Machine Learning Engineer Nanodegree

Capstone Project

Nicholas Low
January 24, 2017

I. Definition

Project Overview

Facial keypoints detection is a field of study that has been relevant for many years due to its potential in creating a non-invasive method of identifying points on a face that can be used to determine any number of details about a person. However, facial features can differ greatly on individuals due to many variations of conditions in gathering images of individuals; therefore, determining accurate information from these features can prove to be difficult.

A well-known beginning to facial recognition systems is due to Tuevo Kohonen, a Finnish academic, who explained that a neural network could perform facial recognition only on aligned and normalized face images utilizing eigenvectors eventually becoming known as eigenfaces seen in Figure 1 [1]. These eigenfaces are representations of what the average face may look like based on a large set of data and allow a computer to determine what parts of the face generally look like. In modern facial keypoints detection, eigenfaces still exist as a primary method of identifying parts of the face although the science has become much more advanced.

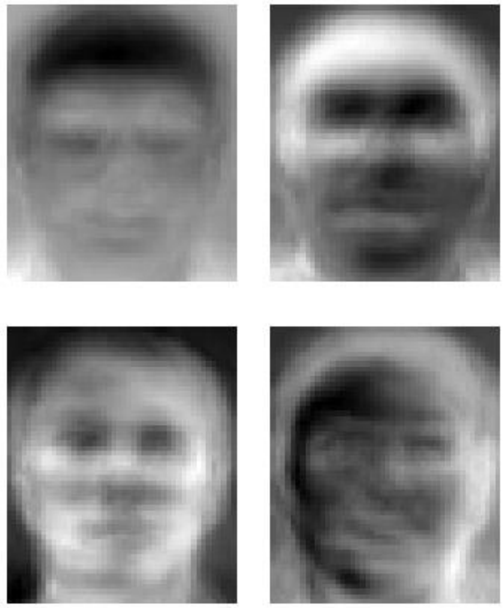

Figure 1: Example of Eigenfaces

By improving the method of facial keypoints detection systems, a number of solutions can occur. People may utilize these systems for lie detectors, medical diagnosis, biometrics, etc. On a bigger scale, every person may eventually have their own Sherlock Holmes, without the personality accompanying it (At least in the show "Sherlock"), to deduce the truth behind the many clues that exist within one's face.

Problem Statement

The problem to be solved is to predict keypoint positions/locations on face images. The goal is to predict the areas and parts where the mouth, eyes, ears, and nose are for all images with high accuracy. By determining the positions of these keypoints, a machine can gather information about the face. This will be a regression supervised learning problem as the data set includes labeled training data that uses continuous values in the form of pixels to learn from. A potential way to solve the facial keypoints detection problem is to utilize neural networks. The accuracy of the solution will be measured utilizing root mean squared error to determine the average of the squares of errors between a set of predicted outcomes and the actual outcomes. The strategy to achieve the desired solution will be:

1. Attempt to filter unnecessary pixels in the image
2. Preprocess outliers by determining quartiles
3. Apply linear regression and determine errors for benchmark model

4. Apply neural network and determine errors to compare to benchmark
5. Utilize k-fold and Grid Search Cross Validation to ensure the model generalizes to new data in the most optimal way

Metrics

The evaluation metric that will be used to quantify performance of both models is going to be the root mean squared error. The root mean square error is relevant for this problem because it works very well for a model whose main purpose is to predict. The root mean square error is the average distance of a data point from the fitted line which will help determine the generalization of the models to new data.

There will be 30 errors displayed for each model representing the spread in the values of the output compared to the model. The errors will be found by determining each of the predicted keypoints' deviation from their actual values, squaring these values, averaging the squared values, and then taking the square root. Mathematically, each of the errors will be determine by the equation:

$$RMSErrors = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

Figure 2: Root Mean Square Error Equation

II. Analysis

(approx. 2-4 pages)

Data Exploration

The details of the dataset to be used in this model were taken from Kaggle's Facial Keypoints Detection competition [2]. The 15 keypoints, to be expected as outputs, being considered that represent elements in the face where left and right refer to the point of view of the subject are:

left_eye_center, right_eye_center, left_eye_inner_corner, left_eye_outer_corner, right_eye_inner_corner, right_eye_outer_corner, left_eyebrow_inner_end, left_eyebrow_outer_end, right_eyebrow_inner_end, right_eyebrow_outer_end, nose_tip, mouth_left_corner, mouth_right_corner, mouth_center_top_lip, mouth_center_bottom_lip.

Each data point for these elements is specified by an (x,y) real-valued pair in the space of pixel indices. Data points that are missing are left blank. The input image is displayed in the last field of the datasets consisting of a list of pixels (ordered by row), as integers between (0,255). A sample of this data set is described in Table 1:

Table 1: Sample of Data Features

left_eye_center_x	left_eye_center_y	Image
66.03356	39.00227	… 238 236 237 238 240 240 239 241 241 243 240 239 231 212 190 173 148 122 …

There are 7049 images in the training set and 1783 images in the testing set and each of these images are 96x96 pixels.

The statistics that were calculated to examine the dataset are the minimum, maximum, mean, median, and standard deviation of the 15 keypoints' x and y values. The statistics will be displayed as in Figure 3:

```
Statistics for left_eye_center_x
Minimum pixels: 22.7633446452
Maximum pixels: 94.68928
Mean pixels: 66.3590212448
Median pixels: 66.4975659574
Standard deviation of pixels: 3.447988302
```

Figure 3: Sample of Statistics Data

The rest of the statistics can be viewed in the code.ipynb file associated with this project under the Data Exploration section. Based on the calculated statistics, there must be outliers in all the keypoints considering standard deviation is fairly small for all the keypoints and the maximum and minimum seem fairly outside the range of the mean. The median and mean are fairly equal for all keypoints. The largest difference between the median and mean is for right_eye_outer_corner_y with a difference of approximately 0.1664; therefore, the data is neither skewed to the right nor the left.

There are also 4909 missing values in the dataset which may make the row of data irrelevant where there is a missing value unless that missing value is also missing its x,y pair.

Exploratory Visualization

In this section, you will need to provide some form of visualization that summarizes or extracts a relevant characteristic or feature about the data. The visualization should adequately support the data being used. Discuss why this visualization was chosen and how it is relevant. Questions to ask yourself when writing this section:

- *Have you visualized a relevant characteristic or feature about the dataset or input data?*
- *Is the visualization thoroughly analyzed and discussed?*
- *If a plot is provided, are the axes, title, and datum clearly defined?*

Algorithms and Techniques

In this section, you will need to discuss the algorithms and techniques you intend to use for solving the problem. You should justify the use of each one based on the characteristics of the problem and the problem domain. Questions to ask yourself when writing this section:

- *Are the algorithms you will use, including any default variables/parameters in the project clearly defined?*
- *Are the techniques to be used thoroughly discussed and justified?*
- *Is it made clear how the input data or datasets will be handled by the algorithms and techniques chosen?*

Benchmark

In this section, you will need to provide a clearly defined benchmark result or threshold for comparing across performances obtained by your solution. The reasoning behind the benchmark (in the case where it is not an established result) should be discussed. Questions to ask yourself when writing this section:

- *Has some result or value been provided that acts as a benchmark for measuring performance?*
- *Is it clear how this result or value was obtained (whether by data or by hypothesis)?*

III. Methodology

(approx. 3-5 pages)

Data Preprocessing

In this section, all of your preprocessing steps will need to be clearly documented, if any were necessary. From the previous section, any of the abnormalities or characteristics that you identified about the dataset will be addressed and corrected here. Questions to ask yourself when writing this section:

- *If the algorithms chosen require preprocessing steps like feature selection or feature transformations, have they been properly documented?*
- *Based on the **Data Exploration** section, if there were abnormalities or characteristics that needed to be addressed, have they been properly corrected?*
- *If no preprocessing is needed, has it been made clear why?*

Implementation

In this section, the process for which metrics, algorithms, and techniques that you implemented for the given data will need to be clearly documented. It should be abundantly clear how the implementation was carried out, and discussion should be made regarding any complications that occurred during this process. Questions to ask yourself when writing this section:

- *Is it made clear how the algorithms and techniques were implemented with the given datasets or input data?*
- *Were there any complications with the original metrics or techniques that required changing prior to acquiring a solution?*
- *Was there any part of the coding process (e.g., writing complicated functions) that should be documented?*

Refinement

In this section, you will need to discuss the process of improvement you made upon the algorithms and techniques you used in your implementation. For example, adjusting parameters for certain models to acquire improved solutions would fall under the refinement category. Your initial and final solutions should be reported, as well as any significant intermediate results as necessary. Questions to ask yourself when writing this section:

- *Has an initial solution been found and clearly reported?*
- *Is the process of improvement clearly documented, such as what techniques were used?*
- *Are intermediate and final solutions clearly reported as the process is improved?*

IV. Results

(approx. 2-3 pages)

Model Evaluation and Validation

In this section, the final model and any supporting qualities should be evaluated in detail. It should be clear how the final model was derived and why this model was chosen. In addition, some type of analysis should be used to validate the robustness of this model and its solution, such as manipulating the input data or environment to see how the model's solution is affected (this is called sensitivity analysis). Questions to ask yourself when writing this section:

- *Is the final model reasonable and aligning with solution expectations? Are the final parameters of the model appropriate?*
- *Has the final model been tested with various inputs to evaluate whether the model generalizes well to unseen data?*
- *Is the model robust enough for the problem? Do small perturbations (changes) in training data or the input space greatly affect the results?*
- *Can results found from the model be trusted?*

Justification

In this section, your model's final solution and its results should be compared to the benchmark you established earlier in the project using some type of statistical analysis. You should also justify whether these results and the solution are significant enough to have solved the problem posed in the project. Questions to ask yourself when writing this section:

- *Are the final results found stronger than the benchmark result reported earlier?*

- *Have you thoroughly analyzed and discussed the final solution?*
- *Is the final solution significant enough to have solved the problem?*

V. Conclusion

(approx. 1-2 pages)

Free-Form Visualization

In this section, you will need to provide some form of visualization that emphasizes an important quality about the project. It is much more free-form, but should reasonably support a significant result or characteristic about the problem that you want to discuss. Questions to ask yourself when writing this section:

- *Have you visualized a relevant or important quality about the problem, dataset, input data, or results?*
- *Is the visualization thoroughly analyzed and discussed?*
- *If a plot is provided, are the axes, title, and datum clearly defined?*

Reflection

In this section, you will summarize the entire end-to-end problem solution and discuss one or two particular aspects of the project you found interesting or difficult. You are expected to reflect on the project as a whole to show that you have a firm understanding of the entire process employed in your work. Questions to ask yourself when writing this section:

- *Have you thoroughly summarized the entire process you used for this project?*
- *Were there any interesting aspects of the project?*
- *Were there any difficult aspects of the project?*
- *Does the final model and solution fit your expectations for the problem, and should it be used in a general setting to solve these types of problems?*

Improvement

In this section, you will need to provide discussion as to how one aspect of the implementation you designed could be improved. As an example, consider ways your implementation can be made more general, and what would need to be modified. You do not need to make this improvement, but the potential solutions resulting from these changes are considered and compared/contrasted to your current solution. Questions to ask yourself when writing this section:

- *Are there further improvements that could be made on the algorithms or techniques you used in this project?*
- *Were there algorithms or techniques you researched that you did not know how to implement, but would consider using if you knew how?*
- *If you used your final solution as the new benchmark, do you think an even better solution exists?*

Before submitting, ask yourself. . .

- Does the project report you've written follow a well-organized structure similar to that of the project template?
- Is each section (particularly **Analysis** and **Methodology**) written in a clear, concise and specific fashion? Are there any ambiguous terms or phrases that need clarification?
- Would the intended audience of your project be able to understand your analysis, methods, and results?
- Have you properly proof-read your project report to assure there are minimal grammatical and spelling mistakes?
- Are all the resources used for this project correctly cited and referenced?
- Is the code that implements your solution easily readable and properly commented?
- Does the code execute without error and produce results similar to those reported?

Work Cited

[1] T. Choudhury, "History of face recognition," 2000. [Online]. Available: http://vismod.media.mit.edu/tech-reports/TR-516/node7.html. Accessed: Jan. 10, 2017.

[2] Kaggle, "Data – Facial Keypoints Detection," 2017. [Online]. Available: https://www.kaggle.com/c/facial-keypoints-detection/data. Accessed: Jan 11, 2017.