BlendED Optimizing Transfer Learning for High-Accuracy Healthcare AI Under Data Scarcity

Authors: Stanley Wijaya, Rumaisa Kashif, LiLeonard

BACKGROUND

This project tackles the **small-data bottleneck**, exploring how transfer learning and augmentation can deliver **accurate, scalable, and trustworthy medical AI** with limited resources

CHALLENGES

- **Scarce Data**: Extremely limited labeled images per rare condition
- **Class Imbalance**: Common diseases dominate datasets, rare ones underrepresented
- **High Variability**: Lesions differ widely in appearance, size, and location
- **Overfitting Risk**: Fine-tuning deep models on tiny datasets often fails to generalize

RESEARCH QUESTION

How can transfer learning and targeted data augmentation be leveraged to build a highly accurate, scalable skin disease classifier from limited data — enabling real-world healthcare AI with minimal resources?

METHODOLOGY

Dataset & Preprocessing:

We utilized the ISIC dermatoscopic image dataset containing nine skin lesion categories. Images were preprocessed with standard augmentations (random crops, flips, rotation ±30°, color jittering) and normalized to 224×224 pixels using ImageNet statistics.

Model Architecture:

Our approach employs a Vision Transformer (**ViT-Base/16**) backbone pre-trained with Masked Autoencoder (**MAE**) self-supervised learning, coupled with a linear classification head for nine-class skin lesion classification.

Two-Stage Training Strategy:

- **Stage 1 - Linear Probing (8 epochs):** Backbone frozen, only classifier trained (lr=3×10^{-4}) to evaluate pre-trained feature quality
- **Stage 2 - End-to-End Fine-tuning (25 epochs):** All parameters unfrozen with reduced learning rate (lr=1×10^{-5}) for gentle feature adaptation

Class Imbalance Handling:

Weighted Random Sampler dynamically balances batch composition by inverse class frequency weighting. AdamW optimizer with weight decay (λ=0.05) and automatic mixed precision training ensure stable convergence.

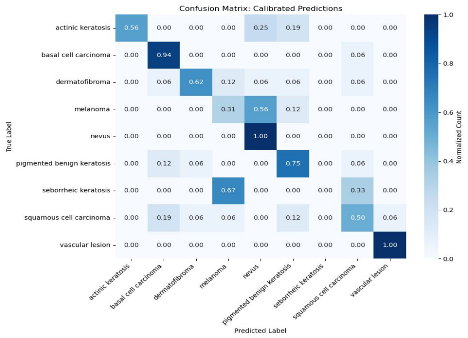
Decision Calibration Framework:

We developed a novel post-hoc calibration technique to enhance classification performance:

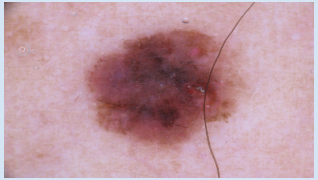
- **Prior Adjustment:** $logits_{adj} = logits - \tau \log prior \quad \text{where} \quad \tau = 0.4$
- **Class-Specific Bias:** Negative bias (1.2) applied to nevus class to reduce systematic misclassifications

RESULTS

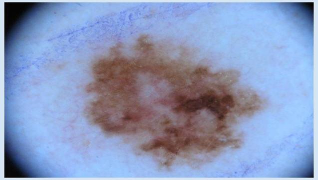
	Precision	Recall	F1-Score	Support
Actinic Keratosis	1.00	0.56	0.72	16
Basal Cell Carcinoma	0.71	0.94	0.81	16
Dermatofibroma	0.83	0.62	0.71	16
Melanoma	0.50	0.31	0.38	16
Nevus	0.53	1.00	0.70	16
Pigmented Benign Keratosis	0.60	0.75	0.67	16
Seborrheic Keratosis	0.00	0.00	0.00	3
Squamous Cell Carcinoma	0.67	0.50	0.57	16
Vascular Lesion	0.75	1.00	0.86	3
Accuracy			0.66	118



DATASET



Training image from class 'Melanoma'



Test image from class 'Melanoma'

CONCLUSION & DISCUSSION

It can be concluded that the modified pre-trained model struggled to detect some cases due to the lack of images to train. To improve further, techniques like data augmentation or specialized loss functions can be implemented. Additionally, a few-shot classification model can be created that only requires a few images per class which solves the lack of images for rare skin conditions

REFERENCES

ISIC dataset
https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic