Skin Cancer Detection from Lesion Images using Deep Learning

Nithya H S
Dept. Of Computer Science And Engineering
BMS Institute Of Technology And Management
Bangalore,Karnataka

Dr Hemamalini B H
Associate Professor
Department Of Computer Science And Engineering
BMS Institute Of Technololgy And Management

***Abstract*—Skin cancer is a prevalent and potentially life-threatening condition that can be effectively treated with early detection. By applying deep learning techniques, this system is designed to identify skin cancer. The initiative is based on a data collection from the International Skin Imaging Collaboration (ISIC), which includes nine classifications of pictures that indicate cancer. Convolutional neural networks (CNNs) and deep learning models such as ResNet-50, VGG16, VGG19, EfficientNet, and DenseNet have been analyzed and compared. A website incorporates the model with the maximum accuracy. There's a chatbot on this website that responds to questions about skin cancer.**

I. INRODUCTION

Detecting skin cancer early is crucial for prevention and effective treatment, given its status as one of the most widespread forms of cancer worldwide. With the rise in skin cancer incidence over the past decade, there's a pressing need for advanced diagnostic tools. Leveraging deep learning, which excels at recognizing intricate patterns in images, we aim to develop a system for automated skin cancer detection. Skin cancer, encompassing basal cell carcinoma, squamous cell carcinoma, and melanoma, stands out as the most prevalent type, considering the skin's expansive coverage of the body. Melanoma, while relatively rare, poses a significant threat due to its potential fatality. However, when detected early, melanoma boasts an impressive 99% 5-year survival rate.

II. RELATED WORKS

[1] The study underscores the importances of early skin cancer detection in light of its rapid global proliferation. It stresses the significance of leveraging deep learning algorithms and dermoscopy techniques for precise, automated diagnosis of skin lesion. The challenges in distinguishing between malignant lesions and benign are highlighted, alongside the correlation between UV radiation exposure and skin cancer incidence. The study recommends the utilization of convolutional neural networks for lesion classification, noting their superior performance compared to human experts. Ultimately, the study underscores the critical importance of combating skin cancer through timely diagnosis and effective treatment strategies.

[2] The study presents a lightweight skin cancer segmentation and recognition algorithm that is tested using data from the ISIC-2016. With CNNs and a U-Net architecture, it focuses on distinguishing minute changes between lesions, exhibiting better results than previous deep learning techniques in recognizing different skin types.

[3] Utilizing the International Skin Imaging Collaboration 2019 dataset, the study introduces an AI-powered system for categorizing skin lesions, integrating “deep learning” and “LIME” explanation techniques. Achieving an accuracy rate of 94.47%, the system aids dermatologists by offering transparent insights into model predictions, thereby augmenting early-stage skin cancer detection.

[4] The research combines VGG, CapsNet, and ResNet models to boost skin cancer detection accuracy, outperforming individual methods through automatic feature extraction. This approach shows promise for broader medical applications and warrants exploration in larger datasets.

III. OBJECTIVES

To detect skin cancer at an early stage: The aim is to create a robust deep learning model capable of detecting the skin cancer earlier, with a specific focus on melanoma. This initiative seeks to improve treatment effectiveness and potentially prevent loss of life.

To increase accuracy: The deep learning model is being trained to accurately classify skin lesions, ensuring consistent performance across diverse instances and datasets. Validation of the model's efficacy across various scenarios is crucial.

To make a comparative study: In the quest to identify the optimal model for skin cancer diagnosis, an evaluation of deep learning methodologies is underway. This includes comparing custom convolutional neural networks (CNNs) against transfer learning models like ResNet50, VGG19, VGG16, DenseNet, and EfficientNet.

IV. PROPOSED METHODOLOGY

A. DEEP LEARNING

Deep Learning, a subset of Machine Learning, employs neural networks to address intricate problems. These networks comprise interconnected layers of nodes, mimicking the structure and functionality of the human brain. Essential to deep learning is the utilization of deep neural networks, featuring multiple layers of interconnected nodes. These networks excel at recognizing hierarchical patterns and features in data, enabling them to learn sophisticated data representations. Deep Learning systems leverage data to autonomously improve over time, negating the necessity for manual feature engineering by humans.

B. APPLIED MODELS

ResNet50: ResNet50 is a variation of the residual network architecture (ResNet), bypasses one or more layers by using identity connections, often referred to as shortcut connections. This helps solve the issue that deep networks have with disappearing gradients. The model consists of 50 layers: one average pool layer, one MaxPool layer, and 48 convolutional layers.

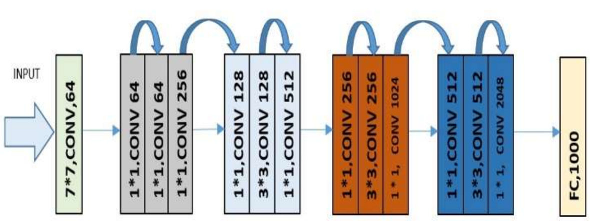

Fig 1: Architecture model of ResNet50

VGG-16: Typically, the VGG-16 design has 16 layers: 3 fully linked layers and 13 convolutional layers. These layers are arranged into blocks, with a max-pooling layer for down sampling coming after each block that contains several convolutional layers.

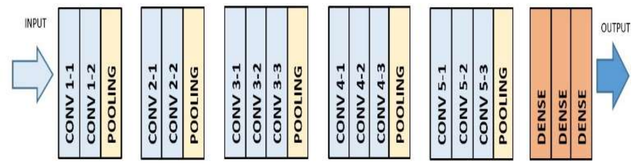

Fig 2: Architecture model of VGG16

VGG-19: There are 16 convolution layers in VGG-19, organized into 5 blocks. Following each block is a Maxpool layer which reduces the input picture size by 2 and raises the convolution layer's number of filters by 2. The final three dense layers in block 6 have the following dimensions: 4096, 4096, and 1000, in that order.

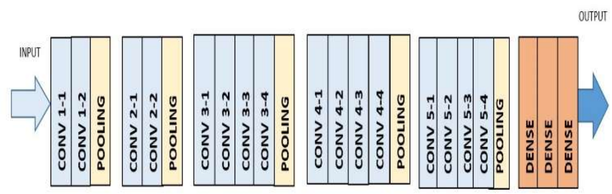

Fig 3: Architecture model of VGG19

DenseNet201: DenseNet's design consists of dense blocks and transition layers. In a dense block, all the convolutional layers are connected to all of the other layers. This is achieved by creating a "shortcut" link by joining each layer's output to the layer that follows.

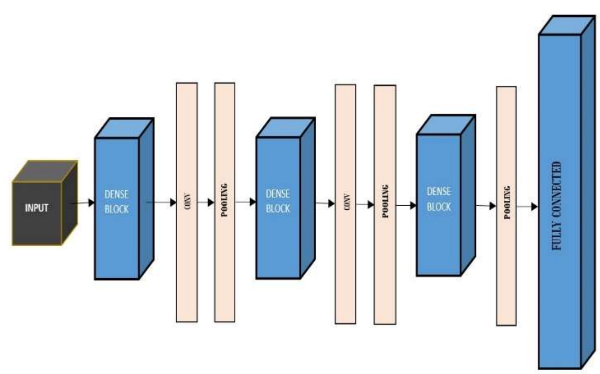

Fig 4: Architecture model of DenseNet201

EfficientNet: EfficientNet utilizes stacked convolutional layers in blocks, employing the mobile inverted bottleneck convolution (MBConv) as its basic unit. Through compound scaling, it adjusts width, depth, and resolution with coefficients α, β, and γ. This balance optimizes model computational efficiency and complexity, achieving top-tier performance in image classification tasks.

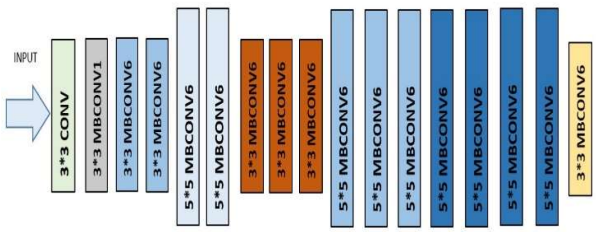

Fig 5: Architecture model of EfficientNet

Dataset: The dataset provided by the "International Skin Imaging Collaboration (ISIC)" comprises 2357 images showcasing various benign and malignant dermatological conditions. While melanomas and moles are somewhat more prevalent in the photos, all images were categorized according to the ISIC classification, with each subset containing an equal number of images. The dataset encompasses the following conditions:
- Actinic keratosis
- Basal cell carcinoma
- Dermatofibroma
- Melanoma
- Nevus
- Pigmented benign keratosis-seborrheic keratosis
- Squamous cell carcinoma
- Vascular lesion

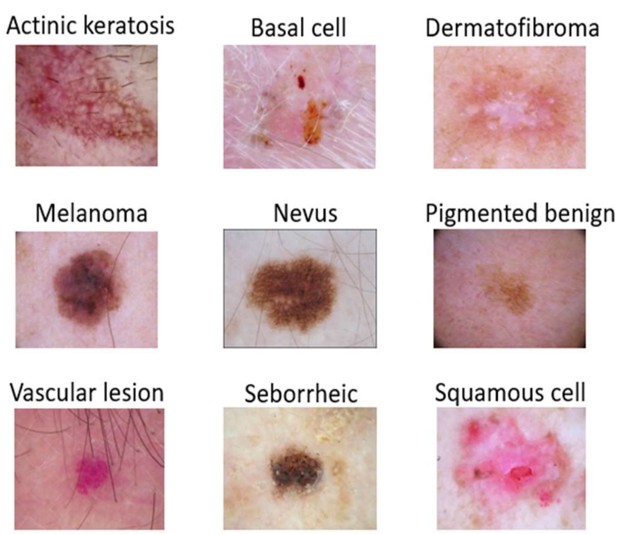

Fig 6: Nine classification of skin cancer

Augmentation: Data augmentation, a machine learning technique, involves applying various transformations to existing data samples to increase the diversity and size of a dataset artificially. In the realm of deep learning, this approach proves highly beneficial. In the context of identifying skin cancer from lesion images, augmentation entails producing new images with alterations in brightness, contrast, noise, translation, rotation, flipping, scaling, and cropping. By generating these modified versions of the original image, the model learns to generalize more effectively and becomes resilient to variations in real-world data, thereby improving its ability to detect skin cancer lesions.

Preprocessing: Preprocessing involves employing diverse methods to obtain images, aiming to enhance image properties such as clarity and quality by reducing or removing unwanted components like backdrops. Grayscale conversion, noise reduction, and image enhancement are the primary preprocessing techniques. Initially, all images are converted to grayscale. Subsequently, the images undergo enhancement and noise reduction using two filters: the median filter and the Gaussian filter.

Segmentation is the process of dividing the region of interest in an image, assigning similar properties to each pixel to facilitate distinction. This enables individual image segments to be processed independently, rather than the entire image.

Classification and clustering algorithms share similarities, yet clustering methods are considered unsupervised. Clustering involves identifying distinctive segments or clusters in the data, separate from the background. Using k-means clustering, the data is typically partitioned into k clusters based on k centroids. This technique is particularly helpful when data lacks labels, allowing the formation of distinct groups based on data similarities. Selecting the appropriate number of clusters, k, is a crucial step in this process. Initially, k centroids are randomly chosen. Each data point is then assigned to the nearest cluster centroid. Subsequently, the new centroids for each cluster are calculated and updated.

Feature extraction: It is widely regarded as the pivotal phase in the overall categorization process. It involves extracting pertinent features from the input dataset to facilitate subsequent computations, such as detection and classification. In our proposed methodology, two approaches, namely ABCD and GLCM, are employed for extracting skin lesion features, which are then recorded in an excel sheet..

Classification: In conducting a comparative analysis, five distinct deep learning algorithms are deployed. These algorithms are tasked with categorizing lesion images into one of nine groups. The selected algorithms comprise VGG-19, VGG-16, ResNet-50, DenseNet, and EfficientNet.

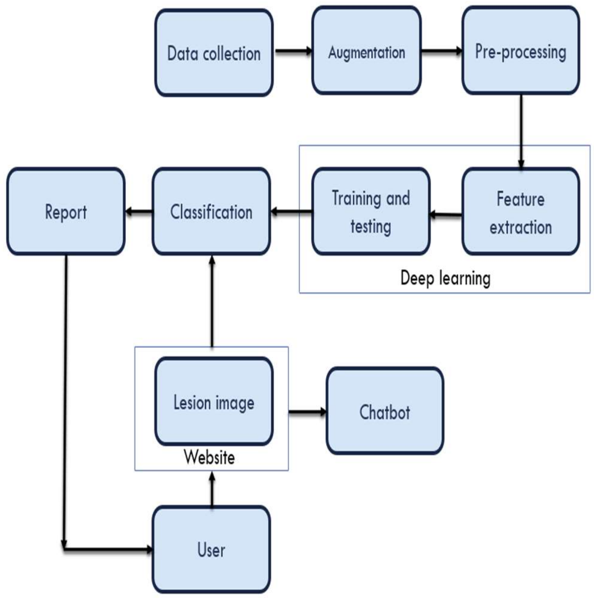

Fig 7: The overall process

V. RESULTS

Fig. 8 displays the comparison study between five different models, which include VGG16, VGG19, ResNet50, EfficientNet, and DenseNet. Through this table, we will come to know the accuracy, f1-score, and precision for each model.

Algorithm	Accuracy	F1-Score	Precision
VGG-16	90.47%	90.25	90.32
VGG-19	91.33%	91.29	91.30
ResNet-50	88.38%	88.35	88.37
EfficientNet	91.11%	91.07	91.12
DenseNet	90.89%	90.73	90.86

Fig 8: Evaluation metrics

A. The Influence of Model Architecture, Overfitting and Model Complexity on Performance.

Differences in model architecture are primarily responsible for the performance variations among the models (VGG-16 VGG-19 ResNet50 EfficientNet and DenseNet). Stacked convolutional layers are the foundation of the comparatively simple VGG-16 and VGG-19 architectures which are followed by fully connected layers. The primary distinction between the two is that in theory VGG-19 can extract more intricate features from the input images because it has more convolutional layers (19 as opposed to 16). Nevertheless the model may find it difficult to acquire valuable representations without overfitting to the training data which could result in diminishing returns. VGG-16 does fairly well despite the additional layers with an accuracy of 90. 47 percent and an F1 score of 90. 25 percent. This suggests that when the dataset is too small or diverse to justify the additional capacity of deeper networks like VGG-19 simpler architectures can occasionally outperform more complex ones.

Overfitting is another significant factor that causes performance discrepancies especially in more complex models like VGG-19. Overfitting is the term used to describe when a model learns to memorize the training data instead of generalizing to new data. More-layered models like VGG-19 and ResNet50 are better at remembering patterns but they may perform worse on test data if regularization techniques like data augmentation are not applied correctly. Despite having more parameters VGG-19 performs little better than VGG-16 which may be less prone to overfitting due to its simpler structure. DenseNet and EfficientNet seem to lessen overfitting more effectively with their innovative structures by encouraging feature reuse (DenseNet) and employing compound scaling (EfficientNet)

Model complexity also plays a significant role in the model discrepancies as well. ResNet50 is intended to train deeper networks by resolving the vanishing gradient issue through the use of residual connections. Although the models deeper (50 layers) architecture enables it to learn more intricate features it can also be vulnerable to underfitting if not properly trained or if the dataset lacks enough variability. Here ResNet50 outperforms VGG-16 and VGG-19 in terms of accuracy (88. 38 percent) and F1 score (88. 35 percent). This suggests that the model might not be making the most of its depth for the skin cancer classification task despite its intricate architecture. However EfficientNet and DenseNet provide higher accuracy and F1 scores (EfficientNet at 91. 11 percent and DenseNet at 90. 89 percent) because they can efficiently scale both depth and width while preserving computational efficiency. This suggests that the models ability to balance efficiency and complexity can improve their ability to generalize to the dataset

Overall, when the model is less prone to overfitting, simpler architectures like VGG-16 might perform on par with more complex models like ResNet50. On the other hand, architectures built for efficiency and scalability, such as EfficientNet and DenseNet, show better overall generalisation to new data, which improves accuracy and F1 scores.

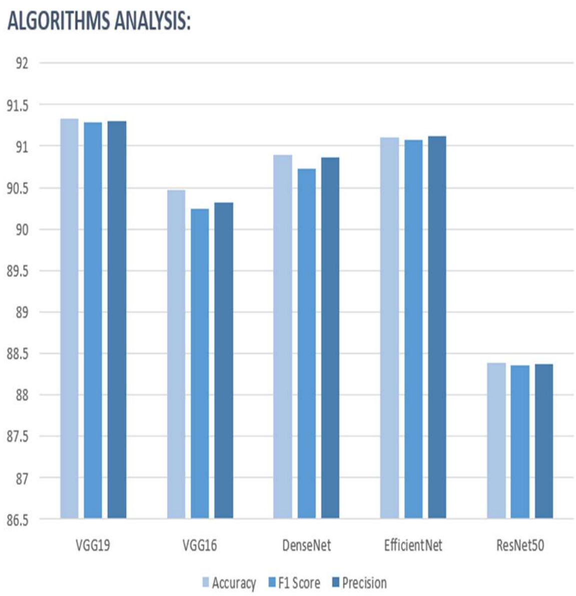

Fig 9: Algorithm analysis

A. Different Measures Used to Evaluate Performance

1) Accuracy

In machine learning and classification tasks, accuracy refers to the percentage of correctly classified cases among all instances evaluated. It measures the model's ability to predict class labels accurately based on the input data. The calculation of accuracy involves dividing the total number of predictions made by the model by the number of accurate forecasts, while accuracy is a fundamental metric for assessing classification model performance, it may not always be the most informative, especially in datasets with imbalanced class distributions where certain classes are significantly more prevalent than others.

$$\text{Accuracy} = \frac{TN+TP}{FN+FP+TN+TP}$$

2) Precision

The degree of exactness, correctness, or refinement in a procedure, measurement, or result is typically referred to as precision. In manufacturing, craftsmanship, language, and scientific activities, it indicates the degree to which a work satisfies a particular aim or standard, emphasizing the precision and attention to detail needed to achieve desired results.

$$\text{Precision} = \frac{TP}{FP+TP}$$

3) *F1-Score*

The F1 score assesses a model's accuracy in binary classification tasks by computing the harmonic mean of precision and recall. It offers a balanced evaluation of these two metrics, assigning equal importance to both, which proves advantageous in scenarios with imbalanced class distributions. Ranging from 0 to 1, with 1 representing perfect precision and recall and 0 indicating poor performance, the F1 score is a crucial metric for evaluating classification models. Its utility becomes evident when there's a need to account for both false positives and false negatives.

$$F1 = 2 \times \frac{\text{Precision} + \text{Recall}}{\text{Precision} \times \text{Recall}}$$

VI. CONCLUSION

In conclusion, this project presents a holistic and innovative solution to skin health concerns. By combining deep learning for skin cancer detection, a chatbot, and a user-friendly interface, it addresses a wide range of healthcare objectives. Has the potential to make a significant positive impact on early diagnosis, user awareness, and healthcare accessibility in the field of dermatology.

Using deep learning to identify skin cancer from image skin imperfections presents exciting new opportunities for accurate, accessible, and early diagnosis in the medical field. These technologies provide the highly precise diagnosis of worrisome lesions through sophisticated image processing and machine learning algorithms, hence facilitating prompt intervention and better patient outcomes. With further research and development, the application for skin cancer detection using deep learning has great promise to improve public health initiatives and lessen the impact of this common disease.

VII. REFERENCES

[1] H. L. Gururaj, N. Manju, A. Nagarjun, V. N. M. Aradhya and F. Flammini, "DeepSkin: A Deep Learning Approach for Skin Cancer Classification," in *IEEE Access*, vol. 11, pp. 50205-50214, 2023, doi: 10.1109/ACCESS.2023.3274848.

[2] L. Wei, K. Ding and H. Hu, "Automatic Skin Cancer Detection in Dermoscopy Images Based on Ensemble Lightweight Deep Learning Network," in IEEE Access, vol. 8, pp. 99633-99647, 2020, doi: 10.1109/ACCESS.2020.2997710.

[3] 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," in IEEE Access, vol. 10, pp. 113715-113725, 2022, doi: 10.1109/ACCESS.2022.32172

[4] A. Imran, A. Nasir, M. Bilal, G. Sun, A. Alzahrani and A. Almuhaimeed, "Skin Cancer Detection Using Combined Decision of Deep Learners," in IEEE Access, vol. 10, pp. 118198-118212, 2022, doi: 10.1109/ACCESS.2022.3220329.