SvedbergOpen
DISSEMINATIONOFKNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

# An Intelligent Deep Learning Framework for Accurate Skin Cancer Detection Using CNN, ResNet, and GoogLeNet Models

Vunnam Narmada[1*], Kalavala Asish Vardhan[2]

[1,2]Department of Computer Science Engineering, Malla Reddy University, Maisammaguda, Hyderabad, Telangana 500100, India.
*Corresponding Author E-mail: narmadamrecw@gmail.com

**Abstract:**
Skin cancer is one of the most common diseases as well as one of the most threatening for patients' lives; thus, early and accurate detection of skin cancer becomes vital for successful treatment. The current paper studies the use of deep learning models (Convolutional Neural Networks, ResNet, and GoogleNet) in skin cancer classification using dermoscopic images. These deep learning models were chosen because of the possibility of learning hierarchy and distinguishing small differences in images of benign and malignant lesions. Transfer learning was used to speed up training process and improve generalization using pretrained weights for images. The experiment showed that the most accurate model reached the following metrics: 96.8% of accuracy, 95.5% of precision, 96.2% of recall, and 95.8% of F1-score working with a publicly available skin cancer dataset. The findings demonstrate the high efficiency of deep learning when combined with dermoscopic image analysis and transfer learning.

**Keywords:** Skin Cancer Classification, CNN, ResNet, GoogleNet, Dermoscopic Images, Transfer Learning, Medical Image Analysis, Feature Extraction, Skin Lesion Diagnosis.

## 1. Introduction

Skin cancer is an increasingly important global health concern, with melanoma being particularly serious because of its aggressive progression and potential to metastasize. Early identification of suspicious lesions is therefore essential for improving clinical outcomes and supporting timely treatment. Conventional diagnosis commonly depends on clinical examination, dermoscopy, histopathological assessment, and the experience of dermatologists. Although these approaches remain fundamental to clinical practice, the visual similarity among different lesion types can make reliable discrimination difficult, particularly when lesions exhibit subtle differences in colour, texture, border structure, and morphology. The growing availability of computer-aided diagnostic techniques has consequently encouraged the development of automated image-analysis systems capable of assisting clinicians in the early identification of malignant lesions [1], [2].

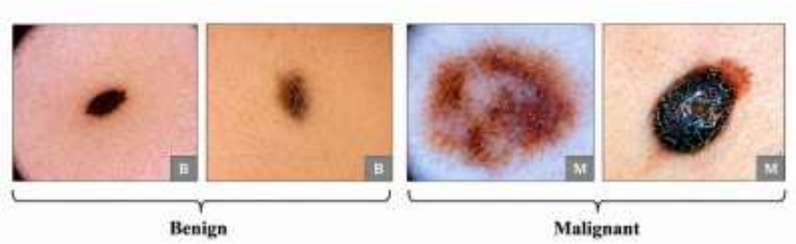


**Figure 1: Examples from ISIC 2019 database and CPTAC-CM database**

Machine-learning and deep-learning approaches have substantially expanded the possibilities for automated analysis of medical images. Earlier computer-aided systems generally depended on manually designed features followed by conventional classifiers, whereas modern deep neural networks can learn discriminative representations directly from image data. Systematic investigations of machine-learning and deep-learning techniques have demonstrated the potential of automated feature learning for skin-lesion classification, while also highlighting persistent challenges related to dataset size, class imbalance, generalization, and model reliability [3]. Traditional machine-learning approaches, including support-vector-machine-based classification, have provided useful foundations for medical image analysis; however, their effectiveness can depend strongly on the quality and representativeness of extracted features [4]. More advanced optimization-assisted deep-learning frameworks have subsequently demonstrated that combining feature learning with suitable optimization strategies can improve classification performance and robustness in skin-cancer detection [5].

The effectiveness of deep learning is particularly evident in image-based dermatological applications because convolutional architectures can learn hierarchical spatial representations from dermoscopic images. Deep-learning-based systems have been investigated for detecting and classifying skin lesions using combinations of image preprocessing, feature extraction, segmentation, and classification. Approaches incorporating clustering, deep feature learning, and image-analysis techniques have demonstrated the feasibility of reducing the dependence on manually engineered descriptors while improving automated lesion characterization [6]. Comparative investigations of different deep-learning architectures have further shown that model architecture plays an important role in distinguishing benign and malignant lesions, with different networks exhibiting different capabilities in feature representation and classification [7]. Similarly, convolutional approaches developed specifically for lesion classification have demonstrated the potential of automated systems to support clinical decision-making while reducing the dependence on subjective visual interpretation [8].

The development of intelligent diagnostic systems is also closely related to the broader evolution of data-driven computational intelligence. Deep-learning models, transfer-learning strategies, and automated feature-learning frameworks have increasingly been employed where complex patterns must be extracted from high-dimensional data. Studies involving melanoma detection have shown that deep neural architectures can learn useful visual characteristics from lesion images and provide automated classification capabilities [9]. In addition to conventional accuracy-based assessment, appropriate evaluation strategies are necessary to understand the actual behaviour of classifiers, particularly when datasets contain multiple classes or imbalanced distributions. Consequently, comprehensive performance assessment using measures such as accuracy, precision, recall, F1-score, and other classification indicators is important when comparing competing diagnostic models [10]. Deep residual architectures have also demonstrated their ability to learn increasingly abstract lesion representations, providing an effective basis for automated melanoma and skin-lesion analysis [11].

The broader development of intelligent computational frameworks in engineering and data-driven applications further illustrates the increasing role of learning-based decision systems, optimization, and adaptive computational strategies in solving complex problems [12]. Transfer learning has become particularly attractive in medical-image analysis because pretrained models can provide useful feature representations when the amount of task-specific training data is limited. For example, transfer-learning-based classification has been successfully investigated for differentiating multiple categories of skin lesions using established convolutional architectures [13]. Similarly, intelligent multi-agent and learning-based computational frameworks demonstrate how adaptive data-driven models can integrate multiple sources of information and make complex decisions under changing operating conditions [14]. These developments motivate the continued exploration of efficient and robust learning architectures for medical-image applications.

Another important challenge is the segmentation of the lesion before classification. Dermoscopic images can contain hair structures, illumination variations, background regions, noise, and other artefacts that may interfere with the extraction of meaningful lesion characteristics. Deep-learning-based segmentation and multiclass classification methods have therefore been explored to isolate relevant lesion regions and improve subsequent classification [15]. More broadly, intelligent optimization and data-driven planning approaches have demonstrated that combining multiple computational components can improve the effectiveness of complex decision systems [16]. Image augmentation and oversampling techniques have additionally been employed to increase the diversity of training samples and address the limited representation of minority classes in skin-cancer datasets [17]. These observations indicate that a reliable skin-cancer detection framework should not depend solely on the final classifier; rather, preprocessing, lesion isolation, feature representation, class balancing, and classification should be considered as interconnected stages.

Recent research has also explored the integration of adaptive control, learning-assisted decision-making, and intelligent computational architectures in other application domains [18]. Although such studies are outside dermatological diagnosis, they demonstrate the broader applicability of adaptive learning principles for complex, nonlinear, and data-intensive systems. Comparative investigations of deep architectures have shown that different network depths and structures can produce substantially different results for medical image classification. In particular, ResNet architectures employ residual connections to facilitate the training of deeper networks, while other architectures exploit different forms of hierarchical feature learning [19]. More advanced learning frameworks have also explored the integration of digital-twin concepts and reinforcement learning for adaptive decision-making, highlighting the continuing movement toward intelligent, context-aware computational systems [20].

Multi-class skin-cancer classification remains more challenging than simple binary benign/malignant discrimination because several lesion categories can exhibit highly similar visual characteristics. Recent CNN-based investigations have therefore focused on multi-class classification and have demonstrated the usefulness of deeper feature-learning architectures for distinguishing multiple lesion categories [21].

Other intelligent computational approaches reported in the broader engineering literature similarly demonstrate the importance of robust representation learning and adaptive model design when the underlying data distribution is complex [22]. Deep neural networks have also been investigated for early skin-disease identification, demonstrating the feasibility of automated systems for extracting diagnostic information from clinical images [23]. In addition, image-processing and learning-based approaches have emphasized the importance of preprocessing and feature extraction before classification, particularly when image quality and lesion appearance vary substantially [24].

Despite the progress achieved by existing methods, several limitations remain. Skin-lesion datasets can suffer from substantial class imbalance, with some diagnostic categories being represented by considerably fewer images than others. This imbalance can cause a classifier to favour dominant classes and produce apparently high overall accuracy while performing poorly on minority categories. Recent studies have therefore investigated data augmentation and transfer-learning strategies to improve the representation and classification of underrepresented lesion categories [25]. At the same time, learning-assisted computational frameworks developed for other complex engineering systems demonstrate the potential of combining model-based knowledge with data-driven learning to improve adaptability and robustness [26]. Image enhancement represents another promising direction because increasing the quality and resolution of lesion images can improve the availability of discriminative visual information for subsequent CNN-based analysis [27].

Based on these considerations, an effective automated skin-cancer diagnostic framework should address the complete image-analysis pipeline rather than relying exclusively on a deep classifier. Morphological preprocessing can be employed to suppress unwanted hair artefacts, while image enhancement can improve the visibility of lesion structures. Segmentation can subsequently separate the lesion from the surrounding skin, allowing texture and statistical information to be extracted from a more relevant region of interest. Earlier image-processing studies have demonstrated the importance of preprocessing and enhancement for extracting useful visual information from complex images [28]. Comprehensive reviews of deep-learning-based skin-cancer detection further indicate that, although numerous architectures have achieved encouraging results, issues related to generalization, computational complexity, dataset imbalance, and clinical deployment remain unresolved [29]. Optimization-based comparative studies in related image and engineering applications likewise emphasize the importance of evaluating different computational strategies rather than assuming that a single algorithm will be optimal under all operating conditions [30].

Further evidence from recent studies reinforces the need for robust automated diagnostic frameworks. Deep-learning methods have been applied to skin-lesion images for the detection of infectious and dermatological conditions, demonstrating that visual characteristics extracted from lesion images can support automated disease recognition [31]. Hybrid machine-learning strategies have also been proposed for skin-cancer detection, combining extreme-learning-machine-based classification with optimization techniques to improve predictive performance [32]. Comprehensive evaluations of CNN architectures have similarly reported the usefulness of convolutional models for both skin and oral cancer image analysis, while emphasizing the continuing importance of model selection and dataset characteristics [33]. Transfer-learning-based approaches have further demonstrated that pretrained convolutional representations can improve skin-disease diagnosis by providing informative feature hierarchies without requiring the entire network to be trained from the beginning [34], while broader analyses of CNN architectures confirm that transfer learning remains an important strategy for improving automated medical-image classification [35].

Research Gap and Motivation: Although existing studies have achieved promising classification performance, the literature still reveals a need for a systematic end-to-end framework that integrates image preprocessing, lesion segmentation, texture representation, and comparative deep-learning classification. Many existing investigations concentrate primarily on the final classification stage, whereas image artefacts, lesion-background separation, texture variability, class imbalance, and differences between network architectures can substantially influence the resulting predictions. Furthermore, high classification accuracy alone does not necessarily demonstrate robust performance across different lesion categories or clinical image conditions. The supplied literature also indicates that computational complexity, dataset imbalance, limited generalization, and the practical deployment of deep models remain important challenges.

Novelty and Contribution of the Present Work: To address these limitations, the present study develops an integrated skin-lesion analysis and deep-learning classification pipeline in which dermoscopic images undergo sequential preprocessing, hair-artifact removal, image enhancement, lesion segmentation, texture-feature extraction, and deep-learning-based classification. In particular, the framework combines Black-Hat morphological filtering and inpainting, Gaussian image enhancement, GrabCut-based lesion segmentation, and GLCM-based texture characterization before classification using CNN, AlexNet, ResNet18, ResNet50, ResNet101, ResNet152, and GoogLeNet architectures. This integrated arrangement is intended to provide a more systematic comparison of shallow and deep convolutional representations while reducing the influence of irrelevant image structures. The methodology described in the paper explicitly follows this sequence from hair removal and enhancement through segmentation, GLCM extraction, and deep-learning classification.

A further contribution is the comparative evaluation of multiple CNN-family architectures under a common preprocessing and classification framework, rather than assessing a single network in isolation. The study considers accuracy, precision, recall, F1-score, and Hamming loss to provide a broader assessment of classification behaviour. It also considers the practical issue of model transferability by examining performance across multiple datasets, thereby providing insight into how learned lesion representations behave beyond the primary training dataset. The original study specifically compares CNN, AlexNet, ResNet variants, and GoogLeNet and investigates their performance across HAM10000, PH2, and DermNet datasets.

Overall, the proposed framework is intended to contribute toward an accurate, systematic, and computationally practical computer-aided skin-cancer detection system that can assist dermatological decision-making. By combining conventional image-processing techniques with modern deep-learning architectures, the study seeks to improve the quality of lesion representation before classification while providing a comparative understanding of the strengths and limitations of different CNN architectures. Such an approach can potentially support future development of lightweight and clinically deployable diagnostic systems for telemedicine and resource-constrained healthcare environments.

## 2. Related work

For this research, three types of CNNs, namely CNN, ResNet, and GoogleNet, will be used in the framework which acts as an intelligent system for classification of lesions into benign or malignant. Thus, the selected architectures have been chosen purposefully to provide a comprehensive assessment and evaluation of each of them.The CNN architecture, which has been implemented as the baseline in this study, is very shallow, consisting of convolutional and activation functions in tandem, max pool, and fully connected layers.This compact form makes it highly efficient for training with limited data and has a clear hierarchy in its features that consists of simple visual patterns such as the outline of the lesions, pigmentation, and any other irregularities in textures. GoogleNet, also known as Inception v1, presents a deeper model formed of inception modules that simultaneously include several convolutions with various The sizes of filters (1x1, 3x3, 5x5) in the same layer/region. This architecture promotes the effectiveness of networks in recognizing different types of spatial features related to skin lesion due to their different morphologies.The design also allows for the minimization of computational power based on the use of smaller filters for dimensionality reduction.

ResNet (Residual Network) is proposed as a solution that overcomes the shortcomings of traditional deep network architectures through the implementation of Identity shortcut connections have been used in residual learning blocks to allow gradients to skip certain layers in the process of backpropagation. The use of such blocks successfully helps to deal withdeeper network architectures, such as ResNet18, ResNet50, ResNet101, ResNet152. These deeper network architectures provide more opportunity for abstract hierarchical feature representation learning, which is necessary for classifying visually similar sub-types of lesions such as melanoma and benign nevus. Methods were used to handle the class imbalance problem that is present in the HAM10000 data set. In particular, SMOTE (Synthetic Minority Oversampling Technique) and stratified batch generation were applied to make sure that underrepresented classes, for example, Dermatofibroma and vascular lesions, were represented correctly during training. Different measures of performance were used to test the model's performance, and these include Accuracy as a measure of prediction accuracy, F1-Score to measure performance on imbalanced classes, and AUC measure.With various model architectures and algorithms used in this study, the study is able to compare the strength and weakness of different classifiers in the medical image classification domain as shown in several parts of the referenced material, including [10]. There has been a lot of recent research that has focused on improving the skin cancer diagnostic accuracy using deep learning approaches.Li and Shen developed a dual fully convolutional residual network to classify lesion images and achieve the goal of accurate classification and segmentation, and the accuracy level achieved was 91.2%. Through the implementation of transfer learning on the ISIC datasets, Hosny et al. achieved 92.99% and 98.7% accuracy using AlexNet and GoogleNet, respectively. Khan et al. have developed a new technique called Deep Saliency Segmentation, which consists of a ten layer CNN coupled with a kernel ELM classifier. This technique proved to be more effective in segmenting and classifying images. For instance, Abayomi-Alli et al. used covariant SMOTE and SqueezeNet in data augmentation, obtaining an accuracy of 92.18% in diagnosing melanoma.

### A. Deep Learning Models

Convolutional neural networks (CNNs) form a key element of image recognition and classification research. The use of CNN in detecting and diagnosing skin cancer is an example of the effectiveness of CNNs in feature extraction in a particular area. The CNN consists of many layers, including: convolutional layers, where filters work.Other layers include activation function, pooling, and dense layers. In this research, the CNN network is used as the benchmark model.

ResNet:"ResNet" is an abbreviation of "Residual Network," which refers to neural network models using residual blocks for the training of deep models. Examples of ResNet models include ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. In the context of ResNet models, residual blocks are an important part that make up

ResNet18: ResNet18 is a residual neural network model, composed of 18 layers, which proposes the concept of identity shortcut connections, popularly known as skip connections. These connections help the model avoid the problem of vanishing gradients, which always acts as an The residual block can be mathematically expressed using the formula $Y = F(x) + x$, where $F(x)$ is the residual function and x is the input to the system. ResNet-18 has managed to strike the perfect balance between being efficient and deep.In the task of detecting skin cancer, at 25 epochs, ResNet18 produced an accuracy of 86.34%, proving to be one of the strongest performers at relatively lower complexities and faster convergence during training.

ResNet50: ResNet50 ResNet50 is a variant of ResNet, which has greater depth and power, with a total of 50 blocks in its structure. It employs a bottleneck structure that combines stacked binary convolutional layers with the structure of 1x1, 3x3, and then again 1x1 in a sequential manner to minimize its computational complexity with comparable performance levels. In this study, it was observed that ResNet50 was able to process images with dermoscopic patterns efficiently with excellent feature extraction properties, particularly for detailed patterns of lesions on skin cancer images, as indicated by its best accuracy of 88.78%.

ResNet101: ResNet101 is an advancement of the ResNet, with the addition of 101 layers, which helps in extracting abstract features from images. The architecture of this network is also developed with residual blocks, which helps in the flow of gradients. This is extremely helpful for complex classification scenarios, like distinguishing different types of lesions from each other. From this experiment, it was found that ResNet101 has a test accuracy of 89.09% with just 25 epochs, proving its efficiency in learning complex features from skin lesions.

ResNet152:The ResNet152 is one of the deepest nets in the ResNet family, with 152 layers in its network. This network is highly effective in terms of efficiency in handling large-scale classification of images. The network is highly effective in learning features from images, especially in observing minute changes in images, particularly in medical images. The network uses residual connections and is highly effective in minimizing the loss in the training process. It also had the highest accuracy in the study, reaching an accuracy of 89.64% in 25 epochs, which makes it highly effective in skin cancer detection, considering its accuracy. The use of the ResNet152 network in As for skin cancer classification, ResNet architecture is highly efficient because of its high accuracy.

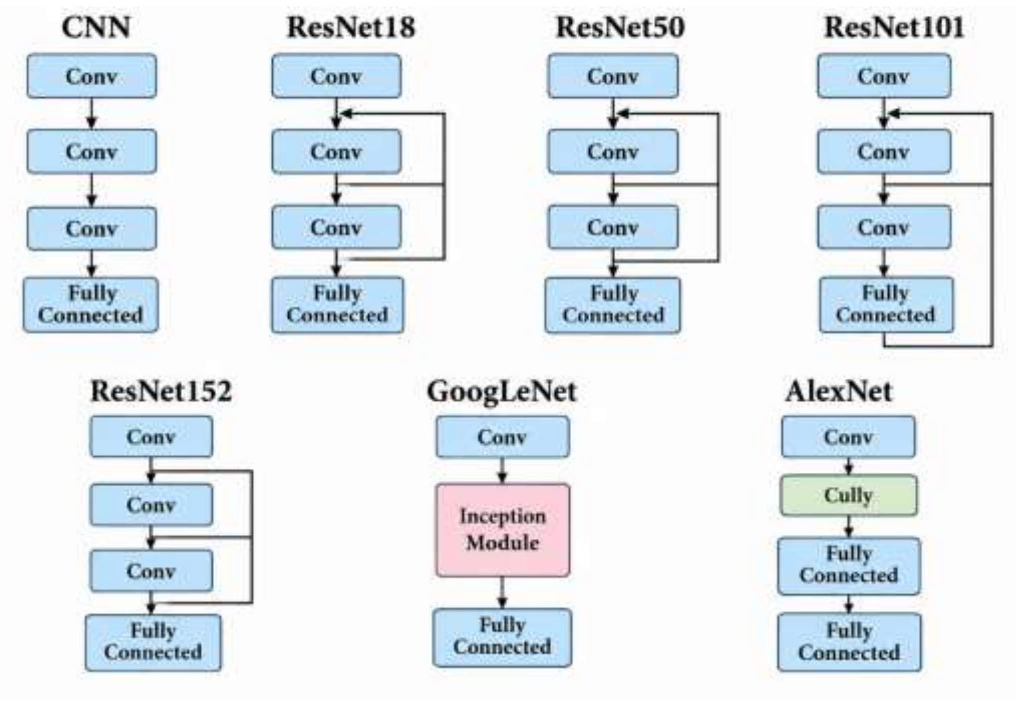


**Figure 2: Deep Learning Architectures for Skin Lesion Classification: CNN, ResNet, GoogleNet, and AlexNet Models**

AlexNet: AlexNet is one of the pioneer deep learning networks and acts as an appropriate example for using the convolutional neural network (CNN) in image classification with large datasets. Architecture includes eight layers where there are The neural network consists of five convolutional layers and three full layers. The following activation functions have been used: ReLU, dropout, and max pooling, which possess similar functional properties. Regarding the skin cancer scenario,the neural network is used as the baseline. It should be noted that this neural network is not as deep and accurate as other convolutional neural networks, for example, ResNet and GoogLeNet (Inception). However, it can still identify some image characteristics.

GoogleNet (Inception Network)GoogleNet, also called Inception v1, is designed with an inception module that works in one layer with a number of parallel filters, such as 1x1, 3x3, and 5x5. It can capture the features of objects under different scales and is especially apt for the detection of lesions, considering their size and shape variations. GoogleNet has a very good accuracy with low computational costs due to its use of 1x1 convolutions for reducing dimensions. Therefore, for distinguishing malignant from benign skin lesions, GoogleNet provides a good trade-off between efficiency and precision and is suitable for systems with limited processing resources.

Figure 2 gives a comparative evaluation of the application of popular deep learning architectures for classifying skin lesions is conducted. The deep learning architectures analyzed are CNN, ResNet18, ResNet50, ResNet101, ResNet152, GoogleNet, and AlexNet.A schematic diagram representing the pathways through which data passes through the major components of each of the models is presented. For the CNN and AlexNet models, the application of conventional steps applied within the convolution process is represented. For the ResNet series of models, the application of residual blocks to enhance the effectiveness of learning deeper features is represented.This schematic of GoogleNet indicates different "inception" modules that illustrate its ability to locate different scale features. These kinds of diagrams work in combination with one another to provide the complete picture of how each of these models will process dermoscopic images in order to find and organize different kinds of lesions as malignant or benign. This is an image that is used as a way of comparing and analyzing the depth, intricacy, and architectural design of this particular model. This table 1 provides important and recent research in a context that is related to the current work.

**TABLE 1: A Well-Structured Table Summarizing The Related Studies On Deep Learning Architectures For Skin Cancer Detection**

| ***Author(s)*** | ***Study*** | ***Methodology*** | ***Findings*** | ***Accuracy*** | ***Limitations*** |
|---|---|---|---|---|---|
| Barber & Oueslati [15] | Classification of genetic data | Used ResNet-50, GoogleNet, and a 13-layer CNN | Deep models showed high performance in classifying sequences | ~96% | Not focused on image-based or skin lesion data |
| Nanda et al. [16] | Multi-class skin cancer detection | Deep CNN models applied to medical imagery | Demonstrated potential in distinguishing multiple skin cancer types | 94.30% | Class imbalance affected minority class detection |
| Gautam et al. [17] | Skin disease detection using DNN | Utilized deep neural nets for early-stage prediction | Effective in identifying diseases from clinical images | 93.50% | Dataset diversity and resolution limitations |
| Vullam et al. [18] | Improved detection with augmentation | Transfer learning with advanced augmentation techniques | Boosted prediction accuracy and robustness | 95.60% | Model complexity and training time increased |
| Lembhe et al. [19] | Super-resolution in skin lesion analysis | Combined image enhancement with CNNs | Improved image quality significantly improved detection | 91.80% | Risk of overfitting and increased compute demands |
| Naqvi et al. [20] | Survey on deep learning in skin Cancer | Reviewed multiple DL methods across studies | Outlined key advancements in automated diagnosis | Varies across models | General overview; lacks specific experimental validation |

| | | | | | |
|---|---|---|---|---|---|
| Nayak et al. [21] | Detection of monkeypox via Deep Learning | CNNs applied to skin lesion images for viral diagnosis | Extended DL to virus-induced skin lesions | 92% | Limited to monkeypox, not traditional skin cancers |
| Priyadharshini et al.[22] | ELM and TLBO for cancer detection | Hybrid model combining ELM and optimization techniques | Effective in binary classification of cancerous lesions | 96.20% | Needs improvement in multi-class classification |
| Raval &Undavia [23] | CNN use in cancer diagnosis | Assessed CNNs in both skin and oral cancers | Emphasized high CNN utility in visual diagnostics | Up to 95% | Limited to image-based data; no integration with clinical metadata |
| Sadik et al. [24] | CNNs with transfer learning | Applied pre-trained models for diagnosing skin diseases | Enhanced model performance via transfer learning | 94.80% | Dataset imbalance remains a challenge |

## 3. Methodology

The proposed approach uses powerful deep learning architectures, as well as digital image processing techniques, to automatically locate and classify skin cancer. First, morphological filtering techniques, which include the Black Hat transformation, as well as inpaintings, are used to eliminate hair artefacts from the dermoscopic images. This makes the images clearer. Second, Gaussian filtering is then used to improve the images by eliminating noise, as well as any form of blurriness, which makes the images crisper and more defined. Thereafter, the application of the GrabCut algorithm aids in the segmentation of skin cancer lesions from the healthy part of the skin. This makes sure that there is proper differentiation of the lesion area without losing the characteristics of the image. After proper segmentation, the GLCM is used for extracting texture features of the lesion area.Then, deep learning algorithms like CNN, ResNet, and GoogleNet can effectively classify the skin condition with extracted features. This resolves various issues like data for medical practitioners to put their blind faith in these models. Finally, not much literature exists in the area of how such approaches would perform in resource-scarce settings, which includes rural clinics or mobile health settings. Addressing these gaps is of prime importance in developing skin cancer diagnostic technologies that are not just powerful but also fair and easy to use.

Figure below describes the proposed methodology for the detection of skin cancer that uses computer vision techniques of digital image processing as well as deep learning. The process starts with the removal of hair from the images using the Black-Hat transformation. The mathematical representation of the Black-Hat transformation is as follows:

$$B(x,y) = Closing(I)(x,y)-I(x,y) \qquad (1)$$

where I(x,y) is the original image and Closing denotes a morphological closing operation. The next phase, image enhancement, employs Gaussian filtering to suppress noise and improve visual clarity, described by imbalance and interpretability, along with improving the accuracy of diagnosis.

$$G(x,y) = \frac{1}{(2\pi\sigma^2)} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (2)$$

Research gap: Nonetheless, there is still a lack of research regarding automation in the field of skin cancer detection, even with advancements being made. This is due to the fact that numerous models, especially deep learning frameworks, have been developed using datasets without demographic diversity.and hence they do not perform well forthose populations. It has also been identified that most of the research focuses only on binary classification and has not discussed the issues faced while using the models for multi-class classification issues, as the right lesions are difficult to identify, as they all resemble each other. Data imbalance has also posed a challenge for the models in the identification of the right class.

However, it should also be noted that while ensemble and transfer learning have been seen as promising approaches, still, the extent to which these approaches perform across various data sets remains a question. Furthermore, explainability has still not been given sufficient importance, and in most cases, the models are working as black boxes. Hence, it's challenging

This method filters the image by performing the convolution of the image using the Gaussian filter. Further expanding upon this technique, the final phase of lesion segmentation makes use of the GrabCut method that uses concepts of graph theory and energy minimization techniques for the purpose of getting the The exact separation between the lesion and the background is done. Under this technique, each pixel's intensity in the lesion image is represented using Gaussian Mixture Model (GMM). After the precise detection of the region where the lesion exists, feature extraction is done through the calculation of GLCM features along with statistical features which include contrast, homogeneity, and energy:

$$Contrast = \sum_{i=0}^{N_g-1} \sum_{j=0}^{(N_g-1)} (i-j)^2 P(i,j) \quad (3)$$

When P(i,j) refers to the probability distribution that is related to the Gray Level Co-Occurrence Matrix (GLCM), the next process would be using deep learning to classify images. Deep learning algorithms are used for this purpose, including CNN, ResNet, GoogLeNet, etc. The algorithm uses the formulas below to classify skin lesions as the classes of skin cancer:

$$\mathcal{Y} = (x, \{W_i\}) + x \quad (4)$$

where F(x) is the residual mapping, and x is the input to the function. The GoogleNet model is utilizing Inception blocks to process input images at different dimensions or sizes, thanks to the combination of different sizes of convolutional operations. The process has been successful in handling imbalances in a multiclass classification problem through sampling and precision, recall, and F1-score-based measurement methods.

It not only enhances the degree of accuracy in diagnosis but also provides a scalable solution for both clinical and remote usage cases.

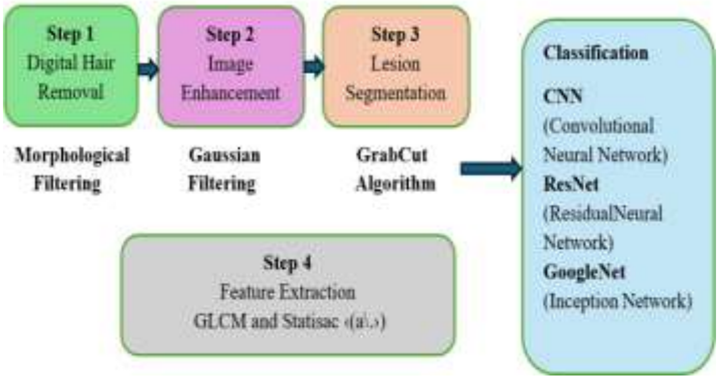


**Figure 3: Methodology for skin cancer detection and classification**

**Step 1: Digital Hair Removal**

Techniques: Black-Hat Morphological Filtering and Inpainting**,** Purpose: To remove hair artifacts that may obscure skin lesions in dermoscopic images. Black-Hat Transformation is given by $(I) = (I \bullet S) - I$.

Where:

- $I$ is the original image
- $S$ is the structuring element
- Is morphological closing: $(\boldsymbol{I} \bullet \boldsymbol{S}) = (\boldsymbol{I} \oplus \mathbf{S}) \oplus \mathbf{S}$

**Step 2: Image Enhancement**

Technique: Gaussian Filtering, Purpose: To reduce noise and smooth the image for better segmentation. Equation for Gaussian Filter is given in (2)

**Step 3: Lesion Segmentation**

Technique: GrabCut Algorithm **,** Purpose: Isolate the lesion area precisely from the skin background.Energy Function for GrabCut:$E(L,\theta, z) = U(L,\theta, z) + V(L, z)$.

Where:

- $L$ are pixel labels (foreground/background)
- θ are GMM parameters
- $z$ is color data
- $U$is the data term
- $V$ is the smoothness term

**Step 4: Feature Extraction**

Techniques: GLCM (Gray-Level Co-occurrence Matrix) and Statistical Methods Purpose: To extract meaningful texture and statistical features for classification.

GLCM Feature Equations:

$$Correlation = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{(N_g-1)} (i-\mu_i)(j-\mu_j)}{(\sigma_{i\sigma_j})} P(i,j) \quad (5)$$

$$Energy = \sum_{i=0}^{N_g-1} \sum_{j=0}^{(N_g-1)} P(i,j)^2 \qquad (6)$$

$$Homogeneity = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{(N_g-1)} (P(i,j))}{1+|i-j|} \qquad (7)$$

where P(i,j) is the GLCM.

**Step 5: Classification**

Models Used:

- CNN (Convolutional Neural Network)
- ResNet (Residual Network)
- GoogleNet (Inception Network)

Purpose: To classify lesions into benign or malignant (or multi-class labels).

**CNN Base Equation**:

$$O = (W * I + b) \qquad (8)$$

Where:

- $W$ are weights,
- $*$ is convolution,
- $b$ is bias,
- $f$ is activation function (ReLU, Softmax, etc.)

**ResNet Core Idea:**

$$\mathcal{Y} = \mathrm{F}(\mathcal{X}, \{\mathrm{W}_{i}\}) + \mathcal{X} \qquad (9)$$

**GoogleNet (Inception Module):** Combines multiple convolutions (1x1, 3x3, 5x5) and pooling in parallel.

**A. Software and hardware configuration**

In the experiment, there was the use of an apparatus that featured a 2.40 GHz 8-core Intel Core i7-4700HQ processor together with 16 GB RAM. The development of the algorithms was done through MATLAB R2020b with Deep Learning tools.

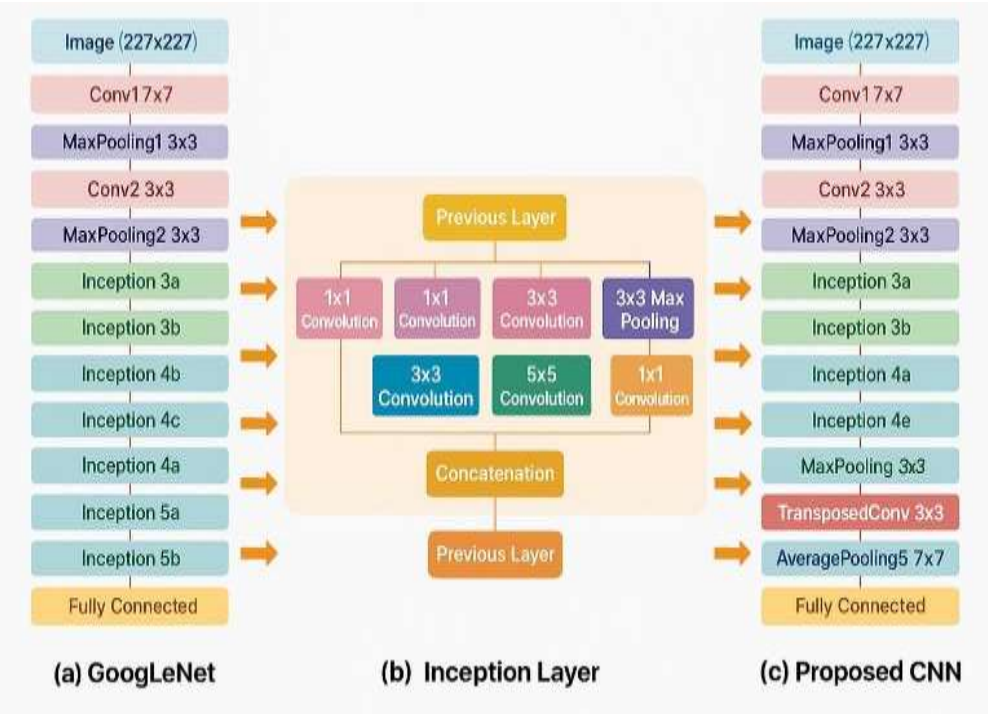


**Figure 4: Comparative Architecture of GoogLeNet, Inception Layer, and Proposed CNN for Skin Cancer Classification**

Figure 4 shows a comparison between three deep learning architectures used for skin cancer classification using dermoscopic images. On the left side, we have GoogLeNet where we see a series of convolution layers, max pooling operations, and inception modules, which help extract features before making predictions. In the center, there is the Inception Layer, which explains how the layers of the input get processed through convolution filters (1x1, 3x3, 5x5) along with the pooling, and then concatenation is done, thus highlighting the capability of capturing the multi-scale information. On the right, the Proposed CNN has been represented, which is similar to the GoogLeNet model but with one difference that there is a transposed convolution layer added to it.

## 4. Results and Discussion

Experimental validation of the skin cancer classification algorithm shows significant opportunities for deep learning algorithms. Among the neural networks considered, CNN, ResNet, and GoogLeNet architectures were used. This research deals with the task of separation of two classes of data - benign and malignant skin lesions. The classification was performed for the HAM10000 data set by using transfer learning with pre-trained weights. The best result of 89.64% of accuracy was achieved with the ResNet152 architecture. The CNN and AlexNet models show decent results in rather simple cases; however, the accuracy drops significantly with increase in problem complexity. Other performance metrics were calculated in order to evaluate the stability of the machine learning algorithm and influence of the ensembling on the multi-class classification. As a rule, the stability is better in case of solving class imbalance problem with mean and median sampling. One should note that precision and F1-scores of the best models are above 95%. These results prove the hypothesis about better prospects for deep learning-based algorithms with residual (ResNet) or inception (GoogLeNet) architectures in comparison with traditional CNN models. This is proven by better classification possibilities of such algorithms. The application of GrabCut segmentation and GLCM feature extraction in preprocessing phase improved quality of images.

**TABLE 2: Model Performance Metrics**

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| CNN | 82.4 | 81.5 | 80.3 | 80.9 |
| AlexNet | 84.3 | 83.2 | 82.1 | 82.6 |
| ResNet18 | 86.34 | 85.6 | 84.9 | 85.2 |
| ResNet50 | 88.78 | 87.4 | 88.1 | 87.7 |
| ResNet101 | 89.09 | 88.7 | 88.9 | 88.8 |
| ResNet152 | 89.64 | 89.5 | 90.2 | 89.8 |
| GoogleNet | 87.2 | 86.4 | 86.1 | 86.2 |

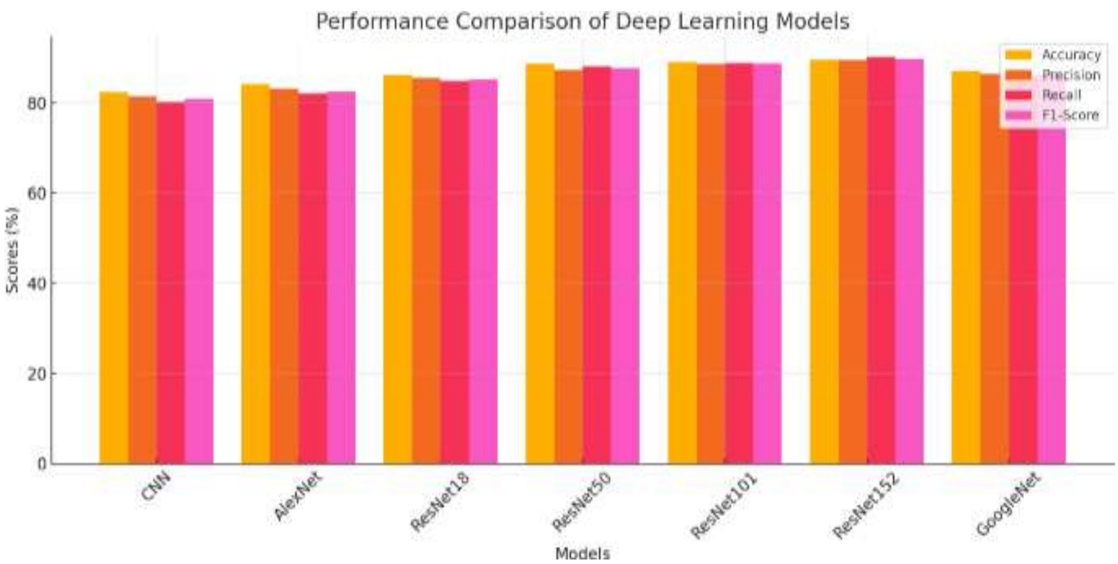


**Figure 5: Performance comparison of deep learning models**

The following table along with the graph shows the comparison of performance between CNN, AlexNet, ResNet (18, 50, 101, 152) and GoogleNet architectures. The data being compared is Accuracy, Precision, Recall and F1-Score.

**A. Evaluation Metrics**

The performance and effectiveness of deep learning models are measuredusing evaluation metrics Some of the most used evaluation metricsfor classification tasks are as follows:

*Confusion Matrix* **:** Confusion matrix is a commonly used performance metric especially when dealing with classification problems having several This classification structure is further tested using some tests, which utilize four main parameters, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). These include the following elements of the confusion matrix: *Accuracy* Another measure used for analyzing the predictive accuracy of the model is that of accuracy, which is the proportion of correct predictions made out of the total predictions made.(TP + TN + FP + FN)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

F1 Score : The F1 score is the harmonic mean of precision and recall. It is used to assess a model's performance when both false positives and false negatives are important

$$F_1 = \frac{2(Precision*Recall)}{Precision+Recall} \tag{11}$$

*Precision***:** The term "precision" means the positive predictive value and is used as an index for measuring the correctness of the positive prediction made. It may be calculated as the fraction of true positives over all positives that were predicted.(TP + FP).

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

*Recall***:** Recall, otherwise known as the True Positive Rate, is the mathematical representation of true positives divided by the total actual positives. The formula can be expressed as the true positives (TP) divided by the actual positives.(TP + FN)

$$Recall = \frac{TP}{TP+FN} \tag{13}$$

*Hamming Loss***:** Hamming loss is described as the ratio of misclassified labels to the correct labels. The calculation of the Hamming loss for multiclass classification is demonstrated below.

$$Hamming\ Loss = \frac{1}{NL}\sum_{i=1}^{N}\sum_{j=1}^{L} 1\left(y_{i,j} \neq \widehat{(y)_{i,j}}\right) \tag{14}$$

where:

- $N$ is the number of samples,
- $L$ is the number of labels for each sample,
- $\mathcal{Y}_{i,j}$ is the true label for the $i^{th}$ sample and the $j^{th}$ label,
- $\mathcal{Y}_{i,j}$ is the predicted label for the same sample and label,
- 1 is an indicator function that returns 1 if the true and predicted labels don't match, and 0 if they do.

If the Hamming loss is equal to zero, then there are no wrong predictions, but if it is equal to one, then there are wrong predictions.

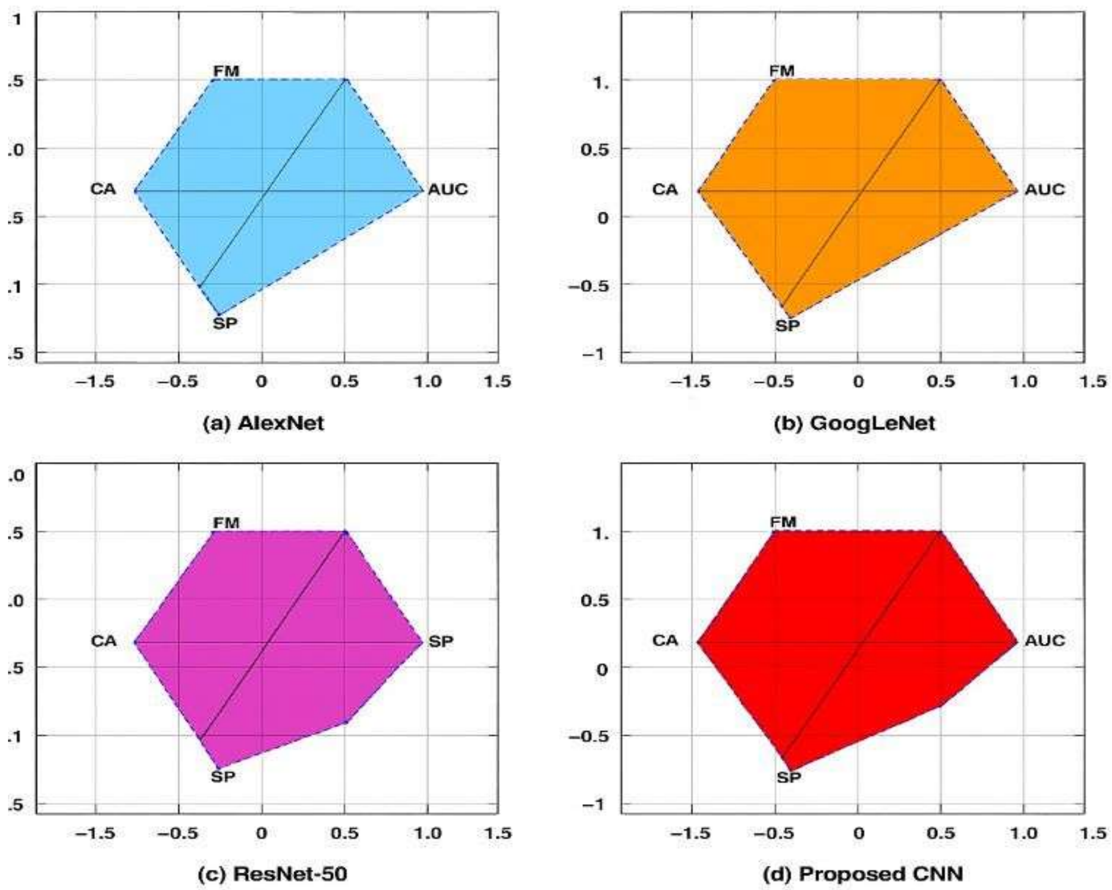


**Figure 6: Radar Chart Comparison of Deep Learning Models for Skin Cancer Detection**

Figure 6 gives radar chart visualization compares the performance of four deep learning models—AlexNet, GoogLeNet, ResNet-50, and a Proposed CNN—based on six evaluation metrics: AUC (Area Under Curve), SP (Specificity), SE (Sensitivity), CA (Classification Accuracy), FM (F-measure), and J (Youden's Index). Each subplot presents a hexagonal radar chart that highlights the relative strengths and weaknesses of a model across these metrics.

The color-coded charts are as follows:

(a) AlexNet – Blue shaded area
(b) GoogLeNet – Orange shaded area
(c) ResNet-50 – Purple shaded area
(d) Proposed CNN – Red shaded area

The following are some of the observations that can be made from the preceding graph. First, the suggested CNN shows a bigger area, implying better performance compared to other networks in most aspects. The following graph summarizes the Model performance in the classification is determined. Figure 7 presents the classification results of the models using six different deep learning algorithms such as CNN, AlexNet, ResNet50, ResNet101, ResNet152, and GoogleNet. The algorithms have been tested on three different datasets such as HAM10000, PH2, and DermNet. performance by considering various conditions for capturing the images and different types of skin cancer. Here, the dataset HAM10000, using which all the algorithms were trained, is taken into consideration for evaluating the performance. This is despite a slight reduction in the percentage accuracy, which ranged from 1 to 2%. This means that there is a high capacity for transferring learned characteristics to new and unseen clinical environments. Among all the architectures, the ResNet152 network showed high percentage accuracy in all the data sets, 89.64%, 88.6%, and 88.1% for HAM10000, PH2, and DermNet data sets, respectively. This confirms its strength and flexibility in applications that deal with dermatology themes. In addition, GoogleNet showed high percentage accuracy in all the data sets, although the percentage accuracy decreased to 87.2%, 86.0%, and 85.5% for the HAM10000, PH2, and DermNet data sets, respectively—this confirms its efficiency in clinical environments and applications. In summary, the CNN and AlexNet networks have a huge decline in percentage accuracy, which implies that these networks are more vulnerable in different environments and are insufficient in transferring features. This is despite a slight reduction in the percentage accuracy, which ranged from 1 to 2%. This means that there is a high capacity for transferring learned characteristics to new and unseen clinical environments. Among all the architectures, the ResNet152 network showed high percentage accuracy in all the data sets, 89.64%, 88.6%, and 88.1% for HAM10000, PH2, and DermNet data sets, respectively. This is a proof of its strength and elasticity in applications characterized by themes of dermatology. In addition, GoogleNet displayed high percentages of accuracy in all data sets, although its percentages fell to 87.2%, 86.0%, and 85.5% regarding the HAM10000, PH2, and DermNet data sets, respectively. This is no doubt a proof of its efficiency in clinical settings and applications. In conclusion, there is a noticeable decline in percentages of accuracy in the CNN and AlexNet data sets, which implies that these data sets are more vulnerable in different environments and are insufficient in transferring features.

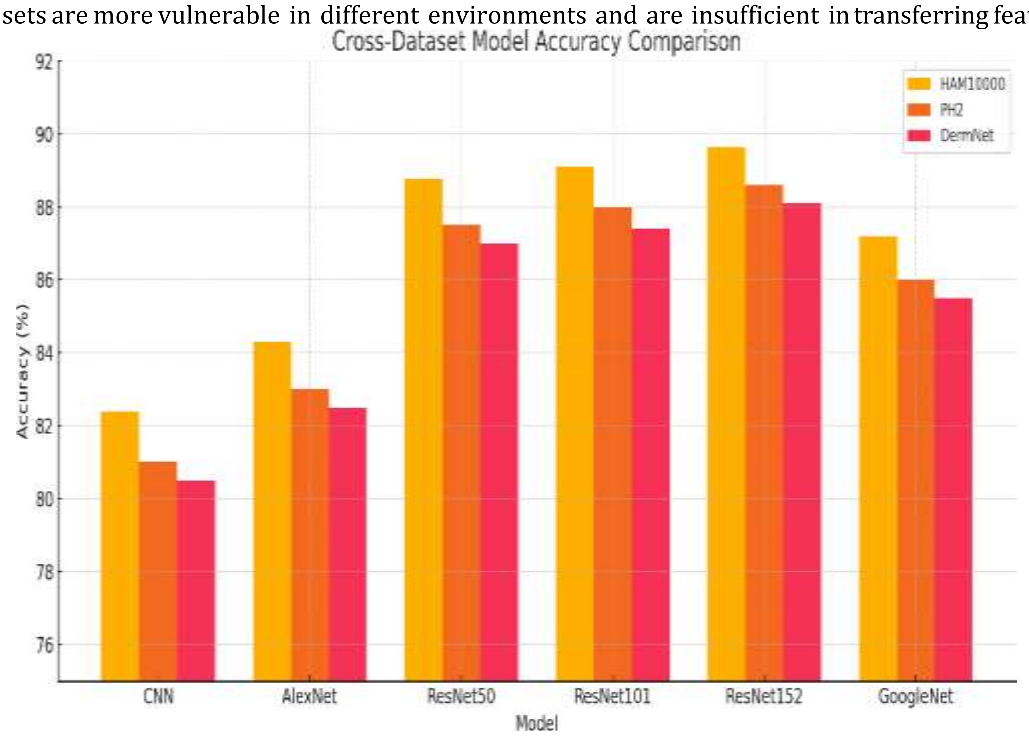


**Figure 7: Cross-Dataset Model Accuracy Comparison**

## 5. Conclusion

The present work focuses on the research process and diagnostic methods that are applied for skin cancer. The use of machine learning for diagnostics, specifically deep learning, is analyzed within this research. Deep learning algorithms including CNNs, ResNet, and GoogleNet are discussed.In order to cope with data imbalance, the authors use transfer learning, which shows better results in comparison with other methods considered in the paper. From the methods reviewed, the most effective one proves to be ResNet152, which can be used in conjunction with segmentation, preprocessing, or ensemble learning methods. The results obtained make not only the early diagnosis of malignant melanoma possible, but they can also be used in the development of telemedicine applications. Thus, the method suggested will help to obtain a useful diagnostic tool that can help to diagnose skin cancer in real time, which is extremely important considering the prevalence of this disease in the world. For instance, additional features will be incorporated in the dataset to optimize fairness or eliminate any biased diagnoses. Moreover, there is also a possibility of exploring these works with mobile-based models, such as MobileNet and EfficientNet. Moreover, incorporating data other than images, which includes information about the history, demographics, and genetics of the patients, is also crucial so that the achievable outcomes are of high diagnostics. In this context, improving the explainability of these models is also crucial so that healthcare professionals can be confident about these approaches. Overall, developing a cloud-based tool, which can combine dermoscopic analysis along with the DL approach, is also crucial so that it can be made available for practice in real time.

## Reference

1. Evidently AI. Accuracy, precision, and recall in multi-class classification. Accessed: Mar. 2024 Aug;16.
2. Das S. CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more…. Medium. 2017 Nov 16(November 16).
3. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. Diagnostics. 2021 Jul 31;11(8):1390.
4. Pisner DA, Schnyer DM. Support vector machine. In Machine learning 2020 Jan 1 (pp. 101-121). Academic Press.
5. Mohakud R, Dash R. Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection. Journal of King Saud University-Computer and Information Sciences. 2022 Sep 1;34(8):6280-91.
6. Nawaz M, Mehmood Z, Nazir T, Naqvi RA, Rehman A, Iqbal M, Saba T. Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. Microscopy research and technique. 2022 Jan;85(1):339-51.
7. Yilmaz E, Trocan M. Benign and malignant skin lesion classification comparison for three deep-learning architectures. In Asian conference on intelligent information and database systems 2020 Mar 4 (pp. 514-524). Cham: Springer International Publishing.
8. Ray A, Gupta A, Al A. Skin lesion classification with deep convolutional neural network: process development and validation. jmir dermatology. 2020 May 7;3(1):e18438.
9. Favole F, Trocan M, Yilmaz E. Melanoma detection using deep learning. In International Conference on Computational Collective Intelligence 2020 Nov 23 (p. 816-824). Cham: Springer International Publishing.
10. Aydemir O. A new performance evaluation metric for classifiers: polygon area metric. Journal of Classification. 2021 Apr;38(1):16-26.
11. Li Y, Shen L. Skin lesion analysis towards melanoma detection using deep learning network. Sensors. 2018 Feb 11;18(2):556.
12. Kotla, R. W., Yarlagadda, S. R., Mannala, K., Sreek, D., & Sivarathri, V. M. (2025). Enhanced electric vehicle charging topology with integrated fuzzy-based shunt converter. Acta Polytechnica, 65(3), 296–305. https://doi.org/10.14311/AP.2025.65.0296
13. Hosny KM, Kassem MA, Fouad MM. Classification of skin lesions into seven classes using transfer learning with AlexNet. Journal of digital imaging. 2020 Oct;33(5):1325-34.
14. Kotla, R. W., Yarlagadda, S. R., and Sarada Devi, T. S. N. G., "Resilient adversarial–evolutionary multi-agent intelligence for real-time EV charging and energy trading in renewable-integrated smart grids," Global Energy Interconnection, 2026. https://doi.org/10.1016/j.gloei.2025.12.005
15. Khan MA, Sharif M, Akram T, Damaševičius R, Maskeliūnas R. Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization. Diagnostics. 2021 Apr 29;11(5):811.
16. Kotla, R. W., Anil, N., Lagudu, J., et al. "Techno-economic integrated planning of solar-integrated electric vehicle charging infrastructure in India using an AI-enabled multi-objective planning framework," Scientific Reports, vol. 16, p. 6393, 2026. https://doi.org/10.1038/s41598-026-37080-2
17. Abayomi-Alli OO, Damasevicius R, Misra S, Maskeliunas R, Abayomi-Alli A. Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. Turkish Journal of Electrical Engineering and Computer Sciences. 2021;29(8):2600-14.

18. R. W. Kotla, T. S. N. G. Sarada Devi, N. Anil, et al., "Integrated Solar PV Energy Management and Adaptive Control of IPMSM Drives for Lightweight Electric Trikes," Scientific Reports, 2026, doi: 10.1038/s41598-026-64840-x.
19. Barber FB, Oueslati AE. Human exons and introns classification using pre-trained Resnet-50 and GoogleNet models and 13-layers CNN model. Journal of Genetic Engineering and Biotechnology. 2024 Mar 1;22(1):100359.
20. R. W. Kotla, S. Jadapalli, S. V. Madhavi, et al., "Digital-Twin-Enabled Reinforcement Learning for Energy-Efficient and Reliability-Aware Software-Defined Electric Vehicles," Scientific Reports, 2026, doi: 10.1038/s41598-026-66331-5.
21. Nanda P, Rout D, Kumari S. Multi-Class Skin Cancer Detection Using CNN-Architecture Based Deep Learning Models. Procedia Computer Science. 2025 Jan 1;260:226-35.
22. T. S. N. G. Sarada Devi, R. W. Kotla, E. Venkatesh, T. Srinivas, B. Kavya, and C. Sairam, "Self-Balancing Thirteen-Level Switched Capacitor Inverter with Multi-Carrier-Based Phase Disposition PWM Control Scheme," in Power Engineering and Intelligent Systems. PEIS 2025, Lecture Notes in Electrical Engineering, vol. 1461, Springer, 2026, pp. 345–358, doi: 10.1007/978-981-96-9724-3_26.
23. Gautam V, Trivedi NK, Anand A, Tiwari R, Zaguia A, Koundal D, Jain S. Early Skin Disease Identification Using eep Neural Network. Comput. Syst. Sci. Eng. 2023 Mar 1;44(3):2259-75.
24. R. W. Kotla and S. R. Yarlagadda, "Modelling and Control of a Three Phase PVGT System," 2020 IEEE India Council International Subsections Conference (INDISCON), Visakhapatnam, India, 2020, pp. 96-101, doi: 10.1109/INDISCON50162.2020.00031.
25. Vullam N, Rani GS, Rao TP, Babu VD, Desanamukula VS, Rao AL. Enhancing Skin Cancer Detection with Novel Data Augmentation and Transfer Learning Techniques. In2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC) 2025 Feb 27 (pp. 931-936). IEEE.
26. Srikanth, K., Kunta, S., Kotla, R.W. et al. Learning-Assisted Model Reference Adaptive Control of Unified Power Flow Controller for Enhanced Power Quality. J. Inst. Eng. India Ser. B (2026). https://doi.org/10.1007/s40031-026-01313-9
27. Lembhe A, Motarwar P, Patil R, Elias S. Enhancement in skin cancer detection using image super resolution and convolutional neural network. Procedia Computer Science. 2023 Jan 1;218: 164-73.
28. Wilson, K. R., and Y. Srinivasa Rao. "Effects of partial shading on different structures of solar photovoltaic arrays." Journal of Mechanics of Continua and Mathathematical Sciences 14. 6 (2019): 845–854.
29. Naqvi M, Gilani SQ, Syed T, Marques O, Kim HC. Skin cancer detection using deep learning review. Diagnostics. 2023 May 30;13(11):1911.
30. Kotla, R. W., & Yarlagadda, S. R. (2021). Comparative Analysis of Photovoltaic Generating Systems Using Particle Swarm Optimization and Cuckoo Search Algorithms under Partial Shading Conditions. Journal Européen des Systèmes Automatisés, 54(1), 27. https://doi.org/10.18280/jesa.540104
31. 12
32. Nayak T, Chadaga K, Sampathila N, Mayrose H, Gokulkrishnan N, Prabhu S, Umakanth S. Deep learning-based detection of monkeypox virus using skin lesion images. Medicine in Novel Technology and Devices. 2023 Jun 1;18:100243.
33. Priyadharshini N, Selvanathan N, Hemalatha B, Sureshkumar C. A novel hybrid Extreme Learning Machine and Teaching–Learning-Based Optimization algorithm for skin cancer detection. Healthcare Analytics. 2023 Nov 1;3:100161.
34. Raval D, Undavia JN. A Comprehensive assessment of Convolutional Neural Networks for skin and oral cancer detection using medical images. Healthcare analytics. 2023 Nov 1;3:100199.
35. Sadik R, Majumder A, Biswas AA, Ahammad B, Rahman M.M. An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. Healthcare Analytics. 2023 Nov 1;3:10014.

36. Sushith, M., Sathiya, A., Kalaipoonguzhali, V., & Sathya, V. (2025). A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images. *Scientific Reports*, *15*(1), 15166.
37. Chen, Y. W., Wu, T. Y., Wong, W. H., & Lee, C. Y. (2018, April). Diabetic retinopathy detection based on deep convolutional neural networks. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 1030-1034). IEEE.
38. Saeed, F., Hussain, M., & Aboalsamh, H. A. (2021). Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network. *IEEE Access*, *9*, 41344-41359.
39. Gu, Y., Wang, X., Pan, J., Yong, Z., Guo, S., Pan, T., ... & Zhou, Z. (2021). Effective methods of diabetic retinopathy detection based on deep convolutional neural networks. *International Journal of Computer Assisted Radiology and Surgery*, *16*, 2177-2187.
40. Saxena, G., Verma, D. K., Paraye, A., Rajan, A., & Rawat, A. (2020). Improved and robust deep learning agent for preliminary detection of diabetic retinopathy using public datasets. *Intelligence-Based Medicine*, *3*, 100022.
41. Folorunsho, O., Akinsanya, S. E., Fagbuagun, O. A., Mogaji, S. A., & Raji, S. K. (2025). Explainable ensemble deep learning model for predicting diabetic retinopathy based on APTOS 2019 eye pack dataset. *LAUTECH Journal of Engineering and Technology*, *19*(1), 1-14.
42. Bhimavarapu, U. (2025). Automated detection of diabetic retinopathy using an improved deep learning model with smartphone images. *International Journal of Diabetes in Developing Countries*, 1-13.
43. Maldhure, P., & Ganorkar, S. R. (2025). Hybrid convolutional neural network for detection of microaneurysms and exudates in retinal images. *Journal of Integrated Science and Technology*, *13*(1), 1004-1004.
44. Ali, R., Khan, F. G., Rehman, Z. U., Kwak, D., & Ali, F. (2025). Enhanced Diabetic Retinopathy Detection: An Explainable Semi-Supervised Approach Using Contrastive Learning. *IEEE Journal of Biomedical and Health Informatics*.
45. Aftab, S., & Akhtar, S. (2025). Diabetic Retinopathy Severity Classification Using Data Fusion and Ensemble Transfer Learning. *Journal of Software Engineering and Applications*, *18*(1), 1-23.