Exercise: Train a classifier on the dataset waferImg26x26.pkl.

Download the dataset :
https://drive.google.com/file/d/1JnXAzHxgM10lQjsFRirFz5QoIg40iQGK/view?usp=sharing

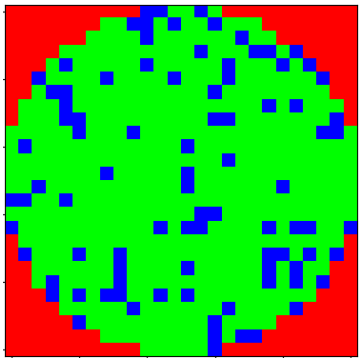
DESCRIPTION

Imagine I am a client of Lynceus and I have wafers that pass a visual inspection, after the inspection they are classified in a few categories. At the moment there is a person doing the inspection, but I would like to automate the process. For this reason, I am sending you a dataset collected during the last month of production: there are pictures of the wafers and their classification.

- Could you tell me if it is possible to automate the test and with what accuracy?
- Could you show me what results you obtain?
- Could you propose strategies to improve the result and give an estimate of how much you expect the results to improve?

ATTENTION

Please be careful with the reshaping of the dataset images. Here is how an image from the dataset should look like:



TIPS

To quickly load the data you can use either Pandas or Pickle. Here is the code for loading the data in python.

Pandas:

- import panda as pd
- import numpy as np
- df = pd.read_pickle(‘waferImg26x26.pkl’)
- images = df.images.values
- labels   = df.labels.values
- labels = np.asarray([str(l[0]) for l in labels])

Pickle:

- import cPickle as pickle
- import numpy as np
- with open('waferImg26x26.pkl', 'rb') as f:
- df = pickle.load(f)
- images = df.images.values
- labels   = df.labels.values
- labels = np.asarray([str(l[0]) for l in labels])