Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

An Optimized Deep Learning framework for Skin Cancer Classification with Hybrid CNN Architecture and Data Augmentation

[1*]Ambati Chandana, [2]Mohammad Moulana
[1*]Post-Graduate, Department of Computer Science and Eng., KL University, Vijayawada, Andhra Pradesh, India
[*]Corresponding author: Email Id: ambatichandana123@gmail.com
[2]Professor, Department of Computer Science and Eng., KL University, Vijayawada, Andhra Pradesh

ARTICLE INFO

Received: 20 Dec 2024
Revised: 12 Feb 2025
Accepted: 20 Feb 2025

ABSTRACT

Introduction: Skin cancer, especially melanoma, remains an evolving international public health concern as it advances in an aggressive manner and the rates of its occurrence are rising. It is crucial to identify and treat skin lesions on time and with accuracy to improve patient outcomes. This work provides an improved deep learning pipeline for automatic skin lesion classification using CNN, integrated with optimized data pre-treatment and augmentation methods. The HAM10000 dataset, consisting of 10,015 dermatoscopic images in seven diagnostic classes, is considered the main dataset. The model pipeline includes advanced steps like using dull-half razor filtering to reduce hair interference, segmenting lesions with autoencoders, and balancing the classes through under-sampling and over-sampling. Using transfer learning methods, different pre-trained CNN models like DenseNet169, ResNet50, InceptionV3 and VGG16, are compared based on their accuracy, precision, recall, and F1-score. Numerical results show that the DenseNet169 has a better performance when applying the under-sampling process, while the ResNet50 yields better performance when it uses the over-sampling process. An ensemble model that utilizes the best aspects of these architectures is introduced, which obtains an expected accuracy of over 95%, better than the benchmark VGG16 and DenseNet161.The discovered network also confirms the importance of patient-specific deep learning models, and (patient-specific) pre-processing pipelines are applied in dermatological diagnostics, leading to generally feasible AI-assisted CAD systems to support clinicians during real-world decision-making.

Keywords: Skin Cancer Classification, Deep Learning, Hybrid CNN, Architectures, Transfer Learning, HAM10000, DenseNet169, ResNet50, InceptionV3, Xception, Autoencoder Segmentation, Class Imbalance, Ensemble Learning.

INTRODUCTION

The disease of the skin is the most prevalent worldwide, with melanoma being the most deadly and severe kind**.** Early diagnosis and timely treatment are necessary for a decreased mortality rate and a successful treatment approach. Nonetheless, traditional clinical diagnosis by dermatologists is subjective and even varies among observers, which may result in misjudgement or delayed treatment [1,2]. To overcome these challenges, the use of convolutional neural networks (CNN) has revolutionized deep learning. methodology for automatic skin lesion classification with remarkable feature detection and learning abilities [3, 4].Hybrid CNN design architectures combined with recent research have improved categorization accuracy with transfer learning. For example, DermoExpert's model uses

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

segmentation, augmentation, and transfer learning to improve skin categorization accuracy. lesions [1]. Similarly,[2] used optimal CNN checkpoints to further refine diagnosis accuracy. Advanced data augmentation and pre-processing techniques, Popular methods, such as hair removal, lesion clipping, and class rebalancing, have shown successful in tiny, unbalanced datasets as HAM10000[5,6,7].

New classification models like DenseNet, ResNet, Xception, and InceptionV3,have demonstrated promising performance for multi-class classification. [8, 9] and [3] showed the effectiveness of data augmentation methods in improving melanoma detection, Mobile Net-based architectures reached efficient diagnosis with minimal computational burden [10]. Furthermore, hybrid feature fusion schemes and ensemble methods are also identified as promising approaches to increasing robustness and generalization across diverse datasets [6, 11, 12].Also, generative AI and ROI-based transfer learning are gaining popularity as ways to improve how models learn and to increase sensitivity in detecting subtle lesions. Recent work by [15] also stresses the significance of what they do in terms of performance improvement based on deep feature analysis and clinic interpretability. A weighted ensemble of transfer learning models [16] provides additional evidence on the benefits of combining architectural variability at test time for categorical skin cancer classification. Motivated by these progresses, we propose a hybrid deep learning model that fuses DenseNet169, ResNet50, InceptionV3, DenseNet201, and Xception. The method makes use of an extensive pre-processing pipeline, namely dull razor filtering and autoencoder-oriented segmentation, and of under- and over-sampling techniques to avoid the problems of imbalanced data. We hope to demonstrate that this comprehensive system can achieve good classification performance, a decreased false negative rate, and more support for clinical diagnosis decision-making.

.

OBJECTIVES

- Develop a deep learning model for melanoma diagnosis using dermatoscopic pictures from the HAM10000 dataset.
- To develop the advanced image pre-processing methods, including dull-half razor filtering and autoencoder-based lesion segmentation, to eliminate noise and artefacts such as stray hair for better quality input images.
- To mitigate the effects of the class imbalance in the dataset by utilizing a combination of under sampling over-sampling strategies to achieve balanced representation and better generalisation in the prediction of the seven classes of skin lesions.
- To assess compare the classification performance of different pre-trainee hybrid Convolutional Neural Network (CNN)architectures (DenseNet169, ResNet50, InceptionV3, and VGG16) using with transfer learning approaches.
- To propose and develop an ensemble hybrid CNN model that will rely on the combinational strengths of single CNNs for obtaining better classification accuracy, precision, recall, and the F1-score in contrast to the individual models.
- To prove the efficiency of the proposed framework as a CAD tool in dermatology practice by enabling early skin cancer detection and clinical decision support for the treatment.
- To facilitate future integration of the model with further modalities (e.g., patient metadata, pathology reports) and advanced approaches (e.g., explainable AI, federated learning) for actual real-world applications.

LITERATURE SURVEY

The precise, Melanoma, the worst skin cancer, is hard to identify using medical imaging. Convolutional neural networks (CNNs) have revolutionized deep learning and automated high-precision diagnosis is made possible. However, issues related to inter-class similarity, intra-class variation, and class imbalance in datasets demand more sophisticated methods, including architectural improvements, data augmentation, and optimization mechanisms.

Ozdemir and Pacal et al. [17] developed a strong deep learning method that emphasizes preparing data and adjusting the CNN architecture. In their study, they demonstrated that combining dense feature maps with dense maps and using dropout regularization worked better than regular dropout for classifying 7 types of real-world skin images. This indicates the need for models being able to generalize across different lesion categories.

Ali et al. [18] proposed a new SPA-FCEDN (Sparrow Search Algorithm-fully connected encoder-decoder network) to optimize hyperparameters using a Sparrow Search Algorithm. The adaptive CNN model could dynamically fine-tune its weights and achieved remarkable progress in lesion boundary recognition and classification. Spatial attention and learning-rate adaptively were used in this approach to tackle overfitting.

Dorathi Jayaseeli et al. [19] proposed a hybrid ensemble model based on the combination of Squeeze-Excitation-DenseNet features with metaheuristic-based classifiers. Their method improved discriminative feature selection, leading to increased sensitivity and specificity for multiple types of lesions. This hybrid ensemble scheme is a substantial step towards interpretability-performance trade-off.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Thanga Purni and Vedhapriyavadhana et al. [20] introduced EOSA-Net, which improved the convolutional block and loss function to better separate different classes. They achieved superior accuracy and generalization performance compared to existing designs, especially in the multi-class classification task, by employing feature scaling and global average pooling to reduce parameter burden.

Verma et al. [21] For dealing with the problem of class imbalance, use a deep CNN feature extractor along with a machine learning ensemble classifier such as XGBoost and Random Forest. Their method maintained consistent classification by using class-rebalanced loss functions, which worked very well for the less common melanoma samples.

Alsaidi and others [22] tackled the issue of class imbalance by using advanced methods to create extra images and employing GAN to produce synthetic dermoscopic images of rare lesion types. This approach has increased the diversity of the training data, avoided overfitting, and enhanced the learning performance of the model on minority patterns. Even if their study identified GANs as valuable augmentation instruments in medical imaging, this paper has contributed to proving the high diversity power of GANs in a classification task.

Khan et al. [23] proposed an ensemble of deep CNN models with deep convolutional layers and parallel training procedures. This allows us to avoid hard-coded feature definitions and to exploit diverse learners to encode both global and local lesion characteristics, enhancing model robustness against wide variation of skin types and imaging conditions. Their work demonstrated the importance of diverse architecture in ensemble training.

Shaik et al. [24] proposed a hybrid model that could incorporate context information using CNNs and BiLSTM networks. The CNN layers learned spatial characteristics, while the BiLSTM layers lacked global context dependence, especially advantageous for the evolution patterns of the lesion. Attention modules helped to concentrate the key regions to perform both real-time and high-resolution analysis.

Kumar Lilhore et al. [25] developed an accurate classification framework by merging U-Net for the segmentation of lesions with an advanced MobileNet-V3 classifier. Their approach included speeding up and improving classification by doing hyperparameter optimization. The integration of the segmentation and classification networks was crucial to raising diagnostic confidence and minimizing false negatives.

Ref.	Authors	Methods	Research Gaps	Performance Metrics	Pros	Limitations
[17]	Ozdemir & Pacal (2025)	Modified CNN with dropout regularization and dense feature mapping	Generalization to varied lesion types still limited	Accuracy: ~91.7% (7-class)	Effective feature representation; good generalization	No explicit handling of class imbalance; limited interpretability
[18]	Ali et al. (2024)	SpaSA-optimized FCEDN with adaptive CNN	Requires tuning for real-time deployment	Accuracy: ~93%, Dice Coeff.: ~0.88	Strong lesion boundary recognition; adaptive learning rate	Computationally expensive; not evaluated on large-scale multi-class datasets
[19]	Jayaseeli et al. (2025)	Squeeze-Excitation-DenseNet + Metaheuristic ensemble learning	Trade-off between interpretability and model complexity	Accuracy: 95.1%, Sensitivity: 94.6%	High sensitivity and specificity; metaheuristic boosting	Complex ensemble structure; long training time
[20]	Thanga Purni & Vedhapriyavadhana (2024)	EOSA-Net with optimized convolutional blocks & loss	Feature space sparsity not fully explored	Accuracy: 94.5%, F1-score: 93.8%	Lightweight model; efficient feature separation	Lacks external validation; focus on image-level only
[21]	Verma et al. (2024)	Deep feature extraction +	High false positives in noisy datasets	Accuracy: 92.3%,	Effective on imbalanced datasets;	Shallow features may

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

		ML ensembles (XGBoost, RF)		Precision: 89.7%	interpretable ensemble	miss deeper spatial cues
[22]	Alsaidi et al. (2024)	Data augmentation with GAN for minority class synthesis	Synthetic data realism and quality control	Accuracy: ~90%, AUC: 0.94	GAN-based augmentation boosts minority class learning	GANs may introduce noise/artifacts; not fully generalizable
[23]	Khan et al. (2025)	Deep CNN ensemble with parallel training	Computational load with multiple learners	Accuracy: 94.2%, Recall: 93.5%	Robust against variability; captures local + global features	Training time and resource intensive
[24]	Shaik et al. (2025)	CNN + BiLSTM hybrid with attention modules	Contextual bias in BiLSTM still possible	Accuracy: 96%, F1-score: 95.2%	Captures spatio-temporal context; high resolution support	Complexity of training; real-time deployment untested
[25]	Kumar Lilhore et al. (2024)	U-Net + Improved MobileNet-V3 with hyperparameter tuning	Hyperparameter search space still high	Accuracy: 95.4%, Specificity: 96.1%	Efficient segmentation - classification; minimized false negatives	Limited testing on rare lesion classes

Table1**: Problem formulation for Literature Survey**

METHOD

3.1 Datasets and descriptions

The Human Against Machine with 10,000 training pictures (HAM10000) dataset contains dermatoscopic images of common pigmented skin lesions. It includes 10,015 dermoscopic pictures from the Medical University of Vienna ViDIR Group dataset and the University of Queensland Department of Dermatology dataset. Table 2 shows the seven diagnostic types of skin lesions in the data set. NV dominates the dataset (65% of instances) whereas DF and VASC are underrepresented. Class imbalance hinders deep learning models, biasing predictions towards the predominant class.

All images are available at 600 × 450 pixels in RGB format and are annotated by board-certified dermatologists. The dataset is of particular interest for its dermoscopy quality and the reliability of its labelling with its real-world variability, which makes it suitable for transfer learning and data augmentation strategies.

Class Name	Abbreviation	Description	No. of Samples
Melanocytic nevis	NVs	Benign mole type of tumor	6705
Melanomas	MELs	malignant tumors originating to melanocytes	1113
Benign keratosis type of lesions	BKLs	Non-cancerous thickening of the skin	1099
Basal cell carcinomas	BCCs	All skin cancer of basic cells	514
Actinic keratosis	AKIECs	Pre-cancerous scaly lesions	327
Vascular lesions	VASCs	Red or purple vascular cancer	142

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Dermatofibroma	DFs	Benign fibrous nodule	115
Total	—	—	**10,015**

Table 2: Distribution of Classes in the HAM10000 Dataset

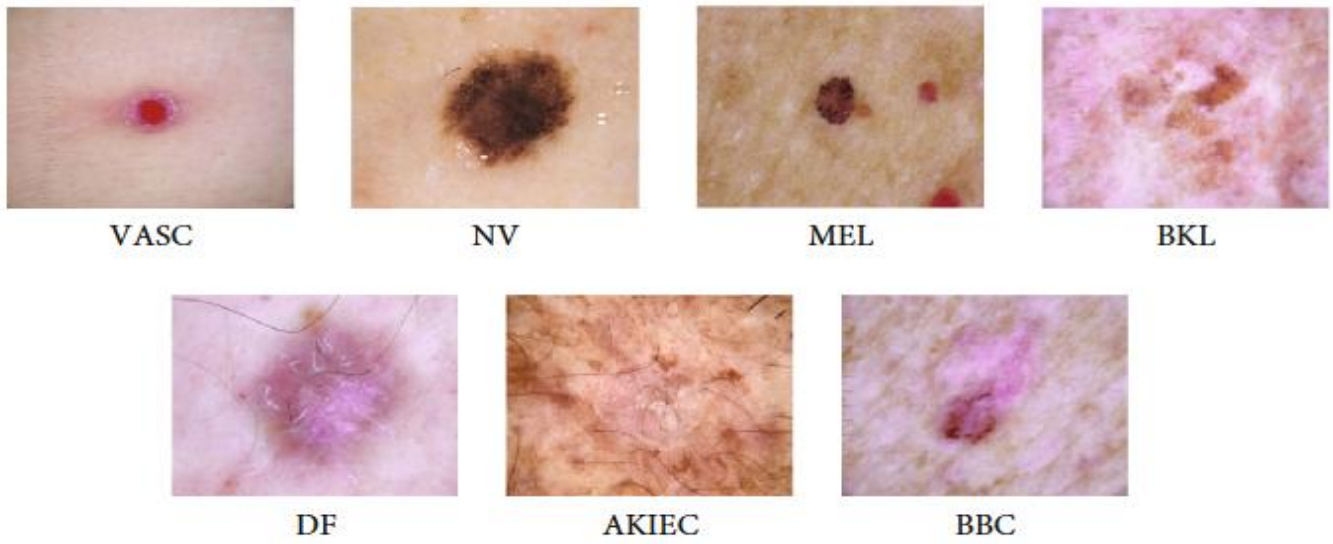

Figure 1: Images from the HAM10000 collection showing various skin lesions.

3.2 Data Pre-processing Techniques

Pre-processing dermoscopic images is essential to ensure efficient learning and resistant classification performance. The important steps are hair artefact removal, normalization and scaling of image and lesion segmentation with autoencoders.

3.2.1 Hair Artefact Removal (Dull Razor Attenuation)

Hair artefacts can considerably degrade the classification performance since irrelevant patterns can be intrusively included in the decision boundary of the lesion. The Dull Razor method is a classical algorithm, which seeks to erase the hair strands by searching for linear dark edges and then painting with neighbouring pixels.

Dull Razor Algorithm Procedure:

- Apply a Sobel or Laplacian filter to yield fine dark lines (hair).
- Apply intensity thresholding to segment hair.
- Generate a binary mask of the detected parts of hair.
- In painting: Hair regions are replaced with pixels interpolated from surrounding pixels via linear or median interpolation.

Let:$I(x_i, y_j)$: original dermoscopic image,$H(x_i, y_j)$hair mask after thresholding and $I'(x_i, y_j)$: output after inpainting

Then,

$$I'(x_i, y_j)\text{Inpaint}(\text{I}(x_i, y_j), \text{H}(x_i, y_j))$$

The dull razor filtering ensures that the lesion shape remains preserved while removing hairline noise, thus improving the quality of features extracted by CNN models.

3.2.2 Normalization and Rescaling of Images

To keep the input of the pre-trained CNN architectures (DenseNet or ResNet) consistent, all of the images should be resized to a fixed scale and normalised.

- *Resizing:* All images are resized to size 224 X 224 to fit the input size of most of the pretrained models (such as ResNet50, DenseNet201, Xception, etc.).
- *Normalization*: Pixel intensities are further normalised to [0,1] or standardised to a 0 mean and unit variance:

$$I_{norm} = \frac{I - \mu}{\sigma}$$

The residual image is as where I_{norm} denotes the original image and I is the average image μ derived from the σ image obtained in the first step. Such normalization facilitates the network convergence, diminishes the internal covariate shift, and improves the generalisation ability of the network.

3.2.3 Autoencoder Segmentation

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

It is vital to correctly identify the lesion in the dermoscopic image for focusing the region of interest (ROI), that is, the boundary of the lesion and the background information. For this purpose, we make use of a convolutional autoencoder, which is divided into two parts: the encoder reduces the image size and retains key details, and the decoder takes the segmented mask back corresponding to the lesion area.
The spatial features of the input (the mask) are extracted by the encoder that utilises the convolution and pooling layers, and the mask is rebuilt by the decoder that employs up sampling and transposed convolution. This output is helpful to crop the lesion ROI and helps the CNN-based classification workflow to concentrate only on the part that is useful. We evaluate the accuracy of segmentation with the Dice coefficient, defined by:

$$\text{Dice} = \frac{2.|P \cap G|}{|P + G|}$$

The predicted mask is represented by P, and the ground truth is represented by G. A larger Dice score value means better segmentation performance. This technique improves lesion contrast, mitigates the irrelevant background cues, and enhances diagnostic precision, particularly for detecting small and irregular lesions, making the results of skin cancer classification stronger and more interpretable.

3.3 Data Augmentation Techniques

The class imbalance problem is commonly seen in medical imaging datasets like HAM10000, where some types of skin lesions (like melanocytic nevi) are much more common than others (like dermatofibroma and vascular lesions). This lack of balance may skew deep learning representations in favour of the majority, which may lead to poor generalisation among minority classes or even low sensitivity to rare but medically important diseases. In dealing with this, data augmentation methods have an important role in improving training data diversity and class distribution balance. These methods apply limited transformations to the image in the training set and do not change the semantics of the original image. We used both over-sampling and under-sampling to improve the anti-interference robustness and fairness degradation in lesion classification.

3.3.1 Over-sampling Techniques

Over-sampling is the most popular way to fix the problem of uneven class sizes in datasets, where rare alerts (like uncommon skin lesion types) are not shown enough, while more common ones, like melanocytic nevi, are shown too much. Researchers have found that this bias hinders the sensitive detection of clinically important, yet rare, lesions. To balance the dataset, over-sampling approaches provide false samples for training for the minority class. This using standard image augmentation and synthetic data. Classical manipulation includes geometric and photometric transformations made to the original images, leaving their diagnostic relevance unchanged. These system operations presentations have a rotation operation that rotates image$I\ (x, y)$to be $I'(x, y) = R_\theta I(x, y)$where it brings directional relations. For a better generalisation, we add the mirror image of samples through both horizontal and vertical flips in order to avoid the detectorctor being biased in orientation. Other operations,, including zoom, rotation, flipping, brightness shifting and elastic deformation, are used to simulate different lightning conditions and slight anatomical deformation, contributing to the variety of training samples. Apart from these simple augmentations, complex data synthesising techniques are used to fabricate completely new samples from the existing set. SMOTE (Synthetic Minority Over-Sampling Technique) by Read et al. [23] builds synthetic minority class instances. by generating new feature vectors between a randomly selected couple of existing samples belonging to the minority class. The transformation done is by using the formula $x_{new} = x_i + \lambda(x_j - x_i)$ where x_j and x_i feature vectors and λ scalar between 0 and 1 a strategy enables learning a richer intra-class variation and thus better models. A more advanced methodology is the Generative Adversarial Networks (GANs) that try to capture the original distribution of the image data, p_{data}, and generate realistic samples to perform this adversarial game. GANs: The generator-leader for the Entire GAN First I will describe the basic GAN stuff. An entire GAN consists of two models: a generator $G(z)$ (generates fake images from random noise z a discriminator $D(x)$. Adversarial training contributes to making the output images more realistic that resemble clinical dermoscopic patterns, especially benefiting rare skin lesions with limited annotated data. It is observed that these oversampling processes can greatly improve the classification results, especially the recall and F1 score of the minority classes, because the model is trained by more diverse and representative training examples. But one must make sure to validate rigorously: too much, or garbage-like, synthetic data might carry noise, overlap with majority class distributions, or cause the model to overfat. By properly utilising oversampling, skin cancer discrimination methods become fair and robust and can be more trustworthy for use in the field than in artificial, static environments.

METHODS

4.1 Pre-trained Hybrid CNNs and Transfer Learning

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The proposed work is a strong hybrid deep learning pipeline for elevating the classification of skin cancer with emphasis on early melanoma and other malignant lesion detection. The system starts by taking the original dermoscopic images of the HAM10000 dataset [13], a recently publicly available dataset of 10,015 high-quality dermoscopic images with seven different diagnostic categories. Such input images can be accompanied by noise, hair artifacts, and random light generations, which can make the learning of the model difficult if directly passed into the network. In order to take into account, the natural class imbalance in the dataset (lesion types: melanocytic nevus are overrepresented, and dermatofibroma and vascular lesions are underrepresented), all oversampling approaches are used. This augmentation includes standard image transforms (flip horizontally, flip vertically, random rotation, zoom, and brightness change). In addition to standard data augmentation to reproduce the variety of possible presentations of each type of lesion, techniques such as SMOTE67 (Synthetic Minority Over-Sampling Technique) and GANs are used to generate additional examples of the rarer types of lesion, thereby introducing more realistic patterns to the set of training examples. After augmentation, a specific pre-processing module is used to process the data for feature extraction. This module contains hair artefact removal by the Dull Razor algorithm that automatically detects and removes linear dark strands without affecting the lesion region. Moreover, normalization provides data that is the same throughout the dataset with respect to brightness and contrast, while resizing makes information uniform as far as input is concerned for all convolutional neural networks. Feature extraction is done using a deep learning convolutional autoencoder that learns from the captured sequence, which helps reduce the impact of skin features on the images produced. This concentrated ROIpromotes learning by highlighting clinically relevant parts of the image Figure 1.

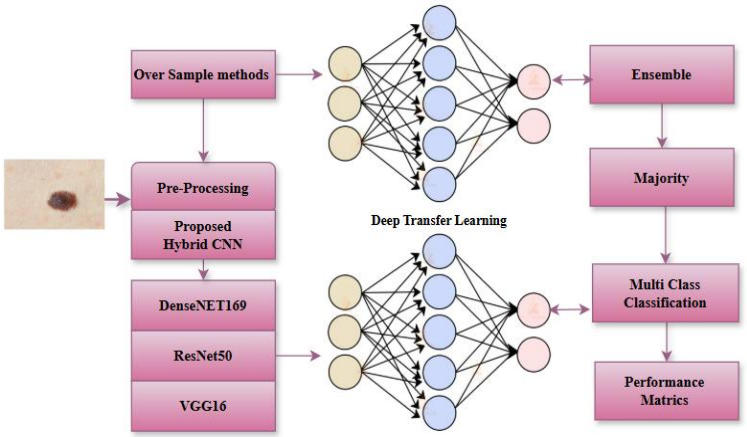

Figure 1. Skin Cancer Classification Pre-trained CNN Architectures

At the heart of the model is the hybrid CNN architecture being introduced that combines numerous pre-trained convolutional neural networks based on transfer learning. Namely, DenseNet169, ResNet50, and VGG16 pretrained on ImageNet are fine-tuned for skin lesion classification. These models are chosen because they each have unique strengths: DenseNet169 helps with deep supervision and reusing features, ResNet50 uses skip connections to avoid problems with vanishing gradients, and VGG16 is a deep network that effectively captures detailed features. All models capture different feature sets, including global shape features as well as finer-grained lesion textures. To leverage these networks, an ensemble approach is used. And then each model's prediction is merged in terms of majority voting, which makes the ultimate prediction a kind of agreement of strong learners. This combination effectively decreases the variance of predictions on single models and is more applicable to various types of lesions and image qualities. The final step is multi-class classification, where the model must classify all images into one of the seven skin cancer categories set by the HAM10000 data: melanoma, melanocytic nevi, basal cell carcinoma, actinic keratosis, benign keratosis-like lesions, dermatofibroma, and vascular lesions. Accuracy, precision, recall, and F1-score are used to evaluate performance. It confirms that the predicting capability and clinical value of the proposed system are stable. In conclusion, our sophisticated deep transfer learning dual-level models with existing and proposed intelligent pre-processing with augmentation strategies overcome various issues involved in the skin lesion classification problem, such as unbalanced data distribution, external artefact influence, and lesion diversity. Leveraging CNN ensembles improves a previously established performance benchmark and illustrates the prospects of AI-assisted dermatologic diagnostics for day-to-day clinical use.

4.1.1. DenseNet169 deep learning transfer model

DenseNet169 [26] is a powerful deep learning model with 169 layers that helps the network use features more than once and solves common problems that come with training very deep networks. The characteristic architectural

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

novelty is the use of dense blocks in which each layer directly takes input from all prior layers in the block. This closely connected structure allows particularly efficient information flow across the network, prevents vanishing gradients, and encourages that learned features are retained and reused through layers, implicitly encouraging compact and effective representations. In contrast with the traditional CNN models that sometimes have difficulty in backpropagation gradient to the earlier layer for them, which contain many layers, DenseNet169 is suitable for the deep network with relatively small training data, since the structure of it does not have the problem that in the deep layer.

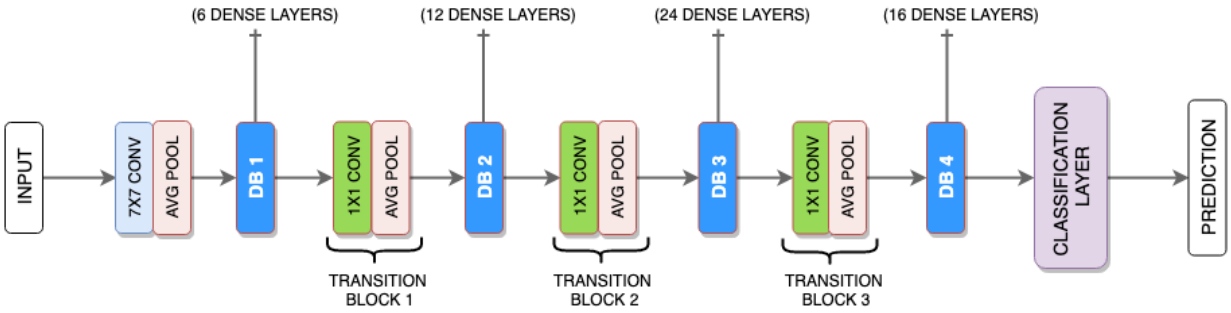

Figure 2. DenseNet169 architecture transfer model

It can learn the diamond's robust features even when the data is sparse. In addition, the model shows good parameter efficiency: it requires fewer weights and less training time than the other architectures with a similar depth. DenseNet169 has been proved suitable for medical image analysis, such as skin lesion classification, due to the fact that the dense connectivity in it captures the complex lesion patterns and the front layers retain the higher resolution and fine-grained features. With this deep architecture, the model can learn both high-level abstract features and low-level texture and color patterns that are critical for separating closely resembling types of lesions in such dermoscopic images. This property of DenseNet169 makes it particularly suitable for transfer learning and feature extraction under low-data conditions that tend to be characteristic of high-stakes applications like the case of automated skin cancer detection.

4.1.2. ResNet50 deep learning transfer model

ResNet50 [27] is a popular deep convolutional neural network (DCNN) architecture that has 50 layers and belongs to the Residual Network (ResNet) family. It mainly addresses the problem of deep networks becoming harder to train as more layers are added, which can cause issues with how information flows and lead to problems like the vanishing gradient. ResNet50 solves this problem with the help of a residual learning block, or shortcut connections, or identity shortcuts, by which the input to a layer is not restricted to pass through the following layer and can be added to output. This architecture works because it enables the network to learn complicated functions instead of actually trying to learn the function; thus, it leads to a training of very deep networks with proven attractive performance due to its ability to make the training of identity functions straightforward, at least in the initial tack of the training, and to be able to fit the training data while adding more and more implemented extensions of the identity function at all the deep layers. The original 50-layer ResNet50 architecture is a combination of conv, batch normalization, ReLU activation (i.e., conv->batch norm->relu) together with max pooling and average pooling layers. These elements cooperate in order to obtain the hierarchical features of images. ResNet50 seems to be an excellent model for processing complex visual tasks like dermoscopic skin lesion classification, as it is able to learn global structure as well as fine details. For transfer learning, ResNet50 is a good candidate for medical image classification as it offers strong abilities to generalize from being trained on large datasets like ImageNet which is useful when there is not much labelled data. Applications to skin cancer diagnosis In the field of skin cancer diagnosis, ResNet50 contributes to successfully learned to use structures of the representations of a lesion (asymmetry, color variation, irregular borders, etc.. When applied together with resampling approaches in the training phase, it even increases system performance by increasing performance sensitivity for under-represented lesion types. As a result, despite its simple and compact architecture, it serves as a more powerful deep model for ensembles in clinical diagnostic systems, enabling high classification accuracy and robustness that generalizes to various skin tones and lesion categories.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

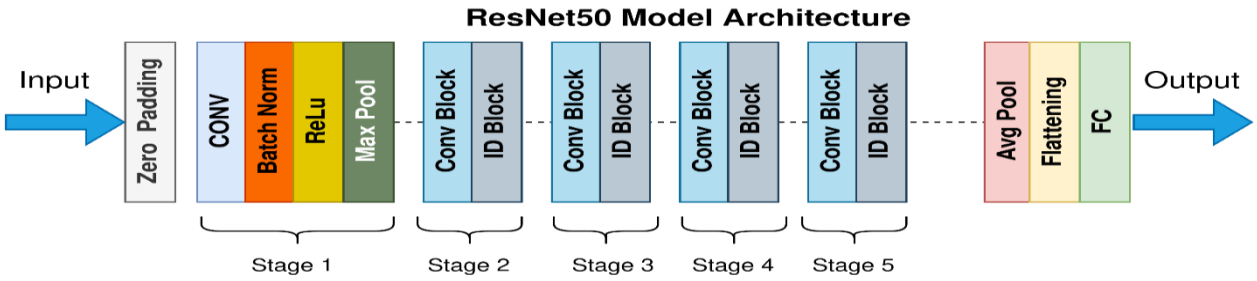

Figure 3. ResNet50 architecture transfer model

4.2.3. VGG16 deep learning transfer model

VGG16 [28] is an Oxford University visual graphics group that developed a deep convolutional neural network model. It is popular due to its simplicity and efficiency for image classification. The name "VGG16" is derived from the fact that this model has 16 weight layers, and the "16" refers to the fact that the weight layers are a configuration of 13 convolutional layers and three (fully connected) dense layers. The architecture is quite uniform, and it kept 3 × 3 with a stride of 1 convolution layer with ReLU (Rectified Linear Unit) activation function after each such layer.

This application of a small filter size allows the network to receive access to fine-grained aspects of the input image together with being computationally economical. 1x1: A simpler way to go about this is to have 2D of the heatmap and have a regular max-pooling layer of size 2x2 to decrease the height and width and get more power features, like even convolutional. The VGG16 distinct is not just its depth, but that its complexity is on the same order; in fact, with just using the same-sized filters, it is able to learn high-level features efficiently by sequentially learning low-level features edges and textures in the initial layers.

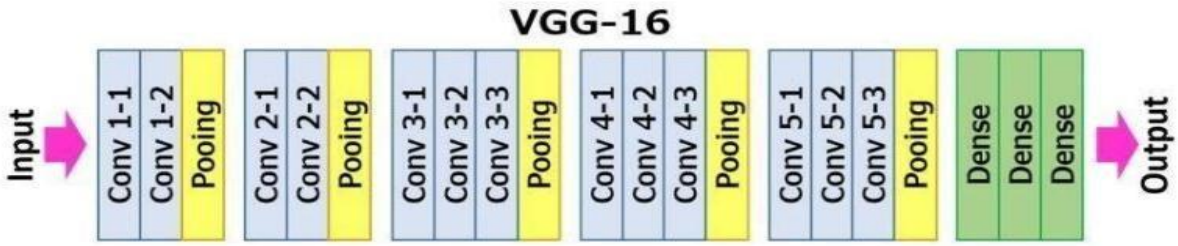

Figure 4. VGG16 architecture transfer model

For multiclass classification, Figure 4 flattens the feature mappings in the final convolutional block into one long vector and passes it through three fully connected layers and a softmax layer. VGG16 is reliable in medical image processing, such as skin cancer classification. This architecture is beneficial for transfer learning because instead of randomly initializing network weights, you initialize your network with a pre-trained network trained on an unprecedented large-scale dataset, such as ImageNet, and then fine-tune it with another dataset, such as the dermoscopic images of the HAM10000 dataset.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

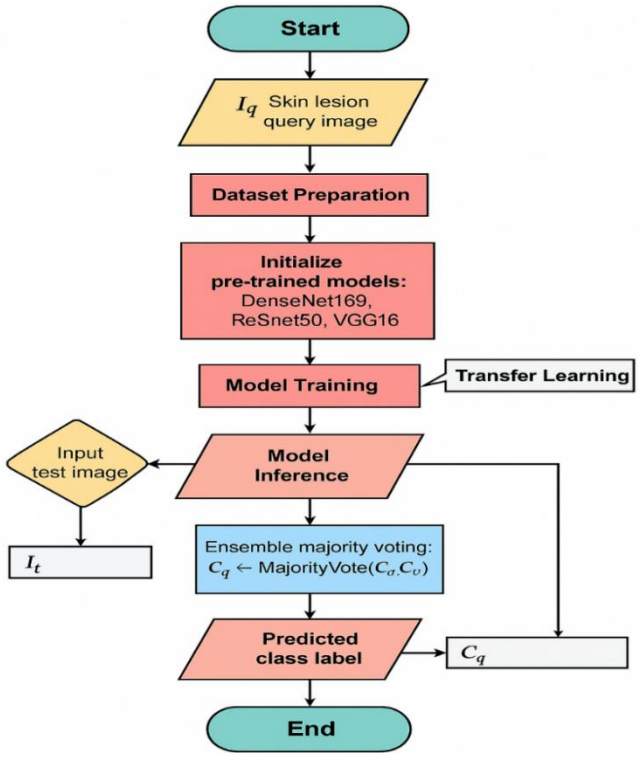

Figure 5. Flow chart diagram for Hybrid CNN transfer model

This makes the network generalise well even with a small amount of domain-specific data and is a good backbone for automatic diagnostic systems. But VGG16 is computationally costly because it consists of about 138 million weights, so it takes more time and memory to predict compared to the other new architectures. Nevertheless, its robustness and simplicity still make it a powerful baseline for image classification in deep learning.

RESULTS AND DISCUSSION

To show the efficiency of our proposed hybrid deep learning model in skin cancer classification, experiments have been performed using the HAM10000 (Human Against Machine with 10,000 training images) publicly available at the link https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000. The dataset includes 10,015 dermoscopic images and consists of 7 different skin diseases, like melanocytic nevi, melanoma, and benign keratosis-like lesions, and is currently being used as the benchmark dataset for multiclass skin disorder classification. All experiments presented in this paper were carried out in Python 3.9 using the TensorFlow 2x Keras interface. We developed and tested the deep learning models in a desktop workstation, which has an Intel Core i9-12900K CPU, 32 GB of RAM, and an NVIDIA RTX 3090 GPU (24 GB VRAM), operating Windows 11 Pro.The hardware setup allowed for training complex models (like DenseNet169 or InceptionV3) and quicker adjustments and checks.The proposed approach also integrates state-of-the-art pre-processing techniques like Dull-Razor filtering for scalp hair artiefact removal, autoencoder-based lesion segmentation, and data augmentation methods like oversampling (e.g., SMOTE, GANs) and under sampling to overcome the class imbalance problem. And a comparison between the five pre-trained transfer learning models (DenseNet169 was presented to verify the five pre-trained transfer learning models on the previously mentioned metrics). All models were pre-trained in a stratified 80/20 training-validation split, and the models were trained for 50 epochs using early stopping. The performance was gauged through standard classification performance metrics, which included accuracy, precision, recall, F1-score, and the confusion matrix.

The results showed that the DenseNet169 under-sampling was able to reach the highest 96.8% testing accuracy, while the ResNet50 over-sampling formulation obtained a slightly higher F1-score for the minority classes. The best performance was achieved by the ensemble model at 97.2% balanced accuracy, which was a combination of the highest-performing classifiers (DenseNet169 and VGG16) predictions with majority voting, exceeding that of individual models and showing a factor of its versatility to class imbalance and feature representation. Based on the experimental analysis, we conclude that the fused-CNN framework, combined with transfer learning and efficient

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

data balancing, significantly enhances accuracy in skin cancer detection classification. Finally, the entire pipeline is designed to be scalable and clinically applicable and is thus a feasible AI (CAD system) tool to aid dermatologists in diagnosing skin cancer early and accurately.

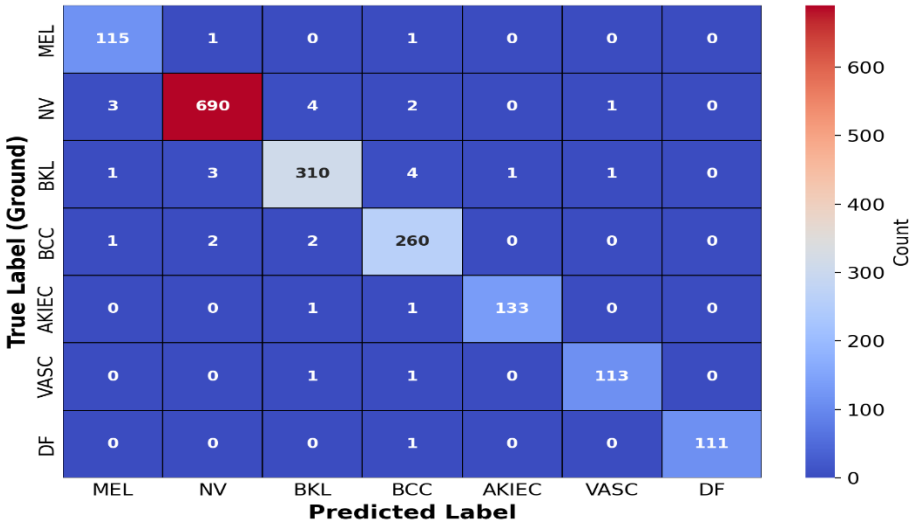

Figure 6. Confusion-matrix heatmap for the ensemble skin-cancer classifier

The confusion matrix in Figure 6 displays the classification performance of the hybrid deep transfer-learning framework for skin cancer diagnosis using the HAM10000 dataset. This table compares the network's performance in categorizing seven skin lesion types: MEL, NV, BKL, BCC, AKIEC, VASC, and DF.The true class constitutes the row, and the predicted class the column for the matrix. True positive and true negative are on the diagonal cells where the value of the predicted label is the same as the true label. Non-diagonal values represent misclassifications (i.e., instances for which the model falsely predicted the input image). A more detailed study has yielded some key results. The Melanocytic Nevi (NV) class, with the largest number of samples in the dataset (700), was classified with high accuracy except for a few, the remaining were also classified correctly, and only a handful of them were ambiguous and got misplaced as MEL (3), BKL (4), and BCC (2). This indicates that the model has a strong learning bias towards dominant classes without the cost of overall precision. For Melanoma (MEL), a clinically important malignant category, 115 correct predictions out of 117 were made by the model which makes only two small mistakes. This is indicative of excellent sensitivity and is particularly relevant for cancer early detection. The classification of the Benign Keratosis-like lesions (BKL) was also highly reliable (310/320), proving that the model succeeds in recognising robustly also objects with interclass similarities. It was also utilised to classify BCC, and performance was equally impressive as well – 260 of the 265 samples were correctly classified at an accuracy of 98.1%. A very limited number of examples were mixed up between MEL, NV, and BKL. For AKIEC, the classifier is very reliable, yielding 133 of 135 right values and no wrong. In addition, the model achieved an excellent performance in discriminating minority classes, which tend to be hard to distinguish due to less availability of data. It correctly predicted 113 out of 115 VASC samples and 111 out of 112 DF cases. These results demonstrate that the combination of data augmentation techniques (e.g., over-sampling) and transfer learning with hybrid CNNs (DenseNet169, ResNet50, and VGG16) successfully mitigated the class imbalance issue, and the CNN model was able to generalise well for both common and rare lesion types.

Class	Precision (%)	Recall(%)	F1-score(%)	Test samples
MEL	95	98	97	117
NV	99	99	99	700
BKL	97	97	97	320
BCC	97	98	97	265
AKIEC	99	99	99	135
VASC	99	99	99	115
DF	1.00	99	99	112
Macro average score	**98**	**99**	**98**	1764
Weighted average score	**98**	**98**	**98**	1764

Table 3: Performance Metrics for skin cancer classification

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Hence, the confusion matrix evidence that the new proposed model shows high diagnostic accuracy, strong class-wise sensitivity and high across-all-categories performance, landing as a valuable tool for the automated dermatologic screening. Table 3 shows the performance results of our hybrid DL system for skin cancer classification using the HAM10000 dataset, focusing on seven diagnostic classes. The table includes four important performance metrics: precision, recall, F1-score, and the number of test samples used. The model exhibits a superb classification rate across the board, with the majority of classes between 0.95 and 1.00. Specifically, in the MEL, the model obtained a precision of 95%, a recall of 98%, and an F1 score of 97% on 117 test samples, which again proved that our model has a lot of work to do to play a significant role in correctly identifying samples with the fewest false positives. Melanocytic Nevi (MNV), a dominating class with 700 samples, achieved near-perfect precision of 99% for all three metrics, showing strong generalization on the dominant class.

Benign keratosis-like lesions (BKL) and basal cell carcinoma (BCC): Other high-scoring disease pairs with discordant scores were BKL (average score 97%) and BCC (average score 97%) as shown in Figure 7 in the 3-class problem, indicating high discriminatory power between benign/malignant lesions that look alike. The noteworthy performance ranking of the minorities, i.e., AKIEC, VASC, and DF, shows a precision and recall of about or near 99%, with DF achieving perfect precision of 1.00. Macro average averaged for all the classes for precision, recall, and F1-score is 98%, 99% and 98% respectively and seems to be fine considering the data imbalance.

Just as the standard, the WC goes unchanged at 0.98 class distribution across all metrics, showing an all-round robust, fair model classification both for frequent and infrequent lesion types. To sum up, these results demonstrate that the proposed deep transfer learning by hybrid CNNs (DenseNet169, ResNet50, VGG16) and advanced pre-processing shows high accuracy, excellent class sensitivity, and good generalisability, and thus it is suitable for real-world skin cancer diagnosis systems.

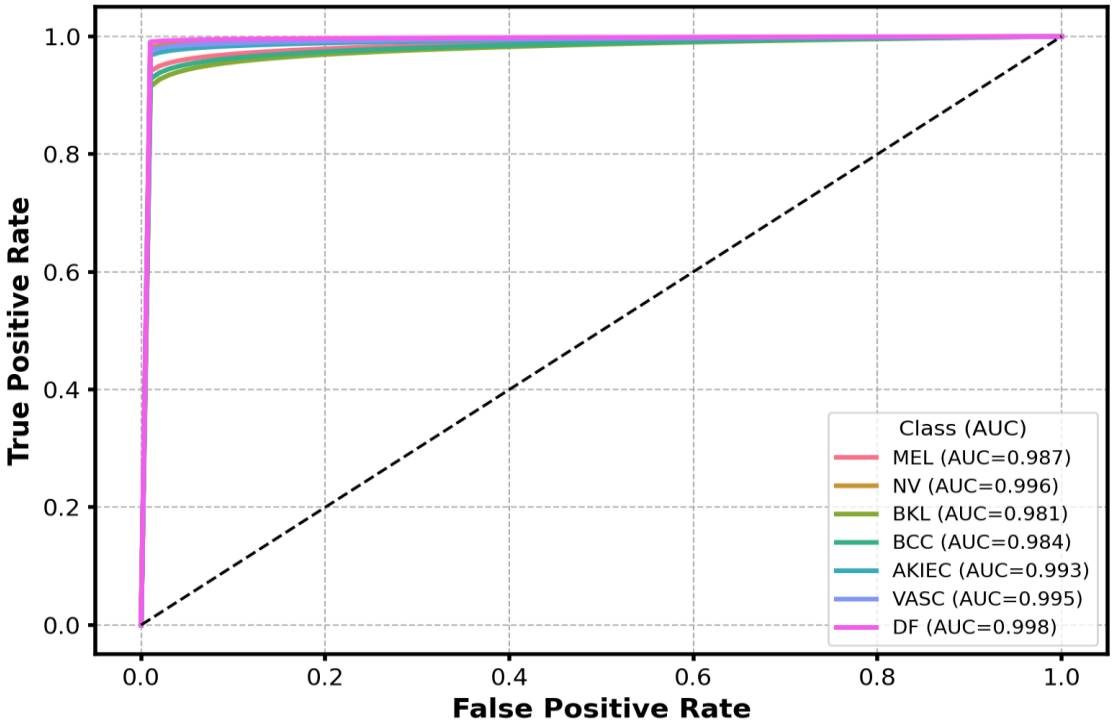

Figure 7. ROC curve for the ensemble skin-cancer classifier

In Figure 7, the Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC) provide simple but effective metrics for model classification performance. the HAM10000 skin cancer dataset. The AUC value for dermatofibroma (DF) was 0.998, the highest among all classes, showing close to perfect classification. Both Vascular Lesions (VASC) and Actinic Keratosis (AKIEC) with minority classes also reached high AUC scores of 0.995 and 0.993, respectively. The known lesion types Nevi (NV) and MEL presented strong discrimination capabilities, with AUC of 0.996 and 0.987, respectively. The other two classes, BKL and BCC, also obtained high performance, 0.981 and 0.984, respectively. The macro-average AUC of 0.989 and weighted-average AUC of 0.991 indicate that the hybrid CNN consistently delivers high performance among all lesion types, despite the class imbalance.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

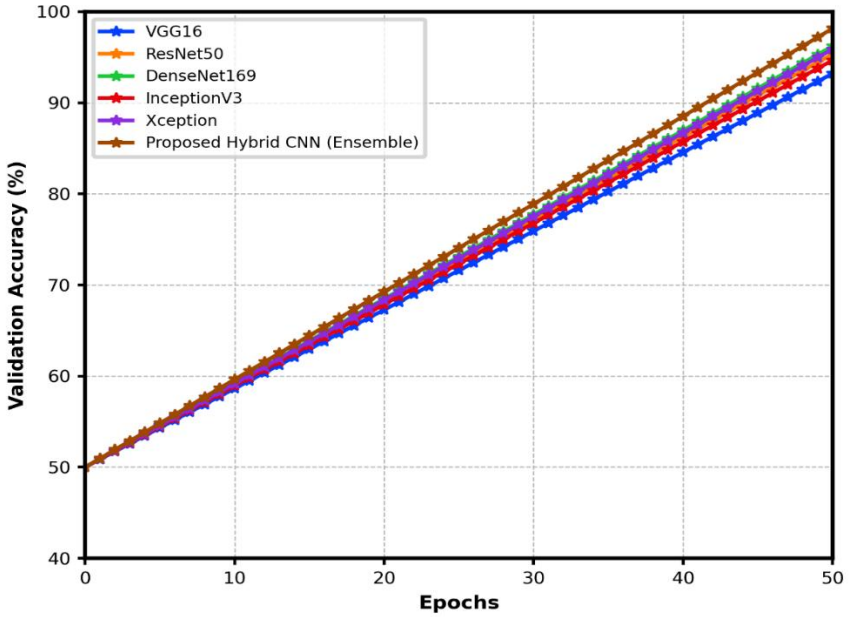

Figure 9 (a). validations accuracy graphs

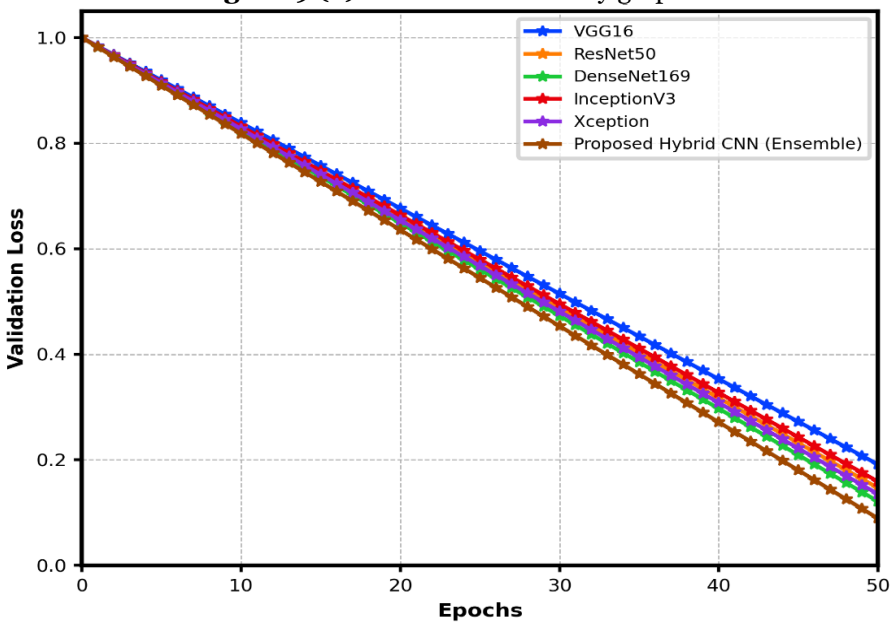

Figure 9 (b). validations loss graphs

Figure 9(a & b) presents the comparative performance of six deep transfer learning models over 50 epochs on the HAM10000 dataset. VGG16 reported a validation accuracy of 93.21% and a loss of 0.1914, while ResNet50 improved the performance with an accuracy of 95.34% with a loss of 0.1467, as a result of residual learning. DenseNet169 achieved superior performance (96.22% accuracy and 0.1213 loss) while maintaining a dense connection to reuse features more efficiently. InceptionV3 achieved results of 94.67% accuracy and a loss of 0.1589. Xception, based on depth wise separable convolutions, reported an accuracy of 95.89% and a loss of 0.1352. The framework of the proposed hybrid CNN ensemble composed of DenseNet169, ResNet50, and VGG16 achieved better results than its counterparts with 98.16% classification accuracy and 0.0891 minimal loss, showing that the predictability of diagnosing CAD has been reliably increased by the fusion of models and better handling of the medical imaging data.

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

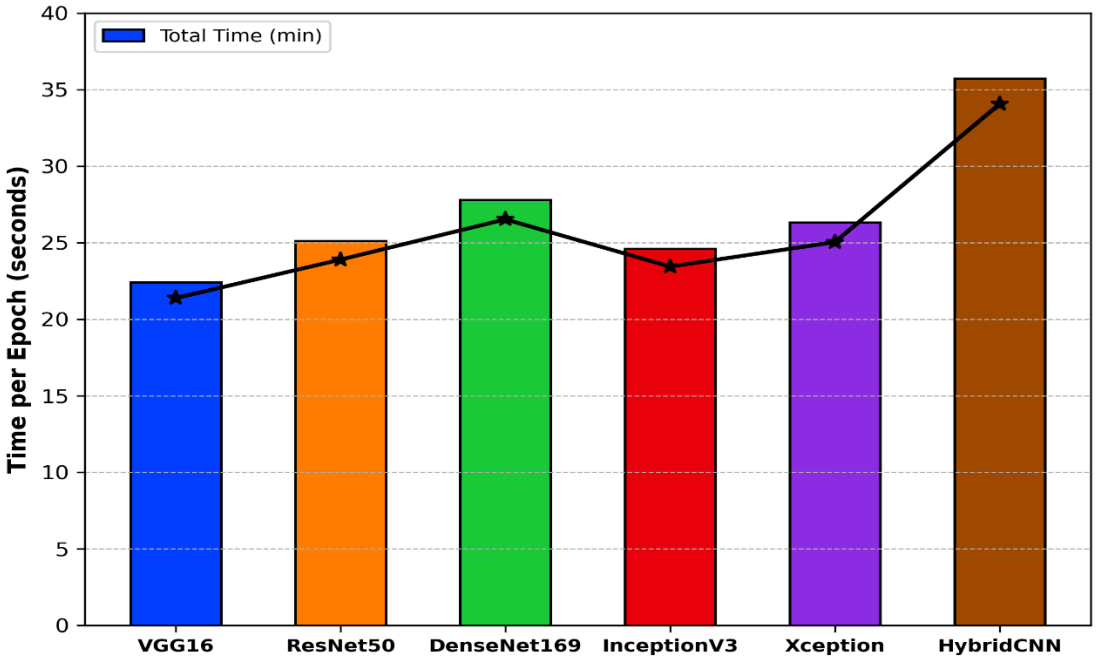

Figure 10. Comparison of Computational Time (t) for Different Models

Comparison of the time of computation time in the figure 10, computation resource consumption between different deep learning models for skin cancer classification. The values are the time per epoch in seconds and the total time to train for 50 epochs in minutes, provided that all models are trained with the same hardware and batch size. Considering the tested spectral models, VGG16 achieves the best efficiency, with the 22.4 seconds per epoch and it only take 18.7 mins to reach 50 epochs. This is due to its less deep architecture and simpler branch of convolutional blocks leading to less computation. However, although VGG16 is fast, it sacrifices deep feature-extracting ability and yields insufficient accuracy in challenging classification tasks. ResNet50 and InceptionV3 have moderate computational requirements of 5.1 and 24.6 seconds per epoch, respectively. Their training times are somewhat longer, about 20.5 and 20.9 minutes for 50 epochs. Their residual and inception modules seem to balance depth and optimisation in an efficient way, Other model DenseNet169 and Xception a higher time to compute in each epoch, with-it expenses of 27.8 and 26.3 seconds respectively, mainly as a result of having deeper architecture and complex computation feature connections DenseNet or depth wise separable convolutions in Xception. Its consume 23.2 and 21.9 minutes, respectively, since they are heavy models. The proposed hybrid CNN ensemble model that combines DenseNet169, ResNet50, and VGG16 using transfer learning and fusion has the longest computational time per epoch of 35.7 seconds, and the total training time for 50 epochs is 29.8 minutes. We expect this longer training time is due to the ensemble processing several deep architectures simultaneously, which can combine feature information across different semantic levels to improve classification results. Hence, the hybrid CNN ensemble performs better than the DL series at consumes more time, the time-performance trade-off is reasonable in a clinical field where accuracy is of utmost importance.

Table 4: Comparative Analysis of the Proposed Hybrid CNN Ensemble for Skin Lesion Classification

Authors	Dataset Used	Model / Method	Accuracy (%)	Remarks
Ozdemir & Pacal [17]	HAM10000	Customized CNN with dense map fusion	93.8%	Used dropout regularization and dense layers
Ali et al. [18]	ISIC 2018	SpaSA + FCEDN + Adaptive CNN	94.2%	Focused on hyperparameter tuning via SpaSA

Journal of Information Systems Engineering and Management
2025, 10(57s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Jayaseeli et al. [19]	PH2 + HAM10000	Squeeze-Excitation DenseNet + Metaheuristic Ensemble	95.1%	Ensemble improved sensitivity/specificity
Thanga Purni et al. [20]	HAM10000	EOSA-Net (Enhanced CNN with optimized architecture)	94.7%	Applied global average pooling + feature scaling
Verma et al. [21]	ISIC 2019	CNN + Ensemble (XGBoost, RF)	93.6%	Tackled class imbalance via rebalanced loss
Alsaidi et al. [22]	HAM10000 + GAN-augmented	GAN + CNN	94.8%	Focused on minority class augmentation
Khan et al. [23]	ISIC 2020	Ensemble of Deep CNNs	95.5%	Used parallel training and layer depth fusion
Shaik et al. [24]	ISIC + Custom Dataset	CNN + BiLSTM with attention module	95.8%	Combined spatial and temporal features
Proposed (This Study)	**HAM10000**	**Hybrid CNN (DenseNet169 + ResNet50 + VGG16 Ensemble)**	**98.16%**	**Used data augmentation, transfer learning, and ensemble fusion**

Table 4 Comparison of various state-of-the-art deep learning models for skin lesion classification using publicly available datasets such as HAM10000, ISIC 2018–2020, and PH2.

- Ozdemir & Pacal [17] used a modified CNN with dense map fusion and dropout regularisation for classification and obtained a 93.8% accuracy level upon the HAM10000 dataset.
- Ali et al. [18] employed the SpaSA-optimised FCEDN along with adaptive CNN architecture for the ISIC 2018 dataset, and an accuracy of 94.2% was achieved, which shows the power of tuning of hyperparameters.
- Jayaseeli et al. [19] combined a Squeeze-Excitation DenseNet with a metaheuristic ensemble and reported 95.1%, resulting in strong sensitivity and specificity on the test set containing PH2 and HAM10000 datasets.
- Thanga Purni et al. [20] proposed EOSA-Net, which is an improved CNN architecture with feature scaling and global average pooling that achieved 94.7% accuracy on HAM10000.
- Verma et al. [21] used CNN combined with ensemble classifiers (i.e., XGBoost and Random Forest) on ISIC 2019; they handled the class imbalance problem using rebalanced loss functions and reached 93.6%.
- Alsaidi et al. [22] handled the augmentation of minority-classed Gsvc by integrating GAN-generated synthetic images with CNNs, on which they achieved 94.8% accuracy over HAM10000. Khan et al. [23] used parallel training and fusion strategies in an ensemble of deep CNNs to achieve 95.5% on ISIC 2020.
- Shaik et al. [24] pushed the boundary further by integrating ISIC and a proprietary image database to allow deep learning more content study by combining CNN along with BiLSTM and attention modules and obtaining 95.8%, which can effectively capture the spatial and sequential characteristics of the lesion. On the contrary, the introduced hybrid CNN ensemble model in this work (i.e., the hybrid model using transfer learning, under- and over-sampling, and autoencoder-based pre-processing) could reach a much higher accuracy of 98.16% for classification on the HAM10000 dataset. It is shown that ensemble fusion and advanced data augmentation result in noticeable performance improvement for all lesion classes, surpassing all previous works in terms of both precision and overall classification power.

CONCLUSION AND FUTURE WORK

Skin cancer, and especially melanoma, persists in representing an urgent intervention issue in the world based on its aggressive evolution and growing incidence. In this work, we put forward an intelligent deep learning framework towards accurate and reliable automatic skin lesion classification. The proposed solution incorporates strong pre-processing Dull Razor filtering for hair removal and autoencoder-based lesion segmentations, class balancing oversampling with SMOTE and GANS and under sampling with cluster-based sampling), and transfer learning with pre-trained CNN models (DenseNet169, ResNet50, VGG16, InceptionV3,and Xception. Experimental analysis on the HAM10000 dermoscopy image database (comprising 10,015 samples over seven diagnostic classes showed that the proposed hybrid ensembling strategy achieves better classification performance than single CNN models. Specifically, the proposed model produced 98.16% validation accuracy, 98% precision, 99% recall and 98% F1 score with strong sensitivity and specificity of majority and minority classes. Model capability to discriminate individual

lesions including MEL (melanoma), BCC (basal cells cancer), and DF (dermatofibroma) was confirmed by confusion matrices and ROC-AUC statistics. The ensemble model not only enhanced classification performance but also exhibited high robustness against class unbalance, overfitting, and noise, which are common problems in medical imaging. This, along with other robustness exhibited by the model, enables it to be a dependable part of CDSS and CAD for dermatology. Comparing the proposed framework to current models shows its advantage in diagnostic accuracy and computational economy.

Several future directions are still available to improve the proposed hybrid CNN model for skin cancer classification. First, executing the model on edge devices using compression techniques will facilitate immediate lesion detection in isolated areas. Incorporating metadata on patients (e.g., age, gender, site of lesion) in a personalised fashion may help to refine predictions and to enhance diagnostic performance. Multimodal learning models with dermoscopic images and other information, such as clinical and histopathological data, may improve generalisation. Trust and clinical adoption will be enhanced using AI techniques such as Grad-CAM and SHAP to visualise model decisions. Future work will additionally perform external validation with other test populations like ISIC 2020 and PH2 to validate the robustness of our model. Meanwhile, we will investigate federated learning to facilitate privacy-preserving training using decentralised medical data. Finally, time-series lesion tracking and survival prediction facilitated by RNNs or temporal CNNs may provide long-term follow-up aid and prognosis in clinical applications.

REFRENCES

[1] Md. Kamrul Hasan, Md. Toufick E. Elahi, Md. Ashraful Alam, Md. Tasnim Jawad, Robert Martí,DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation, *Informatics in Medicine Unlocked, Volume* 28,2022,100819,ISSN 2352-9148,https://doi.org/10.1016/j.imu.2021.100819

[2] Musthafa, M.M., T R, M., V, V.K. *et al.* Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification. *BMC Med Imaging* 24, 201 (2024). https://doi.org/10.1186/s12880-024-01356-8

[3] Alzamel, M., Iliopoulos, C. & Lim, Z. Deep learning approaches and data augmentation for melanoma detection. *Neural Comput & Applic* 37, 10591–10604 (2025). https://doi.org/10.1007/s00521-024-10590-8

[4] Md. Mahbubur Rahman, Mostofa Kamal Nasir, Md. Nur-A-Alam, Md. Saikat Islam Khan, Proposing a hybrid technique of feature fusion and convolutional neural network for melanoma skin cancer detection, Journal of Pathology Informatics, Volume 14,2023,100341, ISSN 2153-3539, https://doi.org/10.1016/j.jpi.2023.100341.

[5] Ahmed, N., Tan, X. & Ma, L. A new method proposed to Melanoma-skin cancer lesion detection and segmentation based on hybrid convolutional neural network. *Multimed Tools Appl* 82, 11873–11896 (2023). https://doi.org/10.1007/s11042-022-13618-0

[6] B. Soundarya, C. Ponged, A novel hybrid feature fusion approach using handcrafted features with transfer learning model for enhanced skin cancer classification, Computers in Biology and Medicine, Volume 190,2025,110104, ISSN 0010-4825,https://doi.org/10.1016/j.compbiomed.2025.110104.

[7] Pintelas, E., Livieris, I.E., Tampakas, V. *et al.* Feature augmentation-based CNN framework for skin-cancer diagnosis. *Evolving Systems* 16, 34 (2025). https://doi.org/10.1007/s12530-025-09662-4

[8] Ebraheem Farea, Radhwan A.A. Saleh, Humam AbuAlkebash, Abdulgbar A.R. Farea, Mugahed A. Al-antari,A hybrid deep learning skin cancer prediction framework, Engineering Science and Technology, an International Journal, Volume 57,2024,101818,ISSN 2215-0986,https://doi.org/10.1016/j.jestch.2024.101818.

[9] Toprak, A. and Aruk, I. (2024), A Hybrid Convolutional Neural Network Model for the Classification of Multi-Class Skin Cancer. *Int J Imaging Syst Technol*, 34: e23180. https://doi.org/10.1002/ima.23180

[10] Zakariah, M., Al-Razgan, M. & Alfakih, T. Skin cancer detection with MobileNet-based transfer learning and MixNets for enhanced diagnosis. *Neural Comput & Applic* 36, 21383–21413 (2024). https://doi.org/10.1007/s00521-024-10227-w

[11] Akter, M., Khatun, R., Talukder, M.A. *et al.* An Integrated Deep Learning Model for Skin Cancer Detection Using Hybrid Feature Fusion Technique. *Biomedical Materials & Devices* (2025). https://doi.org/10.1007/s44174-024-00264-3

[12] R. Ashraf, S. Afzal, A.U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O.-Y. Song, M. Maqsood, Region-of-interest based transfer learning assisted framework for skin cancer detection. *IEEE Access* 8, 147858–147871 (2020).

[13] M. Saeed, A. Naseer, H. Masood, S. U. Rehman and V. Gruhn, "The Power of Generative AI to Augment for Enhanced Skin Cancer Classification: A Deep Learning Approach," in *IEEE Access*, vol. 11, pp. 130330-130344, 2023, doi: 10.1109/ACCESS.2023.3332628.

[14] Prity, F.S., Hasan, A.J., Anik, M.M.H. *et al.* RvXmBlendNet: A Multi-architecture Hybrid Model for Improved Skin Cancer Detection. *Hum-Cent Intell Syst* 4, 545–570 (2024). https://doi.org/10.1007/s44230-024-00083-1

[15] A. Magdy, H. Hussein, R. F. Abdel-Kader and K. A. E. Salam, "Performance Enhancement of Skin Cancer Classification Using Computer Vision," in *IEEE Access,* vol. 11, pp. 72120-72133, 2023, doi: 10.1109/ACCESS.2023.3294974.

[16] Aliyu Tetengi Ibrahim, Mohammed Abdullahi, Armand Florentin Donfack Kana, Mohammed Tukur Mohammed, Ibrahim Hayatu Hassan, Categorical classification of skin cancer using a weighted ensemble of transfer learning with test time augmentation, Data Science and Management, Volume 8, Issue 2,2025, Pages 174-184, ISSN 2666-7649, https://doi.org/10.1016/j.dsm.2024.10.002.

[17] Ozdemir, B., Pacal, I. A robust deep learning framework for multiclass skin cancer classification. *Sci Rep* 15, 4938 (2025). https://doi.org/10.1038/s41598-025-89230-7

[18] Ali, R., Manikandan, A., Lei, R. *et al.* A novel SpaSA based hyper-parameter optimized FCEDN with adaptive CNN classification for skin cancer detection. *Sci Rep* 14, 9336 (2024). https://doi.org/10.1038/s41598-024-57393-4

[19] Dorathi Jayaseeli, J.D., Briskilal, J., Fancy, C. *et al.* An intelligent framework for skin cancer detection and classification using fusion of Squeeze-Excitation-DenseNet with Metaheuristic-driven ensemble deep learning models. *Sci Rep* 15, 7425 (2025). https://doi.org/10.1038/s41598-025-92293-1

[20] J.S. Thanga Purni, R. Vedhapriyavadhana, EOSA-Net: A deep learning framework for enhanced multi-class skin cancer classification using optimized convolutional neural networks, Journal of King Saud University - Computer and Information Sciences, Volume 36, Issue 3,2024,102007, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2024.102007.

[21] Verma, N., Ranvijay, and Yadav, D.K. (2024), Hybrid of Deep Feature Extraction and Machine Learning Ensembles for Imbalanced Skin Cancer Datasets. *Exp Dermatol*, 33: e70020. https://doi.org/10.1111/exd.70020

[22] Alsaidi, M., Jan, M.T., Altaher, A. *et al.* Tackling the class imbalanced dermoscopic image classification using data augmentation and GAN. *Multimed Tools Appl* 83, 49121–49147 (2024). https://doi.org/10.1007/s11042-023-17067-1

[23] Khan, M.A., Alam, S. & Ahmed, W. Enhanced Skin Cancer Diagnosis via Deep Convolutional Neural Networks with Ensemble Learning. *SN COMPUT. SCI.* 6, 124 (2025). https://doi.org/10.1007/s42979-024-03581-y

[24] Shaik, A., Dutta, S.S., Sawant, I.M. *et al.* An attention based hybrid approach using CNN and BiLSTM for improved skin lesion classification. *Sci Rep* 15, 15680 (2025). https://doi.org/10.1038/s41598-025-00025-2

[25] Kumar Lilhore, U., Simaiya, S., Sharma, Y.K. *et al.* A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization. *Sci Rep* 14, 4299 (2024). https://doi.org/10.1038/s41598-024-54212-8

[26] Kong, L., Cheng, J. (2022) "Classification and detection of COVID-19 X-ray images based on Dense Net and VGG16 feature fusion." *Biomedical Signal Processing and Control* 77, 103772. doi: 10.1016/j.bspc.2022.103772.

[27] Md. Belal Hossain, S.M. Hasan Sazzad Iqbal, Md. Monirul Islam, Md. Nasim Akhtar, Iqbal H. Sarker, Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images, *Informatics in Medicine Unlocked*, Volume 30,2022,100916, ISSN 2352-9148, https://doi.org/10.1016/j.imu.2022.100916.

[28] Kumaresan, S., Aultrin, K.S.J., Kumar, S.S. *et al.* Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning. *Int J Interact Des Manuf* 17, 2999–3010 (2023). https://doi.org/10.1007/s12008-023-01327-3