**Report: Skin Cancer Classification using ResNet50**

**1. Introduction**

Skin cancer is one of the most common forms of cancer worldwide. Early detection of malignant skin lesions is critical for effective treatment. In this project, a deep learning approach was applied to classify images of skin lesions into **malignant** and **benign** categories using the **Skin Cancer: Malignant vs. Benign dataset** available on Kaggle. The chosen model was **ResNet50**, a well-established convolutional neural network (CNN) architecture known for its strong performance on image recognition tasks.

**2. Dataset**

The dataset used was [Skin Cancer: Malignant vs. Benign (Kaggle)](https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign). It contains labeled dermoscopic images of benign and malignant skin lesions. The dataset is moderately imbalanced but still provides enough samples for both classes. Images were preprocessed and augmented using techniques such as:

* **Rescaling** (pixel normalization to [0,1])
* **Rotation, width/height shifts, zoom, shear, and flips** for augmentation
* **Batch generation** using Keras ImageDataGenerator

This ensured variability and reduced overfitting during training.

**3. Model Architecture**

The **ResNet50** model was employed with the following setup:

* Pretrained ResNet50 backbone on ImageNet weights
* Fully connected dense layer added at the top for binary classification
* Output layer: Sigmoid activation function
* Optimizer: Adam
* Loss Function: Binary Crossentropy
* Learning rate scheduling to adjust training

ResNet50 was chosen because its residual connections allow effective training of deep networks by mitigating the vanishing gradient problem.

**4. Training Process**

The model was trained for **30 epochs** on the dataset with real-time augmentation. Training logs showed the following performance:

* **Final Training Accuracy**: ~0.86 (86.08%)
* **Final Validation Accuracy**: ~0.83 (82.73%)
* **Final Training Loss**: ~0.3079
* **Final Validation Loss**: ~0.3417

During training, the model showed steady improvement in accuracy up to around **Epoch 23–25**, after which validation accuracy stabilized between **82–83%**, indicating convergence.

**5. Results and Discussion**

The model achieved **83% validation accuracy**, demonstrating its ability to generalize well on unseen data. However, a slight gap remained between training and validation performance, which may indicate mild overfitting. The validation loss plateaued around **0.34**, suggesting the model reached its learning capacity at the chosen settings.

Factors that may have influenced performance include:

* **Dataset Size and Quality**: Limited images compared to real-world variability.
* **Class Imbalance**: Malignant cases were fewer than benign, potentially biasing predictions.
* **Learning Rate**: Gradually decreased to very small values (8e-08), leading to slow convergence at later epochs.

Potential improvements could include:

* Using **class balancing techniques** (e.g., SMOTE, weighted loss).
* Applying **transfer learning with fine-tuning** of deeper ResNet50 layers.
* Trying **ensemble methods** with other CNN architectures such as DenseNet, EfficientNet.

**6. Conclusion**

The ResNet50 model successfully classified skin lesions into malignant and benign categories with an accuracy of **83%** on the validation dataset. While this result demonstrates the feasibility of deep learning for skin cancer detection, further improvements in model tuning, data balancing, and architectural experimentation are recommended for higher accuracy.

This work highlights the potential of deep learning as a supportive tool for dermatologists in skin cancer diagnosis, but deployment in real-world medical practice would require more rigorous validation on larger and more diverse datasets.