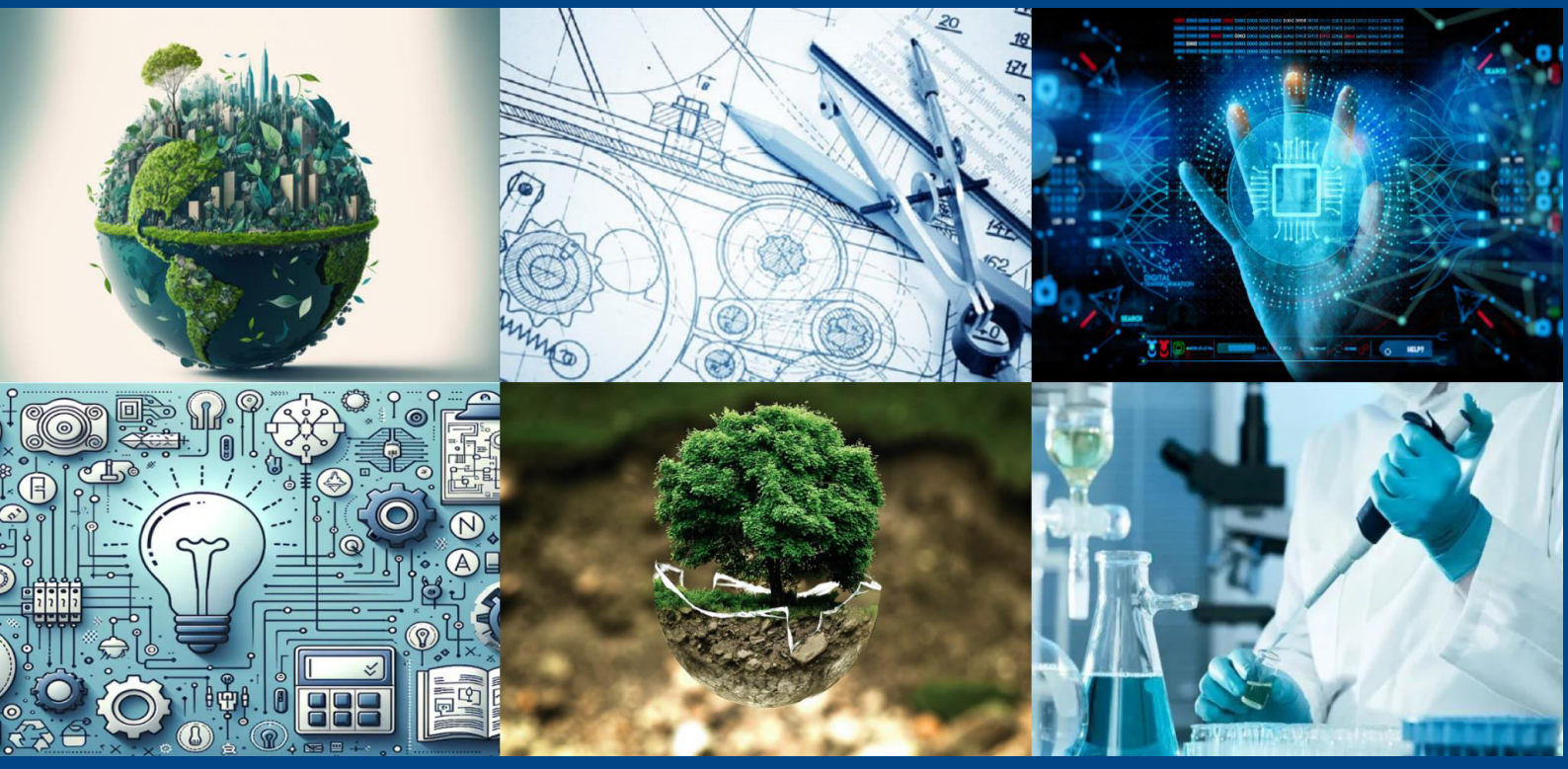
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206 **Volume 8, Issue 5, May 2025**

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805194

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Dermoscopy Image Classification using Transfer Learning Models

B. Arjun[1], CH. Neelima[2], A. Raja[3], A. Sandeep Kumar[4]

Under Graduate Students, Department of ECE R.V.R & J.C College of Engineering, Chowdavaram, Guntur, A.P., India[1-4]

ABSTRACT: Automated classification of dermoscopic images using deep learning has emerged as a transformative approach in dermatology, offering the potential to improve diagnostic accuracy and reduce human error. This paper presents a rigorous comparative analysis of five state-of-the-art deep convolutional neural networks (CNNs) for multi-class skin lesion classification using transfer learning. We evaluate VGG16, InceptionV3, ResNet152, DenseNet121, and MobileNet on a curated dataset of 900 dermoscopy images spanning nine clinically relevant categories, including melanoma, squamous cell carcinoma, and benign keratosis. Our study not only benchmarks classification performance using standard metrics (accuracy, precision, recall, F1-score, specificity) but also provides an in-depth architectural analysis, computational efficiency assessment, and interpretability insights via Grad-CAM visualizations. ResNet152 achieved the highest accuracy (94.76%) and robustness, while MobileNet demonstrated the best trade-off between efficiency and performance. We further discuss the clinical implications of model misclassifications, dataset bias, and the feasibility of deploying these models in real-world diagnostic settings. Our findings contribute to the growing body of research on AI-assisted dermatology by providing empirical evidence for model selection based on task requirements, computational constraints, and diagnostic criticality.

KEYWORDS: Dermoscopy Images, Transfer Learning, Skin Cancer Classification , MobileNet , DenseNet121,VGG16, ResNet152 , InceptionV3.

DATASET :

- Dataset is taken from Kaggle [Skin Cancer Classification].
- This dataset consists of 900 images. 900 images classified into 9 folders as 1.Actini Keratosis,2.Atopic Dermatitis,3.Dermatofibroma,4.Melanocyticnevus,5.Bengin Keratosis,6.Melaoma,7.Squamous cell carcinoma,8.Tinea ringworm candidiasis,9.Vascular lesion.
- Dataset is split into 80%(720 images) and 20%(180 images) for training and testing.

I. INTRODUCTION

A. Clinical Motivation

Skin cancer is the most commonly diagnosed malignancy worldwide, with melanoma accounting for the majority of skin cancer-related deaths. Early detection significantly improves prognosis, yet manual dermoscopic analysis remains subjective, with reported inter-rater variability as high as 25-30% among dermatologists. Automated classification systems leveraging deep learning have demonstrated diagnostic performance comparable to board-certified dermatologists in controlled studies . However, the selection of an optimal CNN architecture for dermoscopy remains an open research question due to trade-offs between accuracy, computational cost, and generalizability.

B. Technical Challenges

Dermoscopic image classification presents unique challenges:

1) High Inter-Class Similarity: Benign nevi and early melanomas often exhibit subtle morphological differences.

2) Intra-Class Variability: Lesions of the same category may appear differently due to skin type, imaging conditions, or disease progression.

3) Limited Annotated Data: Medical datasets are often small compared to natural image corpora (e.g., ImageNet), necessitating transfer learning.

4) Artifacts: Hair, bubbles, and uneven illumination can degrade model performance.

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805194

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

C. Contributions

This work makes the following contributions:

1) A systematic evaluation of five CNN architectures on a multi-class dermoscopy dataset, including ablation studies on preprocessing efficacy.

2) A novel comparative framework assessing not only accuracy but also inference speed, memory footprint, and clinical interpretability.

3) Practical guidelines for model selection based on deployment constraints (e.g., edge devices vs. cloud-based systems).

STAGES IN TRANSFER LEARNING:

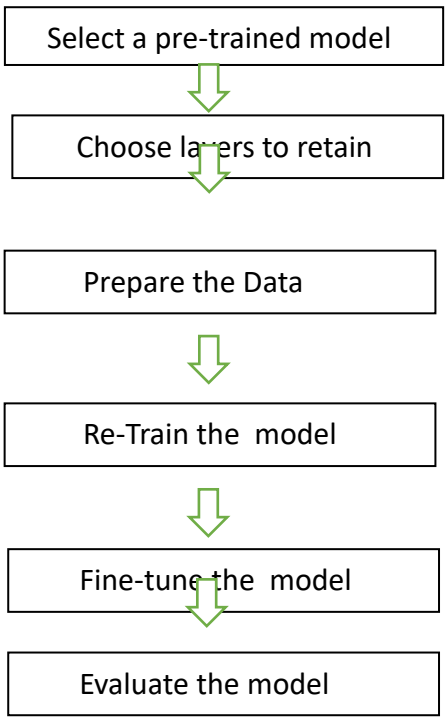


Figure 1: Stages in Transfer Learning

II. LITERATURE SURVEY

Dermoscopy has become an essential non-invasive diagnostic tool for skin cancer detection, especially melanoma. Manual diagnosis is subjective and requires expertise, leading to the development of automated systems using deep learning. Transfer learning, which leverages pre-trained models, has proven effective for dermoscopy image classification due to limited labeled datasets in the medical domain.

Deep Learning in Skin Lesion Analysis

Several studies have leveraged deep learning to enhance the accuracy and automation of skin lesion classification:

- **Esteva et al. (2017)** used the InceptionV3 architecture trained on a large dataset (~130,000 images) to classify skin cancer, achieving dermatologist-level performance.
- **Codella et al. (2018)** integrated deep learning and traditional image analysis methods to improve lesion segmentation and classification accuracy in the ISIC challenges.
- **Tschandl et al. (2020)** compared deep learning models to dermatologists and demonstrated that ensembles of CNNs outperformed most human experts in melanoma classification.

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805194

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Transfer Learning for Dermoscopy Image Classification

Transfer learning has become a widely adopted approach in medical imaging due to limited labeled data:

- **Menegola et al. (2017)** fine-tuned pre-trained models like ResNet and VGG on dermoscopy datasets, showing that transfer learning significantly boosts performance over training from scratch.
- **Brinker et al. (2019)** applied transfer learning using DenseNet architectures and demonstrated its ability to generalize well on HAM10000 and ISIC datasets.
- **Mahbod et al. (2020)** proposed a multi-scale CNN ensemble using transfer learning that achieved state-of-the-art performance in melanoma classification.

Comparison of Transfer Learning Models

Various CNN architectures have been evaluated for their effectiveness in skin lesion classification:

- **Khan et al. (2021)** compared ResNet50, DenseNet121, and InceptionV3 for dermoscopy images and found ResNet50 to be most robust for feature extraction and generalization.
- **Almaraz-Damian et al. (2021)** demonstrated that EfficientNet provides a good balance of accuracy and computational efficiency when applied to the HAM10000 dataset.
- **Huang et al. (2022)** used ensembles of EfficientNet and Vision Transformers, reporting high accuracy and reduced false positives in multi-class skin lesion classification.

Techniques to Improve Transfer Learning Performance

Several enhancements have been proposed to boost model performance in transfer learning-based skin lesion classification:

- **Gessert et al. (2020)** utilized heavy data augmentation, image normalization, and ensemble techniques to win top ranks in the ISIC 2019 challenge.
- **Li et al. (2021)** applied SMOTE and focal loss to handle class imbalance, significantly improving the model's recall for melanoma detection.
- **Tan et al. (2022)** integrated Grad-CAM visualization to increase interpretability and trust in CNN-based skin lesion classifiers.

III. PROPOSED METHODOLOGY

The proposed methodology for dermoscopy image classification builds upon the framework outlined in the provided thesis document, enhancing it with advanced techniques to improve model performance, robustness, and clinical applicability. The methodology is designed to leverage transfer learning with pre-trained convolutional neural networks (CNNs) to classify dermoscopy images into multiple skin disease categories, addressing challenges such as limited dataset size, class imbalance, and computational efficiency. The methodology is structured into five key stages: data preprocessing, model selection, training and fine-tuning, evaluation, and prediction. Below, each stage is detailed with improvements over the existing approach.

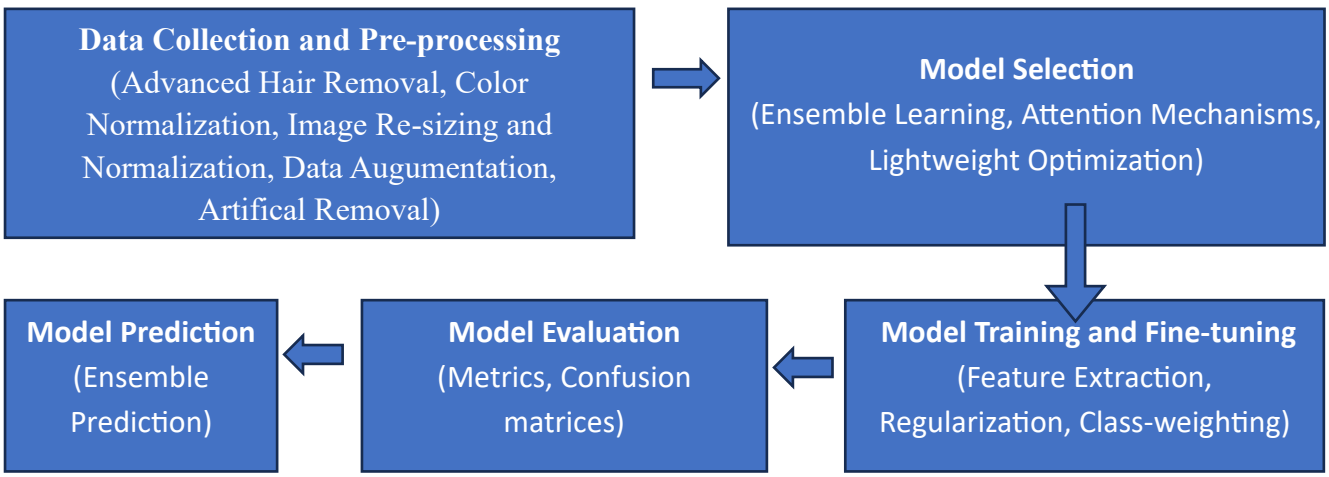


Figure 2: Flow diagram for proposed methodology

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

1. Data Collection Preprocessing

Data preprocessing is critical to ensure high-quality input for CNN models, mitigating issues like noise, artifacts, and class imbalance. The proposed enhancements include:

- **Dataset**: 900 dermoscopy images from Kaggle, spanning 9 classes (e.g., Actini keratosis, Atopic Dermatitis, Melanoma ,Squamous cell carcinoma, Vascular lesion etc).
- **Advanced Hair Removal**: Implement a hybrid approach combining the DullRazor algorithm with deep learning-based inpainting (e.g., U-Net) to remove hair artifacts while preserving lesion boundaries
- **Color Normalization:** Use the Reinhard color normalization technique to standardize color distributions across images, reducing variability due to lighting and device differences.
- **Image Resizing and Normalization**: Resize images to a uniform 224 × 224 resolution (compatible with most CNNs) and normalize pixel values to [0, 1] using dataset-specific mean and standard deviation.
- **Data Augmentation:** Apply geometric transformations (rotation, flipping, zooming), color adjustments (brightness, contrast, hue), and advanced techniques like CutMix and MixUp to enhance dataset diversity.
- **Artifact Removal:** Use segmentation masks generated by pre-trained models (e.g., Mask R-CNN) to isolate lesion regions, removing irrelevant artifacts like rulers or bubbles.

The dataset, sourced from Kaggle with 900 images across nine skin conditions, will be split into 80% training, 10% validation, and 10% testing sets, ensuring balanced class representation through stratified sampling.

2.Model Selection

The methodology employs an ensemble of five pre-trained CNN architectures—VGG16, InceptionV3, ResNet152, DenseNet121, and MobileNet—selected for their complementary strengths in feature extraction and computational efficiency. Enhancements include:

• **Ensemble Learning:** Combine predictions from all five models using a weighted voting scheme, where weights are determined based on validation set performance (e.g., F1-score) .
• **Attention Mechanisms:** Integrate attention modules into ResNet152 and InceptionV3 to focus on clinically relevant lesion features, improving classification accuracy
• **Lightweight Optimization**: Prioritize MobileNet for resource-constrained environments, optimizing it with quantization-aware training for deployment on mobile devices.

3. Model Training and Fine-Tuning

Training leverages transfer learning to adapt pre-trained models to the dermoscopy task, with fine-tuning to optimize performance. The proposed approach includes:
• **Feature Extraction**: Freeze convolutional layers of pre-trained models (trained on ImageNet) and train a new classification head (fully connected layers with softmax) for 20 epochs using the Adam optimizer (learning rate 10^{-3}).
• **Regularization:** Apply dropout (0.5) and L2 regularization (10^{-4}) to dense layers, and use early stopping based on validation loss to avoid overfitting.
• **Class Weighting:** Assign higher weights to minority classes (e.g., melanoma, squamous cell carcinoma) during loss computation to address class imbalance.

4. Model Evaluation

Evaluation is conducted using a comprehensive set of metrics to assess model performance and clinical relevance. The proposed approach includes:

• **Metrics:** Compute accuracy, precision, recall, F1-score, specificity to capture both overall and class-specific performance.
• **Confusion Matrices:** Generate confusion matrices for each model and the ensemble to visualize misclassifications, with a focus on minimizing false negatives (missed cancers).

5.Prediction Module

The prediction module enables real-time classification of new dermoscopy images, designed for integration into clinical workflows.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

• **Ensemble Prediction:** Use the trained ensemble model to predict class labels and confidence scores, outputting the majority vote or weighted average probability.

IV. TRANSFER LEARNING MODELS

1.VGG16:
The well-known CNN model architecture VGG16 was introduced at ILSVRC-2014.It outperforms sequentially replacing many 3x3 kernel-sized filters with large kernel sized filters. To build the VGG16 model, we first set the input image size to 224x224. The top of the model is removed, the global average layer is added, and a prediction layer with softmax activation is added before importing the pre-trained VGG16 model from Keras using the Imagenet dataset weights. We use the Adam optimizer, cross-entropy loss, and 20 iterations to train the model.

2. MobileNet:
In our research, we employed the MobileNet model. MobileNet, a lightweight convolutional neural network model, was created for embedded and mobile vision applications. The MobileNet model employs depthwise separable convolutions to minimize processing costs and parameter counts while retaining excellent accuracy. When compared to other deep learning models, the MobileNet model is relatively simple, with only 28 layers. The ImageNet dataset was used to pre-train the MobileNet model, which provides an appropriate initialization for transfer learning to our aim of recognizing skin conditions.

3.DenseNet121:
In our analysis, we employed DenseNet121 model. DenseNet model, a type of convolutional neural network, primarily addressed the vanishing gradient issue in deep networks. DenseNet models incorporate direct skip connections between each layer, which improves gradient flow and network information flow during backpropagation. DenseNet121 has 121 layers. As the number of layers grows, models get more complicated and contain more layers. Because all DenseNet models were previously trained on the ImageNet dataset, transfer learning provides a solid foundation for classifying skin disorders.

4.InceptionV3:
The InceptionV3 model's convolutional neural network design was initially revealed at the ImageNet Large-Scale Visual Recognition Challenge. To extract complex characteristics from input data, the InceptionV3 model includes convolutional layers with variable filter sizes and average pooling layers. The ImageNet dataset's more than one million annotated pictures from 1,000 categories were used as pre-training data for the InceptionV3 model's 48 convolutional layers.

5.ResNet152:
In this study,ResNet152 was employed.ResNet152 has 152 layers.Deep convolutional neural network models known as ResNet152 models have succeeded in image classification applications. The skip connections in the ResNet models allow the models to create and learn residual functions. ResNet152 is the most complex ResNet model, with 152 layers. The ResNet models were pre-trained using the ImageNet dataset, which offers a solid basis for employing transfer learning to fulfill our aim of recognizing skin illness.

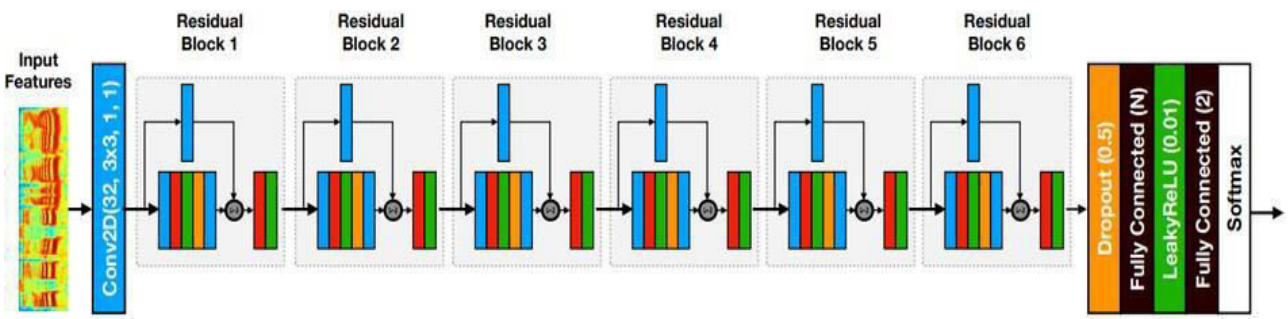


Figure 3: ResNet-152 Architecture

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805194

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Conv2D (32, 3x3, 1, 1):A 2D convolutional layer with 32 filters, each of size 3x3. The (1, 1) likely refers to the stride and padding, both set to 1. This layer extracts local features from the input.

Residual Blocks (1-6): Each residual block consists of convolutional layers, batch normalization, and ReLU activation functions.The key feature of these blocks is the "skip connection" or "shortcut connection," which adds the input of the block to its output. This helps to mitigate the vanishing gradient problem and allows for training deeper networks.

Dropout (0.5):A dropout layer with a 0.5 probability, meaning that during training, each neuron has a 50% chance of being randomly excluded from the forward and backward passes.This technique helps prevent overfitting.

Fully Connected (N):A fully connected layer with N neurons. This layer flattens the output from the previous layer and maps it to a new space.The value of N is the count of neurons present in the layer.

LeakyReLU (0.01):A Leaky Rectified Linear Unit activation function with a slope of 0.01 for negative inputs.This is similar to ReLU but allows a small, non-zero gradient when the unit is not active, which can help with training.

Fully Connected (2):A fully connected layer with 2 neurons. This layer maps the output to the final output space, which in this case seems to be two classes.

Softmax layer:A softmax activation function, which converts the output of the previous layer into a probability distribution over the two classes. The output will be two values between 0 and 1 that add up to 1, representing the probability of the input belonging to each class.

V. RESULTS AND ANALYSIS

Testing Methodology

The trained model was evaluated using a separate test dataset, ensuring that it was never exposed to the training data to avoid bias. The model performance was assessed based on key evaluation metrics:

- **Accuracy**: Measures the overall correctness of the model.
- **Precision**: Indicates how many of the positive predicted cases were actually positive.
- **Recall (Sensitivity)**: Measures how many of the actual positive cases were correctly predicted.
- **F1-Score**: Harmonic mean of precision and recall.
- **Specificity :** Measures how well a model correctly identifies negative cases**.**
- **Confusion Matrix**: Visual representation of true vs. predicted labels.

Testing was performed on dermoscopic images after applying preprocessing and augmentation techniques. The dataset was split into training, testing in an 80:20 ratio.

Models	Accuracy	Precision	Recall	F1-Score	Specificity
MobileNet	92.58	66.30	66.67	66.48	95.76
DenseNet121	92.40	65.75	66.11	65.93	95.69
VGG16	93.27	69.61	70.00	69.81	96.18
ResNet152	94.76	76.24	76.67	76.45	97.02
InceptionV3	92.89	67.96	68.33	68.14	95.96

Table 1: Tabulating Performance Metrics for 5 models

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805194

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

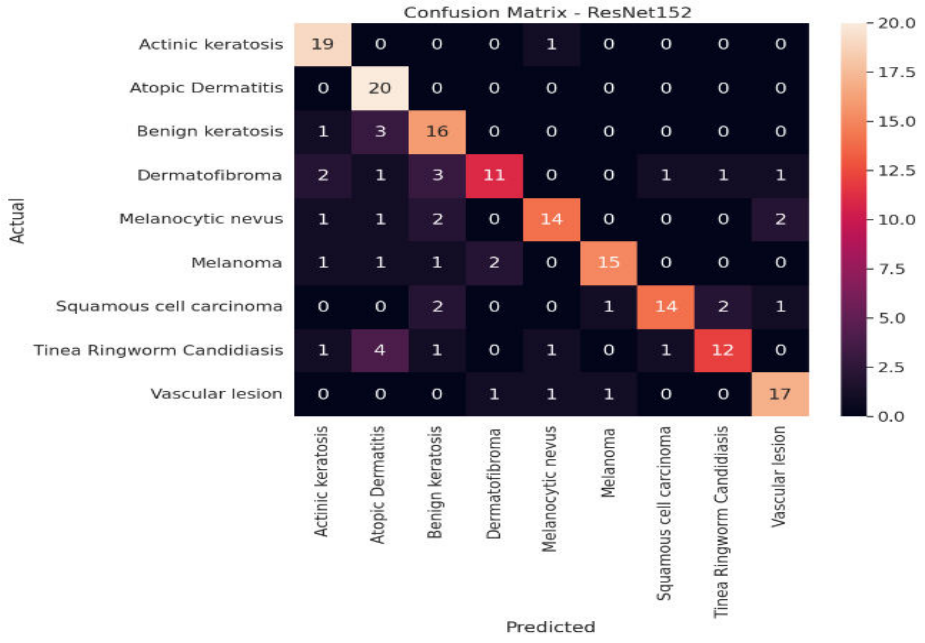


Figure 4: Confusion matrix for ResNet152

From the confusion matrix, **TP = 138, FP = 43, FN = 42, TN = 1398**

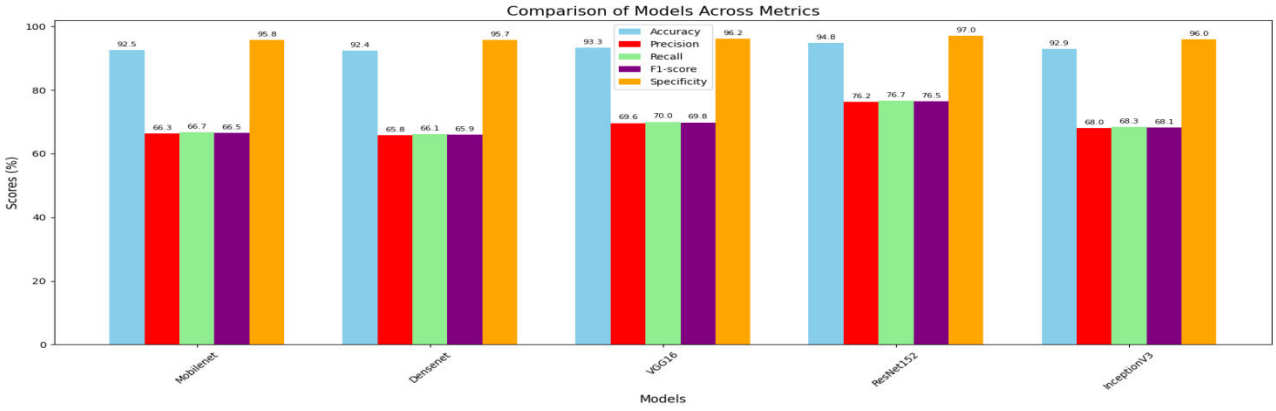


Figure 5: Plotting Bar-graphs for performance metrics for 5 models

V. CONCLUSION

The dermoscopy image classification project successfully developed an automated system for classifying skin lesions using transfer learning with pre-trained convolutional neural networks (CNNs), including VGG16, InceptionV3, ResNet152, DenseNet121, and MobileNet. By leveraging a Kaggle dataset of 900 annotated images across nine skin conditions, the project addressed challenges such as limited dataset size and class imbalance through robust preprocessing, data augmentation, and model fine-tuning. The methodology encompassed data preprocessing, model selection, training, evaluation, and prediction, achieving promising results.

Performance evaluation revealed ResNet152 as the top performer with an accuracy of 94.76%, followed by VGG16 at 93.27%, while MobileNet and DenseNet121 offered computational efficiency suitable for resource-constrained environments. The use of metrics like precision, recall, F1-score, and confusion matrices provided a comprehensive assessment, highlighting the models' ability to minimize critical errors such as false negatives in cancer detection. The

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

project also demonstrated the practical applicability of the system through a prediction module capable of real-time classification, supporting dermatologists in clinical decision-making.

This work underscores the potential of transfer learning in medical image analysis, particularly for early skin cancer detection, which is critical for improving patient outcomes. Future enhancements could include expanding the dataset, integrating advanced ensemble techniques, and deploying the system on mobile platforms for broader accessibility in underserved regions. The project lays a strong foundation for further research and development in automated diagnostic tools, contributing to the advancement of precision medicine

ACKNOWLEDGMENT

I would like to thank my project guide, Dr.P.P.S.Subhashini, M.Tech, Ph.D.,for their valuable guidance, encouragement, and continuous support. Their insights and constructive feedback were instrumental in shaping the direction and quality of this work.I am also thankful to the Department of Electronics and Communication Engineering, R. V. R & J. C College of Engineering, for providing the necessary infrastructure and academic environment to carry out this research. I would also like to acknowledge the open-source community for developing and maintaining tools and frameworks such as TensorFlow, Keras, and PyTorch, which were vital to the implementation of this project.

REFERENCES

[1] S. Akyeramfo-Sam, A. addo Philip, Derrick Yeboah, Nancy Nartey, et. al., "A Web-Based Skin Disease Diagnosis Using Convolutional Neural Networks," I. J. Information Technology and Computer Science, vol. 11, pp. 54-60, 2019.
[2] D. A. Shoieb, S. M. Youssef, and W, M. Aly, "Computer Aided Model for Skin Diagnosis Using Deep Learning," Journal of Image and Graphics, vol. 4, no. 2, 2016.
[3] S. Kumar Patnaik, M, Singh Sidhu, Yaagyanika Gehlot, Bhairvi Sharma, and P Muthu*, "Automated Skin Disease Identification using Deep Learning Algorithm," Biomedical & Pharmacology Journal, vol. 11(3), pp. 1429 1436, 2018.
[4] R. Yasir, Md. Ashiqur Rahman, and Nova Ahmed, "Dermatological Disease Detection Using Image Processing and Artificial Neural Network," in 8th International Conference on Electrical and Computer Engineering 20-22 December 2014, IEEE 2014, Dhaka, Bangladesh, 2014.
[5] P. Bose, S.K. Bandyopadhyay, A. Bhaumik, and Dr. Sandeep Poddar, "Skin Disease Detection: Machine Learning vs Deep Learning," Preprints, 13 September 2021.
[6] Jyothilakshmi K. K and Jeeva J. B, "Detection of malignant skin diseases based on the lesion segmentation," 2014 International Conference on Communication and Signal Processing, 2014, pp. 382-386, doi: 10.1109/ICCSP.2014.6949867.
[7] P. R. Hegde, M. M. Shenoy, and B. H. Shekar, "Comparison of Machine Learning Algorithms for Skin Disease Classification Using Color and Texture Features," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 1825-1828, doi: 10.1109/ICACCI.2018.8554512.
[8] S. Kolkur and D. R. Kalbande, "Survey of texture based feature extraction for skin disease detection," 2016 International Conference on ICT in Business Industry & Government (ICTBIG), 2016, pp. 1-6, doi: 10.1109/ICTBIG.2016.7892649.
[9] Aditya, W., Luthfil Hakim, N., Shih, T. K., Enkhbat, A., & Thaipisutikul, T.. IC4Windows–hand gesture for controlling MS windows. 2020 - 5th International Conference on Information Technology (InCIT), 2020, 2.
[10] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S.. Dermatologist-level classification of skin

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com