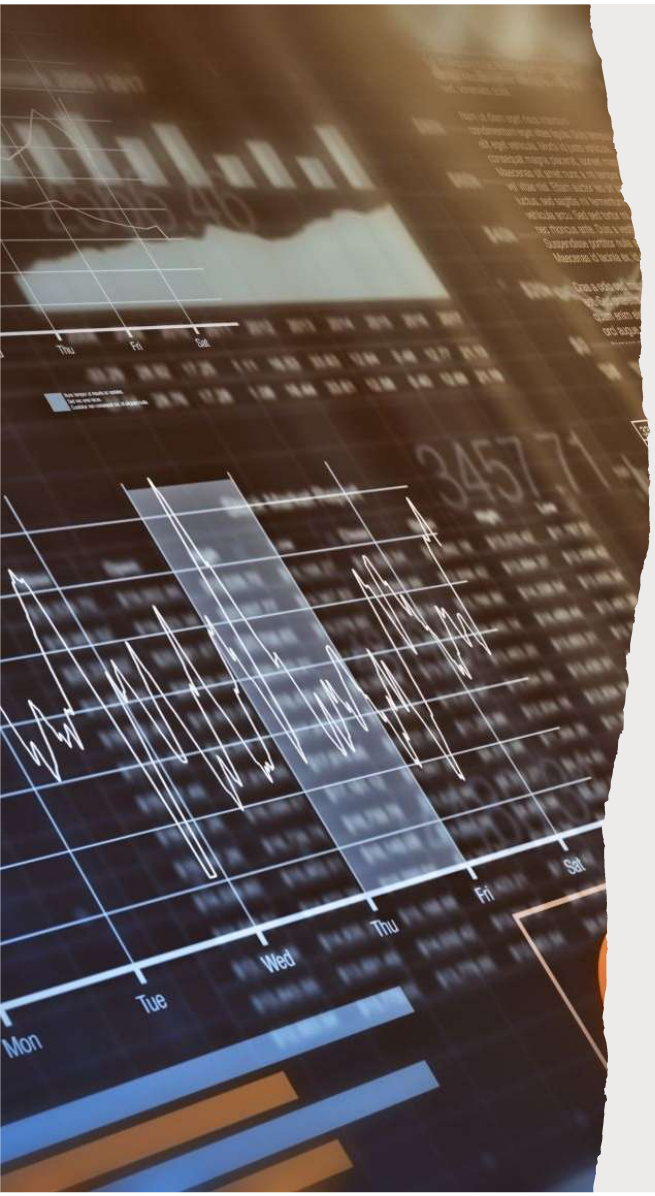
LSMA : Project Tracking - Week 1

RACHIDI Inass

MARONE Mamadou

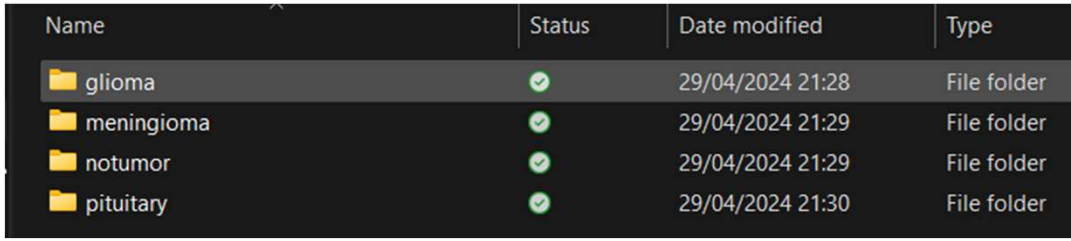
Departamento de Señales, Sistemas y Radiocomunicaciones
E.T.S. Ingenieros de Telecomunicación
Universidad Politécnica de Madrid



TASKS

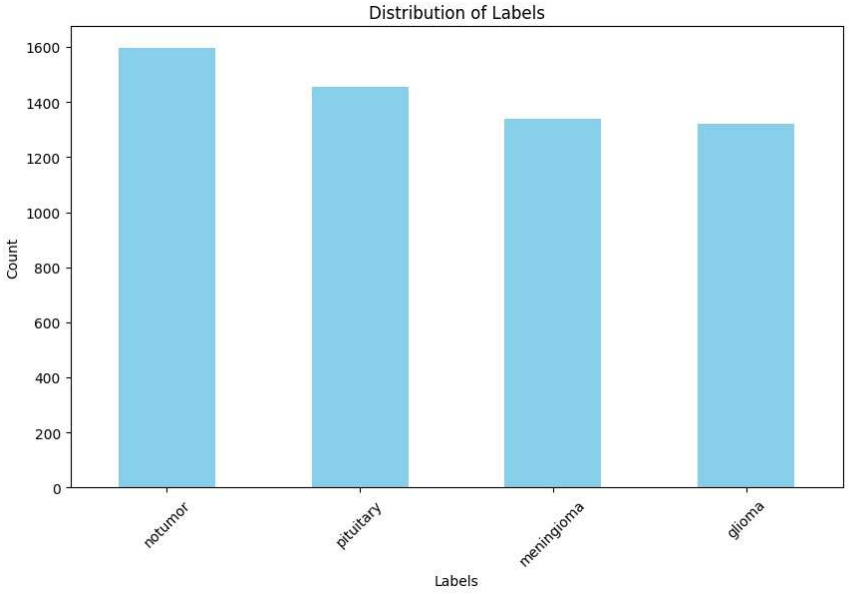
- Data preparation
- Data preprocessing
- Feature extraction
- Scanning type separation

Preparation



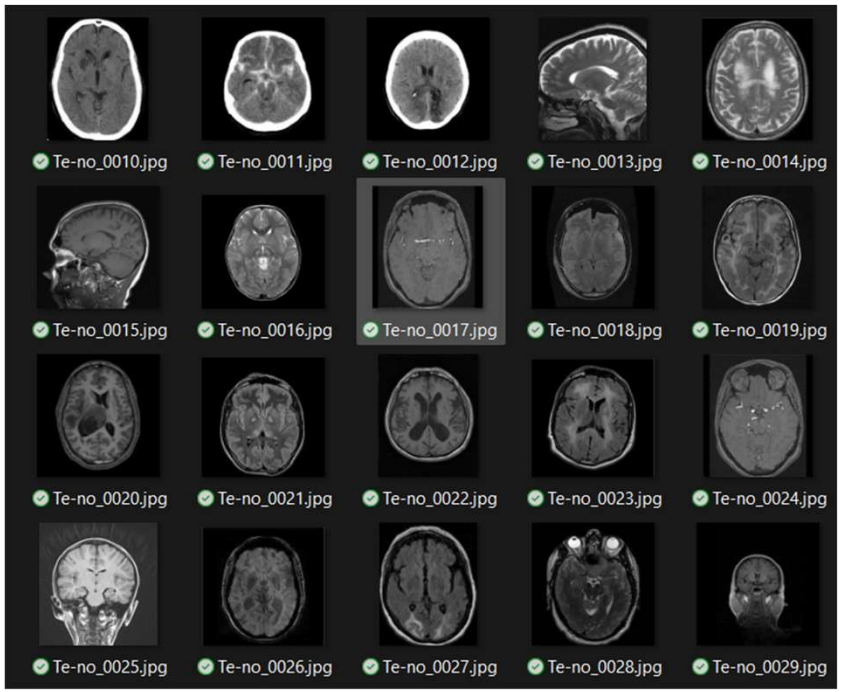
DATA > processed > test_dataset.csv > data

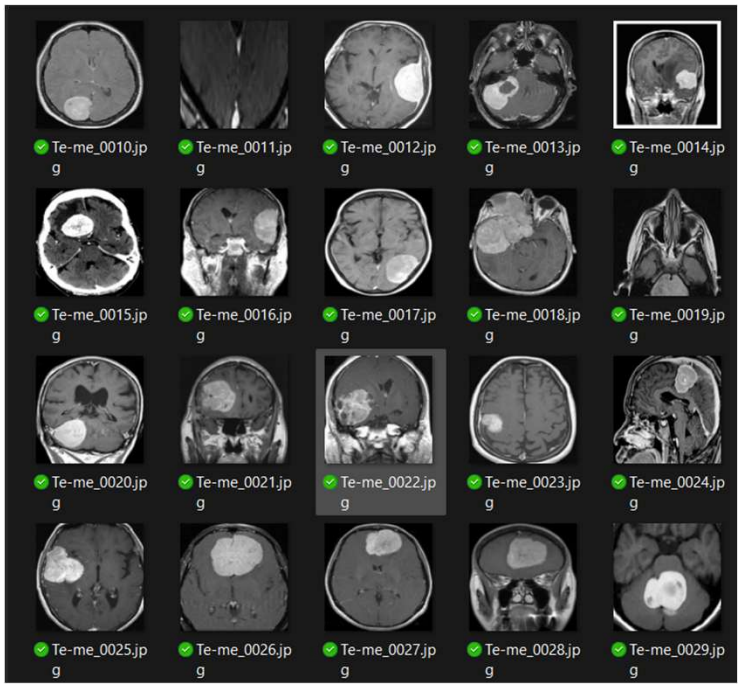
```
filename|path|label
Te-noTr_0000.jpg|DATA\processed\test\Te-noTr_0000.jpg|notumor
Te-noTr_0001.jpg|DATA\processed\test\Te-noTr_0001.jpg|notumor
Te-noTr_0002.jpg|DATA\processed\test\Te-noTr_0002.jpg|notumor
Te-noTr_0003.jpg|DATA\processed\test\Te-noTr_0003.jpg|notumor
Te-noTr_0004.jpg|DATA\processed\test\Te-noTr_0004.jpg|notumor
Te-noTr_0005.jpg|DATA\processed\test\Te-noTr_0005.jpg|notumor
Te-noTr_0006.jpg|DATA\processed\test\Te-noTr_0006.jpg|notumor
Te-noTr_0007.jpg|DATA\processed\test\Te-noTr_0007.jpg|notumor
Te-noTr_0008.jpg|DATA\processed\test\Te-noTr_0008.jpg|notumor
Te-noTr_0009.jpg|DATA\processed\test\Te-noTr_0009.jpg|notumor
```



Data Preprocessing

- ❑ Cropping
- ❑ Resizing





Feature extraction

Features

- HOG:(Histogram of Oriented Gradients)
- Local Binary Pattern
- Lab Histogram
- Color channels statistic

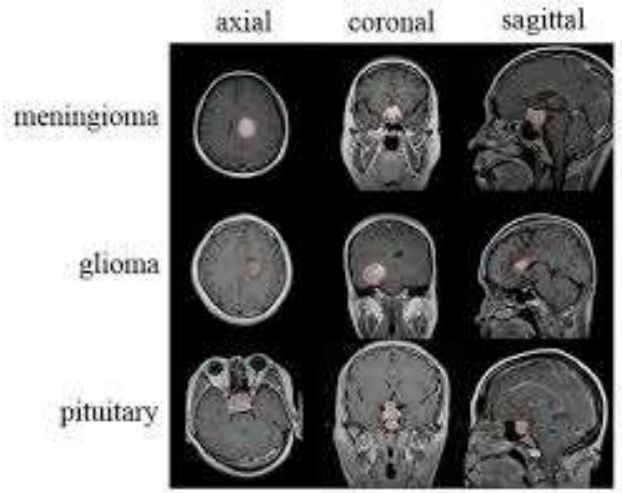
```python
def main(processed_direct):
    # Read the train and test datasets
    df_train = pd.read_csv(os.path.join(processed_direct, "train_dataset.csv"), sep="|")
    df_test = pd.read_csv(os.path.join(processed_direct, "test_dataset.csv"), sep="|")

    hog = HOG()
    lbp = LocalBinaryPatterns(numPoints =24 , radius = 8)
    lab_histogram = LabHistogram()
    color_channel_statistic = ColorChannelStatistic()
```

DATA > processed > test_dataset_feature.csv

```
filename|path|label|color_feature_0|color_feature_1|color_feature_2|color_feature_3|color_feature_4|color_featur
Te-noTr_0000.jpg|DATA\processed\test\Te-noTr_0000.jpg|notumor|110.29023333333333|110.29023333333333|110.2902333
Te-noTr_0001.jpg|DATA\processed\test\Te-noTr_0001.jpg|notumor|93.92197777777778|93.92197777777778|93.9219777777
Te-noTr_0002.jpg|DATA\processed\test\Te-noTr_0002.jpg|notumor|66.17605555555556|66.17605555555556|66.1760555555
Te-noTr_0003.jpg|DATA\processed\test\Te-noTr_0003.jpg|notumor|68.94253333333333|68.94253333333333|68.9425333333
Te-noTr_0004.jpg|DATA\processed\test\Te-noTr_0004.jpg|notumor|109.47771111111112|109.47771111111112|109.4777111
Te-noTr_0005.jpg|DATA\processed\test\Te-noTr_0005.jpg|notumor|68.71336666666667|68.71336666666667|68.7133666666
Te-noTr_0006.jpg|DATA\processed\test\Te-noTr_0006.jpg|notumor|55.34031111111111|55.34031111111111|55.3403111111
Te-noTr_0007.jpg|DATA\processed\test\Te-noTr_0007.jpg|notumor|89.10571111111112|89.10571111111112|89.1057111111
Te-noTr_0008.jpg|DATA\processed\test\Te-noTr_0008.jpg|notumor|138.7379888888889|138.7379888888889|138.737988888
Te-noTr_0009.jpg|DATA\processed\test\Te-noTr_0009.jpg|notumor|115.96314444444445|115.96314444444445|115.9631444
```

MRI type separation



Techniques:

- ❑ Unsupervised learning to cluster:
  - ❑ Feature extraction
  - ❑ Dimensionality reduction
  - ❑ clustering
- ❑ Correct the clustering : handcrafted