Table of Contents

[Task general description 2](#_Toc532828159)

[Data set 2](#_Toc532828160)

[Implemented pipe 2](#_Toc532828161)

[Hyper parameter tuning 2](#_Toc532828162)

[The hyper parameter tuning graphs 4](#_Toc532828163)

[Test results: 13](#_Toc532828164)

[Error visualization: 14](#_Toc532828165)

[Appendices 17](#_Toc532828166)

[Hyper parameter tuning – Linear kernel 17](#_Toc532828167)

[The hyper parameter tuning graphs – Linear kernel 17](#_Toc532828168)

## Task general description

In this task we:

* Implemented an object recognition pipe.
* Trained and tuned it on a subset of the Caletch101 dataset.
* Tested the classifier performance on different subset of the Caletch101 dataset.
* Analyzed and reported our results.

## Data set

In this task we worked with Caltech 101 dataset

* The data contains 101 classes, with 31-800 images each.
* Data homepage: <http://www.vision.caltech.edu/Image_Datasets/Caltech101/>
* We loaded the images, got them to gray scale, and resized them according to the hyper parameter S.
* Data split
* We debugged and tuned the pipe on the first 10 classes (fold 1)
* We ran the algorithm with the best hyper parameter configuration found, on classes 11-20 (fold 2).
* We trained and tested it: training on 20 images per class , testing on 20 others (unless there are less than 40 images for the class – in which case we had less images for test)

## Implemented pipe

The pipe we chose to implement is HOG+ SVM. The SVM classifier was trained with two types of kernel (linear and RBF) as described below.

## Hyper parameter tuning

We tuned the hyper parameters with the two types of kernel separately, and included only the tuning of the RBF kernel in the report – the kernel which led us to the best accuracy rate.

After the tuning of the hyper-parameters on both types of kernel, we got better results on the RBF kernel (0.26 < 0.274). In the body of this report, the results of this kernel are attached, and in the appendices we attached the results of the tuning with the linear classifier.

The hyper parameters that where tuned are:

S - Image size - after resizing the original images of the data set.

Pixels\_Per\_Cell - Spatial cell size - Input parameter for the HOG algorithm. The size of the cell in which the gradient is computed.

Num\_Of\_Bins - Number of orientation bins – Input parameter for HOG algorithm. The number of possible directions of the gradient.

Block norm method – input parameter for the HOG algorithm. The normalization method for each block.

C – The SVM tradeoff parameter between complexity and data fit.

Gamma - RBF kernel. Where

We chose to tune each parameter separately in order to find the best results due to long run-time.

The following table shows the hyper-parameters, training ranges, the selected value and current error rate (validation set) for each parameter (when the previous hyper parameters were fixed to their selected value):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Hyper Parameter | Range | Jumps | Selected Value | Error Rate |
| Gamma | 1.e-07 : 1 | exp |  |  |
|  | 1 : 9 | 1 |  |  |
|  | 1 : 3 | 0.1 | 1.7 | 0.378 |
| C | 1.e-07 : 10 | exp |  |  |
|  | 10 : 50 | 5 | 10 | 0.305 |
| Number of oriented gradient(num\_of\_bins) | 6 : 15 | 1 | 10 | 0.295 |
| Spatial cell size (Pixels per cell) | 5 : 50 | 5 |  |  |
|  | 17 : 33 | 2 | 23 | 0.295 |
| Cells per block | 1 : 5 | 1 | 2 | 0.295 |
| Image size (S) | 70 : 160 | 10 |  |  |
|  | 95 : 115 | 4 | 105 | 0.295 |
| Block norm method | 'L1' 'L1-sqrt' 'L2' 'L2-Hys' |  | 'L1' | 0.29 |

## The hyper parameter tuning graphs

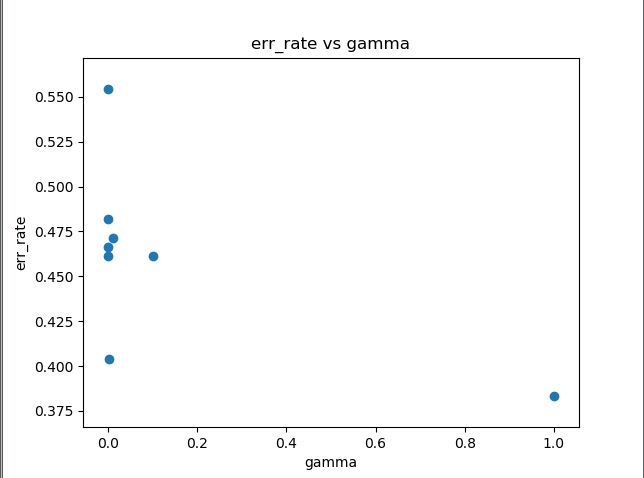
Graphs showing the validation error as a function of hyper parameter value for the hyper parameters that were systematically tuned:

**Gamma**:

Values: 1.e-07 1.e-06 1.e-05 1.e-04 1.e-03 1.e-02 1.e-01 1.e+00

Err rate: 0.48186528 0.55440415 0.46632124 0.4611399 0.40414508 0.47150259

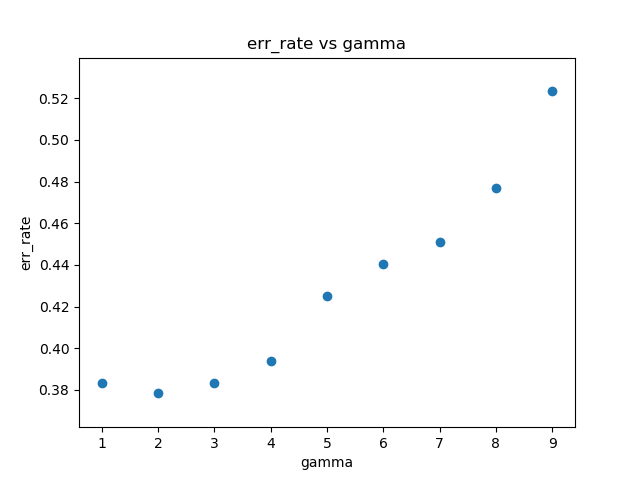
0.4611399 0.38341969



Values: 1 2 3 4 5 6 7 8 9

Err rate: 0.38341969 0.37823834 0.38341969 0.39378238 0.42487047 0.44041451

0.4507772 0.47668394 0.52331606



Values: 1. 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2. 2.1 2.2 2.3 2.4 2.5 2.6 2.7

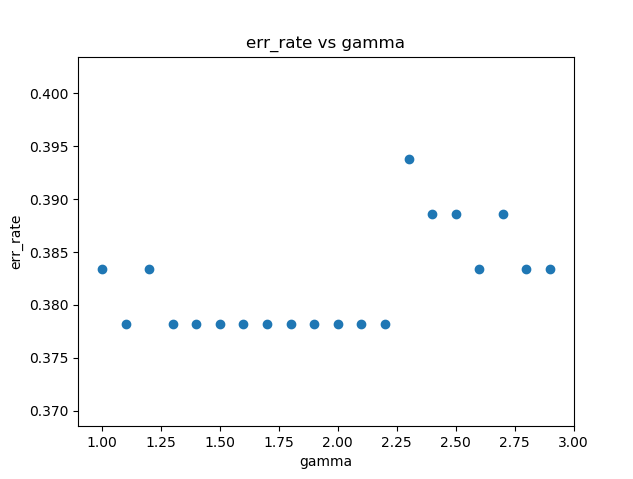
2.8 2.9

Err rate: 0.38341969 0.37823834 0.38341969 0.37823834 0.37823834 0.37823834

0.37823834 0.37823834 0.37823834 0.37823834 0.37823834 0.37823834

0.37823834 0.39378238 0.38860104 0.38860104 0.38341969 0.38860104

0.38341969 0.38341969

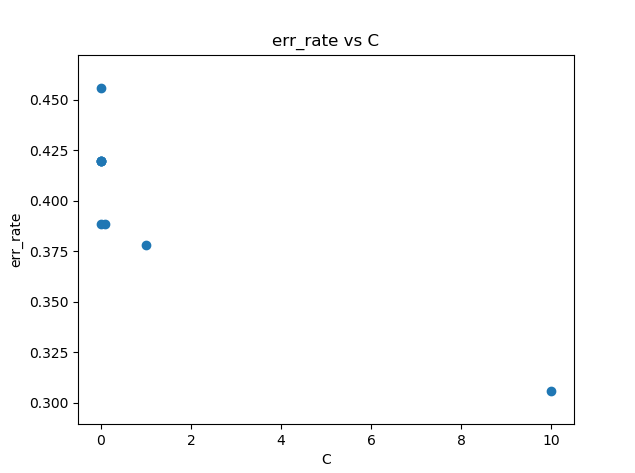
Best gamma is **1.7** with 0.378 err rate

**C:**

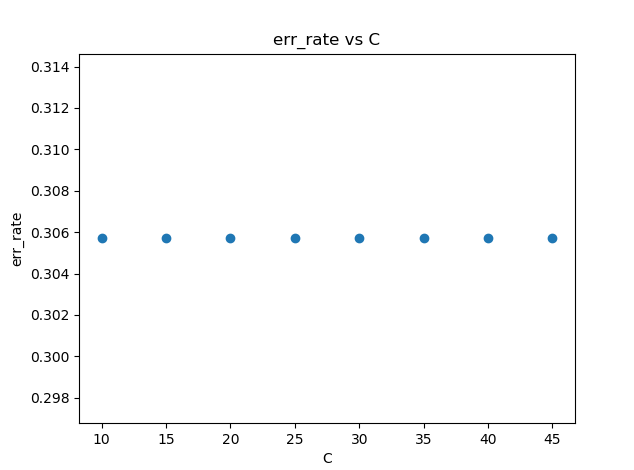
Values for c: 1.e-07 1.e-06 1.e-05 1.e-04 1.e-03 1.e-02 1.e-01 1.e+00 1.e+01

Err rate: 0.41968912 0.41968912 0.41968912 0.41968912 0.45595855 0.38860104

0.38860104 0.37823834 0.30569948



When we checked higher value for C we got the same err rate



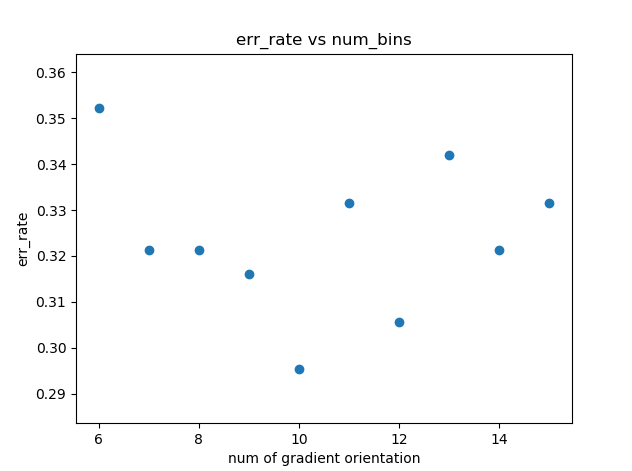
Best C where gamma is set as found before is **10** with err rate of 0.305.

**Number of oriented gradient:**

Values: 6 7 8 9 10 11 12 13 14 15

Err rate: 0.35233161 0.32124352 0.32124352 0.31606218 0.29533679 0.33160622

0.30569948 0.34196891 0.32124352 0.33160622



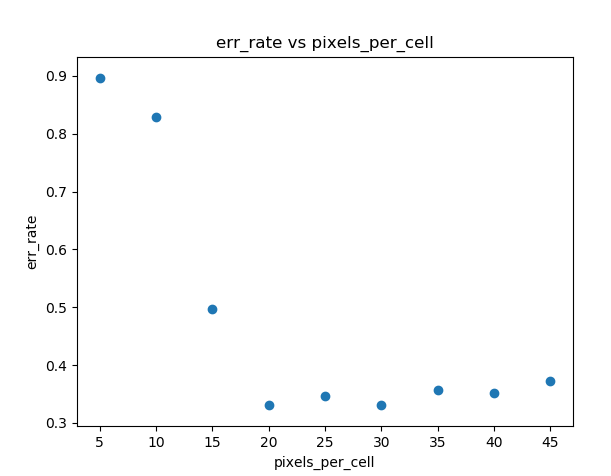
Best num\_of\_bins where gamma and C are as found before is **10** with err rate of 0.295.

**Pixels per cell:**

Values: 5 10 15 20 25 30 35 40 45

Err rate: 0.89637306 0.82901554 0.49740933 0.33160622 0.34715026 0.33160622

0.35751295 0.35233161 0.37305699



Values: 17 19 21 23 25 27 29 31

Err rate: 0.40414508 0.35233161 0.30569948 0.29533679 0.34715026 0.34715026

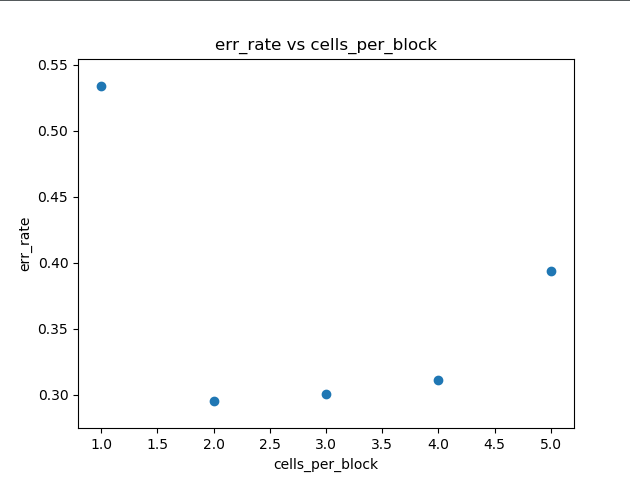
0.31606218 0.30569948

Best pixels\_per\_cell is **23** with err rate of 0.295 where gamma, C and num\_of\_bins as set before.

**Cells per block**

Values: 1 2 3 4 5

Err rate: 0.53367876 0.29533679 0.30051813 0.31088083 0.39378238



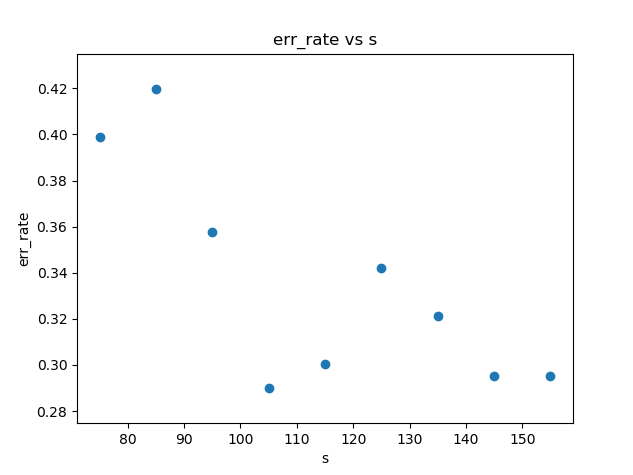
Best cells\_per\_block is **2** with err rate of 0.295 where gamma, C, num\_of\_bins and pixels\_per\_cell as set before.

**S**

Values: 75 85 95 105 115 125 135 145 155

Err rate: 0.39896373 0.41968912 0.35751295 0.29015544 0.30051813 0.34196891

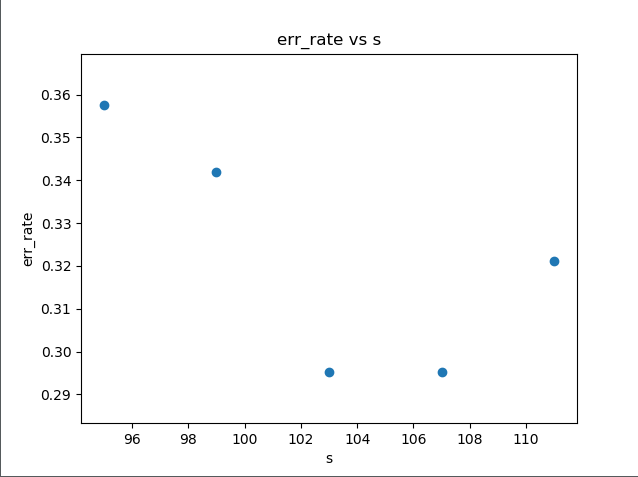
0.32124352 0.29533679 0.29533679



Values: 95 99 103 107 111

Err rate: 0.35751295 0.34196891 0.29533679 0.29533679 0.32124352

For bigger values of S the error rate increased:

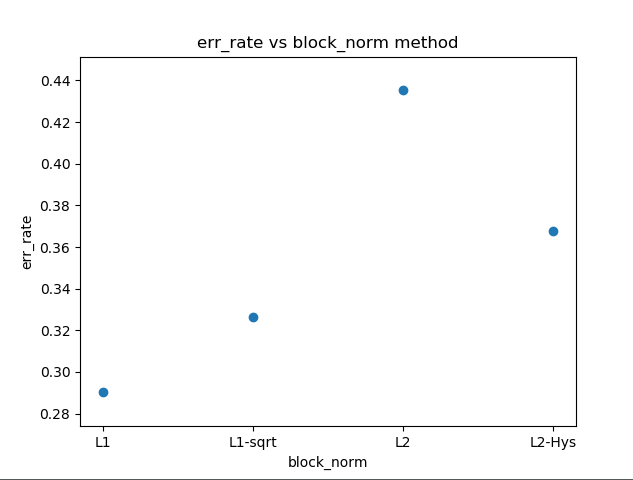


Best S is **105** where with err rate of 0.295 where gamma, C, num\_of\_bins, pixels\_per\_cell and cells\_per\_block are as set before.

**Block norm method**:

methods: 'L1' 'L1-sqrt' 'L2' 'L2-Hys'

err rate: 0.29015544 0.32642487 0.43523316 0.36787565



The best method to normalized is **L1** with all the other hyper parameters as set before.

## Test results

**Error rate:**

The error rate obtained over the test set is 26%

**Confusion matrix:**

The confusion matrix is a matrix of size M×M where M is the number of classes. the Mij cell contains the number of times that an example with label i is predicted to be of label j.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Predicted Label | | | | | | | | | | |
|  |  | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** |
| True Label | **11** | 13 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 3 | 0 |
| **12** | 1 | 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **13** | 1 | 0 | 11 | 2 | 2 | 0 | 0 | 2 | 1 | 1 |
| **14** | 2 | 2 | 0 | 14 | 0 | 0 | 0 | 0 | 1 | 1 |
| **15** | 2 | 0 | 5 | 0 | 11 | 0 | 2 | 0 | 0 | 0 |
| **16** | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 0 | 0 | 0 |
| **17** | 0 | 1 | 1 | 1 | 1 | 0 | 15 | 1 | 0 | 0 |
| **18** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 19 | 0 | 0 |
| **19** | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 14 | 1 |
| **20** | 0 | 3 | 1 | 2 | 1 | 0 | 0 | 0 | 1 | 12 |

## Error visualization:

For each class we show images of the two largest errors on images of the class (i.e. images from the class which were miss-classified).

|  |  |
| --- | --- |
| class 11 – "brontosaurus" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.32.59.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.32.43.jpeg |
| class 12 - "buddha" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.32.29.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.32.16.jpeg |
| class 13 – "butterfly" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.31.38.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.31.17.jpeg |
| class 14 – "camera" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.31.02.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.30.45.jpeg |
| class 15 – "cannon" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.30.18.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.30.00.jpeg |
| class 16 – "car\_side” | |
| None. | None. |
| class 17 – "ceiling\_fan" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.29.28.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.29.09.jpeg |
| class 18 – "cellphone" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.28.34.jpeg | Only one. |
| class 19 – "chair" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.28.14.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.27.57.jpeg |
| class 20 – "chandelier" | |
| C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.27.34.jpeg | C:\Users\Administrator\Documents\שנה ה\סמסטר א\למידה, ייצוג וראייה ממוחשבת\תיקיה חדשה\WhatsApp Image 2018-12-17 at 08.27.16.jpeg |

# Appendices

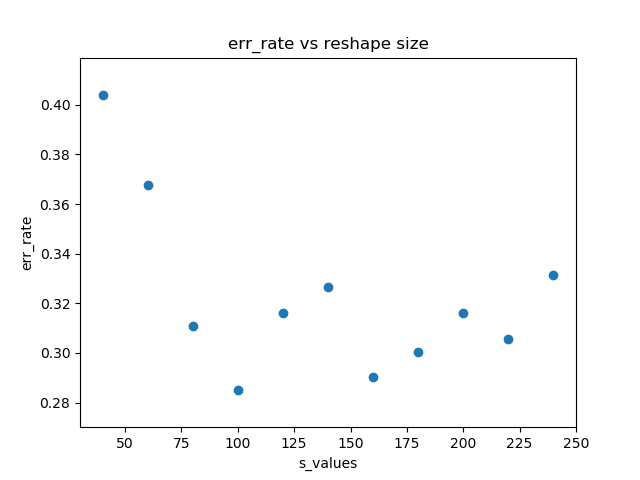
## Hyper parameter tuning – Linear kernel

|  |  |  |  |
| --- | --- | --- | --- |
| Hyper Parameter | Range  Jumps | Selected Value | Error Rate |
| Image size (S) | 40 : 250  20 |  |  |
|  | 240 : 270  10 |  |  |
|  | 75 : 160  10 | 145 | 0. 31 |
| Number of oriented gradient(num\_of\_bins) | 6 : 16  1 | 12 | 0.31 |
| Spatial cell size (Pixels per cell) | 5 : 29  2 |  |  |
|  | 17 : 33  2 | 23 | 0.3 |
| Cells per block | 2 : 6  1 | 2 | 0.3 |
| C | 1.e-05 : 1  exp |  |  |
|  | 0.1 : 1.4 0.1 | 1.1 | 0.274 |
| Block norm method | 'L1' 'L1-sqrt' 'L2' 'L2-Hys' | 'L1' | 0.274 |

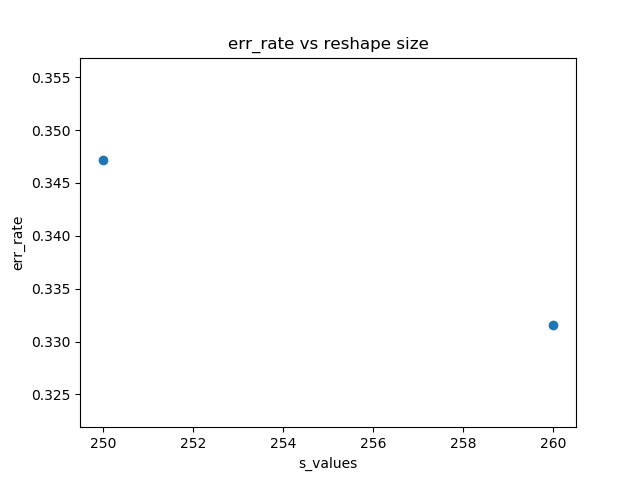
## The hyper parameter tuning graphs – Linear kernel

**S:**

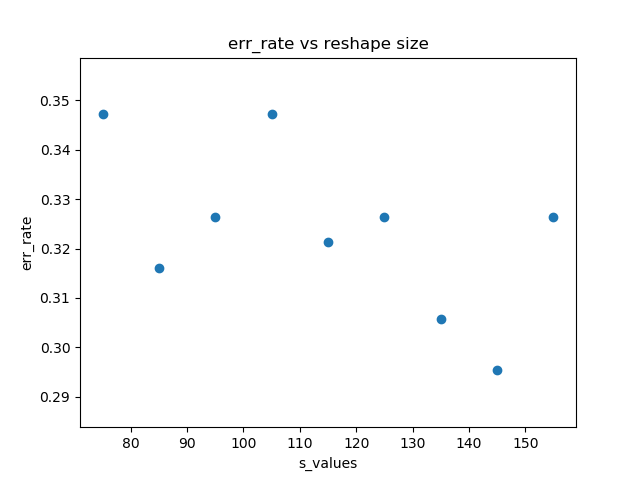
40-250,20



240-270,10



75-160,10



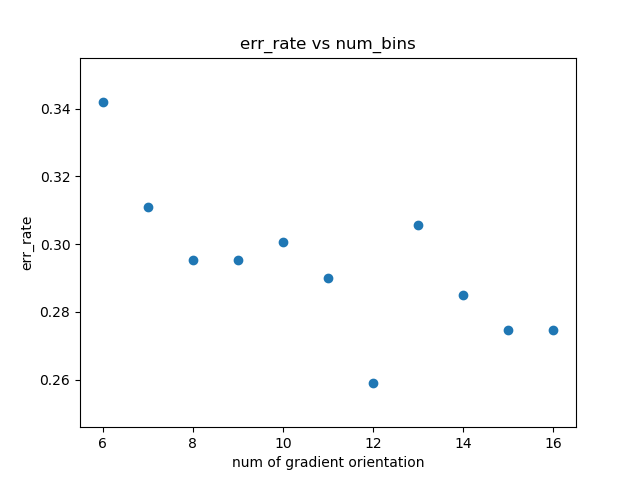
Best S is **145** with err rate of 0.295.

**Num of bins:**

Bins: 6 7 8 9 10 11 12 13 14 15 16

Err: 0.34196891 0.31088083 0.29533679 0.29533679 0.30051813 0.29015544

0.25906736 0.30569948 0.28497409 0.2746114 0.2746114

Where S is fixes as before:

best num of bins is **12** with err 0. 25906736

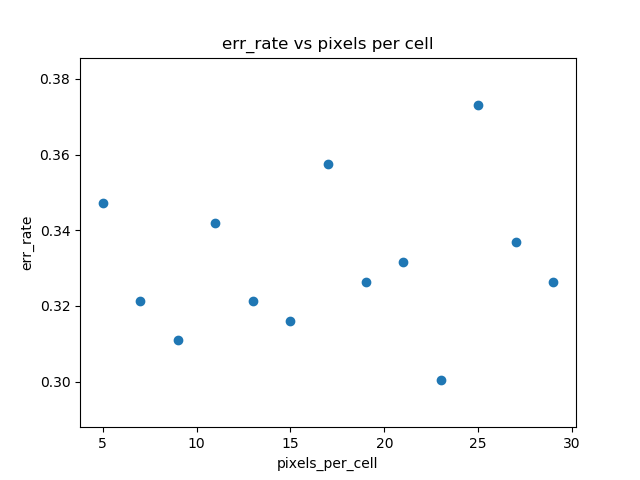
**Pixels per cells**

Pixels: [ 5 7 9 11 13 15 17 19 21 23 25 27 29]

Err: 0.34715026 0.32124352 0.31088083 0.34196891 0.32124352 0.31606218

0.35751295 0.32642487 0.33160622 0.25951813 0.37305699 0.33678756

0.32642487



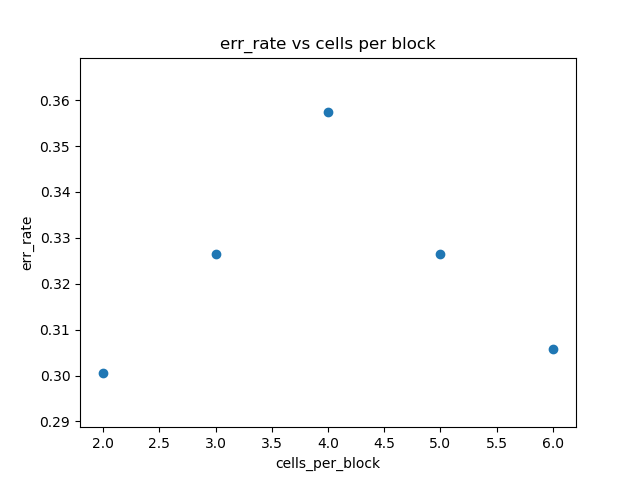
Where S and bins as found before:

Best result for **23** pixels with err rate: 0.3

**Cells per block**

Cells: 2 3 4 5 6

Err: 0.30051813 0.32642487 0.35751295 0.32642487 0.30569948



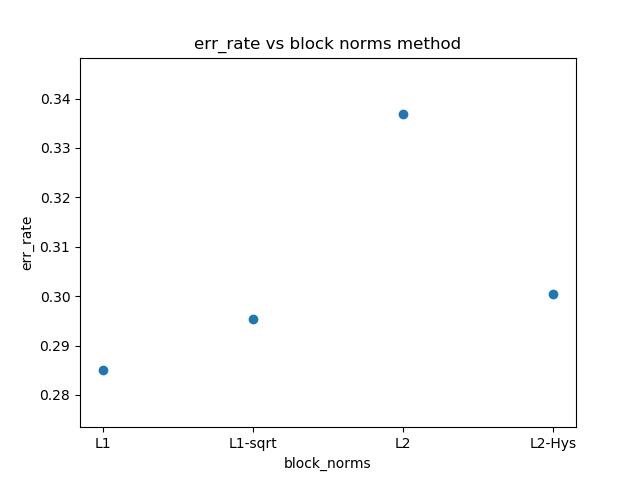
Where S, bins and pixels per cell as found before:

Best cells per block is **2\*2** with err of 0.300

**Block Norm Method:**

Methods: 'L1', 'L1-sqrt', 'L2', 'L2-Hys'

Err: 0.28497409 0.29533679 0.33678756 0.30051813



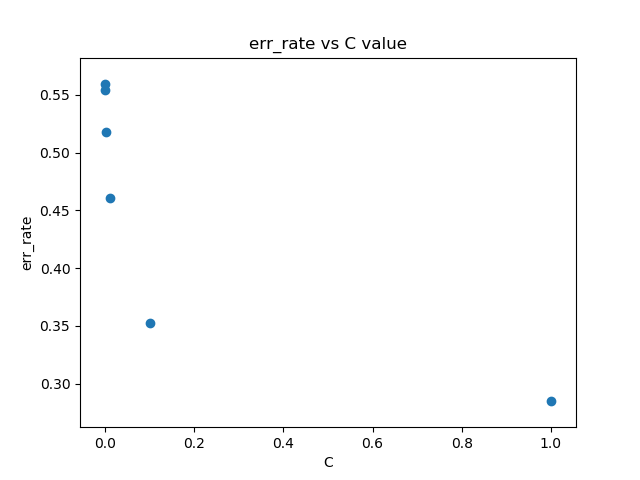
Where S, bins, pixels per cell and cells per block as found before:

Best norm method is **L1** with err of 0.285

**C:**

Values for roughly tuning: 1.e-05 1.e-04 1.e-03 1.e-02 1.e-01 1.e+00

Err: 0.55440415 0.55958549 0.51813472 0.4611399 0.35233161 0.28497409



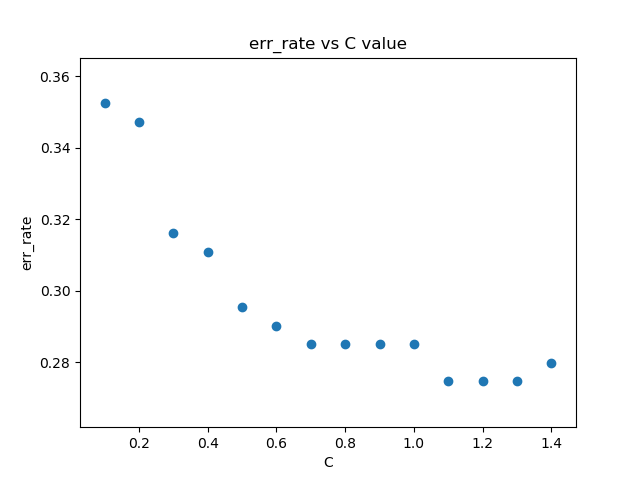
Fine tuning:

Values: 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1. 1.1 1.2 1.3 1.4

Err: 0.35233161 0.34715026 0.31606218 0.31088083 0.29533679 0.29015544

0.28497409 0.28497409 0.28497409 0.28497409 0.2746114 0.2746114

0.2746114 0.27979275



Where S, bins, pixels per cell, cells per block and Norm method were found before:

Best C is **1.1** with err of 0.276