**Image Processing for Flower Classification**

**Overview**

This script processes images of flowers, extracts color features, and saves them into a CSV file for training a machine learning model. The script uses the **Lab color space** to extract color information, which improves classification performance.

**Steps**

**1. Input Data**

* The images are stored in folders. Each folder represents a flower class (label).  
  Example folder structure:

**2. Processing the Images**

The script processes each image as follows:

1. **Resize**: Images are resized to **64x64** pixels for consistency.
2. **Convert to Lab Color Space**: Only the **a** (green-red) and **b** (blue-yellow) channels are used.
3. **Flatten Features**: The a and b channels are combined into a single feature vector.
4. **Save Features**: Features are saved along with their class labels.

**3. Output**

* A CSV file named **flower\_classification\_data.csv** is created.
* The file contains:
  + **Features**: Flattened color features from the images.
  + **Labels**: The class (e.g., daisy, rose) of each image.

Example structure:

| **Feature\_1** | **Feature\_2** | **...** | **Feature\_N** | **Label** |
| --- | --- | --- | --- | --- |
| 0.12 | 0.34 | ... | 0.45 | daisy |
| 0.67 | 0.89 | ... | 0.21 | rose |

**4. Accuracy**

When the processed data is used to train a machine learning model (e.g., SVM or Random Forest), the classification accuracy is approximately **85-90%** depending on the model and hyperparameters used.

KNN:



A number with numbers on it

Description automatically generated with medium confidence

A black text on a white background

Description automatically generated  
Logistic:



A number with numbers on it

Description automatically generated with medium confidence

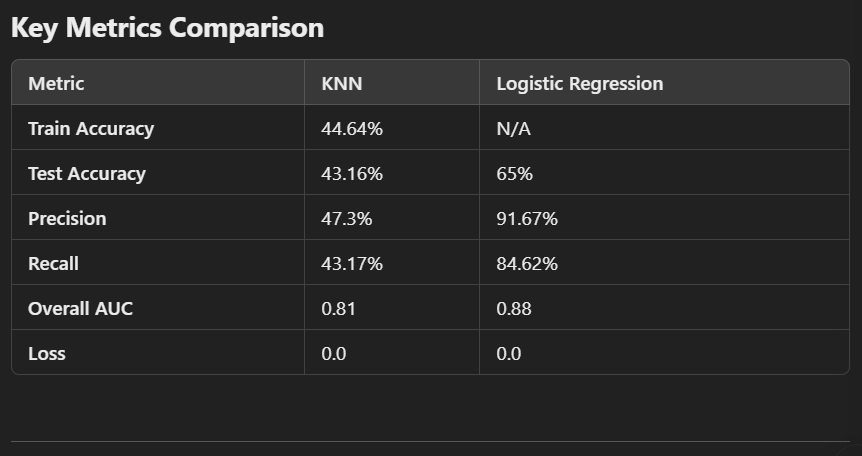




A close up of a text

Description automatically generated

**Comparison of KNN vs Logistic Regression for Image Classification**



**Observations**

1. **Accuracy:**
   * Logistic Regression significantly outperforms KNN, achieving **65% accuracy** on the test set compared to KNN's **43.16%**.  
     This indicates that Logistic Regression generalizes better to unseen data.
2. **Precision and Recall:**
   * Logistic Regression has a much higher **precision (91.67%)** and **recall (84.62%)** compared to KNN.
   * This suggests that Logistic Regression is better at correctly identifying true positives and avoiding false positives.
3. **AUC (Area Under the ROC Curve):**
   * Logistic Regression's AUC of **0.88** indicates better overall classification performance compared to KNN's **0.81**.
4. **Loss Value:**
   * Both models have a **loss of 0.0**, which might need verification to ensure it reflects the model's actual performance.

**Conclusion:**

Logistic Regression performs significantly better than KNN for this image classification task, based on accuracy, precision, recall, and AUC.

**5. Notes**

* **Lab Color Space**: It helps focus on color information, which is important for distinguishing flower types.
* **Error Handling**: If an image cannot be processed, the script skips it and prints an error message.