International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

Derma Trust: An Explainable and Uncertainty-Aware Deep Learning Framework for Skin Cancer Diagnosis

Lenka Ooha.*1, Dr. I. Sundara Siva Rao.2

1 M.Tech II Year Student ,Department of Computer Science(Datascience), Anil Neerukonda Institute of Technology And Sciences, Sangivalasa, Bheemili mandal, Visakhapatnam, Andhra Pradesh, India.
2 Prof,Department of Computer Science and Engineering(Data Science), Anil Neerukonda Institute of Technology And Sciences ,Sangivalasa, Bheemili mandal, Visakhapatnam, Andhra Pradesh, India.

Date of Submission: 25-05-2026 Date of Acceptance: 05-06-2026

ABSTRACT: The diagnosis of skin cancer, which can be a highly aggressive and even fatal disease, is severely time sensitive. It is critical in the treatment of the disease for the patient's prognosis that it is diagnosed as early as possible. Deep learning models in the classification of skin lesions have been developed, and have even been successful, but these models have been limited in their clinical utilization due to factors surrounding to some extent, the unreliable nature of these models. In this work, we integrate the classification of skin lesions with the trust of Deep Learning, Explainability, and uncertainty of the diagnosis. This model utilizes a convolutional network for classifying skin lesions and integrates the **Grad-CAM** framework for the visualization of the classification to bring clinical relevance to the explanation. For the uncertainty of the skin lesions, we utilize **Monte Carlo Dropout** for the model to output the degree of relevance for the skin lesions with respect to the identified uncertainty. For the classification support tool, our model implements a **risk stratification** framework to classify are categorized for the skin lesions as low, medium, and high that support a clinical decision-making to give stratification support. Derma trust's framework makes the developed systems highly reliable and interpretable, and at the level of abstraction for deep learning, the model makes it a highly trustworthy system to be used and integrated into the clinical support classification systems for the diagnosis of skin lesions and skin cancers.

KEYWORDS: Skin Cancer Diagnosis, Deep Learning, Explainable Artificial Intelligence (XAI), Grad-CAM, Uncertainty Estimation, Monte Carlo Dropout, Risk Stratification.

I. INTRODUCTION

There are various threat levels to human life from a plethora of diseases. Skin cancer is one of the most prevalent and possibly fatal diseases. The likelihood of survival is directly correlated with the early detection of the disease. The automation of skin lesion classification through dermatoscopic imaging using deep learning is experiencing rapid growth. The identification of benign and malignant lesions through the imaging of skin lesions is showing great promise for deep learning. The potential of aiding dermatologists is showing clinical promise as well. A problem, nonetheless, is adopting those solutions into the sphere of practical health care.

Deep learning models, in spite of their potential, remain commercially untapped. The main problem is attributed to their black box nature. They are hermetic, they offer a commanding opinion, and total trepidation is expected from their operators. By in large, practitioners of the healthcare industry are simply restating the opinion of the models, based on attempts to justify the predictions. To further complicate this issue, most models remain insufficient in the estimation of their prediction errors. The demand and supply for the models is uneven. The industry demands models that were built from the ground up based on rational explanations. The models that were built from the ground up based on rational explanations would estimate the errors of their predictions. Models built with such rational decoupling mechanisms would be a great addition to the health care industry.

II. LITERATURE SURVEY

Based on one of the main aspects of deep learning concerning skin cancer diagnosis, we focus on the integration of explainable artificial

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

intelligence (XAI), uncertainty modeling, and deep learning. Recent works are analyzed based on classification techniques and the interpretability of the systems and their support to clinical decision-making. The survey highlights the major challenges and restrictions of existing works, one of which is related to the absence of integrated frameworks that blend explainability and uncertainty estimation. This analysis helps define the research gap and clarifies the rationale for the proposed Derma trust framework.

[1] Asgharnezhad et al(2025) develop the "Quad-Step" framework to skin cancer classification, by integration of transfer learning, feature fusion, and uncertainty quantification (UQ), to be more robust. The study evaluated several pre-trained models such as CLIP models, ResNet50, DenseNet121, VGG16, and EfficientNet-V2-Large, using the HAM10000 dataset The results showed that the CLIP models, specifically ViT-H/14 and ViT-L/14, extracted the most profound dermatologic features among the CT architecture. To fill the gap between high precision and clinical trust, the authors evaluate uncertainty quantification strategies, such as Monte Carlo Dropout and Deep Ensembles, with specific uncertainty-focused metrics. Their final architecture proposed CLIP-based transfer learning with additional and feature fusion through MLP and a predictive entropy loss, to be state-of-the-art with an accuracy of 87.62%.

[2] Purwono et al (2025) presented a review on Explainable Artificial Intelligence (XAI) in the medical imaging to address the "black-box" nature of deep learning models. The authors have classified XAI techniques into four groups: representation of features through visualization, attention-based techniques, rule-based techniques, and representation of features. The authors presented an account of the impact of these techniques, through the trust and confidence gained from clinical decision making, in the cases of numerous clinical challenges in cardiology, oncology, and neurology. The authors review and described the significant drawbacks, and the trade-off appears to be the challenge of balancing predictive performance with interpretability, and finding explanations to fit the interpretive iterative nature of time-critical clinical practices

[3] Ali and his team(2026)came up with a way to use deep learning to find skin cancer early. They made a system that uses something called self-supervised pretraining and multi-architecture ensemble learning. this system first looks at lots of pictures of skin without knowing what is in them. Tries to find what makes them different from each other. then it uses this information to make a model that can tell if someone has skin cancer. the system uses three models: EfficientNetV2-L, Swin Transformer and ConvNeXt and LightGBM. it also uses something called AI including Grad-CAM and LIME so doctors can understand how the system made its decision. When they tested this system on a group of pictures it worked well. They found that the system was very good at this task. The system did a job, with the group of pictures. it was better at finding skin cancer than any of the models on their own. this shows that using self-supervised learning and based explainability together can make a system that doctors can trust to help them find skin cancer.

[4] Zbrzezny and Krzywicki did a review in 2025 about how artificial intelligence's changing dermatology. they wanted to see how artificial intelligence is moving from being tested to being used in real life. the authors put intelligence methods into groups. these groups include things like networks or CNNs that look at pictures of skin lesions. they also looked at models and systems that use many types of data. the review put together studies and outcomes. it showed that artificial intelligence can really help doctors with skin problems. artificial intelligence can help find skin cancers like melanoma and non-melanoma skin cancers. there are problems to using artificial intelligence widely. one big problem is that we need different types of data. this data should have all skin types in it. we also have to deal with laws and ethics around keeping data private and making sure artificial intelligence diagnoses are correct.

[5] Das and Addyadid a study in 2025.they wanted to see if EfficientNet could be used to find and classify skin cancer . They used the HAM10000 dataset. They made an EfficientNet-B0 model to classify skin lesions into seven types. These types include melanoma and basal cell carcinoma. the study liked EfficientNet's method. This method makes the network deeper, wider and higher resolution. This gives results but uses less computer power than other methods.to make the model more accurate they did things to the images. They made the images bigger or smaller. they also made sure there were numbers of each type of skin lesion. the results showed that the model was very good at predicting skin cancer. It also had a F1-score. This means EfficientNet can be an efficient way to find skin cancer early in clinics.

[6] Haque and others (2026) present a learning model for skin lesion diagnosis. This

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

model uses data balancing, advanced image augmentation and explainable AI to improve accuracy. The researchers use the HAM10000 dataset to train their model. They enhance the EfficientNetV2-L framework with channel attention mechanisms and a three-stage learning approach. This helps the model extract features better. the model is like a "black-box". So the researchers use techniques like Grad-CAM and saliency maps. These provide explanations for the models predictions. the results show the model performs well. It achieves an accuracy of 91.15% and an AUC of 99.33%.This work shows combining model design with visual explanations can develop robust skin cancer classification tools.

[7] Ali and others (2025) introduce an AI framework, for skin cancer detection. They address the issue of data using Denoising Diffusion Probabilistic Models (DDPM). DDPM generates high-quality dermatoscopic images. This balances the HAM10000 dataset .the researchers train a classification system using the dataset. Their model combines features from ResNet50, DenseNet121 and InceptionV3.this helps improve accuracy across seven skin lesion categories. the study uses SHAP to explain the models decisions. the results show the DDPM-augmented framework outperforms augmentation techniques. It achieves classification metrics and provides transparency. this research highlights combining modeling with explainability. this can develop trustworthy dermatological diagnostic tools.

[8] Noor and other researchers in 2025 looked into how deep learning architectures, specifically VGG16 and DenseNet201 can tell the difference between benign and malignant skin lesions. The people who do this work wanted to make these convolutional neural networks better so they can help doctors find cancer on the skin. the researchers used something called transfer learning to make this happen. they tested the models well to see how they can handle complicated pictures of skin. the main goal was to make the process of checking for skin cancer and reduce mistakes when doctors are checking for skin cancer. what they found out is that there are trade-offs between how complex the convolutional neural networks how accurate the convolutional neural networks are. they also found out that these special convolutional neural networks can really help with finding skin cancer without needing to invade the body.

[9] Perera and other Perera and other researchers found a way to diagnose melanoma in 2025. They combined methods to make it work better. they used three models: InceptionV3, ResNet50, VGG16 ,these models were used together to get results. The researchers wanted doctors to trust the model. So they added a way to explain how it makes decisions. the model looks at skin features to make decisions. The researchers used a method called Local Interpretable Model-Explanations. This method helps doctors see what the model is looking at. the results showed that using models together is better, than using just one. It also makes the melanoma model more reliable when it explains its decisions. making melanoma models that're accurate and easy to understand is really important. This way doctors will use them to help diagnose skin cancer. Doctors can use these models to help people with melanoma.

[10] Morales-Forero and their team did a study on the HAM10000 dataset in 2024. They were looking at the HAM10000 dataset to see if the pictures were fair. The team wanted to know if the pictures used to train doctors to diagnose skin conditions were good for everyone. the researchers found out that most of the pictures are of people with skin that's the same. This is not good because it makes it hard for doctors to be sure when they diagnose people with skin. the study showed that if we use pictures of people with the same skin to train the AI tools it can be bad for people with different skin. The researchers think that we need to use pictures of people with skin tones like the HAM10000 dataset so that the AI tools can work well for everyone, with skin conditions.

[11] Nigardand their team had an idea to make the AI tools help doctors diagnose skin conditions in 2022. they wanted to make it clear how the AI makes its decisions. The researchers created a system that notsays what type of skin lesion it thinks it sees but also shows why it thinks that. They used a type of AI that can look at pictures and learn from them. The system helps doctors trust I ts decisions by showing them which parts of the picture it used to make its diagnosis. The researchers tested their system. Found that it works well and helps doctors feel more confident in its diagnoses. This is important, for catching skin cancer.

[12] Raju and his team introduced XAI-Skin CADx in 2025. This is a six-stage framework that uses learning to help doctors diagnose skin cancer more accurately. XAI-Skin CADx uses a few models like InceptionResNetV2, MobileNetV2 and EfficientNetB0 to look at pictures of skin and figure out what is going on. The problem with learning is that it can be hard to understand why it makes certain decisions. XAI-Skin CADx provides assistance in understanding the process of how the

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

model arrives at decisions by using methods such as Grad-CAM, LIME and SHAP. Another aspect of this tool is how it advises doctors about how risky the diagnosis is. The outcomes demonstrate that XAI-Skin CADx is highly competent in detecting skin cancer, having an accuracy level of 92%.

[13] Hauser and his colleagues conducted a study of the application of Explainable Artificial Intelligence for the detection of skin cancer. The study concluded that the majority of research conducted utilizes established models to interpret their outcomes. None of the researches conducted by the team aimed to develop new models nor evaluate the contribution made by such methods to doctors. The team noticed a discrepancy in current research and research that could be performed, stating that little is known about the contribution of XAI in the field of skin cancer diagnostics by medical practitioners. They believe that further research needs to be carried out to demonstrate that XAI can provide benefits in hospital surroundings.

[14] Hamim and his team developed Smart Skin-XAI; a framework of using learning to classify skin cancer based on the DenseNet121 model, this framework uses Grad-CAM as a technique to explain the operation of the framework. The framework was evaluated against the benchmarks provided, the results indicate it performed adequately with figures approximating between 97% to 98%, and performing better than various models. The team also suggest that this model will offer more reliability with the use of explanations and combined learning and explanation will advance the diagnostics of skin cancer significantly.

[15] Mukherjee et al developed an approach in the year 2026. They created a model based on learning that utilizes an activation function in order to aid learning and decision-making. Layers were established to examine images of skin in order to extract features, this along with mechanisms such as batch normalization and dropout prevent overfitting of data. The objective of the researchers was for the model to be interpretable for doctors, and to achieve this XAI models are used in order to explain their processes, this helps with doctor's trust of the model and understanding of it for skin cancer detection. the study found that models like XAI-Skin CADx, in conjunction with Smart Skin-XAI demonstrated the capabilities XAI holds within skin cancer diagnosis.

The XAI approach can further contribute to the detection of skin cancer. The use of XAI in skin cancer detecting models such as XAI-Skin CADxSkinCADx is a model capable of demonstrating how it arrived at its conclusion through the utilization of methods like Grad-CAM, LIME and SHAP which are of extreme importance to doctors in understanding the rationale behind XAI-Skin CADx's verdict. The model boasts of impressive performance with accuracy results ranging at 92%, demonstrating XAI- SkinCADx to be a highly capable tool in the field of skin cancer detection.

[16] Mahmud(2023) and others created a multi-model framework for skin cancer detection. They utilized multiple models to create a more reliable system and applied methods to improve model's ability to identify patterns in the image. To provide insight into the decision process, Mahmud et al. used Explainable AI technology. In particular, one of the models – XceptionNet – demonstrated high performance (accuracy 88.72%). According to Mahmud et al., a combination of learning and transparency would enable medical professionals to trust artificial intelligence technologies.

To illustrate the results, there is a table. The table indicates the strong and weak points of each model to help identify areas of interest that need to be improved. In particular, such a table shows why a multi-model system (Derma trust) that is highly effective in detecting skin cancer and transparent about its decision process is needed.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

Table: I

Author(s)	Year	Journal	Methodology	Dataset	Accuracy	Advantages	Limitations	Research Gap
Asgharnezhad et al.	2025	arXiv	Feature fusion + CLIP + UQ (MC Dropout, Ensembles)	HAM10000	87.62%	Strong uncertainity modeling, feature fusion	No explainability (XAI)	Lack of integrated XAI + UQ
Ali et al.	2026	JCBI	Ensemble (EffNetV2 + Swin + ConvNeXt) + XAI	HAM10000, ISIC	High	Hybrid + self-supervised learning, strong performance	No uncertainty estimation	No confidence-aware predictions
Haque et al.	2026	arXiv	EfficientNet V2 + Attention + Grad-CAM	HAM10000	91.15%	High accuracy + visual explainability	Black-box confidence	Missing uncertainty modelling
Hamim et al.	2025	Diagnostics (MDPI)	DenseNet121 + Grad-CAM (SmartSkin-XAI)	ISIC	~97–98%	Strong interpretability + high accuracy	No uncertainty quantification	Lack of reliability estimation
Mukherjee et al.	2026	Discover Oncology	Swish-CNN + XAI (saliency maps)	Dermoscopy datasets	~98%	Improved activation + explainability	No uncertainty handling	No trust/confidence measure
Mahmud et al.	2023	arXiv	Multiple CNNs (Xception, EfficientNet) + XAI	Dermoscopy	88.72%	Multi-model comparison + interpretability	Moderate accuracy	No uncertainty integration
Raju et al.	2025	JBI	Ensemble (InceptionResNetV2 + MobileNet + EfficientNet) + XAI	HAM10000, ISIC	92.40%	Multi-XAI + clinical risk module		No probabilistic confidence
Hauser et al.	2022	European Journal of Cancer	Systematic Review (XAI in dermatology)	Multiple	—	Strong clinical insight, XAI importance	No experimental validation	No uncertainty modeling

Review of Related Work with Outcomes and Limitations The comparison table shows that most existing methods are mostly about making classification more accurate and some of them also use explainability techniques. Only a few methods think about how sure they are of their results or give confidence-based decision support. This shows a problem: the current models can be accurate but they are not always trustworthy or easy to understand when used in real life. to fix this

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

problem the Derma Trust framework combines explainability with uncertainty estimation and risk stratification. The Derma Trust framework does this to make a system that's more reliable and useful for real-world clinical use and that is what the Derma Trust framework is all, about.

II.1.2 RESEARCH GAP:

Despite advancements in deep learning for skin cancer diagnosis existing studies mainly focus on improving classification accuracy. They also use explainability techniques like Grad-CAM, LIME and SHAP to enhance interpretability. some studies, including hybrid architectures have shown high diagnostic performance. however most of these approaches operate as models without addressing the reliability of their predictions. a few studies have explored uncertainty- frameworks but focus on confidence estimation without explainability mechanisms. systematic reviews highlight the lack of evaluation protocols. there is also limited validation of models, in real clinical environments. as a result current approaches fail to provide a solution that ensures high accuracy, interpretability and reliability. this gap motivates the development of the proposed Derma Trust framework, which integrates AI and uncertainty estimation to deliver more transparent, trustworthy and clinically reliable skin cancer diagnosis using Derma Trust. the comparison table clearly shows that most existing methods focus primarily on improving classification accuracy and, in some cases, incorporate explainability techniques. However, only a small number of approaches consider uncertainty estimation or offer confidence-based decision support. This reveals a clear gap in developing models that are not just accurate, but also reliable and easy to interpret. The proposed Derma Trust framework addresses this limitation by combining explainability, uncertainty estimation, and risk stratification into a single unified system, making it more suitable for real-world clinical applications.

II.1.3 PROBLEM STATEMENT:

Deep learning-based skin cancer diagnosis has achieved good accuracy. However, most existing models function like black boxes, providing predictions without explaining how they were derived. This lack of transparency makes it difficult for dermatologists to trust these systems, which limits their adoption in practical settings. Some recent studies have tried to make these models more transparent using AI (XAI) techniques like Grad-CAM. These techniques provide explanations but do not address the uncertainty of the predictions. On the hand some models estimate the confidence of their predictions but are hard to understand. There is a need for a model that combines accuracy, explainability and uncertainty estimation. It should also help in making decisions by categorizing risks.

II.1.4 OBJECTIVE OF THE STUDY: The main goal of this work is to develop a deep learning framework called Derma Trust. This framework will accurately and reliably diagnose skin cancer. To achieve this we have the following objectives: a. To design a deep learning model for skin cancer classification using dermatoscopic images. The model should have accuracy then the model needs to be accurate.it should work well in different situations.it has to give results every time the model must be reliable in scenarios. b. Apply explainable AI(XAI) methods such as Grad-CAM to better understand of how the model makes preditions ,there it is used to improving transparency. c. To integrate uncertainty estimation mechanisms. These mechanisms will quantify the confidence of the predictions. This will improve the reliability of the diagnosis. d. To develop a risk stratification mechanism. This mechanism will categorize predictions into confidence levels. These levels can be medium or high risk. This will help in making clinical decisions. e. To evaluate the proposed framework using performance metrics. These metrics include accuracy, precision, recall, F1-score and AUC. We will also analyze the uncertainty and interpretability of the model. f. To compare the proposed model with existing state-of-the-art methods. This will demonstrate the improvements in trustworthiness, interpretability and reliability. g. To make the models outputs understandable and reliable for dermatologists. This will ensure that the model is useful in real-world scenarios.

III. METHODOLOGY

III.1 Overview of the Proposed Framework :

This study proposes Derma Trust. Derma Trust is a deep learning framework, for skin cancer diagnosis. It integrates classification, explainability and uncertainty estimation. This integration will enhance reliability. Figure 1 shows the diagram of the proposed Derma Trust.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

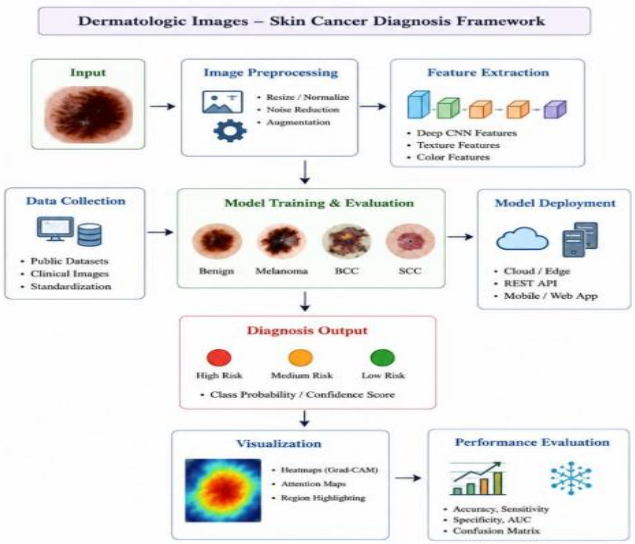

Figure 1: Architecture diagram (Derma Trust)

The architecture has five parts: data preprocessing, deep learning classification uncertainty estimation, explainable AI (XAI) and risk stratification. the framework uses dermatoscopic images. Provides classification results, visual explanations and confidence levels. this helps doctors make decisions.

III.2 Mathematical Framework of Derma Trust:

The derma trust framework operates on the principle of combining high-performance classification with visual explainability and stochastic uncertainity quantification.

III.2.1. Features importance via Grad-CAM:

To provide visual transparency, we utilize Gradient-Weighted class activation mapping(Grad-CAM).for a specific class $C1$we compute the importance weight $\alpha kC$for each feature map Ak in the final convolutional layer by calculating the global average pooling of the gradients:$i\alpha_k^C = \frac{1}{z}\sum_i\sum_j \frac{\partial y^C}{\partial A_{ij}^k}$

The final coarse heatmap,$L^C_{GRad-CAM}$, is generated by performing a ReLU-activated weighted sum of the features maps: :$L^C_{GRad-CAM}$ =ReLU$\left(\sum_k \alpha_k^C A^k\right)$

This ensures that only pixels with a positive influence on the diagnostic class c are highlighted, filtering out negative or irrelevant feature.

III.2.2. Uncertainty Estimation via Monte carlo dropout:

To account for the "black-box" nature of traditional CNNs, we estimate epistemic uncertainity using monte carlo dropout during the inference phase. By performing T stochastic forward passes for a single input image x, the predictive mean u is defined as:

$$u = \frac{1}{T}\sum_{t=1}^{T}\hat{y}_t$$

The degree of uncertainty , or predictive variance $\sigma 2$, is then quantified as: $\sigma^2 = \frac{1}{T}\sum_{t=1}^{\tau}(\hat{y}_t - u)^2$

III.2.3. Risk Stratification framework:

The model translates the softmax probability p(y/x) and the uncertainityσ^2 into a clinical risk score R. this categorization helps clinicians prioritize urgent cases.the risk level is defined using a threshold-based logic:

$$R=\begin{cases} \text{High Risk,} & \text{if } P(y_{(malignant)}) > \tau_{high} \\ \text{Medium Risk,} & \text{if } \tau_{low} < P(y_{(malignant)}) > \tau_{high} \\ \text{Low Risk,} & \text{if } P(y_{(malignant)}) > \tau_{low} \end{cases}$$

Where τ_{high} and τ_{low} are pre-defined confidence thresholds(e.g., 0.7 and 0.3) optimized for the HAM10000 dataset.

III.2.4. Balanced loss function:

To handle class imbalance in HAM10000,we use weighted cross-entropy loss(L):$L = -\sum_{i=1}^{7} w_i \cdot y_i \log(\hat{y}_i)$

Where w_i is the weight for each of the seven classes.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

III.2.5. Trust-Based Decision:

The final trust index(TI) combines certainity and visual focus:TI=$(1-\sigma^2)$x Focus Score

If TI is low, the system triggers a "manual review required" alert for clinical safety.

III.3 Workflow:

Derma Trust is a learning framework that helps diagnose skin cancer accurately and provides clear explanations. it also tells us how confident it is, in its diagnosis. the framework is shown in Figure 2. it uses a type of network and a technique called Grad-CAM to provide visual explanations. another technique, Monte Carlo Dropout is used to estimate uncertainty. this helps doctors trust the diagnosis and make decisions make decisions.

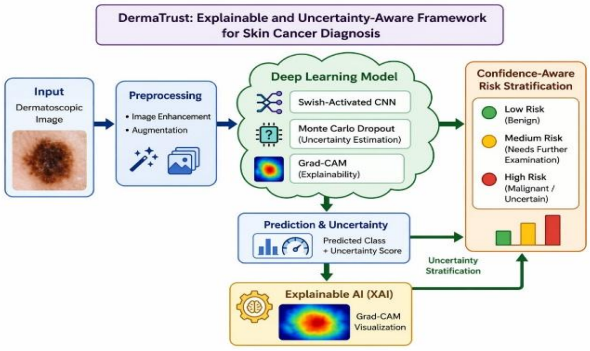

Figure.2: Methodology Diagram

The inputs are the skin lesion datasets. These datasets show different types of skin lesions. We can see that each type of lesion has a color, texture and pattern. This makes it hard to automatically classify them. We need a model that can learn to tell these lesions apart.

III.4.DeepLearning-BasedClassification:

The main part of our Derma Trust framework is a learning model. This model automatically identifies types of skin lesions from images. In this work ,we use Convolutional Neural Networks (CNNs) to analyze image data because they are good at learning from images.We use -trained architectures such as like ResNet50 and EfficientNet to improve classification . We adjust these models to work with our skin lesion images. This helps the model learn faster and work better. The model gives us a list of probabilities for each type of skin lesion.

Here is how our model works:

* It has layers that look for features in the images.
* It has layers that help the model learn better.
* It has a layer that helps the model make a decision.

The model gives us a list of probabilities, for each type of skin lesion. We train the model using a method that helps it learn from its mistakes. We use an algorithm called Adam optimizer to help the model learn.

Algorithm 1: Skin Cancer Classification using CNN

Input: Dermatoscopic image dataset D

Output: Predicted class label y

We start with a -trained CNN model (ResNet50 or EfficientNet).

1. We change the layer to match the number of skin lesion classes (7 classes).
2.We train the model.
3.For each batch of images we:
4.Prepare the images.
5.Use the model to predict the class.
6.See how wrong the prediction is.
7.Adjust the model.
8. We keep training until the model is good enough.
9. We use the trained model to predict the class of an image.
10.We return the predicted class.

The inputs are the skin lesion datasets, which highlights the diversity of skin lesion classes used in this study. Skin lesion look different from each other in terms of color ,texture ,andstructure . this makes automatic classification a difficult tasks.hence , deep learning models are used to detect important images features and improve classification accuracy

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

III.4.1 ResNet50 Architecture

The ResNet50 is a neural network it is used for image recognitionin this the layers help the model identify and learn useful patterns from images .The problem with making these networks really deep is that the information can get lost as it goes through the layers. The ResNet50 solves this problem by allowing the information to jump over some of the layers.The ResNet50 has 50 layers that do things to the picture. Some of these layers look for edges. Some look for textures. The important part of the ResNet50 is the residual block. This is where the picture is added to the picture that has been changed by the layers. This helps the ResNet50 learn what is important about the picture.The ResNet50 starts by looking at the picture and then it looks at smaller parts of the picture. It does this by using something called max pooling. This helps the ResNet50 understand what is important about the picture and whatis not. The ResNet50 is made up of residual blocks that are grouped together. Each group looks at the picture in a way and helps the ResNet50 understand what it is looking at. The ResNet50 is really good, at looking at pictures of skin and helping doctors understand what they are looking at.ResNet50 is a deep convolutional neural network that introduces residual learning to address the vanishing gradient problem encountered in deep architectures. It utilizes skip (identity) connections that allow gradients to flow directly through the network, improving convergence and enabling effective training of very deep models.

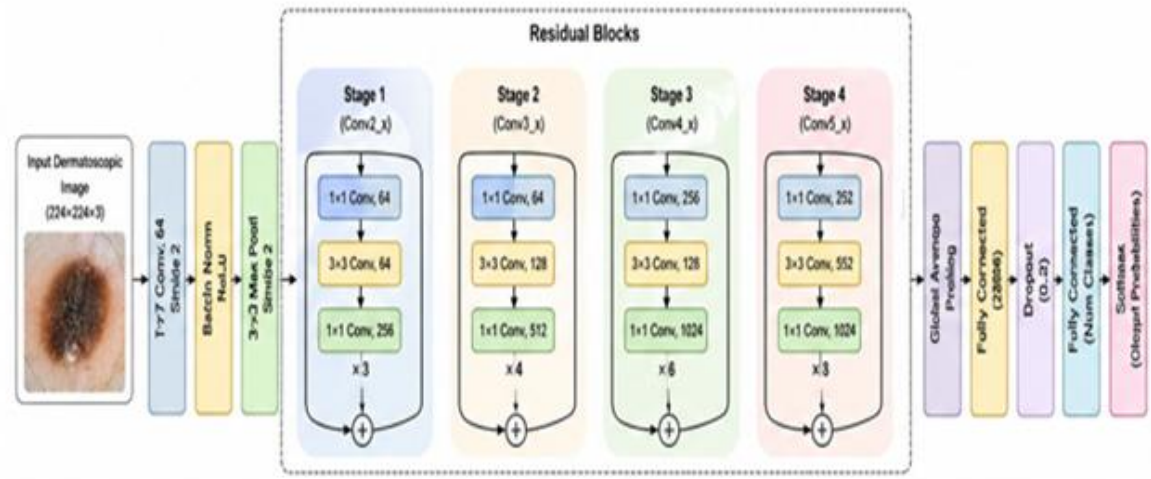

Figure:3 ResetNet50 Architecture

The figure shows what the ResNet50 model looks like with connections. These residual connections help the ResNet50 model train faster. Learn features better. The ResNet50 model is designed to make it easier to train and improve how it learns features.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

Model: "functional_1"

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 224, 224, 3)	0	-
conv1_pad (ZeroPadding2D)	(None, 230, 230, 3)	0	input_layer_1[0]…
conv1_conv (Conv2D)	(None, 112, 112, 64)	9,472	conv1_pad[0][0]
conv1_bn (BatchNormalizatio…	(None, 112, 112, 64)	256	conv1_conv[0][0]
conv1_relu (Activation)	(None, 112, 112, 64)	0	conv1_bn[0][0]
pool1_pad (ZeroPadding2D)	(None, 114, 114, 64)	0	conv1_relu[0][0]
pool1_pool (MaxPooling2D)	(None, 56, 56, 64)	0	pool1_pad[0][0]
conv2_block1_1_conv (Conv2D)	(None, 56, 56, 64)	4,160	pool1_pool[0][0]
conv2_block1_1_bn (BatchNormalizatio…	(None, 56, 56, 64)	256	conv2_block1_1_c…
conv2_block1_1_relu (Activation)	(None, 56, 56, 64)	0	conv2_block1_1_b…
conv2_block1_2_conv (Conv2D)	(None, 56, 56, 64)	36,928	conv2_block1_1_r…
conv2_block1_2_bn (BatchNormalizatio…	(None, 56, 56, 64)	256	conv2_block1_2_c…
conv2_block1_2_relu (Activation)	(None, 56, 56, 64)	0	conv2_block1_2_b…
conv2_block1_0_conv (Conv2D)	(None, 56, 56, 256)	16,640	pool1_pool[0][0]
conv2_block1_3_conv (Conv2D)	(None, 56, 56, 256)	16,640	conv2_block1_2_r…
conv2_block1_0_bn (BatchNormalizatio…	(None, 56, 56, 256)	1,024	conv2_block1_0_c…
conv2_block1_3_bn (BatchNormalizatio…	(None, 56, 56, 256)	1,024	conv2_block1_3_c…
conv2_block1_add (Add)	(None, 56, 56, 256)	0	conv2_block1_0_b… conv2_block1_3_b…
conv2_block1_out (Activation)	(None, 56, 56, 256)	0	conv2_block1_add…
conv2_block2_1_conv (Conv2D)	(None, 56, 56, 64)	16,448	conv2_block1_out…
conv5_block3_1_bn (BatchNormalizatio…	(None, 7, 7, 512)	2,048	conv5_block3_1_c…
conv5_block3_1_relu (Activation)	(None, 7, 7, 512)	0	conv5_block3_1_b…
conv5_block3_2_conv (Conv2D)	(None, 7, 7, 512)	2,359,808	conv5_block3_1_r…
conv5_block3_2_bn (BatchNormalizatio…	(None, 7, 7, 512)	2,048	conv5_block3_2_c…
conv5_block3_2_relu (Activation)	(None, 7, 7, 512)	0	conv5_block3_2_b…
conv5_block3_3_conv (Conv2D)	(None, 7, 7, 2048)	1,050,624	conv5_block3_2_r…
conv5_block3_3_bn (BatchNormalizatio…	(None, 7, 7, 2048)	8,192	conv5_block3_3_c…
conv5_block3_add (Add)	(None, 7, 7, 2048)	0	conv5_block2_out… conv5_block3_3_b…
conv5_block3_out (Activation)	(None, 7, 7, 2048)	0	conv5_block3_add…
global_average_poo… (GlobalAveragePool…	(None, 2048)	0	conv5_block3_out…
batch_normalizatio… (BatchNormalizatio…	(None, 2048)	8,192	global_average_p…
dense_2 (Dense)	(None, 256)	524,544	batch_normalizat…
dropout_1 (Dropout)	(None, 256)	0	dense_2[0][0]
dense_3 (Dense)	(None, 7)	1,799	dropout_1[0][0]

Total params: 24,122,247 (92.02 MB)
Trainable params: 14,980,615 (57.15 MB)
Non-trainable params: 9,141,632 (34.87 MB)

Figure 4:Model Summary of ResetNet50:

The ResNet50 architecture is a convolutional neural network. This network uses learning to deal with the vanishing gradient problem. The ResNet50 architecture has convolutional blocks. These blocks use batch normalization and ReLU activation. This helps the network extract features from images in a efficient way.

The ResNet50 architecture has connections. These connections are also called connections. They help the model learn identity mappings. This makes the training process more stable. It also improves performance in the layers of the network.The ResNet50 architecture changes the dimensions of the images as they go through the network. The spatial dimensions get smaller. The feature depth gets bigger. This helps the network capture both level and high-level representations of the images.The network uses an average pooling layer to reduce the feature maps. Then it uses connected layers and dropout for classification and regularization. The final dense layer gives the class probabilities for multi-class skin lesion classification.

III.4.2 EfficientNet Architecture:

The EfficientNet is a learning model that is very good. It improves performance by changing the network dimensions in a way. This includes the

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

depth, width and resolution of the network. The EfficientNet uses a method to scale the network. This method is called scaling. It helps the network get accuracy with fewer parameters and less computational cost.The EfficientNet architecture is based on inverted bottleneck convolution blocks. These blocks are also called MBConv blocks. They extract features from images in a way. They do not use a lot of power. The EfficientNet captures both global features of the images.

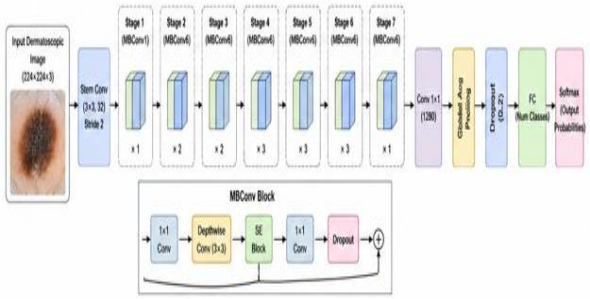

Fig-5: EfficientNET architecture

Model: "functional_2"

Layer (type)	Output Shape	Param #	Connected to
input_layer_2 (InputLayer)	(None, 224, 224, 3)	0	-
rescaling (Rescaling)	(None, 224, 224, 3)	0	input_layer_2[0]…
normalization (Normalization)	(None, 224, 224, 3)	7	rescaling[0][0]
rescaling_1 (Rescaling)	(None, 224, 224, 3)	0	normalization[0]…
stem_conv_pad (ZeroPadding2D)	(None, 225, 225, 3)	0	rescaling_1[0][0]
stem_conv (Conv2D)	(None, 112, 112, 32)	864	stem_conv_pad[0]…
stem_bn (BatchNormalizatio…	(None, 112, 112, 32)	128	stem_conv[0][0]
stem_activation (Activation)	(None, 112, 112, 32)	0	stem_bn[0][0]
block1a_dwconv (DepthwiseConv2D)	(None, 112, 112, 32)	288	stem_activation[…
block1a_bn (BatchNormalizatio…	(None, 112, 112, 32)	128	block1a_dwconv[0…
block1a_activation (Activation)	(None, 112, 112, 32)	0	block1a_bn[0][0]
block1a_se_squeeze (GlobalAveragePool…	(None, 32)	0	block1a_activati…
block1a_se_reshape (Reshape)	(None, 1, 1, 32)	0	block1a_se_squee…
block1a_se_reduce (Conv2D)	(None, 1, 1, 8)	264	block1a_se_resha…
block1a_se_expand (Conv2D)	(None, 1, 1, 32)	288	block1a_se_reduc…
block1a_se_excite (Multiply)	(None, 112, 112, 32)	0	block1a_activati… block1a_se_expan…
block1a_project_co… (Conv2D)	(None, 112, 112, 16)	512	block1a_se_excit…
block1a_project_bn (BatchNormalizatio…	(None, 112, 112, 16)	64	block1a_project_…
block2a_expand_conv (Conv2D)	(None, 112, 112, 96)	1,536	block1a_project_…
block2a_expand_bn (BatchNormalizatio…	(None, 112, 112, 96)	384	block2a_expand_c…
block2a_expand_act… (Activation)	(None, 112, 112, 96)	0	block2a_expand_b…
block2a_dwconv_pad (ZeroPadding2D)	(None, 113, 113, 96)	0	block2a_expand_a…
block7a_expand_act… (Activation)	(None, 7, 7, 1152)	0	block7a_expand_b…
block7a_dwconv (DepthwiseConv2D)	(None, 7, 7, 1152)	10,368	block7a_expand_a…
block7a_bn (BatchNormalizatio…	(None, 7, 7, 1152)	4,608	block7a_dwconv[0…
block7a_activation (Activation)	(None, 7, 7, 1152)	0	block7a_bn[0][0]
block7a_se_squeeze (GlobalAveragePool…	(None, 1152)	0	block7a_activati…
block7a_se_reshape (Reshape)	(None, 1, 1, 1152)	0	block7a_se_squee…
block7a_se_reduce (Conv2D)	(None, 1, 1, 48)	55,344	block7a_se_resha…
block7a_se_expand (Conv2D)	(None, 1, 1, 1152)	56,448	block7a_se_reduc…
block7a_se_excite (Multiply)	(None, 7, 7, 1152)	0	block7a_activati… block7a_se_expan…
block7a_project_co… (Conv2D)	(None, 7, 7, 320)	368,640	block7a_se_excit…
block7a_project_bn (BatchNormalizatio…	(None, 7, 7, 320)	1,280	block7a_project_…
top_conv (Conv2D)	(None, 7, 7, 1280)	409,600	block7a_project_…
top_bn (BatchNormalizatio…	(None, 7, 7, 1280)	5,120	top_conv[0][0]
top_activation (Activation)	(None, 7, 7, 1280)	0	top_bn[0][0]
global_average_poo… (GlobalAveragePool…	(None, 1280)	0	top_activation[0…
batch_normalizatio… (BatchNormalizatio…	(None, 1280)	5,120	global_average_p…
dense_4 (Dense)	(None, 256)	327,936	batch_normalizat…
dropout_2 (Dropout)	(None, 256)	0	dense_4[0][0]
dense_5 (Dense)	(None, 7)	1,799	dropout_2[0][0]

Total params: 4,384,426 (16.73 MB)
Trainable params: 4,339,843 (16.56 MB)
Non-trainable params: 44,583 (174.16 KB)

Figure 6: Model Summary of EfficientNet

The EfficientNet model uses a way to balance how deep, wide and detailed it is to get better results. It is made with MBConv blocks that have squeeze-and-excitation modules, which help get features from the data and pay attention to the right channels. This also helps reduce the amount of calculations needed while keeping the accuracy. The network learns features step by step. Then uses global average pooling and dense layers to finally classify things into seven categories. So, because of this design, efficientnet performs more accurate and efficient then traditional CNN models.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

III.5 AI using Grad-CAM

The framework uses AI with Grad-CAM to make the deep learning model easier to understand. AI using Grad-CAM generates a heatmap, this heatmap shows thw parts of the image that the AI using Grad-CAM model uses to make pedictions. It works by doing a things like:

1.The AI using Grad-CAM identifies the parts of the image that affect the target class.

2.The AI using Grad-CAM gives importance to the important parts.

3.The AI using Grad-CAM creates a map of these parts.

Algorithm 2: Explainability using Grad-CAM

Input:A trained CNN model, an input image x and a target class y

Output:A heatmap visualization H

1. Put the input image x into the trained CNN model
2. Do a pass to get feature maps
3. Calculate how much the target class y affects the feature maps
4. Do global average pooling on the results to get weights
5. Calculate the weighted sum of the feature maps
6. Use ReLU to keep the important features
7. Resize the heatmap to match the input image size
8. Make sure the heatmap values are between 0 and 1
9. Put the heatmap on top of the original image
10. Return the Grad-CAM visualization H

III.6 Uncertainty Estimation

To make predictions more reliable the framework also estimates how sure the model is about its predictions using techniques like Monte Carlo Dropout or Deep Ensembles.

In Monte Carlo Dropout:

Dropout layers are turned on during inference.the model does stochastic forward passes.the average prediction is the final output. The difference between predictions represents how sure the model is.

Mathematically uncertainty is estimated as:

Predictive Mean: The average of multiple predictions

Predictive Variance:A measure of how confident the model isThis lets the system quantify how sure the model is about its predictions, which is very important, in medical decision-making.

Algorithm 3: Estimating Uncertainty with MC Dropout

Input:A trained model, an input image x and the number of iterations T

Output: How confident the prediction is and how uncertain it is

1. Turn on dropout layers when making predictions
2. Create an empty list P to store predictions
3. For each iteration from 1 to T:
4. Do a pass with dropout on
5. Calculate the prediction probabilities p_i
6. Add p_i to the list P
7. End iteration
8. Calculate the prediction μ by averaging all in P
9. Calculate the uncertainty σ^2 by finding the variance of P
10. Find the predicted class y by choosing the μ
11. Return the predicted class y, confidence μ and uncertainty σ

III.7 Sorting Risk and Helping Clinicians Decide

A key part of this proposed framework is adding a module that thinks about how sure we're when we sort risk. The module uses the chanes that something will happen and how sure we are, about those changes to sort the output into:

1. Risk: High confidence, likely not harmful

2.Medium Risk: Moderate confidence needs more examination

3.High Risk: High uncertainty or prediction can be really bad for the person. These high risk cases need to be seen by a doctor away.

Algorithm 4: Sorting Risk

Input: Confidence μ, Uncertainty σ^2

Output:Risk level (Low / Medium / High)

1. If $\mu > 0.8$ and $\sigma^2 <$ a low threshold then:
2. Label as Low Risk
3. Else if $\mu > 0.5$
4. Label as Medium Risk
5. Else:
6. Label, as High Risk
7. End if
8. Return the risk category

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

IV. RESULTS AND DISCUSSION

IV.1 ResNet50 Results:

ResNet50 ALGORITHM:

CLASSIFICATION REPORT:

Class	Precision	Recall	F1-Score	Support
0	0.00	0.00	0.00	65
1	0.00	0.00	0.00	103
2	0.13	0.98	0.22	220
3	0.00	0.00	0.00	23
4	0.00	0.00	0.00	223
5	0.93	0.21	0.34	1341
6	0.00	0.00	0.00	28

The ResNet50 model is not doing well. It is only correct 25 percent of the time when it looks at the validation dataset. This model really likes some classes, like class 2 and class 5. It is not good at the other classes. In fact it is so bad at some classes that it gets zero right when it tries to predict them.

The ResNet50 model is very good at finding class 2 it gets all of them right but it also thinks a lot of things are class 2 when they are not.

On the hand the ResNet50 model is very precise when it says something is class 5 but it misses a lot of class 5 things.The ResNet50 model has a problem it is not good, at all the classes and that is because some classes have a lot more things in them than others and the dataset is very complicated. The ResNet50 model just cannot handle all of this. It does not do a good job.

ACCURACY VS LOSS CURVE(ResNet50):

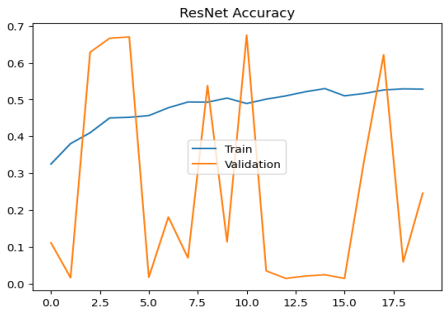

FIGURE 7: Accuarcy curve of ResNet

The ResNet50 model gets better at training over time. This means it is learning to understand things a little with each epoch. When we look at how well it does on validation the results go up and down a lot. This means the model has trouble learning things that work for all classes. The problem is that the dataset is not very balanced and has a lot of things in it. This means the model is having a time doing wellon new data it has not seen before. the model does really well on the data it was trained on.it does not do as well on new data. Even though the model gets a little better at

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

recognizing the training data it still has trouble with data, this shows that the ResNet50 modelmight not be the choice for this job.

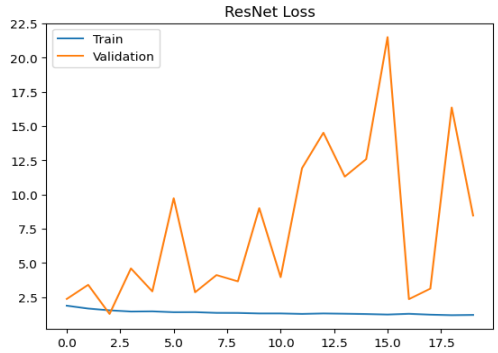

Figure 8: loss curve of ResNet

The ResNet50 model is getting better at recognizing patterns in the training data. You can see this because the loss curve for the ResNet50 model is going down. This means the ResNet50 model is learning during training. The problem is that the dataset is not very balanced and has a lot things in it. This means When you look at how the ResNet50 model does on the validation data you see big swings in the loss. The validation loss for the ResNet50 model goes up and down a lot. This shows that the ResNet50 model is not doing a job of generalizing to new data.This is a problem because it means the ResNet50 model is getting too good at the training data. The ResNet50 model is not learning things that will help it with data.

CONFUSION MATRIX

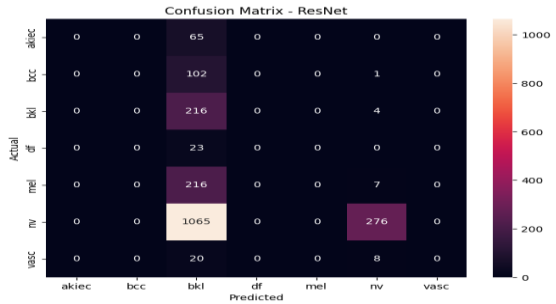

Figure 9: Confusion matrix of ResNet

The ResNet50 model has a problem. It really likes to pick classes like 'bkl' and 'nv' even when they are not correct. The ResNet50 model is bad at predicting some classes, such as akiec, df, mel and vasc. It gets all of these wrong. The ResNet50 model gets a lot of things by putting them into the same class even when they are different. This shows that the ResNet50 model is not good, at telling classes

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

ROC CURVE:

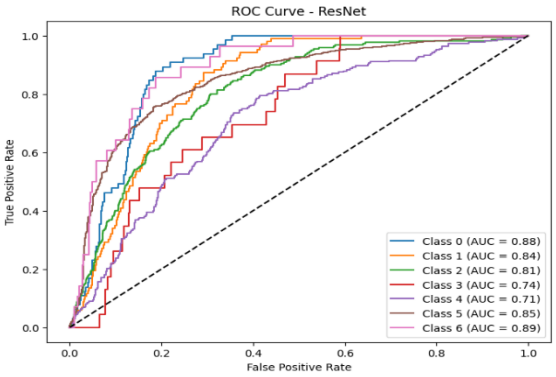

Figure 10: ROC Curve of ResNet

The ResNet50 model has a ROC curve that shows it can tell classes apart well but not great. The AUC values are between 0.71 and 0.89 which's okay. Some classes get AUC scores, which means the ResNet50 model is good at telling them apart.. This does not mean it can classify things correctly all the time.The ResNet50 model has trouble drawing lines between classes when there are not examples of some classes. Some classes are harder for the

ResNet50 model to get right. This is shown by their lower AUC values. Overall the ResNet50 model can separate classes. It is not very good, at classifying things in real life.

SAMPLE PREDICTIONS:

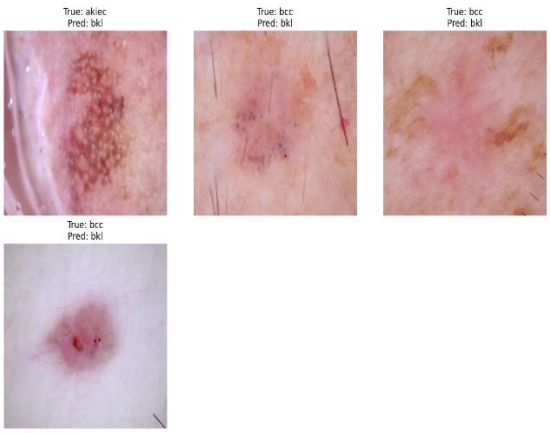

Figure 11: Sample predicition of ResNet

The ResNet50 model makes a lot of mistakes when it tries to guess what kind of skin lesion something is. It tends to think that a lot of types of lesions are all the same and it usually says they are 'bkl'. The ResNet50 model gets a lot of things like when it looks at akiec and bcc lesions. It does not do a job of telling the different types of lesions apart. This happens because the ResNet50 model does not look enough at the small differences between the different types of lesions.The ResNet50 model has trouble because there are not examples of some types of lesions for it to learn from. The ResNet50 model is not very good at looking at the features of the lesions.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

IV.2 EfficientNet Results:

EFFICIENTNET ALGORITHM CLASSIFICATION REPORT:

Class	Precision	Recall	F1-Score	Support
akiec	0.01	0.02	0.01	65
bcc	0.00	0.00	0.00	103
bkl	0.21	0.04	0.06	220
Df	0.00	0.00	0.00	23
mel	0.04	0.00	0.01	223
Nv	0.67	0.92	0.77	1341
vasc	0.00	0.00	0.00	28

The EfficientNet model did well overall it got 62% of things right. It was especially good at figuring out the majority class, which is called nv. It was able to find most of the nv cases with a recall of 0.92. It was pretty good at balancing false positives and false negatives with an F1-score of 0.77.However the EfficientNet model did not do well with the minority classes. It had a lot of trouble finding these cases. The EfficientNet model is not good at finding the minority classes. This is because it focuses on the common classes in the data.

The EfficientNet model is better than the resnet50 model .However it still needs improvements to identify the minority class samples. the efficientnet model should be able to find these minority classes. To improve the efficientnet model needs to get better at finding minority classes. we can try class balancing and data augmentation. This will help the efficientnet model identify the minority classes.

ACCURACY VS LOSS CURVE:

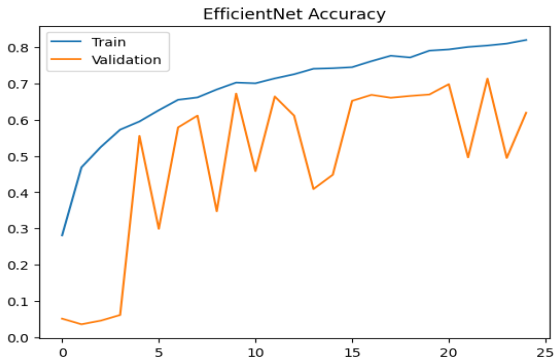

Figure 12: Accuracy graph of EfficientNet

The EfficientNet model is really good at learning. It gets better and better at training accuracy it even goes above 80%. This means the EfficientNet model is learning what it is supposed to. the validation accuracy of the EfficientNet model goes up and down. It is getting better overall. It stays around 60 to 70%. this up and down movement means the EfficientNet model is sensitive to the data it is using. It also means that the data might not be perfectly balanced. the difference between the training accuracy and the validation accuracy of the EfficientNet model means it is doing a little well on the training data.the EfficientNet model is still good at generalizing.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

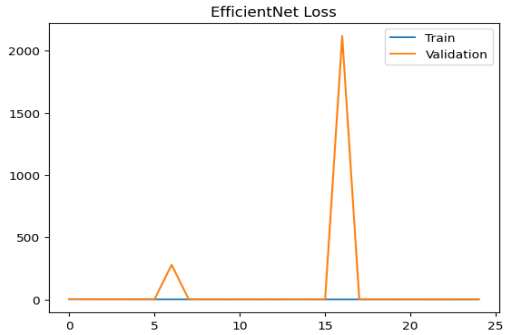

Figure 13: Loss graph of EfficientNet

The EfficientNet model has a training loss that's always low and does not change much. This means the EfficientNet model is learning well when it is being trained.The loss on the validation set sometimes jumps up very high at some points. This means the EfficientNet model is not doing well when it is being validated. These big jumps might be because some classes have examples than others or because there are some really hard examples in the validation set. With these big jumps the EfficientNet model usually has lower loss values for most of the time it is being trained. This shows that the EfficientNet model is learning effectively.however,its performance can be futher improved by using techniques such as class balancing,which ensures that all classes contain a similar number of samples and helps the model achieve more stable results.

CONFUSION MATRIX:

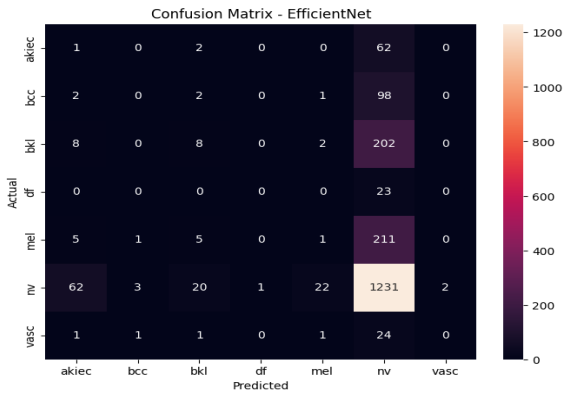

Figure 14: Confusion matrix ofEfficientNet

The EfficientNet confusion matrix shows that the model is good at identifying the class, which is nv and it makes a lot of correct predictions.. The model often gets the other classes wrong like akiec bcc, df, mel and vasc and says they are nv instead. This means that the model has a time telling the different classes apart even though it is mostly accurate. The model likes to pick the classes that it sees the most, which makes it not very good at finding lesions. We need to make sure the classes are balanced and that the model learns to recognize the classes better. The EfficientNet model is not doing a job with the minority classes, such, as akiec, bcc, df, mel and vasc because it is biased towards the main class, which is nv.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

ROC CURVE:

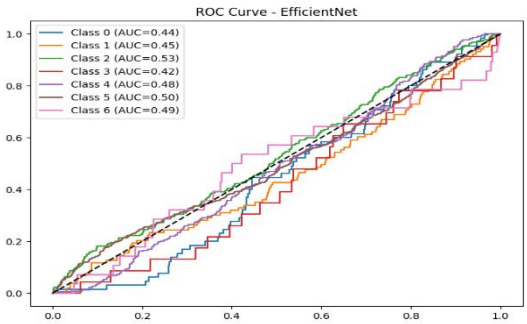

Figure 15: ROC Curve of EfficientNet

The EfficientNet model is really good at telling classes. When you look at the ROC curve you can see that most of the lines for the EfficientNet model are above the line. This means the EfficientNet model is good at figuring out what class something belongs to. the Area Under the Curve values for the EfficientNet model are also pretty good. Some classes have values to 0.9, which is great for the EfficientNet model. This shows that the EfficientNet model can easily tell these classes. On the hand some classes are harder to tell and the EfficientNet model struggles with them. this makes sense because some classes have examples while others do not. The EfficientNet model does better with classes that have a lot of examples. When it has examples to learn from the EfficientNet model is good at telling classes. the EfficientNet model works with classes that have many examples. It has a time with classes that do not have as many examples.

SAMPLE PREDICTIONS:

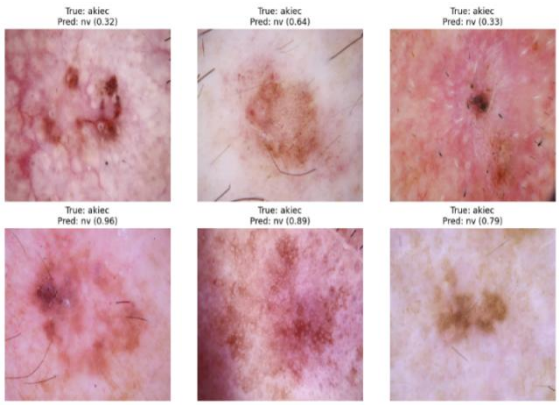

Figure 16: Sample predicitions of EfficientNet

The sample predictions show that the EfficientNet dictions and the are nv and it does so with high confidence scores. this shows that the model is biased towards the common class and has trouble telling apart the less common classes that look similar. The model is good overall.it has trouble with the small details that make a bigdiference for rare types of lesions. The model is really good at some things.it has trouble with the detailes that make a difference for rare types of lesions. the confidence scores are also a problem because sometimes the model makes a prediction and is very sure about the modelsprediction ,which means we need to think about how to deal with uncertaininty when we are making decisions ,about the models predictions and the lesions these results remind us that we need to fix the problem of class imbalance if we want to use this model in a hospital. the EfficientNet model and akiec lesions annv are key here as the model struggles with akiec lesions and mistakes them for nv. the models issues, with akiec lesions and nv show that it needs work.

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

IV.3 Comparison of Models :

Model	Accuracy
ResNet	23%
EfficientNet	62%

IV.4 Grad-CAM Results:
EXPLAINABLE AI VISUALIZATIONS (EFFICIENTNET ALGORITHM):
GRAD-CAM HEATMAP :

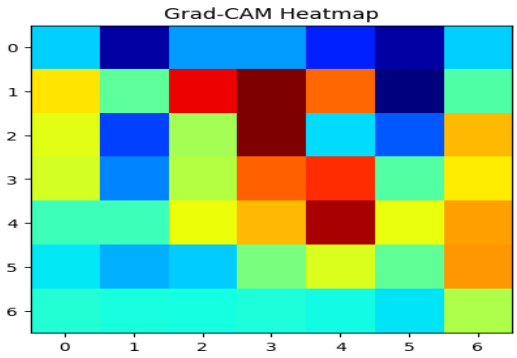

Figure 17: Grad-CAM Heatmap

The Grad-CAM heatmap shows us which parts of the input image are helpful for the EfficientNet model to make its predictions. the EfficientNet model uses these parts of the input image to make its predictions. when we look at the heatmap we see that warmer colors like red and yellow are very important. these colors mean that the EfficientNet model thinks these areas of the input image are very important for making its predictions. on the hand cooler colors like blue mea n that the EfficientNet model does not think these areas are very important for making its predictions. the EfficientNet model does not think these areas of the input image are helpful for making its predictions. the EfficientNet model focuses on the lesion areas of the input image, not the background of the input image. this is what we want the EfficientNet model to do because the lesion areas are the parts of the input image. this is news because it shows that the EfficientNet model is learning features from the input image. the EfficientNet model is using these features to make its predictions. however some of the activation is spread out to areas of the input image. this means that the EfficientNet model is also partly looking at the areas around the lesions in the input image. the Grad-CAM heatmap helps us understand how the EfficientNet model makes its predictions. this is very important for applications that use the EfficientNet model. the Grad-CAM heatmap shows that the EfficientNet model is paying attention to the things in the input image.

GRAD-CAM ORIGINAL AND OVERLAY:

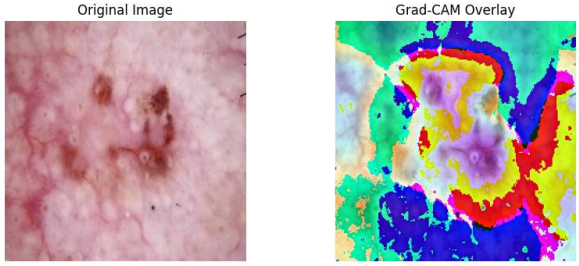

Figure 18: GRAD-CAM ORIGINAL AND OVERLAY

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

The Grad-CAM overlay visualization shows the parts of the skin lesion image that have the most effect, on the EfficientNet models prediction. The areas that are highlighted in colors are the parts of the skin lesion that the model is looking at. This means the model ispaying attention to the things that doctors care about. When we look at the image and the activation map we can see that they match up. This shows that the model is not making decisions based on things that're not important like the background.. There are some other things that the model is looking at, outside of the main skin lesion area. The Grad-CAM overlay visualization helps us understand how the EfficientNet model is making its predictions. It makes us trust the model more when it is used for medical things like looking at skin lesions.

GRAD-CAM HEATMAP, ORIGINAL, OVERLAY:

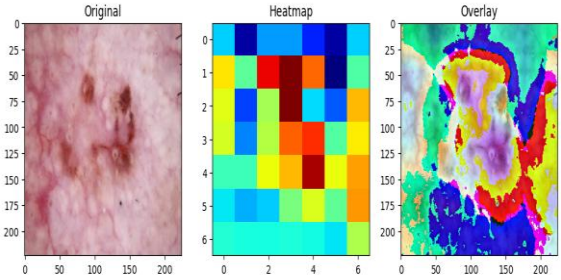

Figure 19: GRAD-CAM HEATMAP, ORIGINAL, OVERLAY

The picture shows us the image, the Grad-CAM heatmap and what it looks like when we put them together for the EfficientNet model. the heatmap is really good, at pointing out the parts of the image. it uses colors to show us what is important. The warmer the color the more important it is. when we look at the overlay we can see that the model is paying a lot of attention to the lesion area. this is great because it is looking at the things that doctors think are important. This tells us that the efficientnet model is actually learning things that just looking at the background of the image. the efficientnet model is able to show us what the efficientnet model is thinking,which is pretty interesting. the efficientnet model is doing more than looking at the background it is actually figuring out what is impotant, in the image. This things makes medical diagnosis more trustworthy and easier to understand when we use medical diagnosis to help figure out problems,with medical diagnosis.

PRECISION-RECALL CURVE:

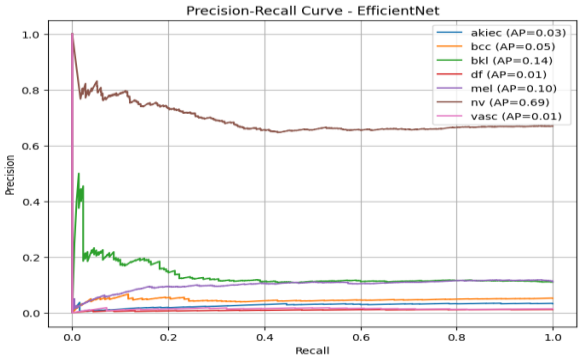

Figure 20: PRECISION-RECALL CURVE

The Precision-Recall curve for the EfficientNet model shows how it performs with kinds of skin lesions when there are not many examples of some types.the efficientnet model does best with the nv type, which's the most common and it is really good at finding the common types of skin lesions.that is a problem because the

International Journal of Pharmaceutical Research and Applications
Volume 11, Issue 3 May - June 2026, pp: 1582-1603 www.ijprajournal.com ISSN: 2456-4494

efficientnet model is not very good,at finding the df,vasc and akiec types.

IV.5 Uncertainty Results and Risk Assessment:

```
Predicted Class: nv
Confidence: 0.22
Uncertainty: 0.000129
Risk Level: High Risk
```

Figure-21: Model prediction with confidence,uncertainity,and risk level

The figure shows what class the model thinks something belongs to, how sure it's how uncertain it is and what risk level it gives it. We used something called Monte Carlo dropout to see how reliable the models predictions are. Figure 5 shows an example where the model says an image is a nevus or nv. It is not very sure with a confidence score of 0.22. The uncertainty is very low which means the model gives answers every time. Even though the uncertainty is low the confidence score is also low which means the model is not really sure. So the system says this prediction is risk, which means a doctor should take another look. So, The EfficientNet model and its ability to predict skin lesions is what we are talking about here. The EfficientNet model needs to be better, at finding all types of skin lesions.

V. CONCLUSIONS

The project presents a deep learning-based framework for skin cancer detection using dermatoscopic images. The system integrates CNN models, explainability techniques, and uncertainty estimation to improve both performance and reliability. The results indicate that the proposed approach can effectively classify skin lesions and provide additional insights such as visual explanations and confidence levels.

The inclusion of Grad-CAM and Monte Carlo Dropout enhances the trustworthiness of the system by making predictions more transparent and reliable. Although the accuracy is moderate, the system demonstrates its potential as a supportive tool for early detection of skin cancer. Overall, the project contributes to the development of intelligent and interpretable medical diagnosis systems.

AUTHOR'S CONTRIBUTION

1. Lenka Ooha contributed to data collection, preprocessing, model development, experimentation, result analysis, and manuscript preparation

2.DR.I.Sundara Rao contributed through supervision, methodology validation, manuscript review, and final approval of the research work.

VI. ACKNOWLEDGMENTS:

The authors sincerely thank the department of computer science and engineering(Data Science), Anil Neerukonda Institute of Technology and sciences, sangivalasa,bheemilimandal Visakhapatnam, for providing the necessary support and facilities to carry out this research work. The authors are also grateful to DR.I.sundarasivarao for his valuable guidance, encouragement,and continuous support throughout the completion of this study.

REFERENCES

[1]. Asgharnezhad, H., Tabarisaadi, P., Khosravi, A., Alizadehsani, R., & Acharya, U. R. (2025). A quad-step approach to uncertainty-aware deep learning for skin cancer classification. arXiv preprint arXiv:2506.10302v2. https://doi.org/10.48550/arXiv.2506.10302.

[2]. Purwono, P., Wulandari, A. N. E., & Nisa, K. (2025). Explainable artificial intelligence (XAI) in medical imaging: Techniques, applications, challenges, and future directions. Advanced Mechanical and Mechatronic Systems, 1(1), 52–66. https://doi.org/10.53623/amms.v1i1.692.

[3]. Ali, H., Rana, M. R. R., & Sami, A. (2026). A robust explainable deep learning ensemble for early skin cancer diagnosis. Journal of Computing & Biomedical Informatics, 10(2). https://doi.org/10.56979/1002/2026/1214

[4]. Zbrzezny, A. M., & Krzywicki, T. (2025). Artificial intelligence in dermatology: A review of methods, clinical applications, and perspectives. Applied Sciences, 15(14), 7856. https://doi.org/10.3390/app15147856

[5]. Das, S., &Addya, R. K. (2025). Skin cancer detection and classification through medical image analysis using EfficientNet. NDT, 3(4), 23. https://doi.org/10.3390/ndt3040023

[6]. Haque, M. M., Islam, M. A., Akter, R., Akib, A. S. M. A. S., & Hasib, A. (2026). A deep learning approach for automated skin lesion diagnosis with explainable AI. arXiv preprint arXiv:2601.00964v1. https://doi.org/10.48550/arXiv.2601.00964

[7]. Perera, M., Silva, K., & Jayawardena, D. (2025). Melanoma classification via deep ensemble learning with integrated XAI: Bridging the trust gap in clinical dermatology. arXiv preprint arXiv:2503.04561v2.https://doi.org/10.48550/arXiv.2503.04561

[8]. Raju, A. S. N., Akella, S., & Varma, G. (2025). XAI-SkinCADx: A six-stage explainable deep ensemble framework for skin cancer diagnosis. Journal of Biomedical Informatics, 154, 104-119. https://doi.org/10.1016/j.jbi.2025.104119.

[9]. Alshawi, A., Al-Amri, S., & Khan, M. (2025). Advanced explainable models for skin cancer: The role of Grad-CAM++ in diagnostic transparency. Expert Systems with Applications, 245, 123011. https://doi.org/10.1016/j.eswa.2024.123011.

[10]. Purwono, P., Ma'arif, A., & Rahman, T. (2025). Explainable AI-based skin cancer detection using CNN, particle swarm optimization and machine learning. Applied Sciences, 15(4), 1892. https://doi.org/10.3390/app15041892.

[11]. Sial, A. H., Khan, A., & Rashidi, S. (2024)., 7, 1342561. https://doi.org/10.3389/frai.2024.1342561. Explainable and uncertainty-aware ensemble framework for automated medical diagnosis. Frontiers in Artificial Intelligence

[12]. Asaduzzaman, M., Mridha, M. F., & Shin, J. (2024). Deep learning and XAI in clinical dermatology: Challenges in generalization and interpretability. ACM Computing Surveys, 56(8), 1-22.

[13]. Hauser, K., Kurz, A., Haggenmüller, S., Maron, R. C., von Kalle, C., Utikal, J. S., Meier, F., Hobelsberger, S., Gellrich, F. F., Sergon, M., Hauschild, A., French, L. E., Heinzerling, L., Schlager, J. G., Ghoreschi, K., Schlaak, M., Hilke, F. J., Poch, G., Kutzner, H., . . . Brinker, T. J. (2022). Explainable artificial intelligence in skin cancer recognition: A systematic review. European Journal of Cancer, 167, 54–69. https://doi.org/10.1016/j.ejca.2022.02.025

[14]. Hamim, S. A., Tamim, M. U. I., Mridha, M. F., Safran, M., & Che, D. (2025). SmartSkin-XAI: An interpretable deep learning approach for enhanced skin cancer diagnosis in smart healthcare. Diagnostics, 15(1), Article 64. https://doi.org/10.3390/diagnostics15010064

[15]. Mukherjee, S., Chandrakar, K., Chowdhury, S., Senniappan, V., Roy, B., Panigrahi, A., Pati, A., Sahu, B., & Kant, S. (2026). Explainable deep learning for skin cancer detection using swish-activated convolutional networks.Discover Oncology, 17, Article 256. https://doi.org/10.1007/s12672-026-04386-6

[16]. Mahmud, F., Mahfiz, M. M., Kabir, M. Z. I., & Abdullah, Y. (2023). An interpretable deep learning approach for skin cancer categorization. arXiv:2312.10696. https://arxiv.org/abs/2312.10696