.

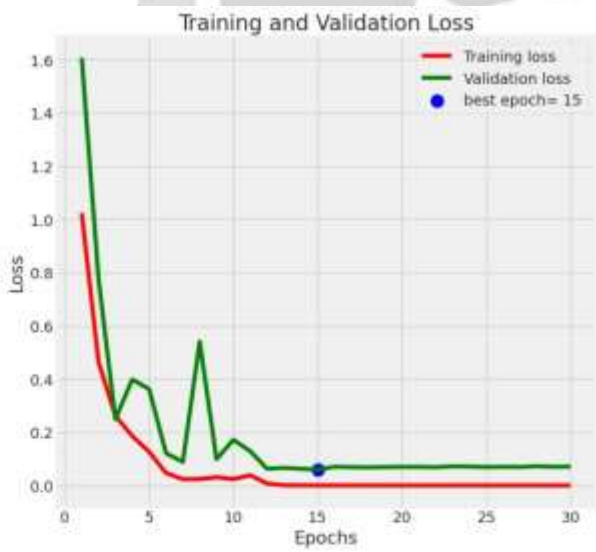

Fig 4: Validation Loss

Fig 5: Validation Accuracy

From Fig 4 the x-axis depicts the number of epochs while the y-axis represents the loss. With increasing epochs, both training and validation losses decrease, indicating the model's learning and enhancement of performance on the training data. It's crucial to monitor validation loss to prevent overfitting. Notably, the validation loss surpasses the training loss, indicating favorable model performance. Additionally, the epoch with the lowest validation loss is observed at epoch 15.

From Fig 5 x-axis illustrates the number of training cycles (epochs), while the y-axis displays the accuracy. As the number of epochs increases, both training and validation accuracy demonstrate an upward trend, indicating the model's learning process. It is essential for the validation accuracy to increase alongside the training accuracy to prevent overfitting on the training data. According to the graph the highest validation accuracy of 0.9 was attained at epoch 22.

6. CONCLUSION

The substantial risk that skin cancer poses to public health highlights the critical need for early detection techniques. Although early research was limited by tiny datasets, automatic detection techniques present a promising way to rapidly diagnose skin cancer. Once this obstacle was overcome, the HAM10000 dataset materialized, offering a big and trustworthy set of 10,015 photos across seven different classifications. Through rigorous experimentation and fine-tuning of our CNN architecture, we have effectively leveraged deep learning technique to accurately classify skin lesions and detect potential instances of cancerous growth. This high level of accuracy promise for improving health care outcomes and saving lives.

7. REFERENCES

[1] Vijayalakshmi M M "Melanoma Skin Cancer Detection using Image Processing and Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-4, June 2019, pp.780-784.

[2] Lakshay Bajaj, Himanshu Kumar, Yasha Hasija . Automated System for Prediction of Skin Disease using Image Processing and Machine Learning. International Journal of Computer Applications. 180, 19.

[3] Pham, TC., Luong, CM., Visani, M., Hoang, VD. (2018). Deep CNN and Data Augmentation for Skin Lesion Classification. In: Nguyen, N., Hoang, D., Hong, TP., Pham, H., Trawiński, B. (eds) Intelligent Information and Database Systems. ACIIDS 2018. Lecture Notes in Computer Science(), vol 10752.

[4] Julie Ann A. Salido and Conrado Ruiz Jr., "Using Deep Learning for Melanoma Detection in Dermoscopy Images," International Journal of Machine Learning and Computing vol. 8, no. 1, pp. 61-68, 2018.

[5] Tembhurne, J.V., Hebbar, N., Patil, H.Y. et al. Skin cancer detection using ensemble of machine learning and deep learning techniques. Multimed Tools Appl 82, 27501–27524,2023.

[6] Kanani, P., & Padole, D.M. (2019). Deep Learning to Detect Skin Cancer using Google Colab. International Journal of Engineering and Advanced Technology.

[7] https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000.

[8] https://researchgate.net/figure/A-generic-CNN-Architecture_fig1_344294512

[9] https://www.educative.io/answers/what-is-leaky-relu

[10]https://www.researchgate.net/figure/ReLU-activation-function-vs-LeakyReLU-activation-function_fig2_358306930

[11] https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score