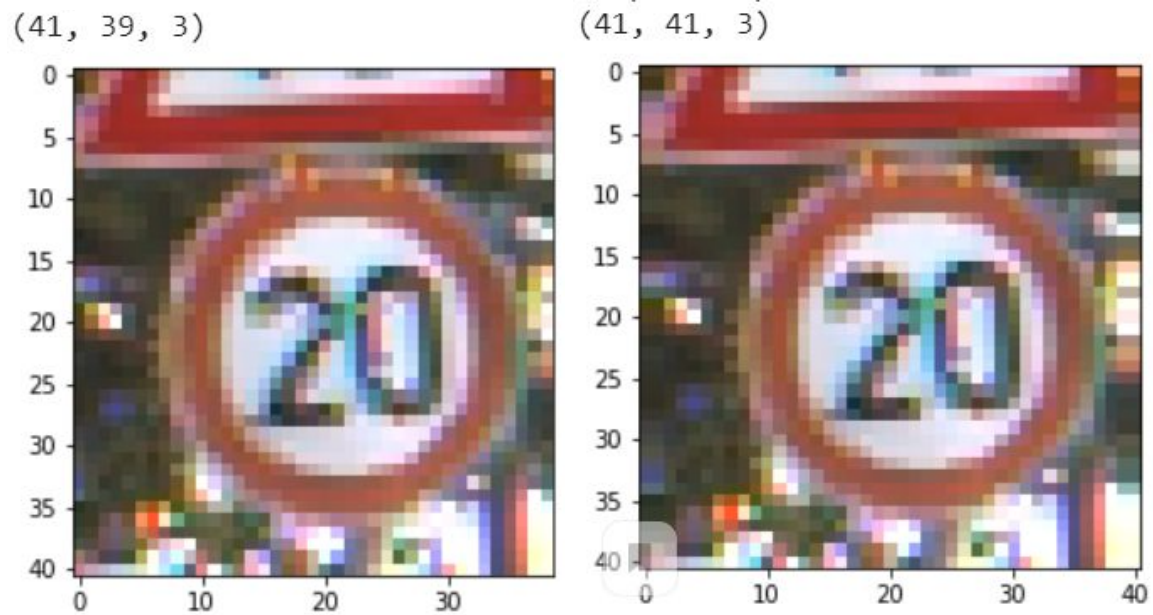


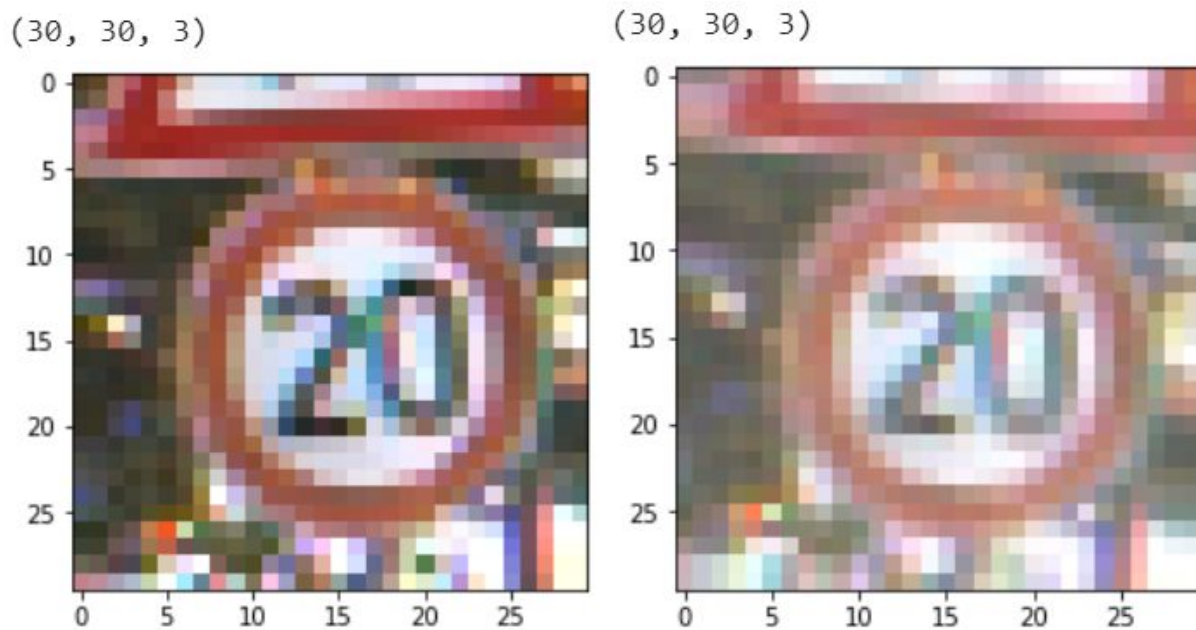
Machine Learning Fall 2019 Assignment 2 Report

Zhuchkov Alexey BS17 -DS-01

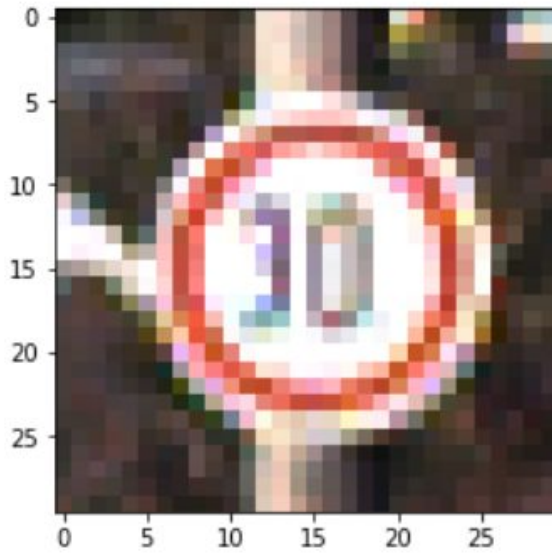
Samples from the dataset (left) and padded version (right)



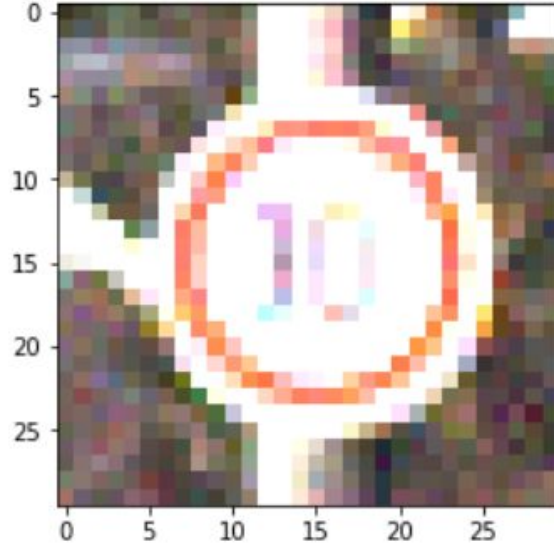
Resized (left) and augmented version (right)



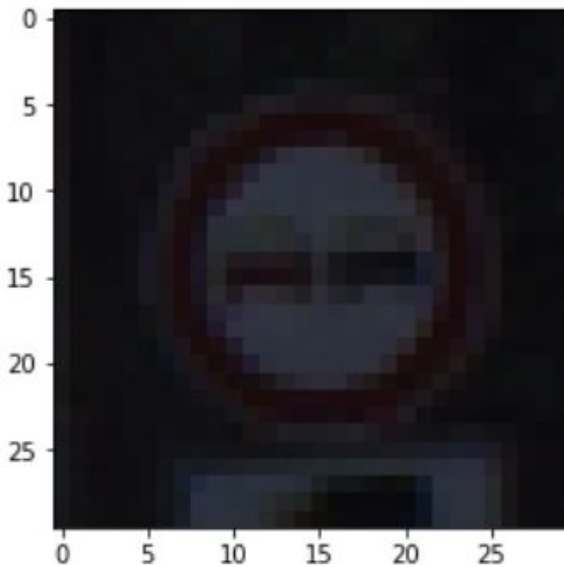
(30, 30, 3)



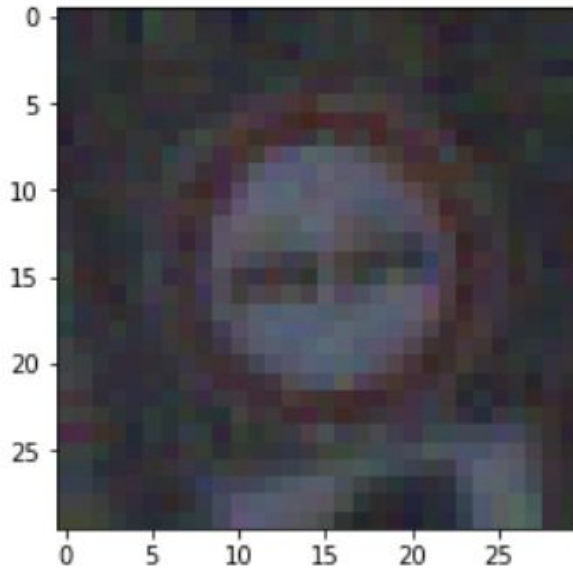
(30, 30, 3)



(30, 30, 3)



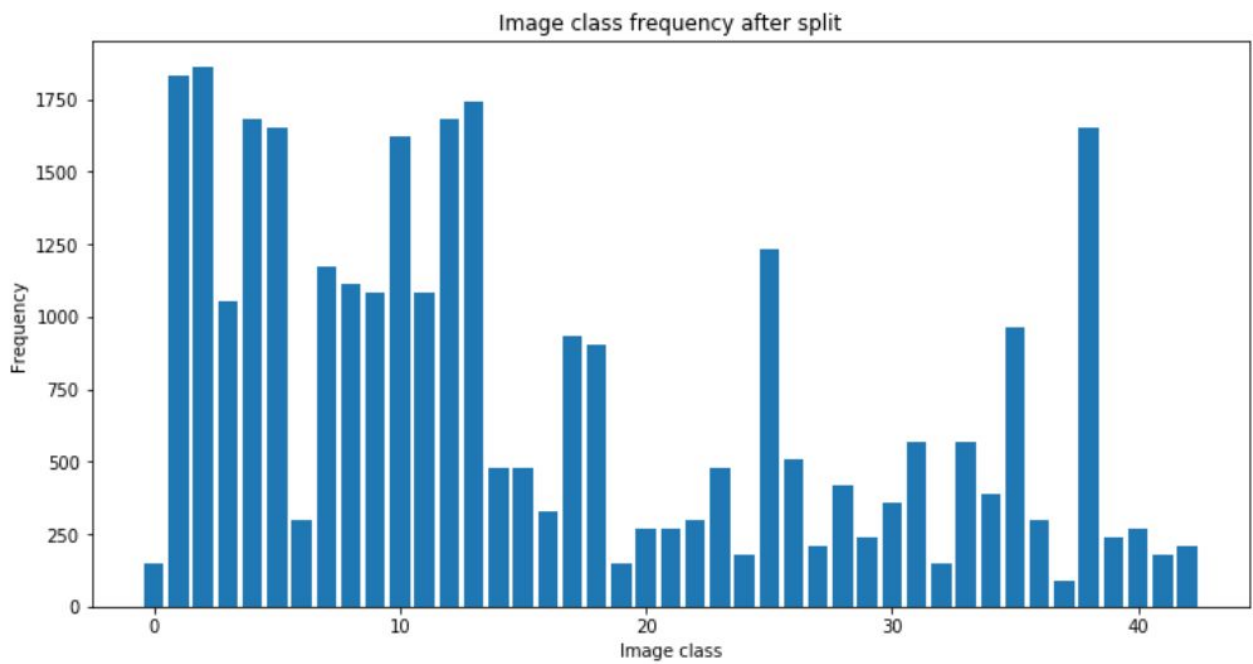
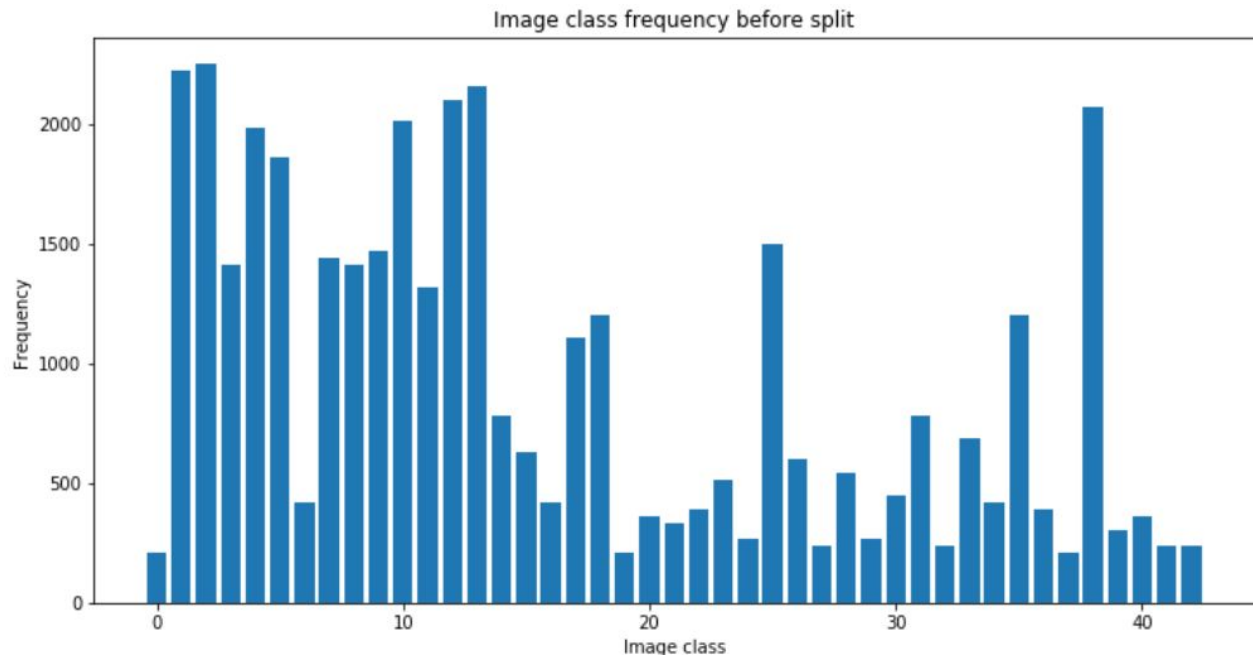
(30, 30, 3)



List of applied augmentations:

1. **Gaussian noise** - Best imitates noise that causes in digital images during acquisition e.g. sensor noise caused by poor illumination or high temperature (probability of applying is 0.8)
2. **Blur** - Imitates the camera vibration during video capturing that results in blur images in a video stream (probability of applying is 0.3)
3. **Brightness and Contrast** - Imitates weather different conditions (probability of applying is 1)
4. **Rotate (limit $\pm 10^\circ$)** - Imitates the perspective of object appearance during the video capturing (probability of applying is 0.2)

Images frequencies



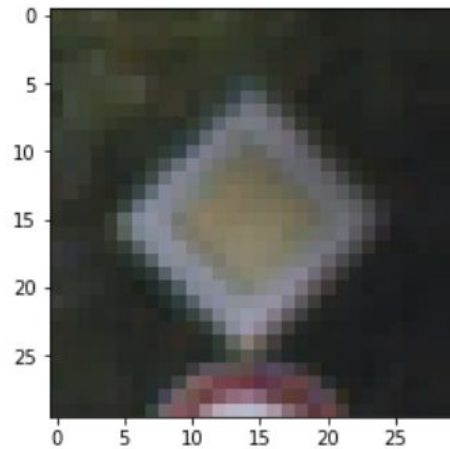
Evaluations

class | precision | recall | f1-score

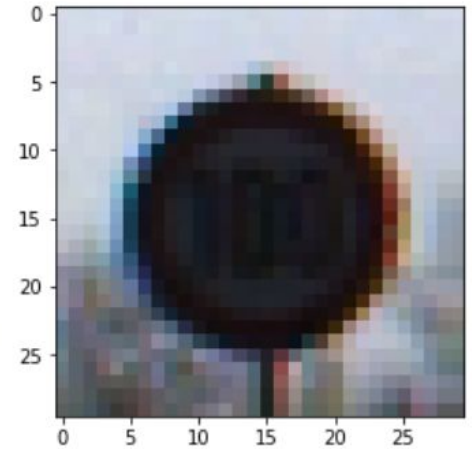
0	0.33	0.43	0.38
1	0.85	0.79	0.82
2	0.73	0.73	0.73
3	0.66	0.75	0.70
4	0.75	0.77	0.76
5	0.66	0.62	0.64
6	0.95	0.55	0.69
7	0.73	0.77	0.75
8	0.75	0.73	0.74
9	0.95	0.87	0.91
10	0.94	0.95	0.94
11	0.94	0.93	0.93
12	1.00	0.92	0.96
13	0.96	0.97	0.96
14	0.96	1.00	0.98
15	0.90	0.89	0.89
16	0.92	0.96	0.94
17	0.99	0.88	0.93
18	0.92	0.59	0.72
19	0.72	0.55	0.62
20	0.41	0.83	0.55
21	0.76	0.68	0.72
22	0.89	0.96	0.92
23	0.68	0.61	0.64
24	0.69	0.78	0.73
25	0.96	0.89	0.92
26	0.66	0.84	0.74
27	0.77	0.88	0.82
28	0.79	0.85	0.82
29	0.61	0.74	0.67
30	0.46	0.53	0.49
31	0.78	0.93	0.85
32	0.64	0.62	0.63
33	0.91	0.98	0.94
34	0.92	0.99	0.95
35	0.98	0.94	0.96
36	0.84	0.97	0.90
37	0.95	0.90	0.92
38	0.98	0.94	0.96
39	0.88	0.88	0.88
40	0.73	0.94	0.82
41	0.49	0.75	0.59
42	0.75	0.89	0.81

Missclasifications visualized

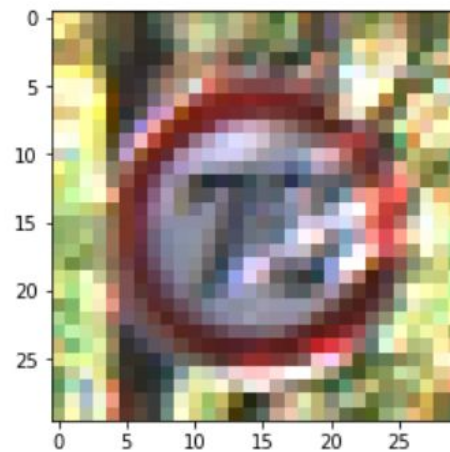
True 12
Predicted 15
(30, 30, 3)



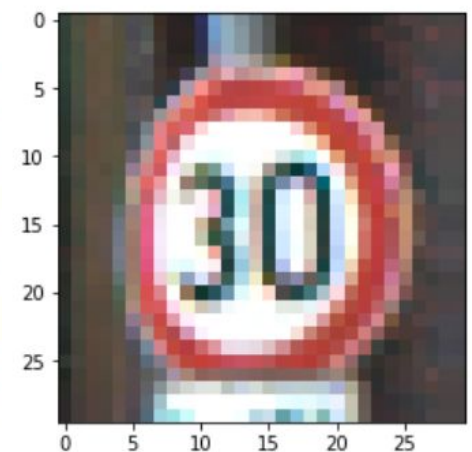
True 7
Predicted 8
(30, 30, 3)



True 4
Predicted 1
(30, 30, 3)



True 1
Predicted 2
(30, 30, 3)



Train accuracy: 0.9999874968742185
Test accuracy: 0.8307996832937451
Validation accuracy: 0.8481622306717364

As we can see from the table, a random forest model (RFM) gives not a too bad result. However, problems arise in the case of classification color and shape similar signs (Ex. sign max speed 100 and 120). RFM cannot easily separate them if a sign is a skewed or bad resolution.

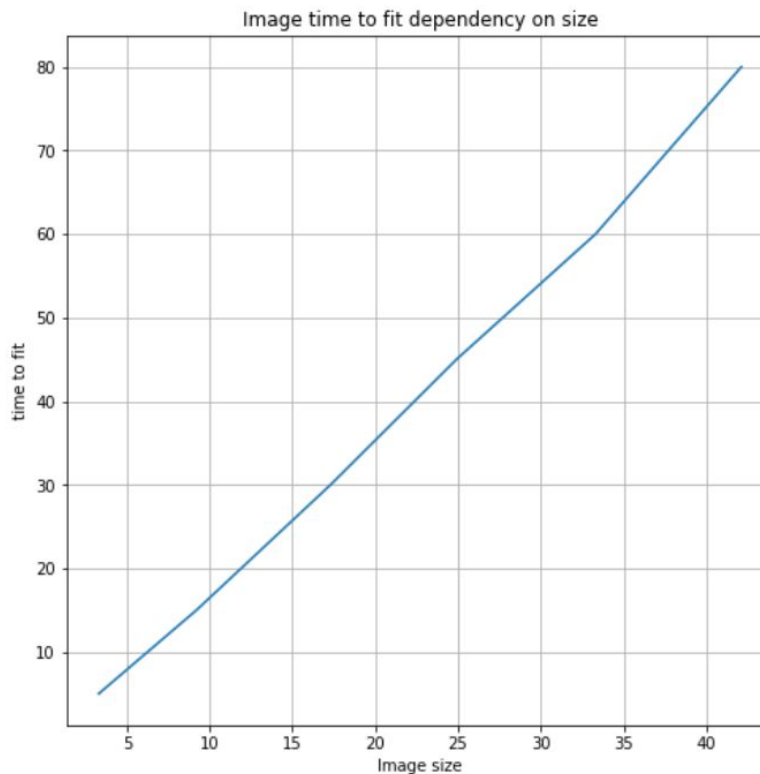
Experiment part

Augmentation vs without augmentations

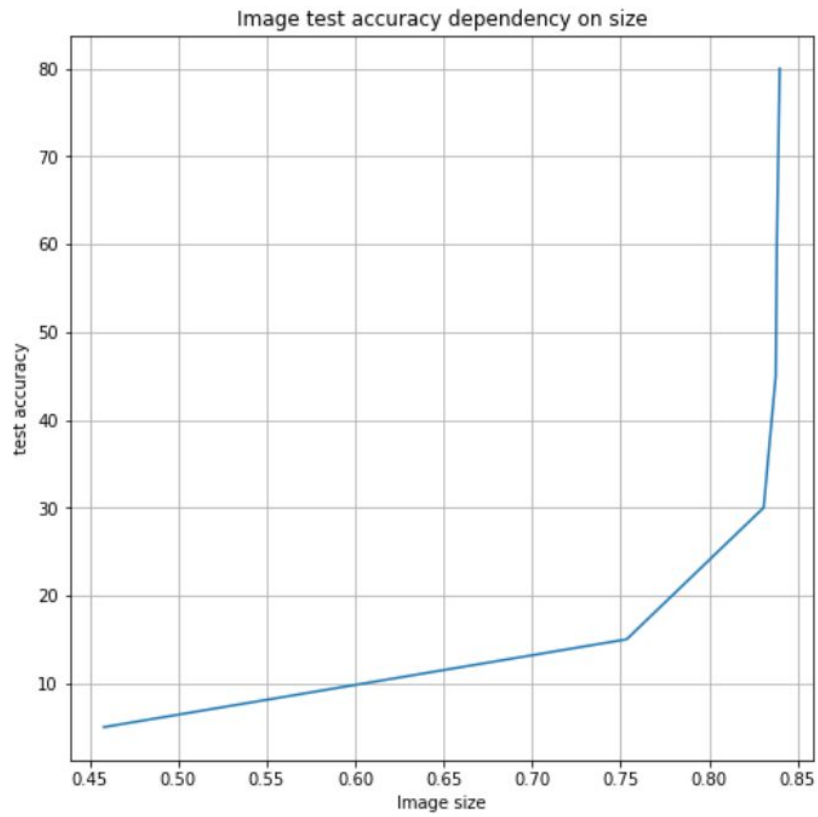
	Accuracy	Macro precision	Macro recall	Macro f1-score
With augmentations	0.83	0.79	0.81	0.79
Without augmentations	0.77	0.80	0.67	0.70

Macro recall and then accuracy in case of absence of augmentations is low because of class imbalance. RFM has no sufficient source to learn patterns

Size influence on fitting time and accuracy



Time is increasing linearly with the increase of size



Accuracy is increasing with the increase in size