|  |
| --- |
| **Team ID :** LTVIP2025TMID32402 |
| Project Name : Enchanted Wings: Marvels of Butterfly Species |

**Split Data and Model Building**

**Train-Test Split**

In this project, we have already separated the dataset into training and testing sets.

* **Training Path:** /content/output\_dataset/train
* **Testing Path:** /content/output\_dataset/test

We use ImageDataGenerator from **Keras** to preprocess and load images from directories.  
This approach allows on-the-fly data augmentation and normalization, which helps improve model performance and generalization.

**Code Explanation**

python

CopyEdit

trainpath = "/content/output\_dataset/train"

testpath = "/content/output\_dataset/test"

# Create data generators

train\_datagen = ImageDataGenerator(rescale=1./255, zoom\_range=0.2, shear\_range=0.2)

test\_datagen = ImageDataGenerator(rescale=1./255)

# Load images from directories

train = train\_datagen.flow\_from\_directory(

trainpath,

target\_size=(224, 224),

batch\_size=20

)

test = test\_datagen.flow\_from\_directory(

testpath,

target\_size=(224, 224),

batch\_size=20

)

**Output**

The result confirms the dataset structure:

* ✅ 3,854 training images across 75 classes
* ✅ 1,332 testing images across the same 75 classes