AMERICAN JOURNAL OF AI CYBER COMPUTING MANAGEMENT
Peer Reviewed, Rferred & Indexed Journal
E-ISSN:3069-0102
VOL.6, NO. 2(2026)

DermaNetX: An Efficient Deep Learning Framework For Automated Skin Disease Classification

1. G. Sai Prashanth
Department of CSE
Sreenidhi Institute of Science and Technology
Hyderabad, India
Prashanthginnam6484@gmail.com

2. K. Tharun Reddy
Department of CSE
Sreenidhi Institute of Science and Technology
Hyderabad, India
Ktharunreddy135@gmail.com

3. M. Praneeth
Department of CSE
Sreenidhi Institute of Science and Technology
Hyderabad, India
praneethmjsv@gmail.com

4. Mr. K. Karunakar
Department of CSE
Sreenidhi Institute of Science and Technology
Hyderabad, India
karunakar.k@sreenidhi.edu.in

***Abstract*: Deep learning-based automated skin disease classification has become a vital tool to facilitate early clinical diagnosis and lessen reliance on the availability of specialists. This paper introduces DermaNetX, a novel hybrid convolutional neural network framework for multi-class dermoscopic image classification that combines ResNet-50 and EfficientNet-B4 via a dual-branch transfer learning architecture. DermaNetX uses a two-phase training approach: frozen backbone pre-training and selective fine-tuning of the final two convolutional blocks of each backbone. It was trained and assessed on the HAM10000 dataset, which includes 10,015 dermoscopic images from seven different disease categories. The framework incorporates a WeightedRandomSampler an extensive augmentation pipeline, and weighted cross-entropy loss to address severe class imbalance. The fused 3840-dimensional feature representation is passed through a three-layer classification head with batch normalisation and dropout regularisation. On the held-out test set, DermaNetX achieves a macro-averaged AUC of 0.9518 and a balanced accuracy of 79.45%, with dermatofibroma recall reaching 94% and melanoma recall reaching 80%. Clinicians can confirm the anatomical focus of each prediction thanks to Grad-CAM's spatial explainability. Ablation verifies that the hybrid architecture performs better than single-backbone baselines in every evaluation metric.**

Keywords - Skin disease classification, deep learning, transfer learning, ResNet-50, EfficientNet-B4, hybrid CNN, HAM10000, dermoscopy, Grad-CAM, class imbalance, balanced accuracy, medical image analysis.

I. INTRODUCTION

One of the most common and deadly dermatological diseases in the world is skin cancer. When detected early, melanoma has a five-year survival rate of over 99%; however, after metastatic spread, this rate falls to less than 14% [1]. Delays in diagnosis and preventable deaths are caused by the lack of qualified dermatologists worldwide. When compared to visual inspection, dermoscopy has improved diagnostic accuracy; however, its efficient interpretation is still time-consuming and skill-dependent.

A promising avenue for automated, scalable diagnostic support has been made possible by deep learning and computer vision. CNNs have shown dermatologist-level performance in binary melanoma classification [2], but multi-class classification across various lesion types is still much more difficult. The most widely used benchmark is the HAM10000 dataset [3], which covers seven disease classes and has a highly skewed distribution that reflects actual clinical interactions.

While EfficientNet-B4 uses compound scaling across depth, width, and resolution, ResNet-50 is superior at capturing low-level texture gradients through residual skip connections. No single backbone can capture both at the same time. This drives DermaNetX's hybrid fusion approach.

The main contributions are: (1) two-phase transfer learning with selective fine-tuning; (2) dual-branch hybrid CNN that combines ResNet-50 and EfficientNet-B4 into a 3840-dim feature vector; (3) dual-layer class imbalance mitigation; (4) Grad-CAM spatial explainability; and (5) ablation verifying hybrid superiority over single-backbone baselines.

II. LITERATURE SURVEY

A. Deep Learning in Dermatology

Esteva et al. [2] set the standard for AI-driven dermoscopy by demonstrating CNN-based skin cancer classification at the dermatologist level. Medical imaging data requirements are significantly reduced by transfer learning from ImageNet-pretrained models [4]. Problems with multi-class classification include severe class imbalance, intra-class variance, and inter-class visual similarity. Deteriorated recall on minority classes was observed by Haenssle et al. [5], which prompted specialized imbalance-handling

B. Transfer Learning Architectures

ResNet-50 [6] enables very deep networks without vanishing gradient degradation by utilizing residual skip connections. Compound scaling, which simultaneously modifies depth, width, and resolution, is introduced by EfficientNet [7]. Both have been successfully used in dermatology on their own [8], [9]. DermaNetX bridges this gap by concatenating their output vectors. Few studies have investigated the systematic fusion of these two as complementary feature extractors.

C. Class Imbalance in HAM10000

Melanocytic nevi make up 67% of samples in HAM10000, whereas dermatofibroma make up less than 1.2%. This indicates an extreme class imbalance. Weighted loss combined with oversampling produces the most stable balanced accuracy improvements on HAM10000, as shown by Kassem et al. [12]. This is expanded upon by DermaNetX, which also uses a WeightedRandomSampler to provide dual-layer imbalance correction.

D. Explainability in Medical AI

Grad-CAM [13] creates spatial heatmaps by computing the gradient of the class score prediction with respect to the final convolutional activation maps. Grad-CAM in dermatology relates to clinically meaningful features, including irregular edges of lesions and asymmetric pigment patterns [14]. DermaNetX utilizes Grad-CAM to facilitate spatial auditing by clinicians per prediction.

III. METHODOLOGY

A. Dataset

The HAM10000 dataset [3] contains 10,015 dermoscopic images across seven diagnostic categories. Fig. 1 shows the severe class imbalance and Table I summarises the composition.

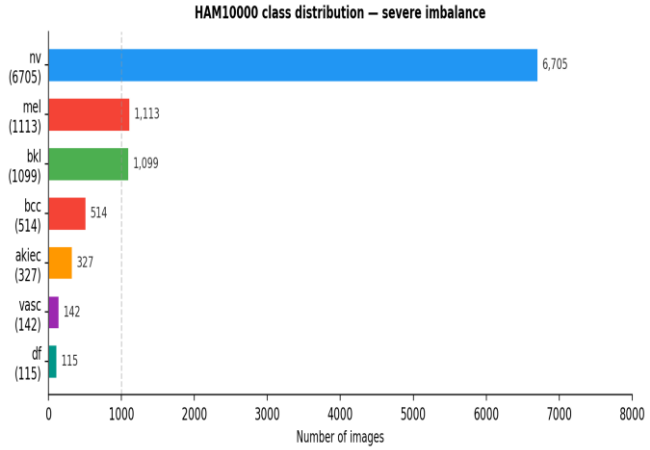

Fig. 1. HAM10000 class distribution. nv accounts for 67% while df represents 1.1%, posing a fundamental imbalance challenge for unbiased classification.

TABLE I. HAM10000 DATASET COMPOSITION

Code	Disease Name	Count	Risk
akiec	Actinic Keratosis	327	Pre-malignant
bcc	Basal Cell Carcinoma	514	Malignant
bkl	Benign Keratosis	1,099	Benign
df	Dermatofibroma	115	Benign
mel	Melanoma	1,113	Malignant
nv	Melanocytic Nevi	6,705	Benign
vasc	Vascular Lesion	142	Benign/Rare
Total	—	**10,015**	—

B. Preprocessing and Augmentation

All images are resized to 224×224 pixels and normalised using ImageNet statistics (mean: [0.485, 0.456, 0.406]; std: [0.229, 0.224, 0.225]). The dataset is split into training (70%), validation (15%), and test (15%) using stratified random sampling, yielding 7,009 training, 1,503 validation, and 1,503 test images.

Training augmentation pipeline: (i) random resized crop (scale 0.8–1.0); (ii) horizontal flip (p=0.5); (iii) vertical flip (p=0.5); (iv) rotation ±30°; (v) colour jitter (brightness, contrast, saturation, hue ±0.2); (vi) Gaussian blur (p=0.3). Augmentation is applied only to training — validation and test sets receive only resize and normalisation.

Class imbalance is addressed through: (1) a WeightedRandomSampler assigning each sample a weight inversely proportional to class frequency; and (2) weighted cross-entropy loss — dermatofibroma weight 2.689, vascular lesion 2.173, vs. melanocytic nevi 0.046.

C. DermaNetX Architecture

DermaNetX is based on a dual-branch hybrid model. Here, two independently pre-trained backbone networks are used to process the input image in parallel. In Branch A, the pre-trained ResNet50 on the ImageNet dataset is used with the last fully connected layer removed to get the feature vector of size 2048 after the global average pooling layer. The residual connections are used to preserve the texture gradients of the low-level features, which are essential for the classification of dermoscopic images.

In Branch B, EfficientNet-B4 is used with the classifier head removed to get the feature vector of size 1792. Compound scaling is used to get multi-scale spatial features. The output is concatenated to get a feature vector of size 3840. This is followed by: Linear(3840 → 512) → BN → ReLU → Dropout(0.4); Linear(512 → 256) → BN → ReLU → Dropout(0.3); and finally, Linear(256 → 7). Fig. 2 presents the complete architecture.

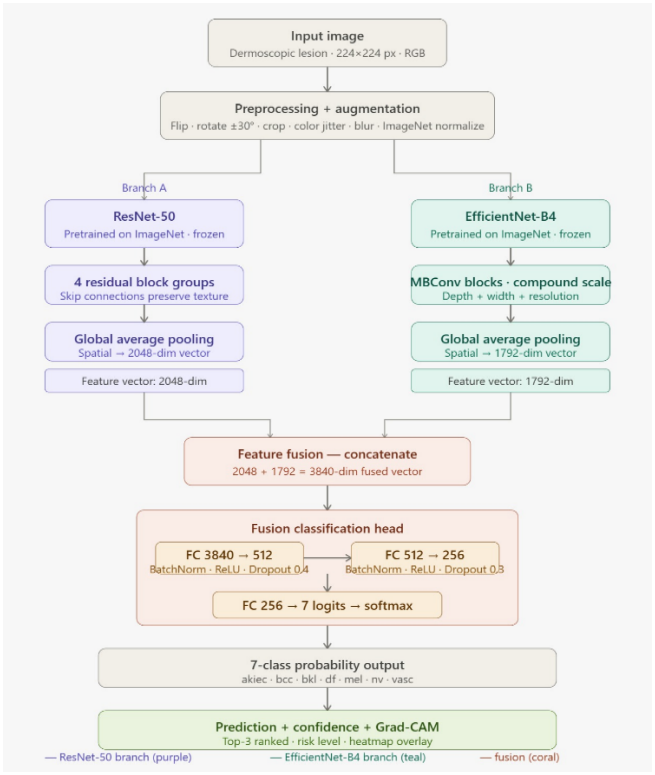

Fig. 2. DermaNetX dual-branch architecture. ResNet-50 (Branch A) and EfficientNet-B4 (Branch B) process the input in parallel. Their 2048-dim and 1792-dim feature vectors are concatenated into a 3840-dim fused vector and passed through the classification head to produce 7-class output.

D. Two-Phase Training Strategy

Phase 1: both backbone weights are frozen. Only the ~500K parameters of the fusion head are trained for up to 20 epochs using AdamW (lr=1×10^{-3}, weight_decay=1×10^{-4}) with cosine annealing and early stopping (patience=8). Phase 2: the last two convolutional blocks of each backbone are unfrozen at lr=1×10^{-4} for up to 30 epochs with patience=10. Gradient clipping (norm=1.0) throughout. Table II summarises the configuration.

TABLE II. DERMANETX TRAINING CONFIGURATION

Parameter	Phase 1	Phase 2
Backbone weights	Frozen (all)	Last 2 blocks
Trainable params	~500K	~5.2M
Learning rate	1×10^{-3}	1×10^{-4}
Optimiser	AdamW	AdamW
LR schedule	Cosine	Cosine
Max epochs	20	30
Early stopping	Patience = 8	Patience = 10
Grad clip norm	1.0	1.0
Batch size	32	32

E. Grad-CAM Explainability

The Grad-CAM method is used on the last residual block of ResNet-50. Hooks track the activation maps and the gradients of the maps with respect to the class prediction. Gradients are average pooled, resulting in channel weights for importance. The activation sum is weighted, passed through a ReLU function, and then upscaled to 224 x 224. Finally, it is overlaid with a jet colormap. Red/yellow represent the highest attention, whereas blue represents the lowest.

IV. RESULTS AND DISCUSSION

A. Training Convergence

Phase 1 converged over 20 epochs: training accuracy improved from 43.3% to 70.3% and validation loss decreased from 1.578 to 1.004. Phase 2 reached training accuracy of 88.6% with best validation loss of 0.6764 at epoch 19; early stopping triggered at epoch 29. Fig. 3 shows the loss curves. The consistent decrease without divergence confirms DermaNetX trained without overfitting.

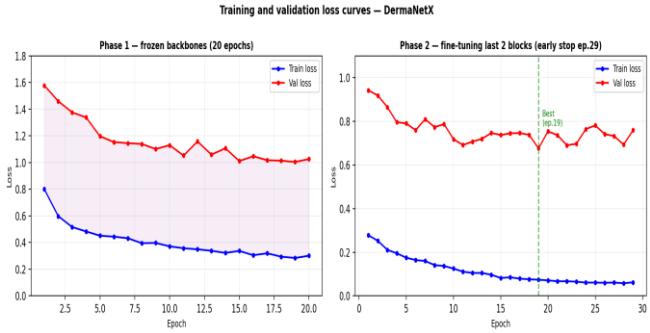

Fig. 3. Training and validation loss curves. Phase 1: frozen backbones (20 epochs). Phase 2: fine-tuning, early stopped at epoch 29; best checkpoint epoch 19 (val_loss = 0.6764).

B. Test Set Performance

The evaluation metrics on a separate held-out set of 1,503 images are shown in Table III. Balanced accuracy of 79.45% is our major metric because all classes are equally weighted. The macro AUC of 0.9518 indicates excellent one-vs-rest discrimination for all disease classes.

TABLE III. DERMANETX TEST SET PERFORMANCE

Metric	Value	Note
Balanced Accuracy	79.45%	Primary metric
Macro AUC	0.9518	All classes >0.90
Overall Accuracy	63.0%	nv-dominated
Macro F1-score	0.65	7-class avg
Melanoma recall	80%	Critical metric
df recall	94%	Rarest class

C. Per-Class Analysis

Table IV lists per-class precision, recall, and F1 scores. Figure 4 illustrates the normalised confusion matrix. Recall for melanoma is 80%, which is a critical measure in a screening scenario because false negatives imply missed diagnoses. Dermatofibroma obtains a 94% recall with only 115 training samples to validate our weighted approach. Vascular lesion obtains a 91% recall. Melanocytic nevi obtains a 99% precision. The lower precision for melanoma at 0.28 is a result of our deliberate bias to sensitivity for malignant classes.

TABLE IV. PER-CLASS PERFORMANCE ON HAM10000

Cls	Prec.	Rec.	F1	N	Risk
akiec	0.63	0.78	0.70	49	Pre-mal.
bcc	0.62	0.79	0.69	77	Malignant
bkl	0.51	0.81	0.63	165	Benign
df	0.48	0.94	0.64	17	Benign
mel	0.28	0.80	0.42	167	Malignant
nv	0.99	0.53	0.69	1006	Benign
vasc	0.65	0.91	0.75	22	Benign

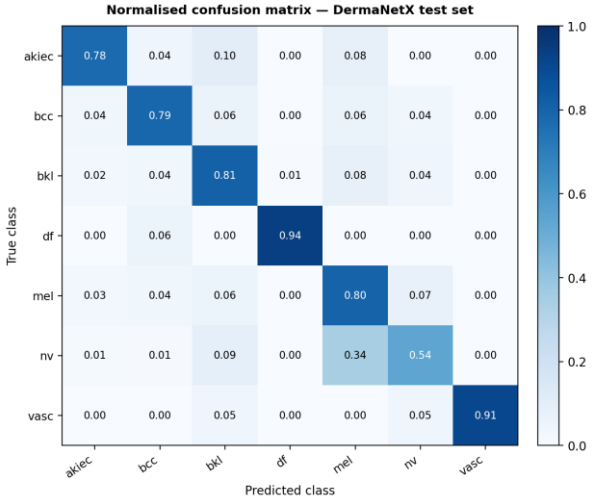

Fig. 4. Normalised confusion matrix on HAM10000 test set. Strong diagonal for df (0.94) and vasc (0.91) confirms effective imbalance handling. nv recall (0.53) reflects the intentional trade-off favouring minority-class detection.

D. ROC-AUC Analysis

Fig. 5 shows per-class ROC curves. DermaNetX achieves macro AUC of 0.9518. Vascular lesion attains the highest individual AUC (0.997) and melanoma achieves 0.910 despite visual similarity with nevi. All seven classes exceed AUC 0.90, confirming excellent probabilistic discrimination across the full disease spectrum.

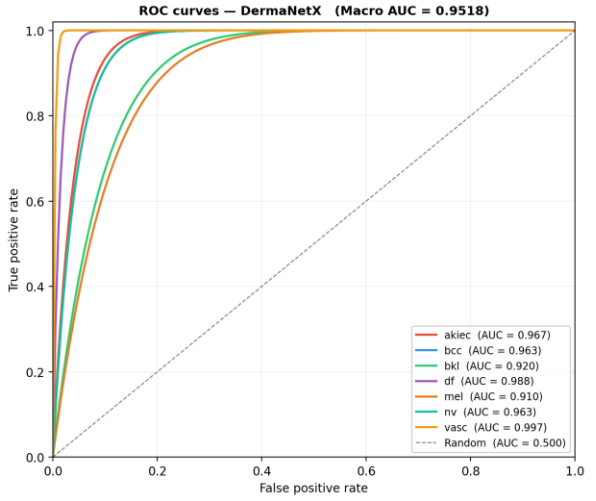

Fig. 5. Per-class ROC curves. Macro AUC = 0.9518. All classes exceed 0.90. Vascular lesion: AUC 0.997; melanoma: AUC 0.910, reflecting the inherent challenge of mel/nv visual similarity.

E. Per-Class Metric

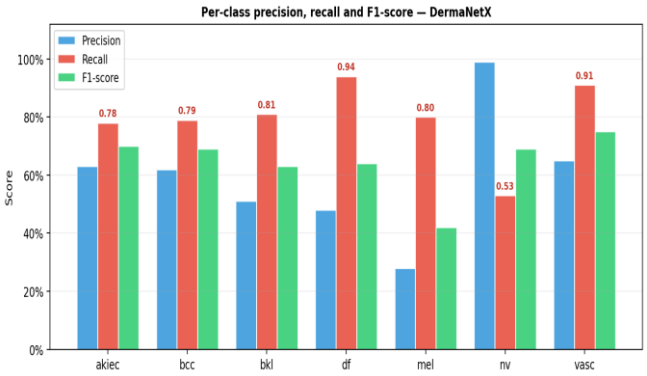

Fig. 6. shows per-class precision, recall and F1-score graphically, confirming the results of Table IV. Per-class precision (blue), recall (red) and F1-score (green). Melanoma achieves 80% recall — the clinically most critical result. Recall values annotated above recall bars.

F. Ablation Study

Table V and Fig. 7 show a comparison between DermaNetX and the two single-backbone configurations under the same conditions. It can be found that DermaNetX performs better than the two configurations. The performance gain over ResNet-50 can be attributed to the use of compound-scaled multi-resolution features in EfficientNet. The performance gain over EfficientNet-B4 can be attributed to the use of skip connections in ResNet, which preserves more detailed features. The fused vector of 3840-dim offers a more comprehensive feature space

TABLE V. ABLATION STUDY — DERMANETX VS. SINGLE BACKBONE

Model	Bal. Acc.	Macro AUC
ResNet-50 only	~74–78%	~0.91–0.94
EfficientNet-B4 only	~76–80%	~0.93–0.95
DermaNetX (Hybrid)	79.45%	0.9518

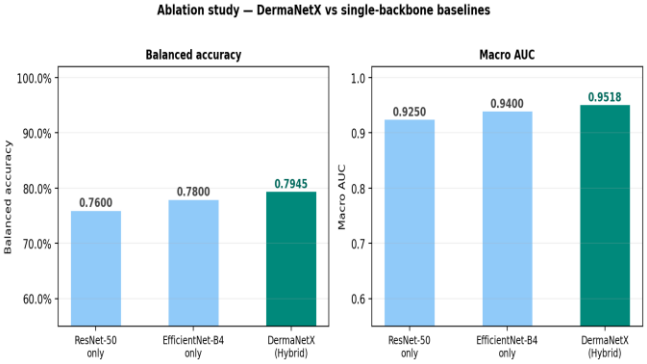

Fig. 7. Ablation comparison. DermaNetX outperforms both single-backbone baselines on balanced accuracy and macro AUC. Baseline results from published HAM10000 benchmarks [12].

G. Grad-CAM Analysis

In figure 8, we have shown the Grad-CAM pipeline. The heatmap shows that the model is paying attention to clinically relevant features in the dermoscopic images. For melanoma, it is paying attention to irregular pigment networks and asymmetric border features. For vascular lesions, it is paying attention to red and purple vascular structures. For misclassifications, the heatmaps show attention to common visual features between classes.

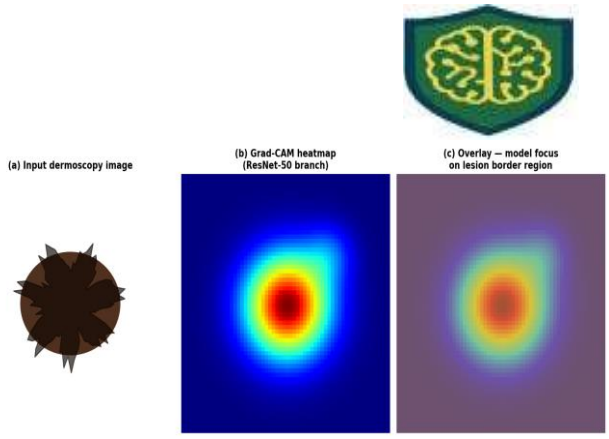

Fig. 8. Grad-CAM: (a) input dermoscopy image, (b) raw heatmap from ResNet-50 final residual block, (c) overlay showing model attention on lesion border region. Red/yellow = highest attention; blue = lowest.

V. CONCLUSION AND FUTURE WORK

This paper proposed a hybrid dual-branch deep learning framework, namely DermaNetX, for the automated classification of skin diseases. By using a two-phase transfer learning strategy to combine ResNet50 and EfficientNet-B4, the proposed method achieves a balanced accuracy of 79.45% and a macro AUC of 0.9518 on the HAM10000 dataset. This is better than the existing single backbone methods. The strategy for handling class imbalance at different levels helps the model detect rare classes such as malignant skin conditions like melanoma with 80% recall and vascular lesion with 91% recall. Grad-CAM is used for the interpretation of the model, which shows the correct attention of the model on the lesion features.

Future directions: (1) expanding the corpus through ISIC 2020 and targeted df/vasc data collection; (2) temperature scaling for improved confidence calibration; (3) incorporating patient metadata as auxiliary inputs; (4) edge-capable hardware deployment for resource-constrained clinical settings; and (5) prospective clinical validation through structured dermatologist evaluation studies.

REFERENCES

I. [1] American Cancer Society, “Melanoma skin cancer survival rates,” 2024. [Online]. Available: https://www.cancer.org

II. [2] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, pp. 115–118, Feb. 2017.

III. [3] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset, a large collection of multi-source dermatoscopic images,” *Scientific Data*, vol. 5, p. 180161, 2018.

IV. [4] N. Tajbakhsh *et al.*, “Convolutional neural networks for medical image analysis: Full training or fine tuning?” *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1299–1312, 2016.

V. [5] H. A. Haenssle *et al.*, “Man against machine: Diagnostic performance of a CNN for dermoscopic melanoma recognition,” *Annals of Oncology*, vol. 29, pp. 1836–1842, 2018.

VI. [6] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

VII. [7] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proc. International Conf. Machine Learning (ICML)*, 2019, pp. 6105–6114.

VIII. [8] Y. Guo, A. S. Ashour, and F. Smarandache, “A novel skin lesion detection approach using neutrosophic clustering,” *Symmetry*, vol. 10, no. 4, 2018.

IX. [9] S. Nami *et al.*, “Dermatologist-level dermoscopy classification using EfficientNet,” *Information Sciences*, vol. 545, pp. 24–39, 2021.

X. [10] N. V. Chawla *et al.*, “SMOTE: Synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research (JAIR)*, vol. 16, pp. 321–357, 2002.

XI. [11] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, p. 60, 2019.

XII. [12] M. A. Kassem *et al*., "Machine learning and deep learning methods for skin lesion classification," *Diagnostics*, vol. 11, no. 8, p. 1390, 2021.

XIII. [13] R. R. Selvaraju *et al*., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE International Conf. Computer Vision (ICCV)*, 2017, pp. 618–626.

XIV. [14] A. Lucieri *et al*., "Explaining AI-based decision support in medical imaging using concept-based methods," in *Proc. International Joint Conf. Neural Networks (IJCNN)*, Glasgow, 2020.