Report: Skin Cancer Detection Using Deep Learning

Internship – DevelopersHub Corporation

Week 3, Project 1

# 1. Introduction

This report presents the development and implementation of a skin cancer detection system using deep learning techniques. The project was undertaken as part of Week 3, Project 1 of the internship at DevelopersHub Corporation. The primary objective of this project is to classify skin lesions as cancerous or non-cancerous using the HAM10000 dataset from the International Skin Imaging Collaboration (ISIC).

# 2. Dependencies

The following libraries were required for implementation:  
- kagglehub – dataset acquisition from Kaggle  
- tensorflow – deep learning framework  
- matplotlib – visualization of training metrics  
- seaborn – visualization and analysis  
- scikit-learn – evaluation of classification performance

# 3. Dataset

The dataset used in this project was the HAM10000 Skin Cancer dataset obtained from ISIC via Kaggle. It contains a large collection of multi-source dermatoscopic images representing different types of skin lesions. The dataset was divided into training and validation subsets to support effective model development and evaluation.

# 4. Data Preprocessing

The preprocessing steps applied were as follows:  
- All images were resized to 128 × 128 pixels.  
- Pixel values were normalized to a range of 0 to 1.  
- Data was split into training and validation sets with an 80-20 ratio.  
- Preprocessing was performed using ResNet50’s preprocessing function to align with the model architecture.

# 5. Model Architecture

A transfer learning approach was employed using the ResNet50 model, pre-trained on the ImageNet dataset. The base model was set to non-trainable to leverage pre-learned features. The final architecture included:  
1. ResNet50 base model (pre-trained, without top layers)  
2. Global Average Pooling layer  
3. Dense layer (128 units, ReLU activation)  
4. Dropout layer (rate: 0.5)  
5. Dense output layer (1 unit, Sigmoid activation)  
  
The model was compiled using the Adam optimizer and Binary Crossentropy loss function.

# 6. Model Training

The model was trained for 10 epochs using the training dataset, with validation performed on the reserved validation dataset. Training accuracy and validation accuracy were recorded across epochs to assess the learning performance.

# 7. Model Evaluation

The trained model was evaluated based on accuracy metrics. Training and validation accuracy curves were plotted across epochs to visually analyze performance trends and model generalization capability.

# 8. Results

- The accuracy curves demonstrated consistent improvement during training.  
- Validation accuracy confirmed the ability of the model to generalize effectively.  
- The ResNet50-based architecture successfully classified skin lesion images with high accuracy.

# 9. Conclusion

The ResNet50-based skin cancer detection system successfully classified dermatoscopic images into cancerous and non-cancerous categories. The project demonstrated the effectiveness of transfer learning in medical image classification tasks, highlighting the potential of deep learning to support dermatologists in the early detection and diagnosis of skin cancer.