



Master in Computer Vision *Barcelona*

Image Classification

Ramon Baldrich

Artificial Intelligence in Computer Vision

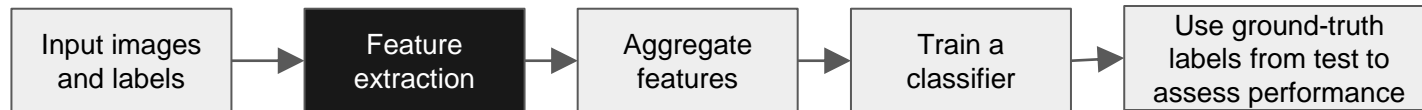
Machine learning for image classification:

Data driven methods: Deep Convolutional Networks: 3 sessions

From hand-crafted to learnt features

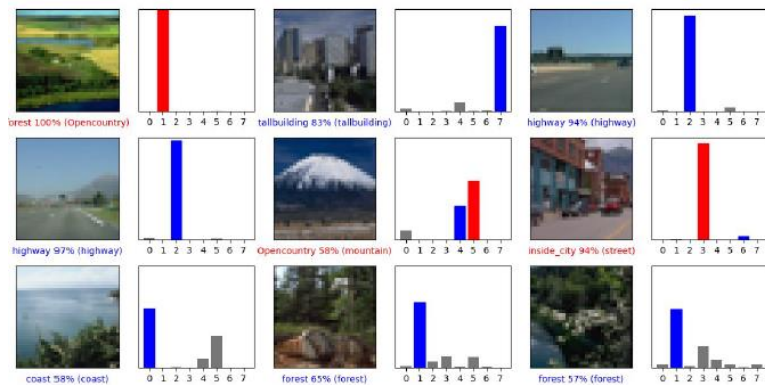
Fine tuning of pre-trained CNNs

Training a CNN from scratch

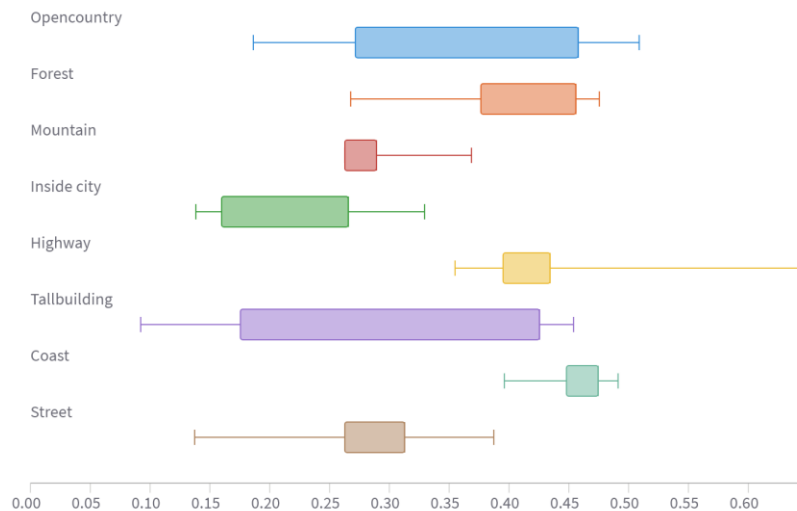


MLP Inference

Inference on test set



Class validation accuracies



Data normalization

No augmentation

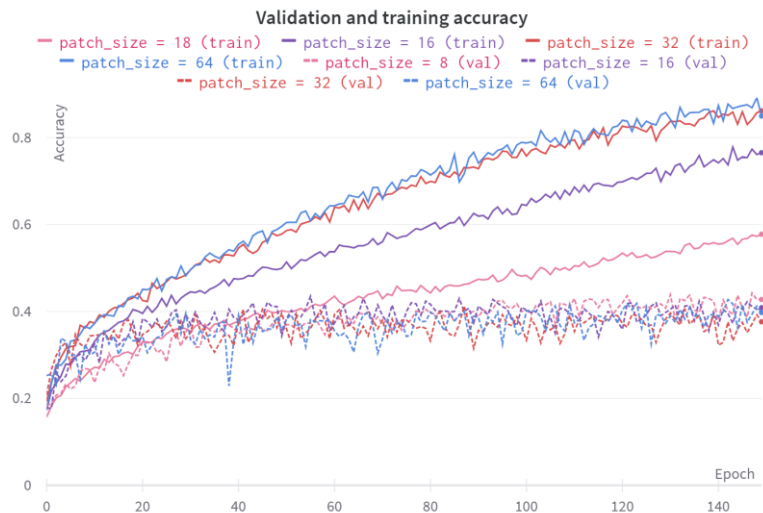
0 10 20 30 40

	train accuracy	test accuracy
x / 255	0.70	0.51
channel-wise standardization	0.996	0.598
pixel-wise standardization	0.996	0.592

Augmentation

	train accuracy	test accuracy
x / 255	0.64	0.56
channel-wise standardization	0.787	0.641
pixel-wise standardization	0.784	0.640

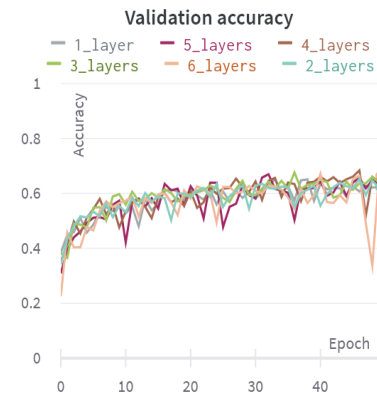
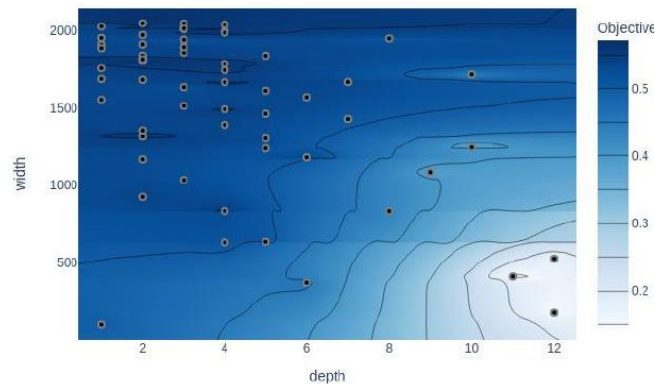
Patch size



Network architecture:

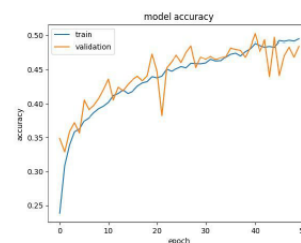
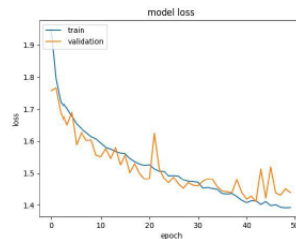
depth

	train accuracy	test accuracy
1	0.894	0.645
2	0.932	0.673
3	0.942	0.686
4	0.942	0.688
5	0.953	0.676



Layer size

	train accuracy	test accuracy
256	0.851	0.851
512	0.85	0.833
1024	0.866	0.844
2048	0.894	0.845
4096	0.901	0.846
8192	0.914	0.854



dense_2	input:	(None, 2048)
Dense	output:	(None, 8)

MLP	16 patches MLP
0.6514	0.7001

MLP+SVM

	test accuracy
end-to-end	0.666
SVM (Dense 1)	0.358
SVM (Dense 2)	0.397
SVM (Dense 3)	0.437

Dense MLP agregation

Patch Size	train accuracy	test accuracy
4x4	0.782	0.751
8x8	0.772	0.760
16x16	0.89	0.701
32x32	0.924	0.645

Overlaped? Dense MLP

Num. patches (Patch size=64)	Accuracy
10	74.6
20	76.1
40	78.6
80	78.4
160	79.8

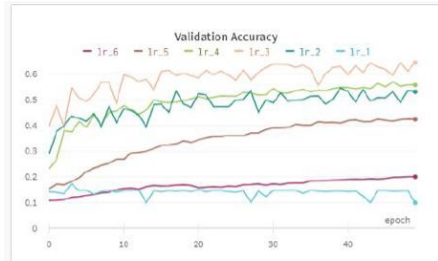
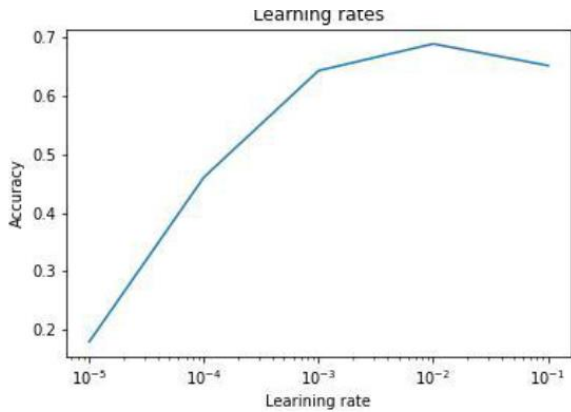
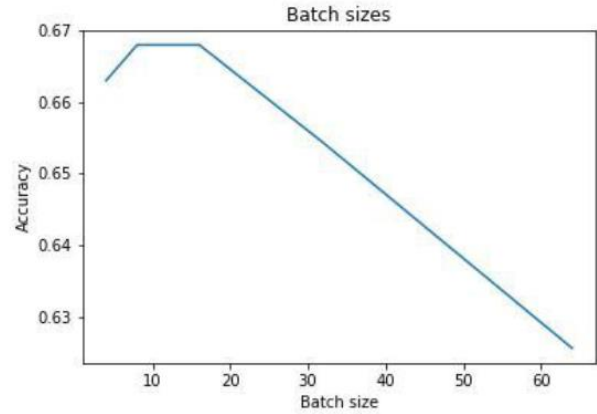
Bow

codebook size	test accuracy
32	0.459
64	0.522
128	0.540
256	0.602
512	0.589
1024	0.574

Test accuracy

	MLP1	MLP2	MLP3
MLP	0.5850	0.6625	0.6525
MLP+ BoW + SVM	0.6245	0.6914	0.6765

Method	Test accuracy
BoVW (SVM classifier)	0.721
MLP end-to-end classifier	0.623
MLP hidden layer as SVM input	0.625
MLP patches (mean)	0.381



Learning Rate

lr_1 = 0.1
 lr_2 = 0.01
 lr_3 = 0.001
 lr_4 = 0.0001
 lr_5 = 0.00001
 lr_6 = 0.000001

Group	grade
1	10
2	9
3	10
4	9
5	10
6	9
7	9
8	8
9	10
10	0. Report?