SKIN LESION DETECTION

Using Deep Learning

Guided by
Dr. Manpreet Kaur

Presented by

Dogga Pavan Sekhar	2K22CSUN01167
Narapureedy Durga Prasad Reddy	2K22CSUN01180
Kesanakurthi Naga Siddhartha	2K22CSUN01172

PROBLEM STATEMENT

- Early detection of skin cancer is important for improving patient survival.
- Diagnosing skin cancer is difficult because lesions often look very similar.
- Lesions vary a lot in size, shape, and color, making detection harder.
- Images contain noise like hair and shadows, which reduces clarity.
- Manual diagnosis is slow and can differ from one dermatologist to another.
- Skin cancer cases in India are increasing every year.
- Around 25,000–30,000 new cases are reported yearly in India.

- Melanoma cases are rising by about 3–5% annually in high-UV areas.
- Late detection is common and reduces chances of recovery.
- Existing AI systems struggle with class imbalance and poor generalisation.
- Many models do not provide clear or interpretable results for doctors.
- This project aims to build an automated system to detect and classify skin lesions.
- The system will address data imbalance, feature variation, and interpretability.

The goal is to achieve high accuracy and support dermatologists in early diagnosis.

DATASET DETAILS

The dataset used in this project is HAM10000.

- HAM10000 stands for “Human Against Machine with 10,000 images.”
- Contains 10,015 dermatoscopic images of pigmented skin lesions.
- The dataset includes 7 classes: akiec, bcc, bkl, df, nv, mel, and vasc.
- Images come from multiple sources and different imaging devices.
- Includes metadata such as age, sex, and lesion location.
- Many cases are histopathology-verified, which increases reliability.
- Images are high-quality and primarily dermoscopic, not clinical photos.
- Designed to support skin cancer classification research.
- One of the most widely used datasets for melanoma detection.
- Contains class imbalance, with “nv” being the largest class.
- Used often for training, benchmarking, and evaluating deep learning models.
- Available through the ISIC Archive and Harvard Dataverse.

Figure 1: *Sample Images in HAM10000*

METHODOLOGY

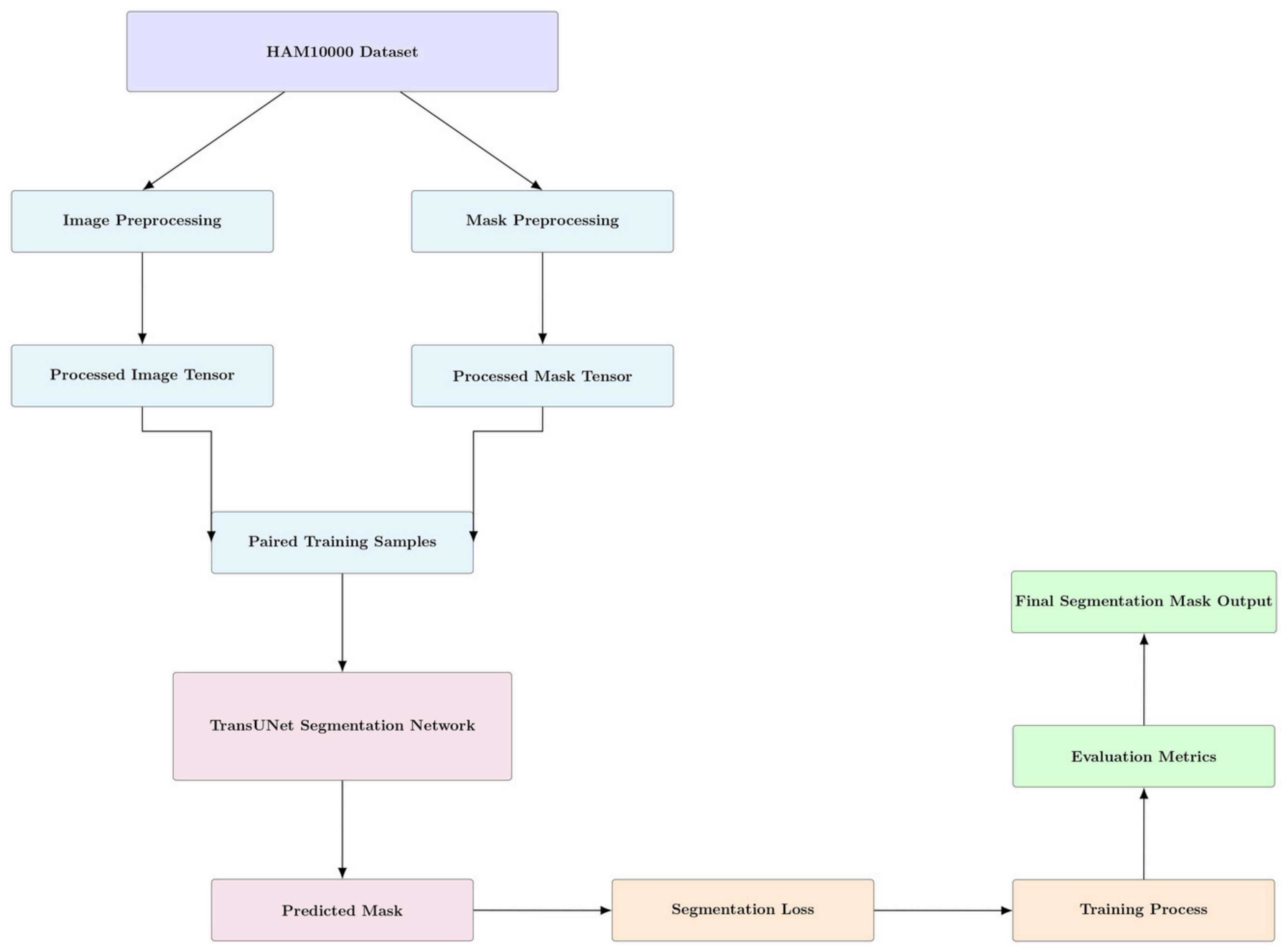


Figure 2: *Segmentation Pipeline Using TransUNet for Automated Skin Lesion Mask Generation*

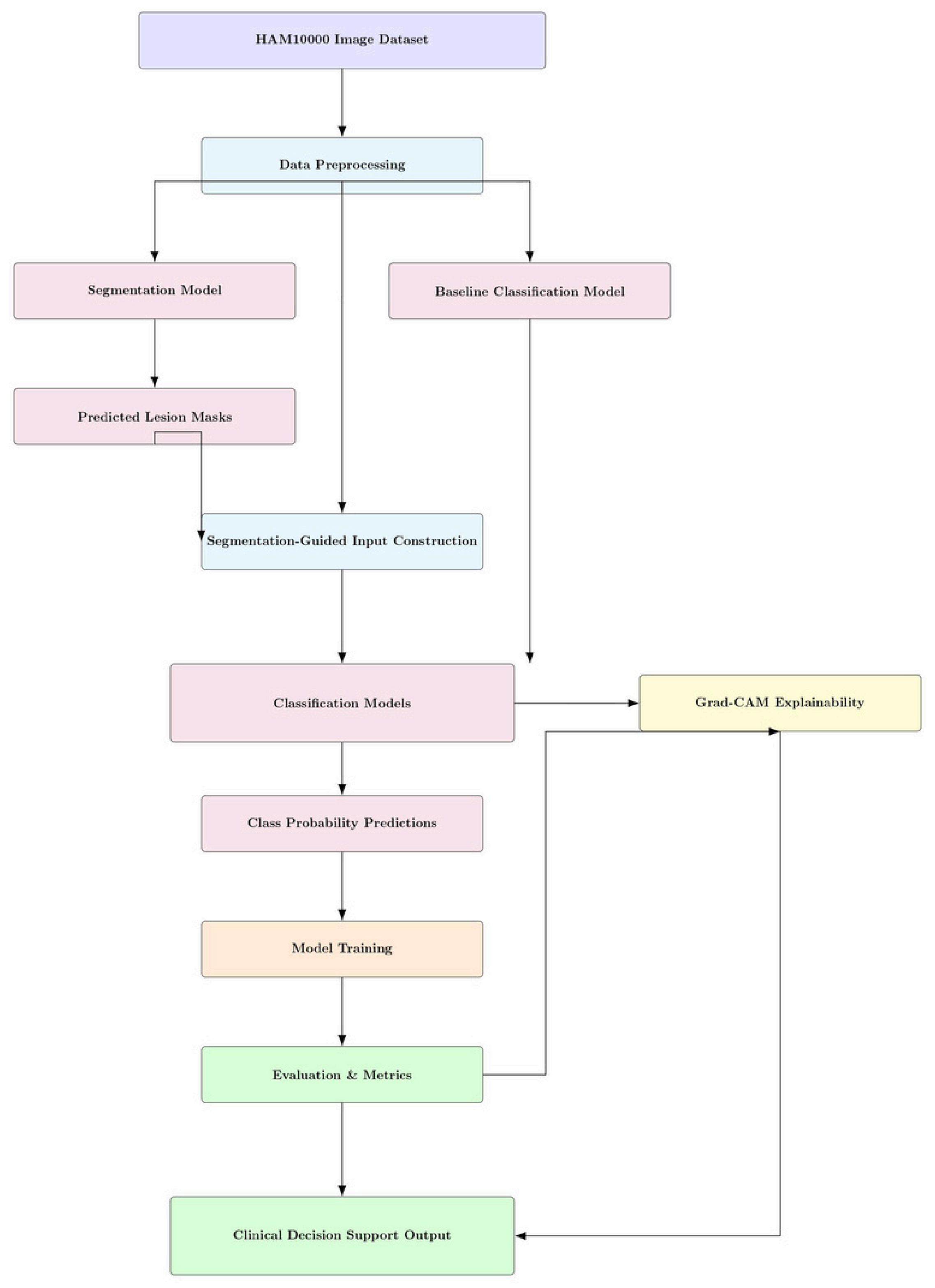


Figure 3: *Combined Segmentation-Guided and Baseline Classification Pipeline for Skin Lesion Diagnosis*

MODELS TRAINED

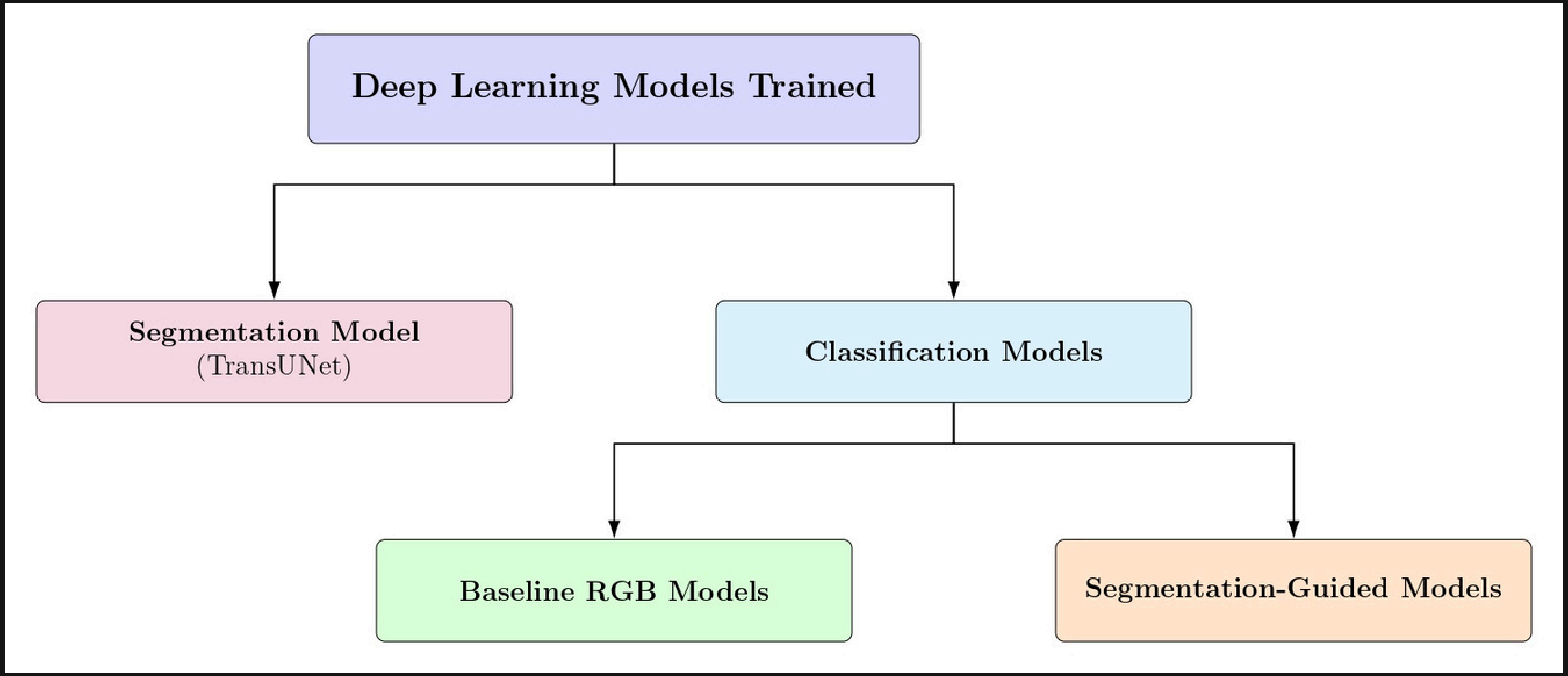


Figure 4: *Hierarchical overview of the deep learning models trained in the study.*

SEGMENTATION MODEL

The TransUNet-AEO model was trained for skin lesion segmentation.

- The model uses a UNet-style encoder–decoder with a transformer bottleneck.
- Dermoscopic images were resized to 256×256 during preprocessing.
- Image normalization was applied using ImageNet mean and standard deviation.
- Binary masks were generated from ground-truth segmentation images.
- Data augmentation included flips, rotations, shifts, scaling, and noise.
- Dice + Binary Cross Entropy loss was used for training.
- The Adam optimizer was used with a learning rate of 1e-4.
- The model includes convolutional encoder layers and transformer-based token mixing.
- The decoder reconstructs segmentation maps with skip connections.
- Dice coefficient was calculated as a primary evaluation metric.
- IoU (Intersection over Union) was computed for segmentation quality.
- Pixel accuracy was measured for overall pixel-level correctness.
- A pixel-wise confusion matrix was generated for background vs lesion.

BASELINE RGB MODELS

Feature	ResNet-34 (RGB Lesion Crop)	DB-LARNet (Dual-Branch RGB)	PLA-MIL (Patch-Level Attention RGB)
Input Type	Single RGB lesion crop (224×224)	Full RGB image + Lesion-crop RGB	RGB lesion crop split into 4×4 patches
Preprocessing	Resize, Normalize, Augmentations (flip, rotate, color jitter)	Same (applied separately to both branches)	Resize, Normalize, Augmentations, Patch extraction (16 patches)
Architecture Type	ResNet-34 backbone	Dual ResNet-18 backbones + Fusion	ResNet-18 Patch Encoder + Attention MIL
Model Complexity	Moderate	High (two CNN branches)	Moderate–High (patch encoder + attention module)
Feature Extraction Strategy	Global features from entire lesion crop	Combines global context + localized lesion region	Local patch-level features aggregated with attention
Lesion Localization Used?	No	Yes (uses explicit lesion crop)	Yes (patch-based spatial reasoning)
Optimizer Used	Adam	Adam	Adam
Learning Rate	0.0001	0.0001	0.0001
Loss Function	Weighted Cross-Entropy	Weighted Cross-Entropy	Weighted Cross-Entropy
Primary Metric Evaluated	Accuracy (Val Accuracy)	Accuracy (Val Accuracy)	Accuracy (Val Accuracy)
Strengths	Simple, fast, strong baseline	Learns global + focused lesion features	Strong for subtle patterns; interpretable attention
Limitations	May miss fine lesion details	More computationally expensive	Slower training due to many patches
Output	7-class probability distribution	7-class probability distribution	7-class probability distribution

Table 1: *Comparision of the Baseline RGB Models in this study.*

SEGMENTATION GUIDED MODELS

The Segmentation-Guided ResNet-34 model was trained for skin lesion segmentation.

- Uses 4-channel input combining RGB image and binary lesion mask.
- Lesion masks are loaded, resized, binarised, and concatenated with the RGB tensor.
- Model architecture modifies the ResNet-34 first convolution layer to accept 4 channels.
- The dataset was split using a stratified train–val split based on label counts.
- Uses data normalisation with ImageNet mean and std for RGB channels.
- Handles class imbalance using WeightedRandomSampler with inverse class frequencies.
- Applies cross-entropy loss with class-wise weights for balanced training.
- Optimised using the Adam optimiser with a learning rate of 1e-4.
- Uses the ReduceLROnPlateau scheduler to adjust the learning rate dynamically.
- Trained for 100 epochs with batch size 32 on GPU when available.
- Primary evaluation metric: validation accuracy.
- Additional performance metrics generated: confusion matrix, classification report, ROC curves.
- Training outputs include loss curves and accuracy curves.
- The best model checkpoint was saved based on the highest validation accuracy.
- Final predictions plotted using one-vs-rest ROC curves for all 7 classes.

RESULTS

Model	Type	Final Val Accuracy
TransUNet	Segmentation	Dice = 0.9142 IoU = 0.8576 Acc = 0.9548
ResNet-34 (Lesion Crops)	Baseline RGB	0.8777
DB-LARNet (Dual-Branch)		0.8762
PLA-MIL (Patch Attention)		0.7304
SG-ResNet34 (RGB + Mask)	Segmentation-Guided	0.8707

Table 2: *Results Comparison Table*

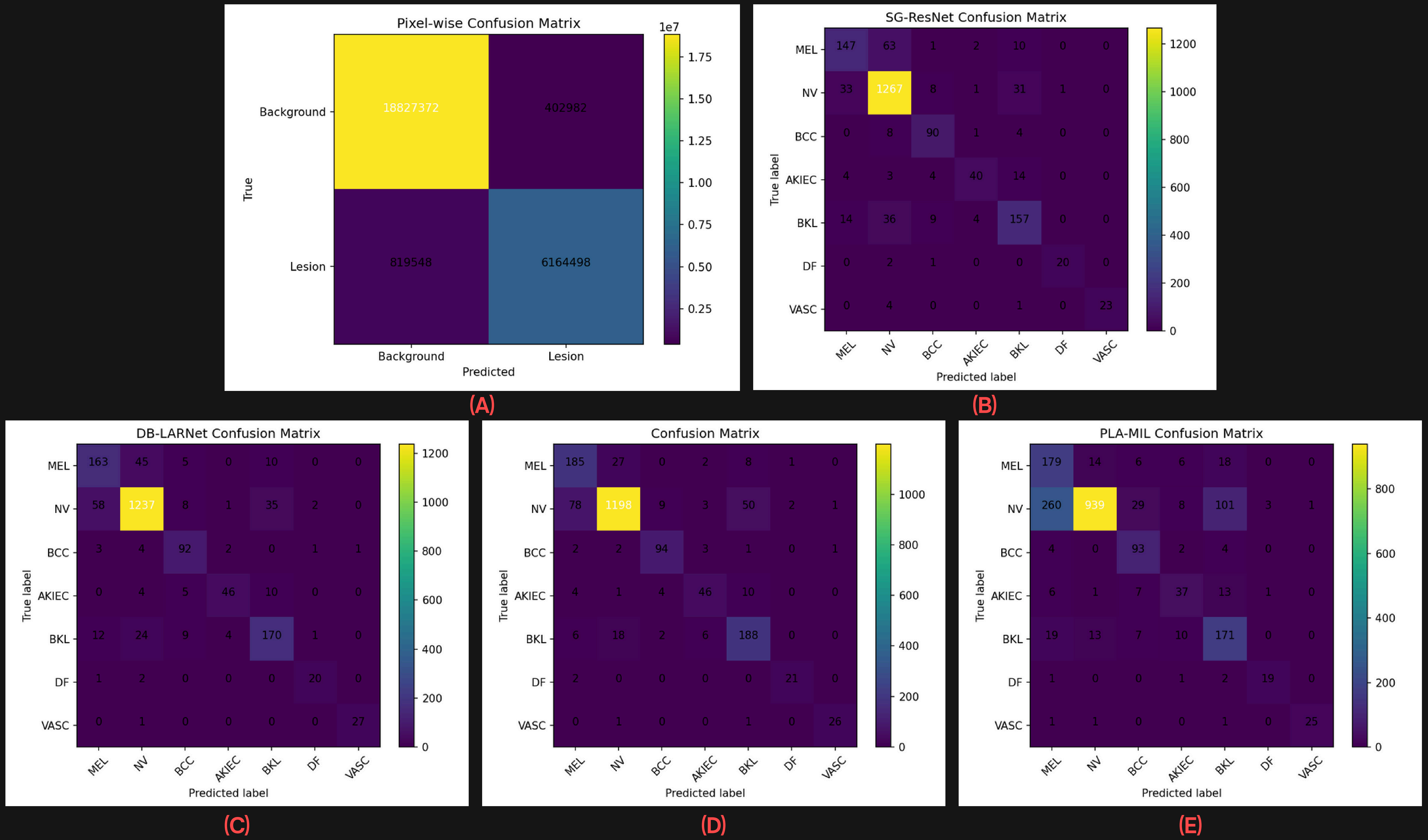


Figure 5: *Confusion Matrices of Proposed Models.*
(A) TransUNet (B) SG ResNet34 (C) DB LARNet (D) ResNet34 (E) PLA MIL

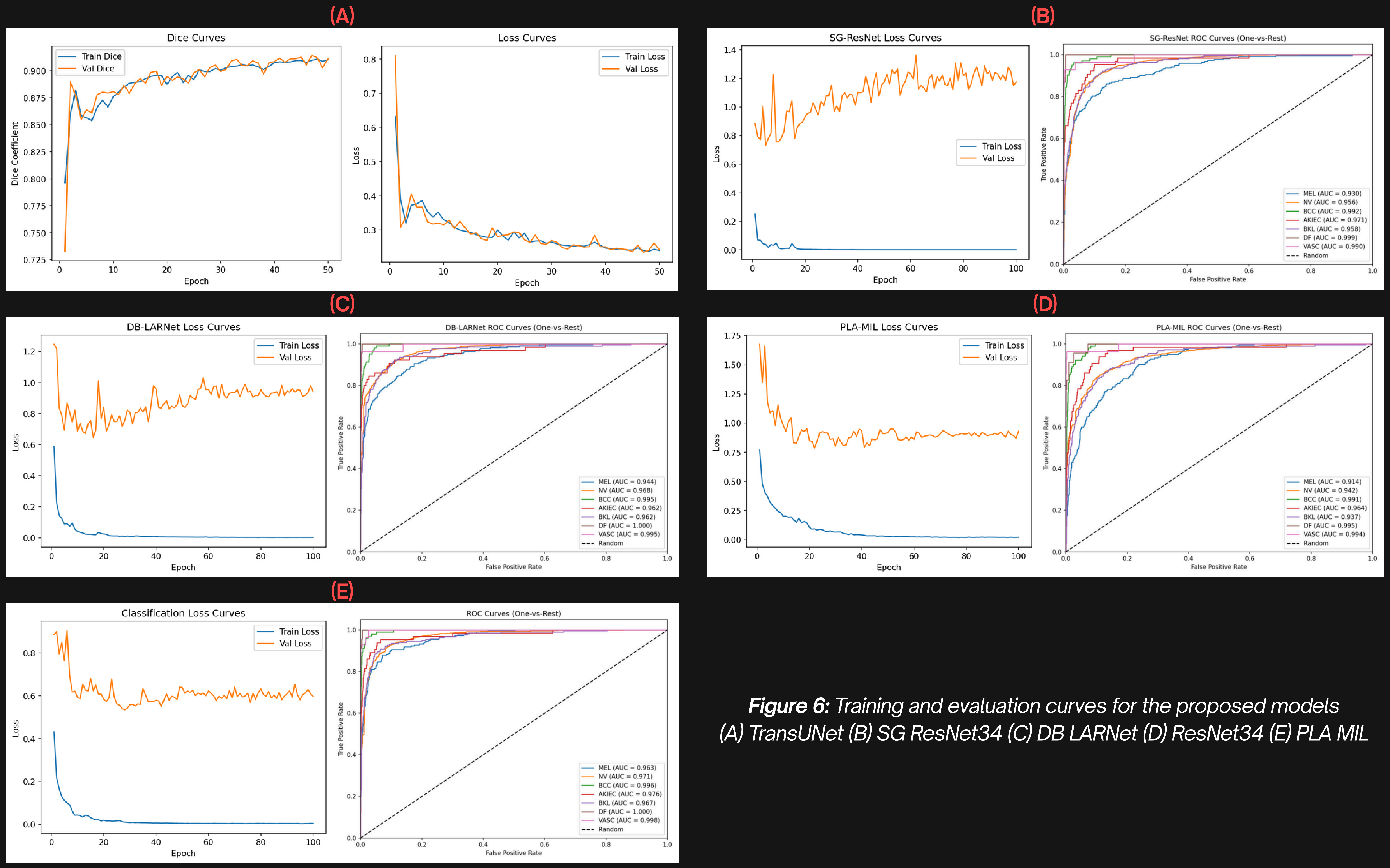


Figure 6: *Training and evaluation curves for the proposed models (A) TransUNet (B) SG ResNet34 (C) DB LARNet (D) ResNet34 (E) PLA MIL*

CLINICAL VALIDATION
GRAD CAM

- Grad-CAM (Gradient-weighted Class Activation Mapping) visualises which regions of an image most influence a model's prediction.
- It uses the gradients of the predicted class flowing into the final convolutional layer.
- Highlights discriminative image regions as a heatmap overlaid on the original image.
- Works with any CNN-based classifier.
- Helps interpret why a model predicted a specific class.
- Useful for detecting model bias and failure cases and improving trustworthiness.
- In medical imaging, it shows lesion-specific areas a model focuses on.

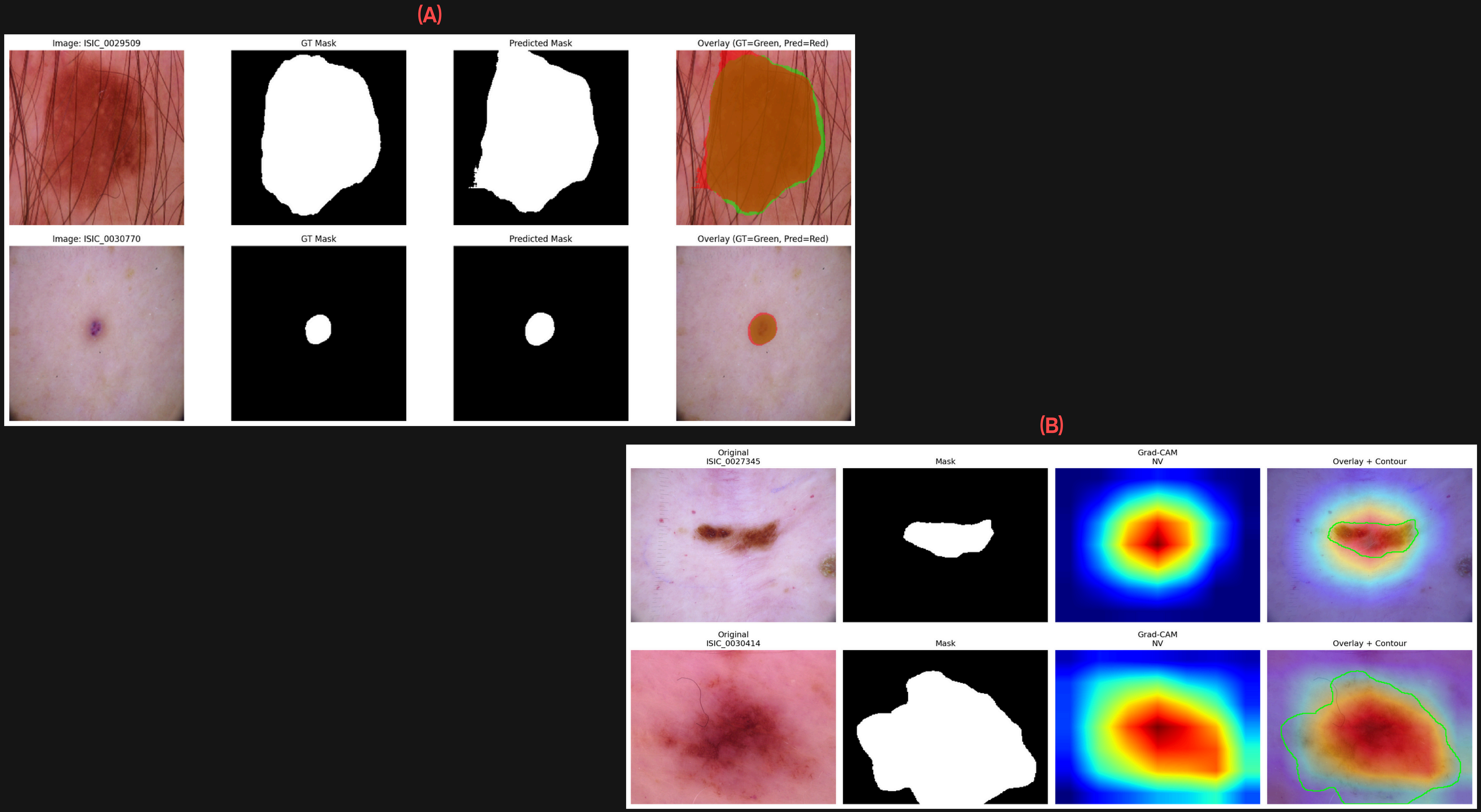


Figure 7.1: Grad-CAM Visualizations

(A) TransUNet (B) SG ResNet34

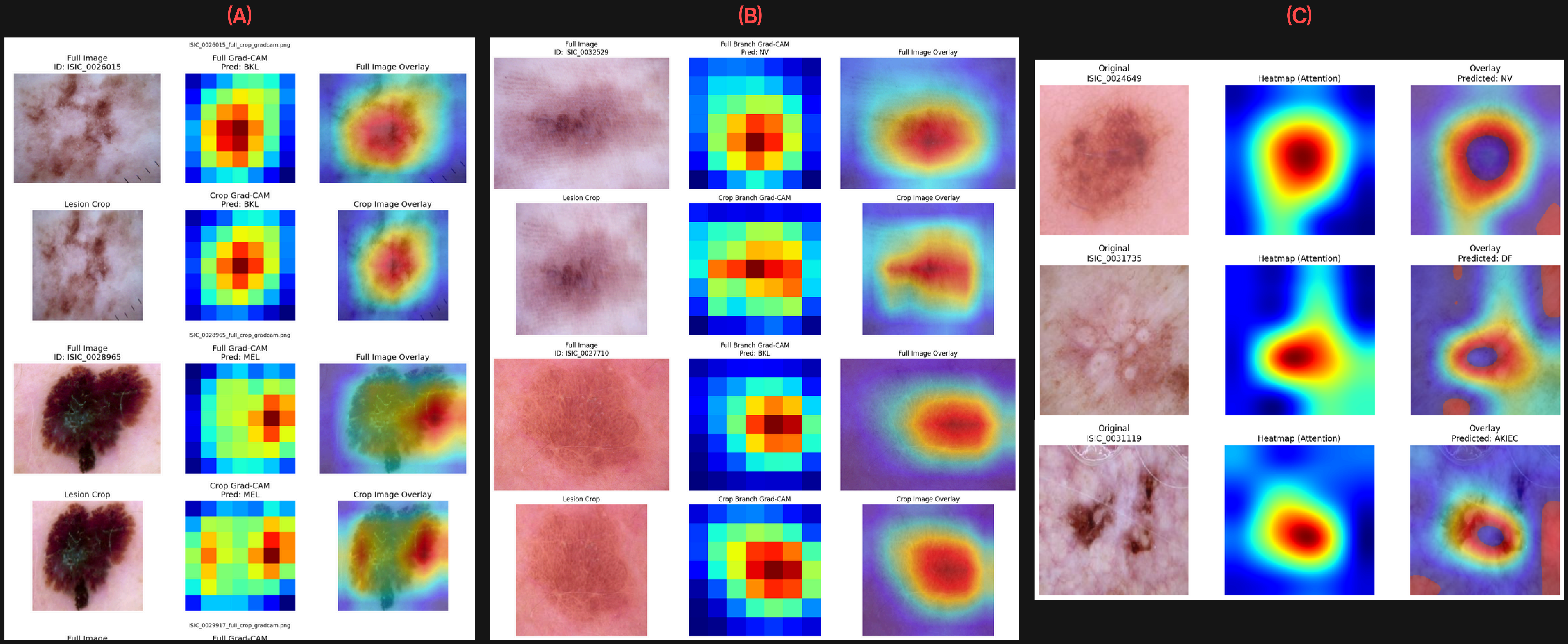
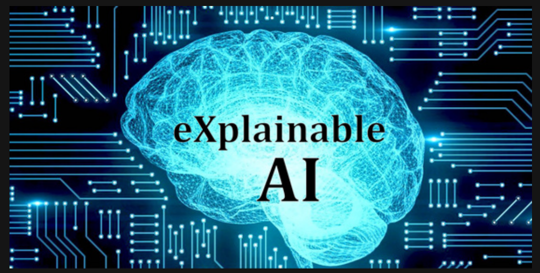


Figure 7.2: Grad-CAM Visualizations

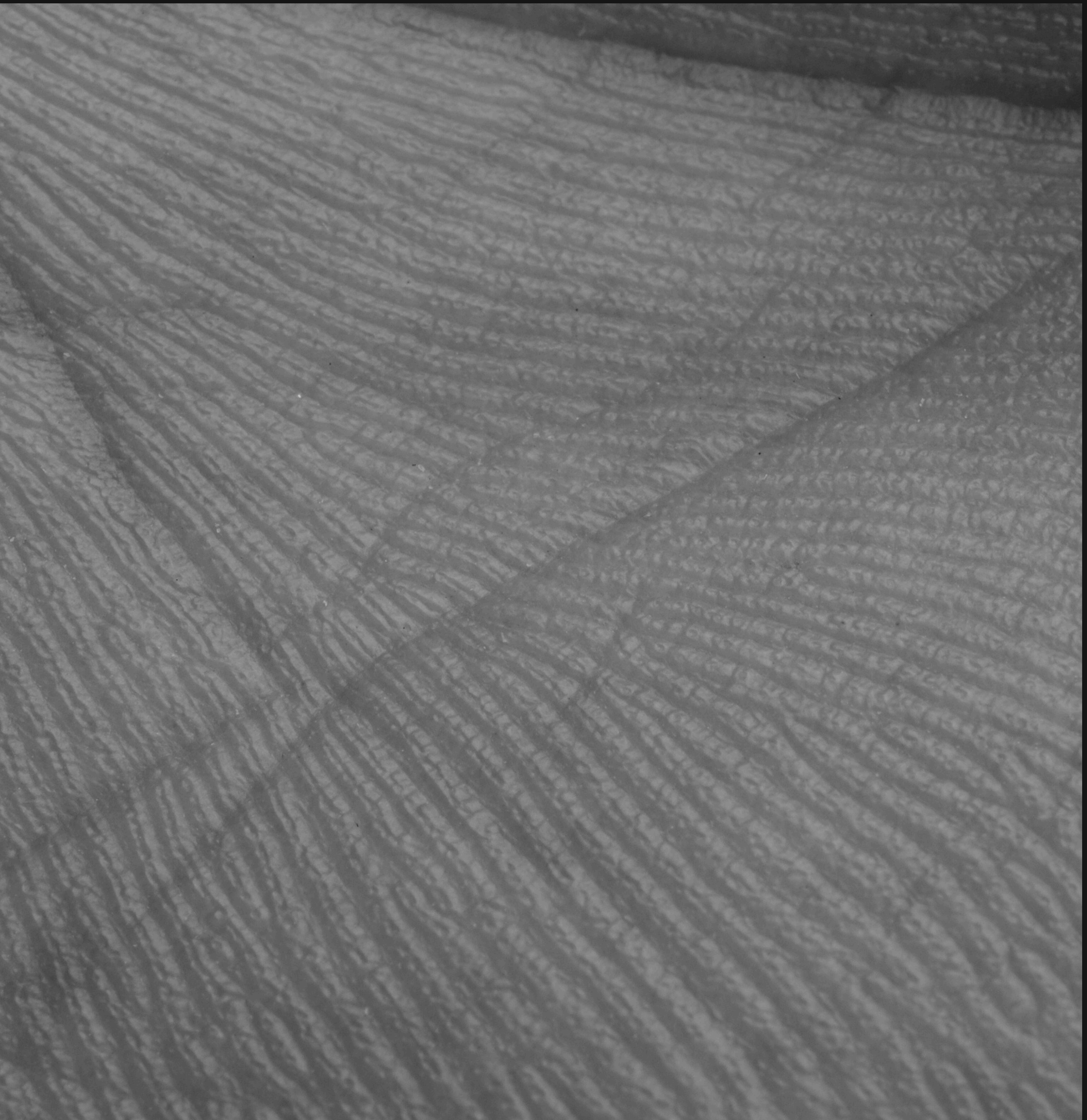
(A) DB-LARNet (B) ResNet34 (C) PLAMIL

TOOLS USED

- OpenCV — color maps, heatmap overlays
- PyTorch — segmentation, classification, MIL, attention
- Pandas / NumPy — data handling
- Matplotlib — graphs, ROC curves, Grad-CAM visualizations
- Overleaf/TikZ — pipeline diagrams
- Kaggle GPU environment — efficient training
- Explainability Tools — Grad-CAM, MIL attention maps



CONCLUSION



- This work successfully developed a complete deep-learning pipeline for skin lesion segmentation and classification, addressing challenges of class imbalance, variability in lesion appearance, and lack of interpretability.
- The TransUNet segmentation model achieved strong performance with a Dice score of 0.9142, IoU of 0.8576, and pixel-level accuracy of 0.9548, enabling reliable lesion-focused region extraction.
- Among classification models, ResNet-34 with lesion crops delivered the highest accuracy (0.8777), showing the effectiveness of focusing on localized lesion regions.
- The Dual-Branch LARNet and Segmentation-Guided ResNet-34 models demonstrated competitive performance, confirming that integrating global and localized features can improve decision quality.
- The PLA-MIL patch-attention model provided strong interpretability, highlighting fine-grained lesion patterns through attention maps, even though accuracy was lower than CNN-based models.
- Grad-CAM and attention heatmaps increased transparency by showing clinically relevant regions influencing predictions, improving trust and suitability for medical decision support.
- Overall, the proposed pipeline shows that combining segmentation, focused lesion crops, and explainability techniques leads to a robust, interpretable, and clinically meaningful system for early melanoma detection.
- This work can support dermatologists by providing consistent, explainable predictions and has potential for future deployment in tele-dermatology and clinical screening tools.

REFERENCES

[1] Groh, M., Harris, C., Soenksen, L., Lau, F., Han, R., Kim, A., Koochek, A., & Badri, O. (2021). Evaluating deep neural networks trained on clinical images in dermatology with the Fitzpatrick 17k dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1820–1828.
[2] Daneshjou, R., Vodrahalli, K., Novoa, R.A., Jenkins, M., Liang, W., Rotemberg, V., Ko, J., Swetter, S.M., Bailey, E.E., Gevaert, O., & Mukherjee, P. (2022). Disparities in dermatology AI performance on a diverse, curated clinical image set. Science Advances, 8(32), eabq6147.
[3] Du, S., Hers, B., Bayasi, N., Hamarneh, G., & Garbi, R. (2022). FairDisCo: Fairer AI in dermatology via disentanglement contrastive learning. In European Conference on Computer Vision, 185–202. Springer Nature Switzerland.
[4] Montoya, L.N., Roberts, J.S., & Hidalgo, B.S. (2025). Towards fairness in AI for melanoma detection: Systematic review and recommendations. In Future of Information and Communication Conference, 320–341. Springer Nature Switzerland.
[5] Zong, Y., Yang, Y., & Hospedales, T. (2022). MEDFAIR: Benchmarking fairness for medical imaging. arXiv preprint arXiv:2210.01725.
[6] Mirikharaji, Z., et al. (2023). A survey on deep learning for skin lesion segmentation. Medical Image Analysis, 88, 102863.
[7] Naqvi, M., et al. (2023). Skin cancer detection using deep learning—a review. Diagnostics, 13(11), 1911.
[8] Al-Waisy, A.S., et al. (2025). A deep learning framework for automated early diagnosis and classification of skin cancer lesions in dermoscopy images. Scientific Reports, 15(1), 31234.
[9] Hosny, K.M., et al. (2024). Explainable deep inherent learning for multi-class skin lesion classification. Applied Soft Computing, 159, 111624.
[10] Sulthana, R., et al. (2024). A novel end-to-end deep convolutional neural network based skin lesion classification framework. Expert Systems with Applications, 246, 123056.
[11] Madarkar, S.S., et al. (2025). DermaCon-IN: A multi-concept annotated dermatological image dataset of Indian skin disorders for clinical AI research. arXiv preprint arXiv:2506.06099.
[12] Wang, J., et al. (2024). Achieving reliable and fair skin lesion diagnosis via unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.
[13] Giavina-Bianchi, M., et al. (2023). Explainability agreement between dermatologists and visual explanation techniques in deep neural networks for melanoma AI classification. Frontiers in Medicine, 10, 1241484.
[14] Zhao, P., & Ding, X. (2024). DFCA: Disentangled Feature Contrastive Learning and Augmentation for Fairer Dermatological Diagnostics.
[15] Aayushman, et al. (2024). Fair and accurate skin disease image classification by alignment with clinical labels. In MICCAI – International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer Nature Switzerland.
[16] Munia, N., & Imran, A.-A.-Z. (2025). DermDiff: Generative diffusion model for mitigating racial biases in dermatology diagnosis. arXiv preprint arXiv:2503.17536.

THANK YOU