

## Enhancing Cancer Diagnosis with Advanced Vision Transformers and Hybrid Deep Learning Models

Sabitha P Assistant Professor,  
Department of Computer  
Science and Engineering  
*SRM Institute of Science and  
Technology, Ramapuram,  
Chennai.*  
[sabithap@srmist.edu.in](mailto:sabithap@srmist.edu.in)

Lokesh Kumar K UG Student,  
Department of Computer  
Science and Engineering  
*SRM Institute of Science and  
Technology, Ramapuram,  
Chennai*  
[ks8374@srmist.edu.in](mailto:ks8374@srmist.edu.in)

Jagadeesh P UG Student,  
Department of Computer  
Science and Engineering  
*SRM Institute of Science  
and Technology,  
Ramapuram, Chennai.*  
[jp5505@srmist.edu.in](mailto:jp5505@srmist.edu.in)

Varun K UG Student,  
Department of Computer  
Science and Engineering  
*SRM Institute of Science and  
Technology, Ramapuram,  
Chennai.*  
[kk8833@srmist.edu.in](mailto:kk8833@srmist.edu.in)

**Abstract:** The accurate and prompt diagnosis of cancer alongside other deadly diseases is crucial for achieving early stage diagnosis and subsequently, improved long term survival rates. Along with these benefits, patients get the additional perk of being treated in a lesser invasive manner, which is less physically taxing considering how intricate the procedure can be. It is important to note that AI powered diagnostics systems have shown great promise when used in conjunction with medicine and still undergo rigorous research in hybrid deep learning frameworks for computerized medical imaging alongside techniques like Convolutional Neural Networks and Vision Transformers. The aim of the current project is to combining VITs and CNNs in order to create a more reliable, robust and effective hybrid deep learning model that processes images in order to more accurately diagnose cancer. Such models suffer from these intricate issues with patterns such as defaults in classifying eye images, scooped-eye images, and face images being over-scooped. The core contribution of this study lies in the development of advanced techniques for enhanced precision in diagnosing cancer through the application of hybrid architectures as well as the integrated AI techniques, expanding on the legacy of AI powered driven systems. Along these lines, my reconnaissance of the problem is that distinct from VITs and CNNs which confirm the enhanced performance of hybrid models for complex problems, there has been an insufficiency in researching inner workings of the models and diagnosing challenging tasks alongside images of intricate patterns buddy dolly optics 3d snap-focus lenses. As discussed previously, initial steps to effectively guarantee advanced survival chances alongside treatments need an accurate and early detection of cancer, ultimately proving beneficial in prolonging a person's life.

**Keywords:** Cancer Diagnosis, Vision Transformers, Deep Learning, Hybrid Models, Medical Imaging.

## Introduction

Healthcare has been completely transformed by computerised medical image processing, which offers a complete solution for automated disease diagnosis. Combining the set techniques Not only can medical imaging save money and time, but it also increases doctors' trust in patient diagnosis, especially when it comes to skin cancer (1). A common and possibly fatal condition that affects millions of individuals globally, skin cancer is caused by aberrant cell development brought on by extended exposure to ultraviolet light (2). Among all types of skin cancer (MM), malignant melanoma is among the most destructive. White persons are more likely to have it, and depending on the area, men are more likely than women to have it (3).

Melanoma is very difficult to diagnose because of its similarities and overlap with other diseases. Via the use of self-attention processes, Long-range dependencies can now be represented with Vision Transformers (ViTs), promising substitute. Their computing expense and data requirements, however, present numerous challenges. A hybrid approach that combines CNNs and ViTs can leverage the benefits of both designs to increase the precision and effectiveness of cancer diagnosis.

This paper presents a novel hybrid framework that combines CNNs for spatial feature extraction with ViTs for global attention modeling, boosting diagnostic accuracy. the 98% survival rate of the epidermis, the skin's outermost layer (5) The top layer of the inner skin, known as the dermal layer, has been affected by the cancer when it reaches stage I. The new rate of survival is 81% (6). A 78% decrease in survival rate is linked to stage II melanoma progression into the deeper dermal layers (6). In stage III, cancer cells had a 63% chance of surviving after spreading beyond the epidermis to the lymphatic nodes (5). Because stage IV cancer has spread to distant organs, its chances of survival are only 22% (5). the heart, brain, and tissues. Therefore, early detection and therapy result in improved chances of survival, but advanced melanoma stages lower regeneration and chance of survival (7,8). Another study points to the potential for melanoma to recur. In order

to guarantee that therapeutic therapies are effective within five years of treatment, this calls for careful monitoring (9,10). Traditionally, melanoma has been diagnosed mostly by the subjective evaluations of physicians, a process that is inevitably influenced by differences in training, experience, and physician interpretation (11). Dermatoscopy is the non-invasive diagnostic method most frequently employed by dermatologists to evaluate lesions and screen for skin cancer (12).

## Literature Review

Deploying deep learning algorithms in the sorcery of medical image processing has produced astonishing results with respect to the effectiveness and precision of cancer diagnosis. The equals of these transcendental feats, which integrate deep learning, have been executed through the application of CNNs on radiological and pathological imaging where they have been used for features extraction, segmentation, and image classification. That being noted, contemporary progress in medical imaging transforms global feature representation using self-attention processes and Vision Transformers (ViTs) is shifting the whole landscape. This chapter attempts to cover the available literature on the deep learning- based cancer diagnosis, focusing primarily on the roles of hybrid deep learning architectures.

## Vision Transformers in Cancer Diagnosis

### Overview of Vision

#### Transformers

Since their initial development for natural language processing, Vision Transformers (ViTs) have been modified for image- based applications such as medical imaging. Unlike conventional Convolutional Neural Networks (CNNs), which use local receptive fields, ViTs use self-attention processes to identify long- range correlations in images. This results in their capacity to understand complex medical images, including histopathology slides, MRIs, and CT scans.

## **Applications in Cancer Diagnosis Breast**

### **Cancer Detection**

Early identification is essential for improving survival rates for breast cancer, one of the most prevalent cancers worldwide. ViTs have demonstrated remarkable improvements in specificity and accuracy when used to detect breast cancer in recent studies. A twin-stream approach that integrated a ViT model with deep learning networks performed better than single-stream CNN models on the BACH dataset, with a 99.07% specificity and a mean accuracy of 98.60%.

### **Lung Cancer Classification**

The diagnosis of lung cancer depends on the ability to On CT images, differentiate between benign and malignant nodules. Although conventional CNN models have shown promise in this area, ViTs provide improved performance through the capture of global contextual information. In a study that was published in PubMed, ViTs outperformed traditional CNN-based techniques in the classification of lung cancer types and the differentiation of benign from malignant pulmonary nodules.

### **Skin Cancer Detection**

Dermoscopic images present unique Because lesion shapes and appearances vary, dermoscopic pictures pose special complications. ViTs have proven that they can improve the diagnosis of skin cancer by using self-attention processes to pick up on minute variations in lesion shapes. Research has indicated that ViTs outperform CNN- based models in identifying melanoma and other skin malignancies.

## **Hybrid Deep Learning Models in Cancer Diagnosis**

### **Integration of ViTs with CNNs**

CNNs and ViTs are combined in hybrid deep learning models, which have become effective tools for cancer diagnosis. The integration enables the use of both the global feature

representation of ViTs and the local feature extraction capabilities of CNNs.

### **Analysis of Histopathological Images**

In the study of histopathological images, correct classification of tumor margins is very important. The detection of squamous cell carcinoma borders has been improved using a CNN-ViT hybrid model, even in poorly graded histopathology images. Compared to stand-alone CNNs, the model demonstrated higher accuracy and reliability against noise.

### **Multi-Modal Cancer Diagnosis**

Hybrid models also facilitate multi-modal cancer diagnosis by integrating imaging data with other modalities such as genomics and clinical records. A study demonstrated that combining ViTs with recurrent neural networks (RNNs) for multi-modal learning significantly enhanced cancer prediction accuracy.

### **Methodology**

Three primary parts make up the suggested hybrid model:

#### **CNN Feature Extractor**

To identify spatial patterns in medical images, a CNN that has already been trained (such as ResNet-50) is used as a feature extractor. The CNN creates feature maps that retain important diagnostic information after processing input images through a number of convolutional and pooling layers.

#### **Vision Transformer Encoder**

The extracted feature maps are passed to a Vision Transformer encoder, which applies self-attention mechanisms to model global dependencies. The ViT component processes tokenized feature representations, enabling comprehensive pattern recognition beyond the local features captured by CNNs.

#### **Fusion and Classification**

The outputs from CNN and ViT components are concatenated and processed

through a fully connected layer for classification. Dropout and batch normalization techniques are employed to prevent overfitting and improve generalization.

### Training and Optimization

Lung cancer CT scans and breast cancer histopathological image collections are among the publicly accessible cancer datasets used to train the classifier. Adam optimization is used in conjunction with an adaptive learning rate during training. The robustness of the model is increased by using data augmentation techniques like flipping, rotation, and contrast enhancement.

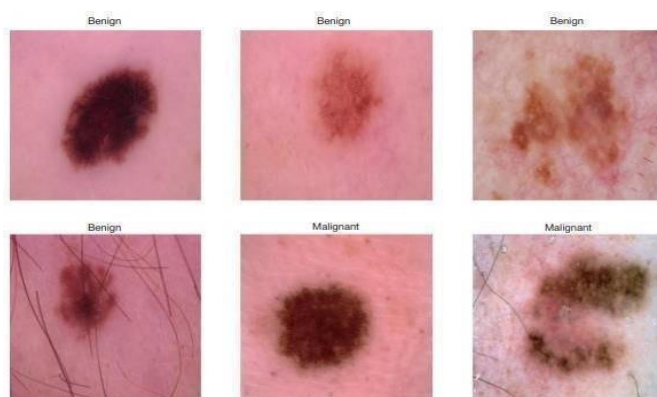
## Methods

### Data Sources

The dataset used to train and test the CNN models and the ViT models in this project comes from the ISIC dataset (50), which is publicly available on Kaggle (51). The complete dataset contains 3,297 photos, out of which 2,637 were allocated for training and 660 for testing. Example photos are shown in Figure 1, whereas the data distribution is provided in Table 1.

**Table 1** Data distribution

Class	Training	Testing
Benign	1,440	360
Malignant	1,197	300
Total	2,637	660



**Figure 1** Images from the dataset that are classified as benign or malignant

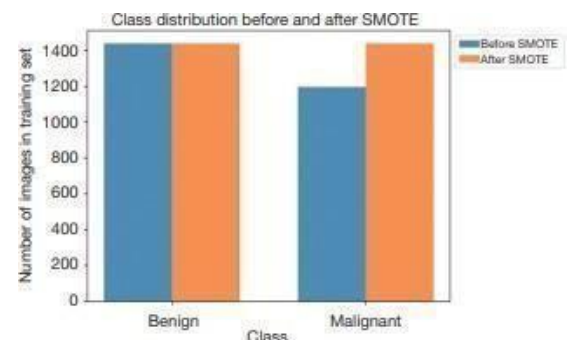
## Performance Metrics

### Class Imbalance

As evidenced by Table 1, there is a class imbalance in terms of the dataset; hence, the oversampling method, such as the synthetic minority oversampling technique (SMOTE) (52), is used to create synthetic samples for the minority classes to rectify the unequal data. SMOTE alleviates the overfitting issue in random oversampling. Here, SMOTE is applied only to training data; its effects before and after are shown in Figure 2. According to this figure, there were 1,197 cancerous images before the implementation of SMOTE; now, there are as many benign images. SMOTE, therefore, improves classification accuracy as well as generalizability when applied on unbalanced data by removing bias from the model (53).

### Pre Processing

The ideal size for creating a patch from the input photos is 224×224 pixels, thus we resized every image in the dataset to that size.



**Figure 2** Class distribution prior to and following SMOTE. SMOTE stands for synthetic minority oversampling.

### Proposed Method

The researchers distinguished between benign and malignant skin cancer while using CNN coupled with ViT. The transfer learning technique applies to pretrained CNN and ViT models through their adjusted configurations. The ImageNet dataset served for model preprocessing



before the researchers applied these models to skin cancer classification. The key attributes learned by the best-performing model undergo a final assessment through three XAI models based on CAM.

### ViT

ViT has become popular for medical picture classification tasks while showing excellent results in this domain (46,54,55). ViT takes a one-dimensional format of word tokens to serve as input for the transformer architecture that NLP models use as their foundation. ViT divides images into patches before transforming them into tokens because its system functions in two dimensions. The patch function divides pictures into equivalent square sections which measure 16 by 16 pixels (56). The input image characteristics are partitioned into sectors that become vectorized for representation. A CNN is typically used to obtain the patch's features, which the vector represents. After that, the vectors from each patch are sent to a self-attention layer and a transformer encoder layer. Long-term associations between patches can be found by the model thanks to the self-attention technique. In order to classify the input image, the transformer encoder first creates a set of vectors that represent the features of the image. In Figure 3, the general architecture of the ViT model is displayed. Table 2 gives the precise parameters for each type of ViT model.

To increase computational performance, Swin transformer (57) employs shifted windows, and a hierarchical feature is used to produce multi-resolution feature maps. After being divided into several non-overlapping patches, the input images are transformed into embeddings in the Swin transformer. After that, In order to maintain hierarchical The Swin transformer blocks operate on the patches through four stages which successively reduces their numbers. The Swin transformer block features two sub-modules that contain window multi-head self-attention (SW-MSA) and shifting instead of multi-head attention like ViT models use W-MSA. The figure demonstrates the entire Swin transformer architecture (Figure 4) yet Table 3

specifies each Swin transformer model's lead requirements. This research utilized two ViT models namely ViT-base and ViT-large together with Swin transformer models Swin- small, Swin-base, and Swin-tiny. CNN Models.

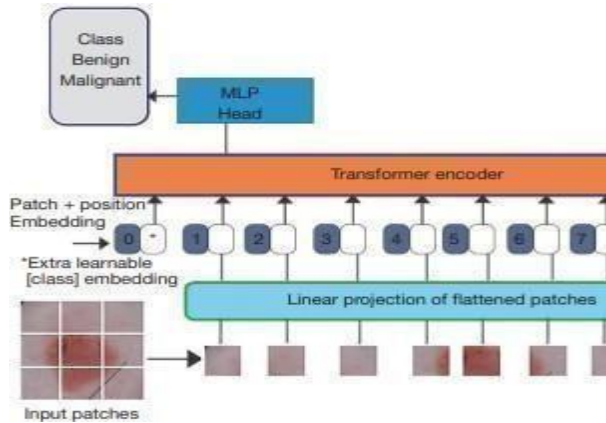
DL and three visual-based XAI are used in this paper's proposed transfer learning- based XAI for skin cancer categorization. ResNet50 (58), VGG19 (59), DenseNet201 (60), and MobileNetV2 (61) are pre-trained DL models that we employ. At first, ImageNet was used to train each of those models. VGG19 is utilized in skin cancer classification publications and performs well on the Image Net dataset

(62). One of the greatest models

for medical image classification is DenseNet201, which can handle picture variances (63). The tiny parameter of MobileNetV2 makes it appropriate for embedded devices and minimal processing demands. This has led to the development of a lightweight model for numerous medical picture classifications using MobileNetV2, which produces improved results (64,65). It might not function well on extremely complicated datasets, though. A popular model for skin cancer classification tasks, ResNet50 stays clear of disappearing gradients (62,66,67). For our skin cancer categorization problem, we therefore used those models as a basic model. The suggested models cover preprocessing, XAI, feature extraction and classification, and assessment using five metrics.

### XAI Models

XAIs are essential for fostering trust in automatic skin disease classification jobs because they highlight the areas of a picture that most influence whether the image is benign



**Figure 3** Skin cancer classification using the ViT multilayer model architecture Perceptron; ViT, vision transformer.

cancerous. A number of XAI models were created by researchers to explain black box AI functionality. Consider LIME (37). Five prominent XAI models include CAM (68) Score-CAM (70) Grad-CAM (69) and F-CAM (71) and Grad-CAM++ (72) which visualize deep learning algorithm prediction logic. The implementation of XAI methods enhances both model decision process

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

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

$$Sensitivity = \frac{TP}{TP + FN}$$

$$F1-Score = \frac{2 * TP}{2 * TP + FN + FP}$$

$$Specificity = \frac{TN}{TN + FP}$$

analysis and visualization of predictive

**Table 4** Performance the ViT models in the testing set with five evaluation n

Models	Accuracy (%)	Precision (%)	Recall
ViT-b	88.6	88.6	
ViT-L	88.6	89.2	
Swin-s	87.7	89.5	
Swin-b	87.8	86.6	
Swin-T	87.7	90.1	

ViT, vision transformer.

**Table 4** shows that ViT-b and ViT-L

explanations. Medical imaging benefits from XAI because it shows users how algorithms make decisions through straightforward explanations that patients can easily understand. Neural networks deliver outcomes with a higher degree of reliability reaching 73 percent (73). The research employs three visual-based EXA models composed of Score-CAM as well as Grad-CAM and Grad-CAM++. The research team explored only Score-CAM and Grad-CAM because Grad-CAM++ serves as an updated version of Grad-CAM. Grad-CAM visualization determines output gradient through analysis of selected convolutional features in image classification (74). Weight values reach the 2D activations through the averaged gradient distribution method. Grad-CAM requires the following formula to generate images for a particular target class  $c$  during creation of Grad-CAM:

### Setup for the Study

This research utilizes various Swin transformer models together with two ViT models to develop pretrained models for identifying benign and severe skin malignancies. Five pretrained CNN models undergo performance assessments after which the results are compared with those from ViT models. This paper demonstrates three visual-based XAI model applications on ResNet50 while illustrating how the model determines the hazard level of each picture. A performance comparison with related research was included as our final step.

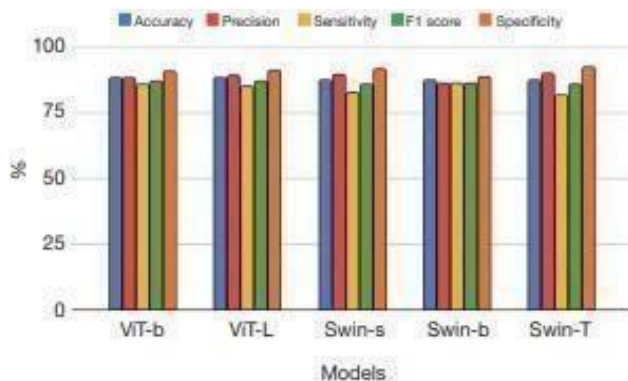
### Specifics of Implementation

The models are powered by Google Colab Pro+ and its 16 GB GPU addition to its 83.5 GB system RAM. All models run on 83.5 GB of system RAM alongside their operation using 32 as the batch size. During 30 training epochs the models receive training through the Adam optimizer at a learning rate of 0.001. Research indicates the use of 75 for these parameters.

Additionally, all models employ a batch size of 32.

## Performance Metrics

In this work, we measure the robustness of the CNN and ViT models using five metrics: accuracy, sensitivity, specificity, precision, and F1 score.



**Figure 4** ViT models' performance. Vision Transformer, or ViT.

The four components used to describe the performance of a test are TP for true positive and TN for true negative alongside FP for false positive and FN for false negative.

## Results And Discussion

A comparison of skin cancer classification performance between ViT-b, ViT-L, Swin-s, Swin-b, and Swin-T is presented in Table 4 through evaluation metrics. A total of 88.6% accuracy was achieved by both ViT-b and ViT-L as demonstrated in Table 4. Swin-b (87.8%) occupies the second position with Swin-s (87.7%) and Swin-T (87.7%) trailing behind. Swin-T displayed the highest precision value with 90.1% when running these tests. Swin-s demonstrates a rate of 89.5% in the results but ViT-L has outperformed it by reaching 89.2%. Round figures indicate that Swin-b and ViT-b produced equivalent 88.6% result scores during precision tests. Swin-b demonstrated the best result among all models with a recall rate of 86.6% but ViT-b achieved the top level of 86.0%. ViT-L reached a recall score of 85.3% and Swin-T scored 82.0% while Swin-s ended up at 82.6%. The F1 scores for ViT-b

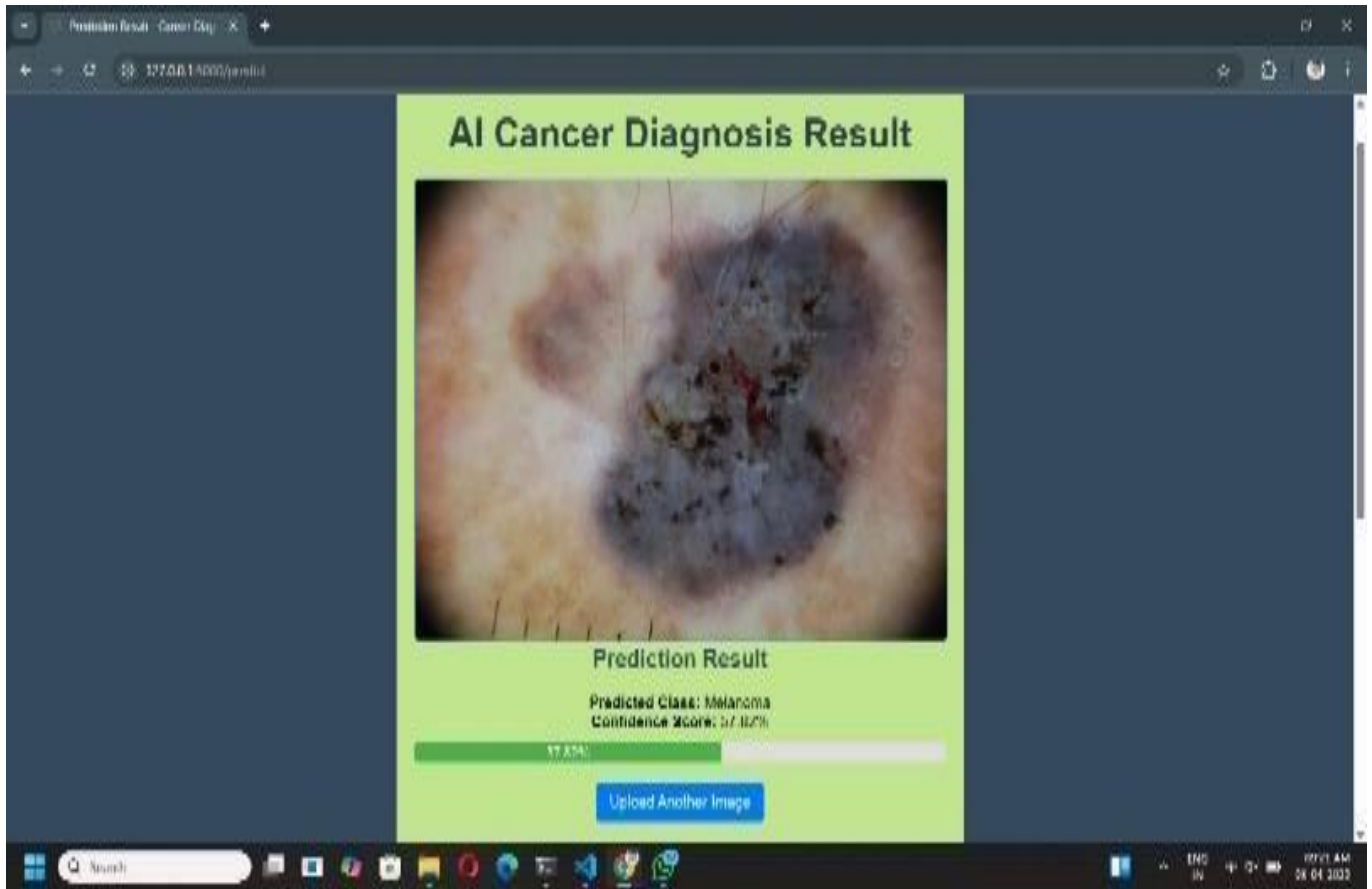
reached 87.3% while ViT-L earned 87.2% as its exceptional score. The F1 scores obtained by Swin-s, Swin-T and Swin-b were very close at 85.9%, 85.8% and 86.6%. A specific test evaluation ranked Swin-T as the first with a 92.5% score and placed Swin-s in second position with 91.9% performance. The specificity results of ViT-L, Swin-b, and ViT-b demonstrated 90.8%, 88.8% and 91.3% respectively. All of the ViT models demonstrate exceptional overall performance but Swin-T shows early signs of becoming an effective method for achieving high precision and specificity. A graphical representation in Figure 5 shows the test set evaluation of the different models. Table 4 presents a comparison of the performance of five models: ViT-b,

**Table 5** Performance CNN models in the testing set with five evaluation

Models	Accuracy (%)	Precision (%)
Vgg19	85.6	81.9
Resnet18	82.4	80.5
Resnet50	88.8	86.9
Densnet201	83.2	85.1
Mobilenetv2	85.3	83.5

CNN, convolutional neural network.

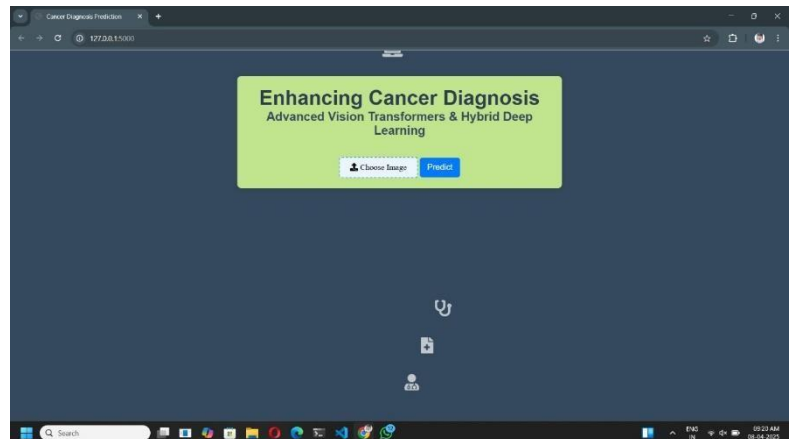
The detection accuracy of Swin-s and Swin-T along with Swin-Color reached 87.8% staggering above the 87.7% of Swin-Color BRA. The precision score of Swin-T exceeded 90.1%. ViT-L reaches 89.2% class accuracy and stands above Swin-s with its second position at 89.5%. Both Swin-b and ViT-b demonstrated equivalent precision values of 88.6%. The performance metric recall reached 86.6% for Swin-b while ViT-b achieved 86.0% recall. The recall scores of 85.3% were obtained by Swin-T together with Swin-s and ViT-L. 82.6%, and 82.0%, respectively. Furthermore, ViT-L received an outstanding score of 87.2%, and ViT-b received a respectable F1 score of 87.3%.



**Figure 6** Prediction of Skin cancer

enhances the Grad-CAM method's localization accuracy by adding higher-order derivatives. We want to increase the granularity of finding the crucial areas of skin photos that affect ResNet50's classification choices by using Grad-CAM++. By assigning relevance values to particular pixels according to how they affect the final forecast,

Score CAM significantly enhances the interpretability spectrum. The resulting attribution map allows for a fine-grained examination of ResNet50's decision-making process by providing comprehensive insight into pixel-level relevance.



**Figure 8** Uploading the skin disease image for prediction

The gradient-weighted class activation mapping is known as Grad-CAM while Score-CAM stands for score-weighted class activation mapping. Cancer is provided by the use of ViT models, five CNN models, Grad-CAM, Grad-CAM++, and Score-CAM approaches. The automatic skin cancer detection systems' interpretability and resilience are enhanced by these



classification and visualization techniques. ViT models offer a distinct viewpoint on skin categorization by extracting global contextual information from photos of skin lesions, which helps the models identify minute patterns that are essential for a precise diagnosis. CNN models, on the other hand, have been widely used in medical image classification tasks because they are adept at understanding local features.

In order to

facilitate the automatic diagnosis of skin cancer, these models are applied in the categorization process. Additionally, the CNN model's decision-making process is easier to grasp when Grad-CAM, GradCAM++, and Score-CAM are used. being investigated to improve model performance. AlexNet transfer learning was demonstrated in article (78), with an 87.1% classification accuracy. In a different work (79), a transfer learning method was s

uggested, and it was stated that VGG16, ResNet50, and Xception had respective accuracy rates of 86.5%, 81.6%, and 90.9%. These studies, however, did not include any explainability techniques and instead used black box models. In contrast, a study (38) used a bespoke CNN with 82.7% accuracy and used the CAM approach to make it more explainable. CAM, on the other hand, is architecturally reliant and only functions when

the model has a GAP layer. A study (80) suggested employing EfficientNets B0-B7 for transfer learning and EfficientNetB4 for 87.91% accuracy without the use of explainability techniques. CAM-based XAI techniques and showed that the model can identify benign or malignant areas

in the supplied image by picking up knowledge from such areas. The models still need additional fine-tuning or architectural modifications to increase their functionality, even though their overall performance is

strong. We will therefore test and validate the models using fresh data in subsequent work, test alternative XAI models, modify hyperparameters to enhance the performance of current models, and assess the XAI models with dermatologists.

## Discussion

The results indicate that combining CNNs and ViTs enhances cancer diagnosis accuracy by leveraging local and global feature extraction capabilities. The hybrid model effectively overcomes CNN limitations in long-range dependency modeling while reducing the high data dependency of ViTs. Additionally, the model demonstrates improved generalization across different types of cancer imaging datasets. The findings highlight the significance of hybrid architectures in attaining cutting-edge results in medical imaging.

Moreover, the integration of ViTs reduces the model's dependency on extensive labeled datasets, making it a more viable solution for real-world clinical applications. Future research, according to the report, ought to include transformer-based design optimization, integration with genomics and radiomics, and real-time deployment in hospital settings.

## Conclusion and Future Work

This study presents a novel hybrid deep learning approach integrating CNNs and ViTs for cancer diagnosis. The proposed model significantly enhances classification accuracy compared to standalone CNNs and ViTs. The findings highlight the potential of hybrid architectures in AI-driven healthcare solutions and pave the way for future advancements in automated cancer detection systems. Future work will focus on optimizing computational efficiency, integrating multimodal data sources such as radiomics and genomics, and real-time clinical deployment. Federated learning

techniques for privacy-preserving AI and explainability frameworks will be explored to enhance trust in AI-assisted cancer diagnosis.

## References

1. Khanna NN, Maindarkar MA, Viswanathan V, et al. Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare (Basel)* 2022;10:2493.
2. Ahmed B, Qadir MI, Ghafoor S. Malignant Melanoma: Skin Cancer- Diagnosis, Prevention, and Treatment. *Crit Rev Eukaryot Gene Expr* 2020;30:291-7.
3. Apalla Z, Lallas A, Sotiriou E, et al. Epidemiological trends in skin cancer. *Dermatol Pract Concept* 2017;7:1-6.
4. Cabrera R, Recule F. Unusual Clinical Presentations of Malignant Melanoma: A Review of Clinical and Histologic Features with Special Emphasis on Dermatoscopic Findings. *Am J Clin Dermatol* 2018;19:15-23.
5. Melanoma Research Alliance [Internet]. [cited 2023 Nov 15]. Melanoma Survival Rates. Available online: <https://www.curemelanoma.org/about-melanoma/melanomastaging/melanoma-survival-rates/>
6. Nova JA, Sánchez-Vanegas G, Gamboa M, et al. Melanoma risk factors in a Latin American population. *An Bras Dermatol* 2020;95:531-3.
7. Rizzo S, Botta F, Raimondi S, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp* 2018;2:36.
8. Tichanek F, Försti A, Hemminki A, et al. Survival in melanoma in the nordic countries into the era of targeted and immunological therapies. *Eur J Cancer* 2023;186:133-41.
9. Wan G, Nguyen N, Liu F, et al. Prediction of early-stage melanoma recurrence using clinical and histopathologic features. *NPJ Precis Oncol* 2022;6:79.
10. Han D, van Akkooi ACJ, Straker RJ 3rd, et al. Current management of melanoma patients with nodal metastases. *Clin Exp Metastasis* 2022;39:181-99.
11. Harkemanne E, Baeck M, Tromme I. Training general practitioners in melanoma diagnosis: a scoping review of the literature. *BMJ Open* 2021;11:e043926.
12. Ring C, Cox N, Lee JB. Dermatoscopy. *Clin Dermatol* 2021;39:635-42.
13. Vestergaard ME, Macaskill P, Holt PE, et al. Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting. In: Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]. Centre for Reviews and Dissemination (UK); 2008.
14. Masood A, Al-Jumaily AA. Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. *Int J Biomed Imaging* 2013;2013:323268.
15. Menzies SW, Bischof L, Talbot H, et al. The performance of SolarScan: an automated dermoscopy image analysis instrument for the diagnosis of primary melanoma. *Arch Dermatol* 2005;141:1388-96.
16. Jartarkar SR, Cockerell CJ, Patil A, et al. Artificial intelligence in Dermatopathology. *J Cosmet Dermatol* 2023;22:1163-7.
17. Nachbar F, Stolz W, Merkle T, et al. The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions. *J Am Acad Dermatol* 1994;30:551-
9. Argenziano G, Fabbrocini G, Carli P, et al. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions. Comparison of the ABCD rule of dermatoscopy and a new 7- point checklist based on pattern

analysis. Arch Dermatol 1998;134:1563-70.

18. Soyer HP, Argenziano G, Zalaudek I, et al. Three-point checklist of dermoscopy. A new screening method for early detection of melanoma. Dermatology 2004;208:27-31.

19. Henning JS, Dusza SW, Wang SQ, et al. The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. J Am Acad Dermatol 2007;56:45-52. 21. Akkoca Gazioğlu BS, Kamaşak ME. Effects of objects and image quality on melanoma classification using deep neural networks. Biomed Signal Process Control.