Department of Electrical and Computer Engineering
North South University (NSU)

# CSE 445: Machine Learning
Section 6

## Project Report

Instructor: Dr. Mohammad Mahmudul Alam Semester: Fall 2025

| | |
|---|---|
| Title: | Skin Cancer detection Model using the Deep learning with transfer learning Custom CNN |
| Name: | Hasnat Karibul Islam |
| NSU ID: | 2211275042 |
| Name: | Mahin Ahmed |
| NSU ID: | 2211772042 |

**Title:** Skin Cancer detection Model using the Deep learning with transfer learning Custom CNN

**Abstract:**

In this Project we have developed a robust deep learning framework for the automated classification system of Skin Cancer, targeting fiver distinct type: Actinic Keratoses, Basal Cell Carcinoma, Benign Keratosis-like Lesion, Melanocytic Nevi, and Melanoma, Recognizing the high mortality rate associated with late-stage skin cancer diagnosis. We have addressed significant challenges such as severe class imbalance and visual similarity between benign and malignant lesion implementing a rigorous data processing pipeline that included undersampling majority classes and augmented minority classes. We have designed and trained two distinct architectures: a custom Hybrid CNN transformer to capture both local features and global context, and a ResNet152 model leveraging transfer learning for deep feature extraction. Our experimental results demonstrated that ResNet152 Model achieved a test accuracy if accuracy 94.21%, while our custom CNN model achieved 92.58% accuracy. To maximize the reliability, we have constructed an ensemble interface system that aggregates the prediction of both models. After Ensemble we got around 95% accuracy.

**Introduction:**

Skin Cancer is one of the most common and deadliest cancer globally, yet it is highly treatable if it is detected at the early. The current standard for diagnosis dermoscopy relies heavily on the subjective visual interpretation of dermatologist, which can be time consuming and prone to human error. With the rapid advancement of architectural intelligence, specially, Deep Learning, there is a huge opportunity to assist medical professionals with automated diagnostic tools. In this project we have focused between five specific categories skin lesion. While Convolutional Neural Networks (CNNs) have historically been the standard for image analysis, they sometimes struggle to capture long range dependencies within an image. To overcome this we have investigated a hybrid approach, Our project integrates a standard deep residual network with a custom built architecture that fuses the CNN blocks with Transformer attention mechanism.

**Problem Statement:**

The primary problem I addressed in this project is the difficulty of accurately classifying skin lesions due to high inter-class similarity and intra-class variation. For example, a benign keratosis can look remarkably similar to a malignant melanoma to the untrained eye. For this project and for training we have used the marmal88/skin_cancer dataset which was heavily imbalanced, with Melanocytic Nevi accounting for the vast majority of samples. Training a model on such data without intervention leads to biased predictions where the model simply guesses the majority class. Additionally, relying on a single model architecture creates a single point of failure; a standard CNN might miss the global context that a Transformer could catch. Therefore, our challenge was to create a balanced training pipeline and an ensemble architecture. So that we can minimizes these biases and errors

**Methodology:**

We have structured our methodology into main three stages: Data augmentation, Model Development , and Ensemble Implementation.

**Dataset and Preprocessing:**

We have utilized a dataset of skin lesion images. To ensure that our models learned meaningful patterns rather that noise, we have implemented following steps:

- **Balance the Data:** We have identified that our dataset has a severe class imbalance where **Melanocytic Nevi** dominated the dataset**.** We addressed this by undersampling the majority class and aggressively augmenting the minority classes.
- **Data Augmentation:** We have used the ***albumentations*** library to apply transformations such as rotation, vertical/horizontal flips, zooming, shearing, and color jitter. This artificially expanded our training set and helped us to prevent the overfitting.
- **Image Enhancement:** We have applied ***cv2.fastNlMeansDenoisingColored*** library function to remove artifacts like hair or air bubbles from the dermoscopy images and resized all inputs to a standard ***224 x 224*** resolution.

**Model Architecture:**

1. **Custom Hybrid CNN-Transformer Model:** We have designed a custom architecture ro combine local feature extraction with global attention mechanism. Our Custom Hybrid CNN-Transformer Model contains total 69 layers and 12 million parameters among them all are trainable.

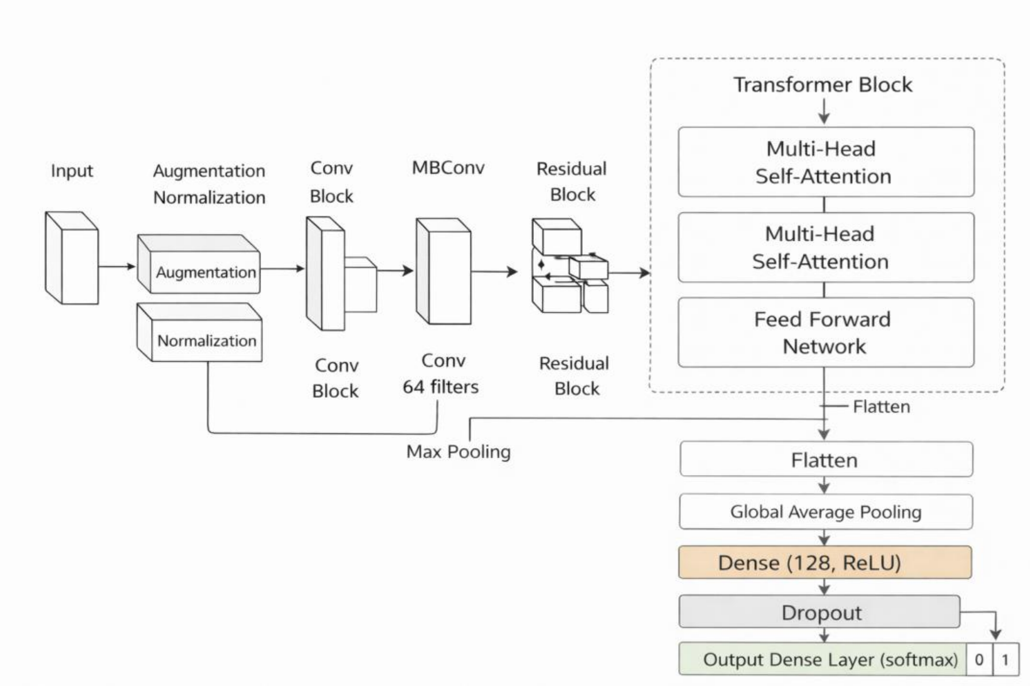



**Fig: CUSTOM Hybrid CNN Transformer Model Architecture**

The Key components are

- **Convolutional Blocks:** Initial layers for low level feature extraction.
- **Residual Blocks:** To prevent the vanishing gradient and allow deeper feature learning.
- **MBConv Blocks:** Efficirnt depthwise separable convolutions for intermediate feature extraction.
- **Transformer block:** Two multi-head self-attention layers at the bottleneck to capture global context and long-range dependencies within the lesion lesion images.
- **Regularization:** We have used the Dropout and L2 weight decay to prevent the overfitting

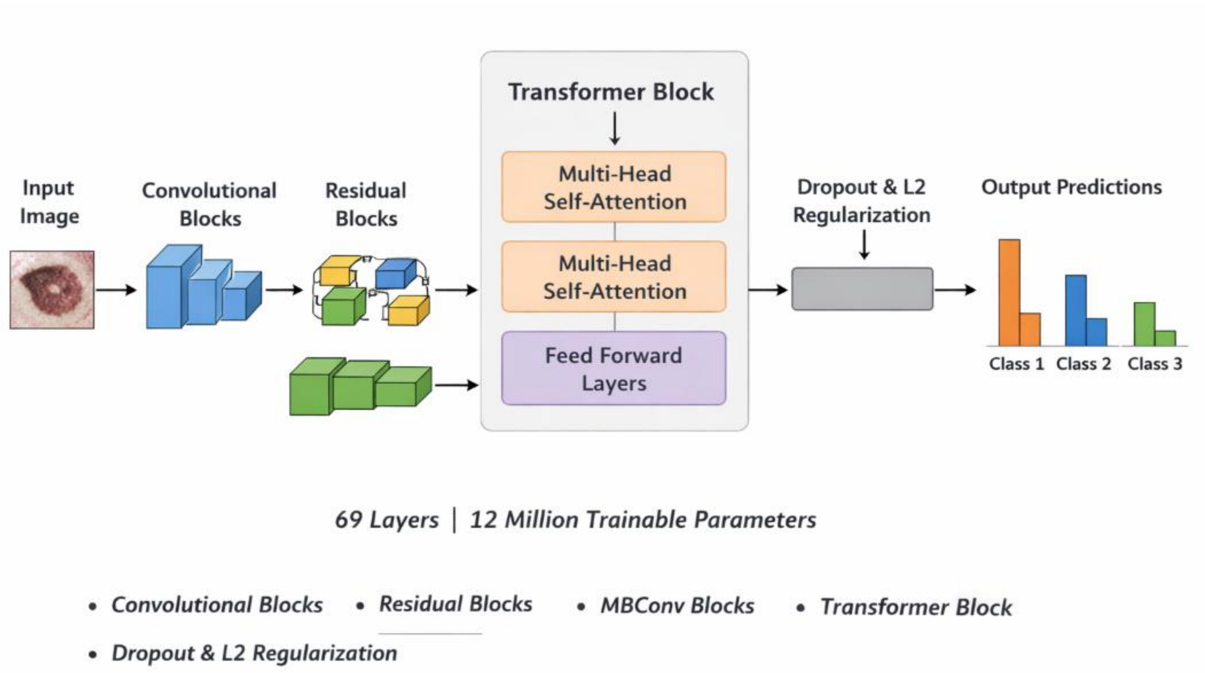



**Fig: CUSTOM Hybrid CNN Transformer Model Block diagram**

2. **ResNat152 Model:**

   In this project we have used the ResNet152 Model which have 152 layers and contains around 795 million parameters and among them 354 million parameters are trainable.

   - We have modified the Fully connected layer head to include a Dropout layer (0.15) followed by a final linear layer for the five output classes.
   - ResNet152 Model leverages deep residual learning to extract highly complex visual features from the dermoscopy images.

**Training & Ensemble:**

- **Optimizer:** We have used the **Adams Optimizer** to train our **Custom Hybrid CNN Transformer Model** and the **ResNet152 Model.**

$$\boldsymbol{\theta(t+1) = \theta(t) - \alpha * \left(\frac{\left(\sqrt{1-\beta 2^{t}}\right)}{(1-\beta 1^{t})} \frac{m(t)}{\left(\sqrt{v_t} + \varepsilon\right)}\right),}$$

$$\boldsymbol{m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t,}$$

$$\boldsymbol{v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2}$$

- **Loss Function:** We have used the Cross Entropy loss function to calculate the loss during the training.

$$L = -\sum_{i=1}^{m} y_i \log(\hat{y}_i)$$

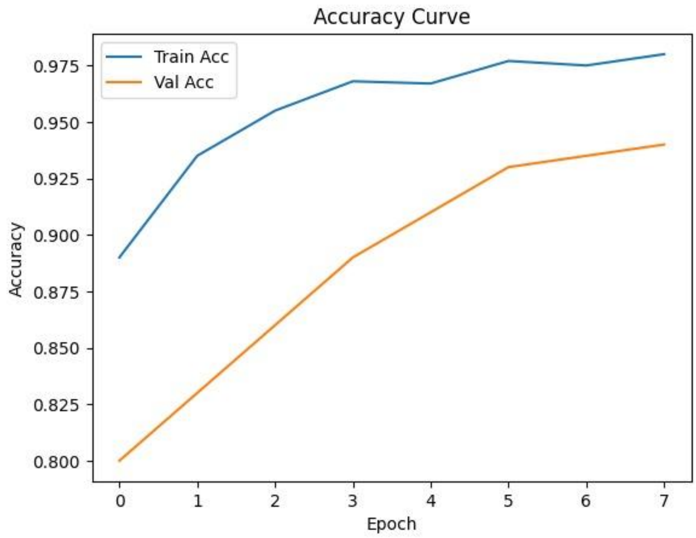
- **Callbacks:** We have used the early stopping and learning rate reduction **(ReduceLRONPlateau)** ensure optimal convergence.
- **Ensemble Strategy:** We have ensembled both of our model get more accurate prediction. The system processes images through both networks and aggregates their prediction through soft voting to determine the final class label, thereby reducing the variance and bias of individual models.

**Results:**

We have evaluated both of the models independently on a test spit which contains around 1253 images to rigorously assess their performance before ensemble them.

**ResNet152 Performance:**

**Accuracy Curve:**



**ResNet152 Detailed Classification Metrics:**

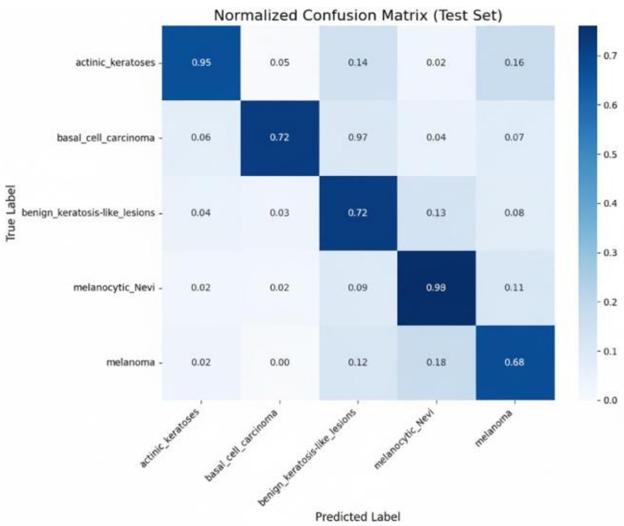
| | precision | recall | f1-score | support |
|---|---|---|---|---|
| actinic_keratoses | 0.91 | 0.88 | 0.89 | 42 |
| basal_cell_carcinoma | 0.93 | 0.91 | 0.92 | 67 |
| benign_keratosis_like | 0.90 | 0.89 | 0.89 | 142 |
| melanocytic_Nevi | 0.97 | 0.98 | 0.98 | 858 |
| melanoma | 0.92 | 0.90 | 0.91 | 144 |
| accuracy | | | 0.94 | 1253 |
| macro avg | 0.93 | 0.91 | 0.92 | 1253 |
| weighted avg | 0.94 | 0.94 | 0.94 | 1253 |

```
===== TEST METRICS =====
Test Accuracy : 0.9421
Precision     : 0.9368
Recall        : 0.9284
F1 Score      : 0.9326
```

From the upper evaluation metrics, we can see that our ResNet152 Performed really well and achieved almost 94.21% accuracy.

**Custom Hybrid CNN Transfer Model Performance:**

**Confusion Matrix:**



**Custom Hybrid CNN Transfer Model Detailed Classification Metrics**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| actinic_keratoses | 0.9250 | 0.8810 | 0.9024 | 42 |
| basal_cell_carcinoma | 0.9385 | 0.9104 | 0.9242 | 67 |
| benign_keratosis-like_lesions | 0.8889 | 0.9014 | 0.8951 | 142 |
| melanocytic_Nevi | 0.9769 | 0.9382 | 0.9572 | 858 |
| melanoma | 0.7167 | 0.8958 | 0.7963 | 144 |
| accuracy | | | 0.9258 | 1253 |
| macro avg | 0.8892 | 0.9054 | 0.8951 | 1253 |
| weighted avg | 0.9333 | 0.9258 | 0.9281 | 1253 |

From the metrics above we can also see that out Custom Hybrid CNN Transfer Model also performed really well and achieved 92.58% accuracy.

As these two models are performing really well that's for more preciseness and accurateness, we have ensembled these two models through soft voting and out ensembled models also performed really well and gained around 95% accuracy.

**Conclusion:**

In this project, we have successfully designed and implemented a comprehensive skin cancer detection system. By creating a custom Hybrid CNN-Transformer and comparing it against a powerful ResNet152, we have gained deep insights into how different architectures handle medical imagery. ResNet152 currently offers higher raw accuracy, but the custom hybrid approach is a viable and promising direction for capturing global image context. The ensemble framework we built have built effectively bridges these two approaches, using the high Recall of the ResNet152 for malignancy screening and the complementary features learned by the Custom CNN to provide a robust final prediction. Future improvements to our work would involve fine-tuning the transformer hyperparameters and incorporating patient metadata to further improve the distinction between visually similar lesions and if possible we will try to build our own custom Dataset which will be more meaningful and will be more balanced. While doing the project we have found that most of the Skin Cancer dataset are imbalanced, biased or low quality. That's why will try build our own dataset.

We would like extend our sincere gratitude to our honorable faculty, **Dr. Mohammad Mahmudul Alam** sir for providing us with the opportunity to work on this project. His guidance and support have been invaluable throughout development, enabling us to deepen our understanding of Machine Learning.

I would also like to thank my project-mates, for his collaborative efforts and dedication to this project. Together, we navigated through various challenges and milestones, leveraging our combined skills and knowledge to create this Robust Skin Cancer Detection model using **Custom Convolutional Neural Network.**