Skin Cancer Detection Using Deep Learning

Ali Jaouni

Introduction

Skin cancer is one of the most common cancers worldwide, and survival rates depend heavily on early diagnosis. Dermatoscopic examination reveals visual patterns that distinguish benign lesions from malignant ones; however, manual diagnosis requires expertise and is prone to variability. This project investigates whether deep learning can automate binary skin lesion classification to support dermatologists and reduce diagnostic delays.

Convolutional Neural Networks (CNNs) are suitable for this task because they learn spatial and textural representations directly from images, unlike traditional models that rely on handcrafted features. To further improve lesion-specific focus, this project incorporates Squeeze-and-Excitation (SE) channel attention, enabling adaptive re-weighting of informative feature maps.

Input: Dermatoscopic RGB image (224×224).
Output: Predicted lesion category (`benign` or `malignant`).

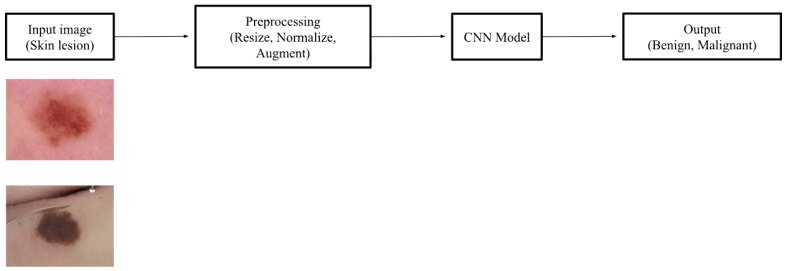

Figure 1: Overview of the binary skin lesion classification system.

Data Processing

This project uses the HAM10000 dataset [4], containing 10,015 dermatoscopic images across seven clinical diagnoses. Labels were converted into a binary classification problem:

Benign (0): `nv, bkl, df, vasc`
Malignant (1): `mel, bcc, akiec`

A stratified 70%/15%/15% split ensured the same class proportions across train, validation, and test sets (80.5% benign / 19.5% malignant). Preprocessing steps included:

- resizing to 224×224
- normalization to [0, 1]
- augmentations: flips, rotation, and color jitter

Challenges: class imbalance, low contrast melanoma images, and visual similarity between nevi and early melanoma.

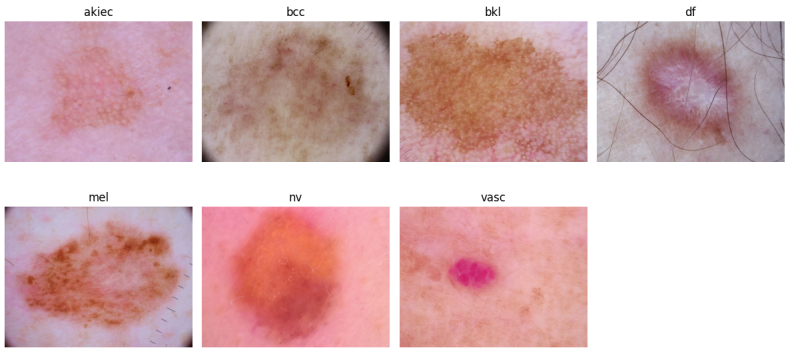

Figure 2: Sample input images from HAM10000 showing benign and malignant lesions.

Model	Train Accuracy	Validation Accuracy
Logistic Regression	0.78	0.69

Table 1: Baseline model quantitative performance.

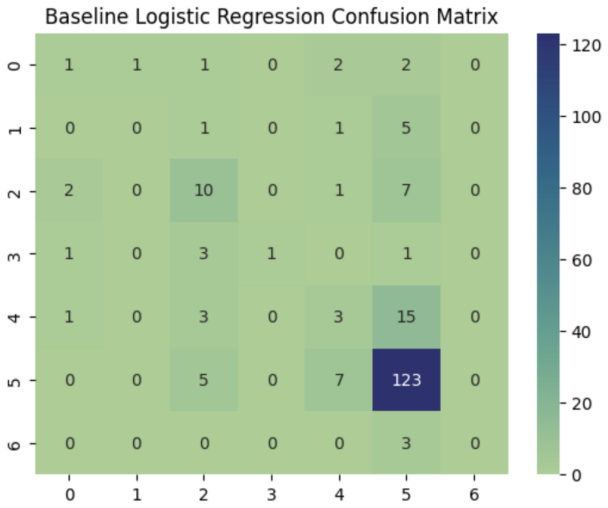

Figure 3: Confusion matrix for baseline model.

Baseline Model

The baseline model used logistic regression trained on handcrafted features consisting of 32-bin HSV histograms and Local Binary Pattern (LBP) texture descriptors.

The baseline struggled with malignant recognition due to limited feature representation capacity, motivating the transition to deep learning.

Primary Model: SE-CNN

The final model is a custom SE-CNN with channel attention applied after each convolutional block.

Architecture Overview:

- Conv (32) + BN + ReLU + MaxPool + SE(32)
- Conv (64) + BN + ReLU + MaxPool + SE(64)
- Conv (128) + BN + ReLU + MaxPool + SE(128)
- Flatten $\rightarrow$ Dense(256) $\rightarrow$ Dropout $\rightarrow$ Dense(2)

Total parameters: $\sim$1.7M.
Optimizer: Adam (LR $= 1 \times 10^{-4}$), batch size 32, 5 epochs, class-weighted loss.

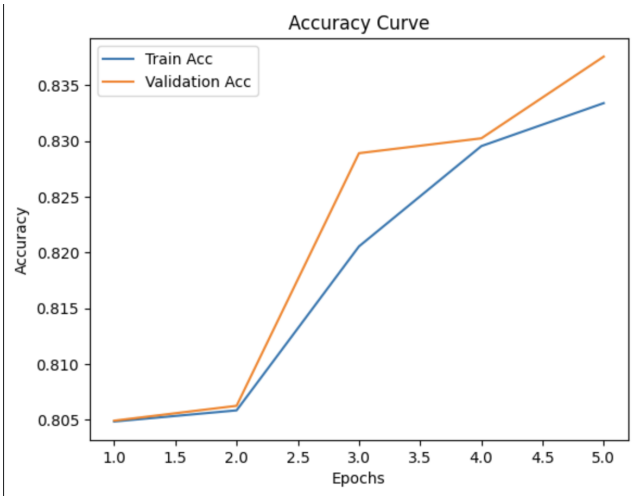

Figure 4: Training and validation accuracy over epochs for the SE-CNN.

Results

The SE-CNN achieved a test accuracy of **0.818** on fully unseen data.

Class	Precision	Recall	F1-score	Support
Benign (0)	0.82	0.98	0.90	1210
Malignant (1)	0.66	0.13	0.22	293
Accuracy		**0.818**		

Table 2: SE-CNN performance on the test set.

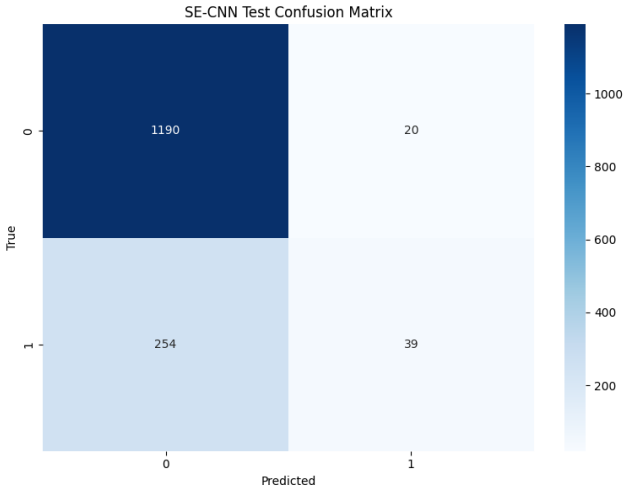

Figure 5: Confusion matrix - most errors result from malignant lesions predicted as benign.

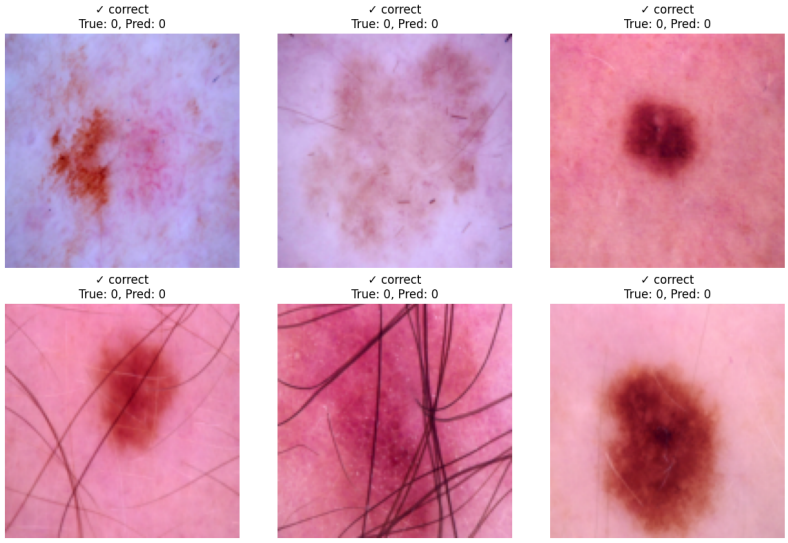

Figure 6: Qualitative predictions: correct benign predictions.

Evaluation on New Data

The 15% test split was held out from the start and never used during model development or hyperparameter tuning. All performance metrics reported above are computed exclusively on this unseen set.

Discussion

The SE-CNN delivers strong overall performance and significantly outperforms the logistic regression baseline. However, malignant recall (0.13) remains critically low, indicating that the model misses subtle melanomas, especially those with low contrast or small lesion size.

Future improvements:

- class-balanced or focal loss to penalize missed melanoma predictions
- multi-scale architectures for fine-grained lesion detection
- domain-specific augmentation (e.g., glare/hair occlusion simulation)

Ethical Considerations

The model must not be used as an autonomous diagnostic tool. Misclassification of malignant lesions can delay treatment and cause significant harm. The dataset lacks global demographic diversity, risking bias across skin tones. Any deployment should occur only as a decision-support tool under expert clinical supervision.

References

[1] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[2] H. A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A. Hassen, L. Thomas, A. Enk, et al., “Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists,” Annals of Oncology, vol. 29, no. 8, pp. 1836–1842, 2018.

[3] T. Pham, T. Luong, and V. Le, “Interpretable deep learning for the classification of skin cancer,” in Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2019, pp. 3539–3542.

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

[5] T. J. Brinker, A. Hekler, A. H. Enk, J. Klode, A. Hauschild, C. Berking, B. Schilling, S. Haferkamp, D. Schadendorf, T. Holland-Letz, et al., “Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task,” Eur. J. Cancer, vol. 113, pp. 47–54, 2019.

[6] Y. Yao, "SE-CNN: Convolution Neural Network Acceleration via Symbolic Value Prediction," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 13, no. 1, pp. 73-85, March 2023.

[7] “SECNN: Squeeze-and-Excitation Convolutional Neural Network for Sentence Classification,” Arxiv.org, 2023.