 # Dataset Access Instructions:

This project uses the **DeepSat-6 dataset,** which is publicly available on Kaggle. Due to its size **(>5 GB),** it cannot be uploaded directly to this repository.

Steps to Access the Dataset:

Go to the Kaggle dataset page: 🔗 **DeepSat (SAT-6) Airborne Dataset**

Log in to your Kaggle account (sign up if you don't already have one).

Click on Download to get the dataset files.

Extract the dataset into a local folder named dataset/ so that the folder structure looks like this:

```
dataset/
├── sat-6-full.mat
├── sat6annotations.csv
├── X_test_sat6.csv
├── X_train_sat6.csv
├── y_test_sat6.csv
├── y_train_sat6.csv
└── ...
```

Notes:

1. Ensure the dataset path matches what is used in the Jupyter notebooks (e.g., scipy.io.loadmat('dataset/sat-6-full.mat')).
2. Run the project scripts (e.g., VGG16.ipynb, ResNet50.ipynb) to use the data and weights.
3. You may need a Kaggle account and API key to download the dataset. The pre-trained weights and histories are provided for immediate use or reproducibility.

Files Included in the Dataset:

sat-6-full.mat: The main dataset in MATLAB format, containing the satellite image data and labels used for training and testing.

sat6annotations.csv: A CSV file mapping class indices to human-readable category names (e.g., "road", "building", etc.).

X_train_sat6.csv / X_test_sat6.csv: CSV files containing the training and test input data (features/images) in a flattened or preprocessed format.

y_train_sat6.csv / y_test_sat6.csv: CSV files containing the training and test labels (targets/classes) corresponding to the input data.

Pre-trained Model Output Files (Included in this Folder):

vgg_baseline.weights.h5: Saved model weights for the VGG16-based model after training.

vgg_history.csv: Training history (accuracy, loss per epoch) for the VGG16 model.

res_baseline.weights.h5: Saved model weights for the ResNet50-based model after training.

res_history.csv: Training history (accuracy, loss per epoch) for the ResNet50 model.

🛠 About DeepSat (SAT-6) Dataset:

The DeepSat (SAT-6) dataset contains 405,000 aerial image patches (28×28 pixels, 4 channels: RGB + NIR) extracted from NAIP imagery over California. Each patch is manually labeled into one of 6 land cover classes, i.e. The dataset includes six classes of terrain: barren land, trees, grassland, roads, water, and buildings.

Sampled from 1500 diverse NAIP tiles (urban, rural, forest, etc.)

Designed for training and evaluating land cover classification models

Available on Kaggle: DeepSat (SAT-6) (Link to dataset mentioned above)