Assignment 04: Data Augmentation and Model Performance Comparison

Objective:
Analyze the impact of **data augmentation** on model performance using different CNN architectures.

Tasks:

1. **Data Preparation**:
   - Load the **animal image dataset**.
   - Split the dataset into:
     - **Training Set** (70%)
     - **Validation Set** (15%)
     - **Test Set** (15%).
   - Resize all images to **150x150** and normalize pixel values to `[0, 1]`.

2. **Implement Data Augmentation**:
   - Use **ImageDataGenerator** to perform the following augmentations:
     - Horizontal flip
     - Random rotation (up to 20 degrees)
     - Zoom-in/out (up to 20%)
     - Width and height shifts (up to 10%).

3. **Model Training**:

   - Train and compare two CNN models on the augmented data:

     a. **Custom CNN**:
        - Build a CNN with 3 convolutional layers, max pooling, and 1 dense layer with softmax output.
     b. **Pre-trained Model**:
        - Use **MobileNetV2** or **InceptionV3** as a pre-trained model (fine-tuning allowed).

   - Train both models for **10 epochs** and monitor validation accuracy.

4. **Model Evaluation**:

- Evaluate both models on the **test set** and report:
    - **Test Accuracy**.
    - **Confusion Matrix**.
    - **Precision, Recall, and F1-Score** (Classification Report).

5. **Visualization**:

- Plot:
    - Training and validation accuracy/loss curves for both models.
    - Examples of **original images** and their **augmented versions**.

6. **Analysis**:

- Write a brief analysis comparing:
    - Model performance with and without data augmentation.
    - The custom CNN vs. the pre-trained model.

Bonus Task (Optional):

- Use **Grad-CAM** to visualize which parts of the images contributed most to the classification decisions.

Submission Requirements:

- Submit a **Jupyter Notebook** containing:
    - Complete code for data preprocessing, augmentation, model training, and evaluation.
    - Final metrics and visualizations.
    - A brief analysis summarizing your findings.