Creating letter datasets with generateDataset.py

Inside of uavf_2023 repo in the letter-model
Navigate to **./imaging/letter-detection/generateDataset.py**

TLDR:

1. Run pickFonts.py
2. Run generateDataset.py

Picking fonts:

To pick the fonts you want to use, run the **pickFonts.py** function. It will create the myfonts.txt file which is used to determine which fonts you are using to generate the dataset. To choose which fonts you want, use "Y" and "N" to either add it to the list or not.

Generating a dataset

Open generateDataset.py and set the size of the generated dataset and (optionally) the distributions of each key in the main().

```
def main():
    d = DataSet(2000) # Change the number here to change the size of the dataset

    custom_distributions = {}
    #add (key, 0 <= dist <= 1) to the custom_distributions
    #TODO Have them read from a file
    for key, dist in custom_distributions:
        d.setDistribution(key, dist)

    d.setRemainingDistributions()
    d.generate()
```

Then, run generateDataset.py to create a dataset. It will put the data and labels in the ./dataset directory.
The labels.txt has the name of each photo and the label of that photo.
Each label corresponds to a different character. Starting at 'a' = 0 and ending at "9" = 34

Note: zero has been left out as a key