Facial Keypoint Detection

Image Processing using Machine Learning

Preview Data

Essentially what we have here is a small preview of the data in visual studio.

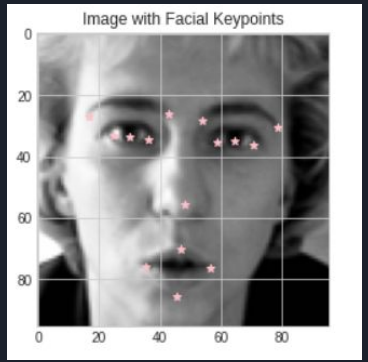
As simple as it is, all of these hard to read numbers are the values for their “corresponding” columns . These values are organized in a special file format known as .csv or comma separated values.

You can see that the first data point we’re trying to identify is the left eye center x-coordinate, which has a coordinate of 66.033..., then we have the left eye center y-coordinate which is 39.002, even though they’re not 100% lined up the the x number of each row corresponds to the same characteristic position on the first row.

```
unleft_eye_center_x,left_eye_center_y,right_eye_center_x,
66.0335639098,39.0022736842,30.2270075188,36.4216781955,5
64.3329361702,34.9700765957,29.9492765957,33.4487148936,5
65.0570526316,34.9096421053,30.9037894737,34.9096421053,5
65.2257391304,37.261773913,32.0230956522,37.261773913,60.
66.7253006135,39.6212613497,32.244809816,38.0420319018,58
69.6807476636,39.9687476636,29.1835514019,37.563364486,62
64.1318657718,34.2900402685,29.5789530201,33.1380402685,5
67.4688932039,39.4134524272,29.355961165,39.6217165049,59
65.80288,34.7552,27.47584,36.1856,58.65216,37.32928,72.95
64.1212307692,36.7403076923,29.4689230769,38.3901538462,5
65.2301886792,34.3426415094,28.8027169811,33.9378113208,5
```

Entries

- According to the data there is 7049 rows which is equivalent to 7049 different images being processed. Or 7049 faces.
- We also have 31 columns of features for each sample. There are 30 columns representing 15 key-features to be identified and 1 column that contains Grey-Scale intensity values for each pixel of each 96x96 image.
- Images are plotted using UV-coordinate systems with the origin in the top left corner.
- Left and right in our samples refers to the subject's point of view



	unleft_eye_center_x	left_eye_center_y
0	66.033564	39.002274
1	64.332936	34.970077
2	65.057053	34.909642
3	65.225739	37.261774
4	66.725301	39.621261
...	...	...
7044	67.402546	31.842551
7045	66.134400	38.365501
7046	66.690732	36.845221
7047	70.965082	39.853666
7048	66.938311	43.424510

[7049 rows x 31 columns]

Missing Data

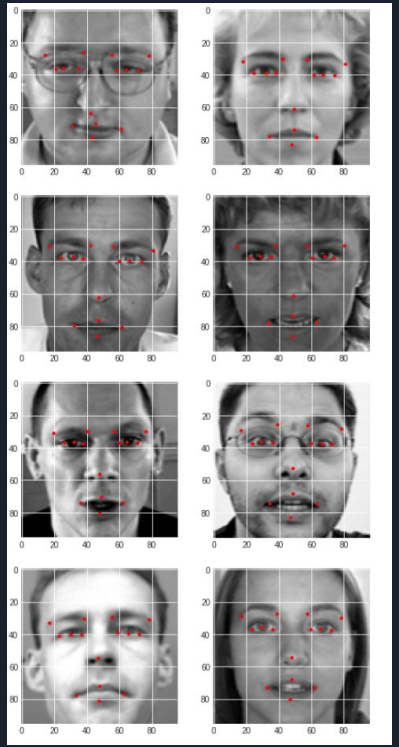
- Not all the data from each face is available or “visible” so some values appear as null/missing.
- From the Non-Null Count column in the image to the right we can observe that approximately 68% of the data for several key points is missing.
- In addition to this all 7049 images are complete with no missing pixel values or corrupted data.
- This data will most likely be split into clean and unclean sets for training.

```
Length of train data: 7049
Number of Images with missing pixel values: 0
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7049 entries, 0 to 7048
Data columns (total 31 columns):
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   left_eye_center_x          7039 non-null   float64
 1   left_eye_center_y          7039 non-null   float64
 2   right_eye_center_x         7036 non-null   float64
 3   right_eye_center_y         7036 non-null   float64
 4   left_eye_inner_corner_x    2271 non-null   float64
 5   left_eye_inner_corner_y    2271 non-null   float64
 6   left_eye_outer_corner_x    2267 non-null   float64
 7   left_eye_outer_corner_y    2267 non-null   float64
 8   right_eye_inner_corner_x   2268 non-null   float64
 9   right_eye_inner_corner_y   2268 non-null   float64
 10  right_eye_outer_corner_x   2268 non-null   float64
 11  right_eye_outer_corner_y   2268 non-null   float64
 12  left_eyebrow_inner_end_x   2270 non-null   float64
 13  left_eyebrow_inner_end_y   2270 non-null   float64
 14  left_eyebrow_outer_end_x   2225 non-null   float64
 15  left_eyebrow_outer_end_y   2225 non-null   float64
 16  right_eyebrow_inner_end_x  2270 non-null   float64
 17  right_eyebrow_inner_end_y  2270 non-null   float64
 18  right_eyebrow_outer_end_x  2236 non-null   float64
 19  right_eyebrow_outer_end_y  2236 non-null   float64
 20  nose_tip_x                 7049 non-null   float64
 21  nose_tip_y                 7049 non-null   float64
 22  mouth_left_corner_x        2269 non-null   float64
 23  mouth_left_corner_y        2269 non-null   float64
 24  mouth_right_corner_x       2270 non-null   float64
 25  mouth_right_corner_y       2270 non-null   float64
 26  mouth_center_top_lip_x     2275 non-null   float64
 27  mouth_center_top_lip_y     2275 non-null   float64
 28  mouth_center_bottom_lip_x  7016 non-null   float64
 29  mouth_center_bottom_lip_y  7016 non-null   float64
 30  Image                      7049 non-null   object
dtypes: float64(30), object(1)
memory usage: 1.7+ MB
```

Duplicate Entries

- As each face is unique and has key-points that correlate to their unique features we end up having no two sample images with exactly the same values for key-points.
- Additionally, by utilizing the Pandas method drop_duplicates() we can see that our data frame remains unchanged and we still have all 7049 rows of data entries.
- This implies that there are 7049 unique photos with unique features that require processing.



Summary of Variables

After taking a look at the coordinates in our data, we found our coordinates were being stored as floating point coordinates.

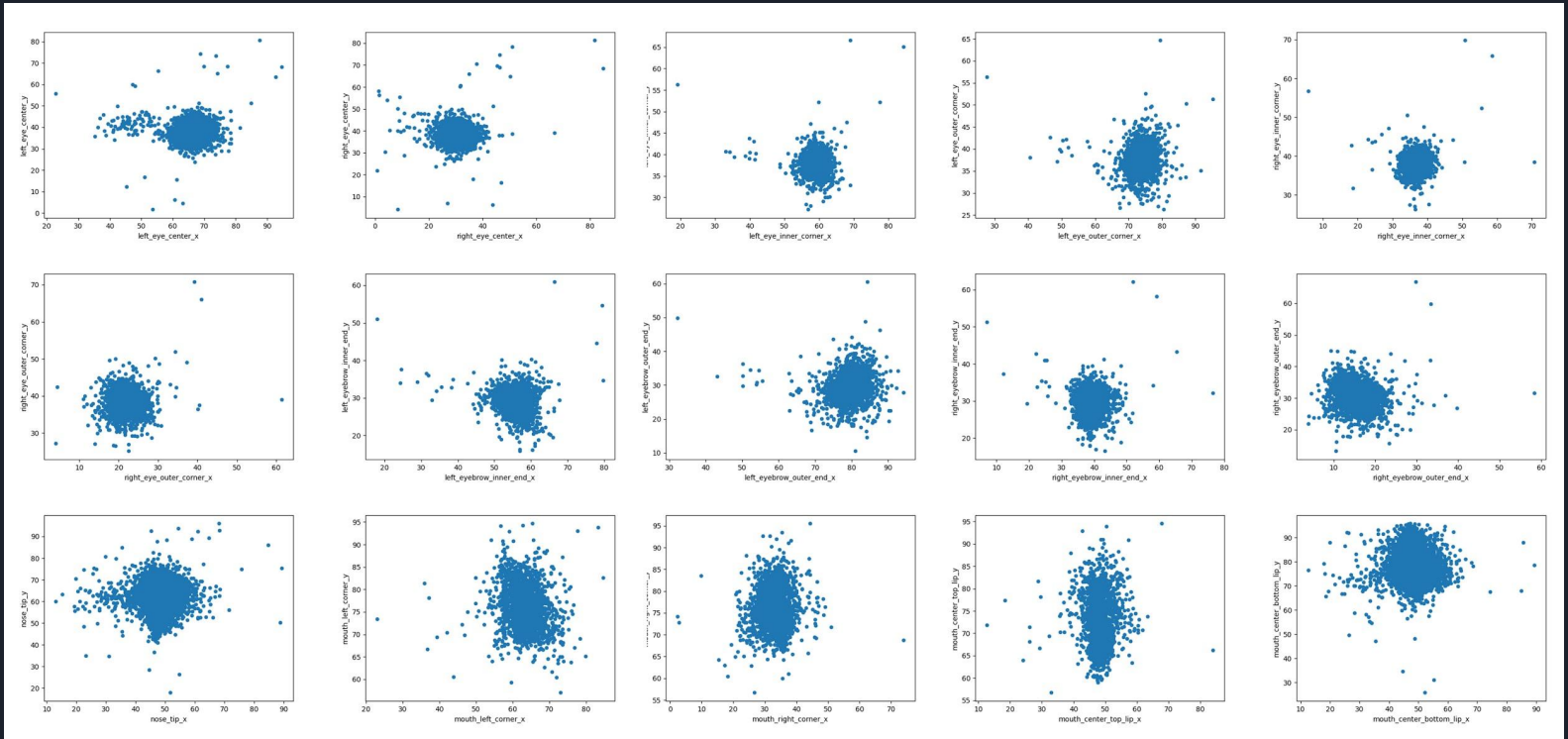
```
unleft_eye_center_x,left_eye_center_y,right_eye_center_x,
66.0335639098,39.0022736842,30.2270075188,36.4216781955,
```

The data is plotted using UV mapping, a modeling process used for texture mapping. This process uses float values in the range of [0,1], and uses transformations, such as multiplication to manipulate the data.

For example if you wanted to map a value onto a certain point on an image, you would take the value and multiply it by the dimensions of the image, which would result in a pixel coordinate.

```
texel_coord = uv_coord * [width, height]
```

Plotted Distribution of Numeric Data



Categorical Data

- Due to the nature of our data there isn't any categorical data to be processed or quantized even though one would initially think that facial key-points themselves could be categorical data.
- Inputs are all continuous real-number values on a scale from [0,96] for features and integer values [0, 255] in the arrays of pixel intensity values in the 31st column of the data
- Because of this we won't have to find ways to quantify our categorical data and we can begin preprocessing and training with the data without any augmentations to the columns of our data to account for categorical data.

Observations and Conclusions

- The dataset we are given is very straightforward and easy to work with, but there is a lot of data missing for our key-points
- Since the calculated (x,y) coordinates are given to us as continuous float values we need to determine the best way to convert them into discrete pixel values through either rounding or truncation
- Because there is missing data we need to determine the best way to handle the pre-processing of our data. There are many ways to do this, but it may be beneficial for us to partition our data into clean and unclean sets, as this may increase our chances of being able to isolate anomalies in the training of our ML model.
- This project has plenty of room for exploration in the field of Image Processing as well as Machine Learning which means we will have many techniques at our disposal, but we need to find a good way to measure which techniques hurt our predictions and which ones help

Things to Investigate

- Since we are given a dataset that is incomplete we need to find ways to make the most of the data we do have.
- We first need to discover what types of Machine Learning is best but we will most likely utilize a convolutional neural network.
- We also need to determine what forms of Data Augmentation will increase the accuracy of our predictions
- Data Augmentation in our case could include, but is not limited to: Deconvolution of images, Brightness/Contrast adjustments, Histogram Equalization, Sharpness, application of filters, alternate LUTs as well as many other image processing techniques
- We would also like to investigate how the large chunk of missing data from our training data will affect our results and what we can do to compensate for this missing data.

Thank You For Your Time!