Skin Lesion Classification Using Deep Learning

Sameer Shaik

Methodology

Pretrained Models

We used pretrained **DenseNet121** and **ResNet50**. These models were originally trained on the ImageNet dataset and have learned to detect general visual features such as edges, textures, and shapes. Using pretrained weights helps our model start from a strong baseline and adapt to the medical domain more efficiently.

Data Augmentation

To improve generalization and prevent overfitting, we applied the following image transformations:

- Resize to 256x256
- Random crop to 224x224
- Horizontal and vertical flipping
- Random rotation and color jitter
- Normalization

These transformations make each image slightly different during training, which helps the model learn general patterns instead of memorizing specific details.

Class Imbalance and Class Weights

The dataset is highly imbalanced:

- Class 0 (benign): 1626 samples
- Class 1 (malignant): 374 samples

This imbalance can cause the model to overpredict benign cases. To handle this, we applied **class weighting** in the loss function, which increases the penalty for misclassifying malignant lesions. This encourages the model to be more sensitive to the minority class.

Learning Rate Scheduler

We used a learning rate scheduler (‘StepLR‘) that reduces the learning rate after a fixed number of epochs. This helps the model:

- Learn quickly in the early epochs
- Fine-tune and stabilize performance in later epochs

Results: Image-Only Models

Best results using only image data with both class weighting and learning rate scheduling:

Model	F1 Score	Recall	Specificity
DenseNet121	0.5149	0.6933	0.7692
ResNet50	0.5070	0.7200	0.7415

Without class weighting, models tend to favor the majority class, leading to poor recall. Adding class weights significantly improved F1 and recall. Combining this with a learning rate scheduler produced the best results.

Results: Including Extra Features

I extended our approach by integrating dermatological features (e.g., pigment network, streaks). These were binary features (0 or 1) and used alongside image data.

Model	**Recall**	**Specificity**
ResNet	0.4267	0.8615
DenseNet	0.6533	0.5477

Conclusion

While the inclusion of features added more information, the performance dropped slightly. This could be due to: The combination of pretrained models, class weighting, and learning rate scheduling resulted in strong performance on this challenging medical task. Adding dermatological features showed mixed results and may require more advanced fusion techniques for consistent improvement.