

Siva Krishna Sirigineedi

Hotel Image Classification

Machine learning project report

# APPROACH

In this part of the project, I would be briefly describing the submitted code’s approach, implementation, and design decisions etc.,

# IMPLEMENTATION FLOW

* Training and testing images are read into dictionaries and stored them with R, G, B and grayscale intensities.
* R, G, B histograms are calculated for every image.
* Custom features have been created by myself. The image is halved into two halves horizontally, Mean intensities of red, green and blue are calculated for both the halves.
* HOG features of the grayscale image are calculated.
* Using PCA R, G, B histograms are compressed into a ten-dimensional array of each dimension.
* Features of each image are stacked horizontally for both test and train images.
* Decisions of training images are retrieved from ‘train.csv’.
* Training data is split in 90:10 ratios for validation set and training set
* Using the training set I have trained the logistic regression classifier.
* Took the best model for which validation accuracy was high
* Posterior probabilities for testing are generated by running through testing datasets.
* Validation accuracies and their log loss values are calculated.
* Posterior probabilities are saved into the CSV files.

DESIGN DECISIONS

* Images read are stored in dictionaries in terms of red, green, blue and grayscale intensities
* Extensive use of dictionaries and lambda functions for clean code and faster execution.
* No of components used in PCA for reducing the dimensions of R, G, B histograms for the image is 10.
* Logistic Regression is the classifier I have used for classifying images.
* One vs Rest Classification technique has been used for multi-class classification.
* Features used are Intensity histograms, Histogram of Oriented Gradients, Color moments
* The invalid images have been assigned equal probabilities since these images cannot be read by the computer. In the testing set, there are 2 error images. The error images in the training dataset are not considered for the training of the classifier.

FEATURES

* HOG FEATURES EXTRACTION: HOG features for the image is extracted from the grayscale image of 200 \* 200px. The approximate no of HOG descriptors in each image is around 288 key points.

Histogram is formed by dividing the image into multiple overlapping blocks of the same size(20\*20) .this histogram represents the weighted gradient distribution inside each block

* HISTOGRAM: Red, Green, Blue intensities of each and every image is compressed into the R, G, B histograms.

SPECIAL FEATURES

I had designed few features myself which have vastly improved the classification accuracy. These features have been designed with respect to the nature of images given in the training and testing dataset.

*IDEA BEHIND THESE FEATURES*: While going through the images of the dataset, I have observed a common pattern between images. Most of the aerial images I have observed are either green or blue in color. I have also observed that in swimming pool images bottom half of the image have higher no of blue pixels due to the pool area. Bathrooms and other indoor rooms have almost uniform color intensities which would be useful for image classification.

*FEATURE EXTRACTION PROCEDURE*:

* Halve the image into two equal parts horizontally.
* Split the images into R, G, B intensities matrices for the images.
* Calculate the mean intensities of both parts of images in Red, Green and Blue dimensions.

*PERFORMANCE:* The performance of classifier has been improved by 0.08 in terms of log loss.

# LIFECYCLE

|  |  |  |
| --- | --- | --- |
| S.NO | modules | log-values |
| 1 | grayscale histogram | 1.87105 |
| 2 | R,G,B histograms dimensions each compressed t0 1 component | 1.82613 |
| 3 | R,GB histograms | 1.70887 |
| 4 | R,G,B histograms dimensions each compressed t0 10 PCA components | 1.67699 |
| 5 | Compressed R,G,B histograms of size 10 with special features | 1.62201 |
| 6 | Above features with HOG features extracted from 100\*100 resized grayscales images | 1.31566 |
| 7 | Same as 5th with HOG features extracted from 200\*200 resized grayscales images | 1.26567 |

From the above graph and table, you can see the procedures I followed to reach the final destination.

We can see that maximum performance has been achieved using HOG features then next major performance has been with the special features I had developed.

# PERFORMANCE ANALYSIS

My best **LOG-LOSS** score is **1.26567** and overall **ACCURACY** is **56.672%**

# CONFUSION MATRIX

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CLASSES** | **PRE-1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** |
| ACT-1 | 2320 | 400 | 21 | 96 | 163 | 236 | 19 | 121 |
| 2 | 540 | 1970 | 15 | 138 | 233 | 311 | 24 | 154 |
| 3 | 53 | 43 | 391 | 41 | 85 | 43 | 141 | 16 |
| 4 | 161 | 197 | 20 | 1235 | 252 | 315 | 26 | 77 |
| 5 | 182 | 210 | 40 | 186 | 2091 | 520 | 62 | 147 |
| 6 | 391 | 339 | 19 | 274 | 618 | 1543 | 36 | 155 |
| 7 | 14 | 23 | 75 | 39 | 72 | 16 | 908 | 7 |
| 8 | 221 | 289 | 4 | 124 | 248 | 251 | 10 | 677 |

# CLASS LABELS

1-Bathroom,2-Guestroom,3-Pool,4-Gym,5-Restaurant,6-Lobby,7-Aerial

View, 8-Business Center

BATHROOM: If the true image is **bathroom** the highest false prediction is **guestroom**. According to my intuition, I think since the bathroom and guestrooms are uniformly colored with similar intensities. The classifier has fooled. The accuracy of predicting bathroom correct is 69% which is very high compared to the average accuracy of 56.672%.

GUESTROOM: If the true image is **guestroom** the highest false prediction is the **bathroom**. According to my intuition, I think since the bathroom and guestrooms are uniformly colored with similar intensities. The classifier has fooled. The accuracy of predicting bathroom correct is 58.19% which is little high compared to the average accuracy of 56.672%.

POOL: If the true image is **pool** the highest false prediction is **aerial view**. According to my intuition, I think since the pool and aerial view are similarly colored since the sky and water color is similar. The classifier has fooled. The accuracy of predicting pool correct is 48.09% which is low compared to the average accuracy of 56.672%.

GYM: If the true image is **gym** the highest false prediction is the **lobby**. The accuracy of predicting gym correct is 54.09% which is little low compared to the average accuracy of 56.672%.

RESTAURANT: If the true image is a **restaurant** the highest false prediction is the **lobby**. The accuracy of predicting restaurant correct is 60.9% which is high compared to the average accuracy of 56.672%.

LOBBY: If the true image is **lobby** the highest false prediction is a **restaurant**. According to my, The accuracy of predicting lobby correct is 45.71% which is low compared to the average accuracy of 56.672%.

AERIAL VIEW: If the true image is **aerial view** the highest false prediction is a **pool**. According to my intuition, I think since the pool and aerial view are similarly colored since the sky and water color is similar. The classifier has fooled. The accuracy of predicting pool correct is 78.68% which is very high compared to the average accuracy of 56.672%.

BUSINESS CENTER: If the true image is **business center** the highest false prediction is a **guest room**. The accuracy of predicting pool correct is 37.19% which is very low compared to the average accuracy of 56.672%.

# FAILED EXPERIMENTS

Experiment 1: Convolutional Neural Networks,

1. all my image data is converted into HDF5 format for running pre-trained models from the caffe library.
2. My pre-trained model was Caffenet.
3. Took the batch size of 250 for training my neural net.
4. Learning rate was around 0.01
5. I have modified the final layer of my pre-trained model into 8 node layer for matching with existing image classification problem.
6. After running for 250 iterations my error loss was showing –NaN

The running time was around 3 days on my laptop and error loss has turned into –NaN, I find it unreliable wasting lot of time, so I stopped my idea of implementing in neural nets since it was unpredictable.

Experiment 2: Linear SVC

1. Linear SVC classifier has been initially used.
2. Predict\_ proba inbuilt function is not present in the LinearSVC classifier.
3. The performance of this classifier is not satisfactory.
4. The training time was about 7 hours and few instances 16gb of memory was not sufficient and the program crashed.
5. So, I abruptly stopped using LinearSVC classifier due to memory and time constraints.