DTU Compute
Department of Applied Mathematics and Computer Science

Methods for image classification

Tim F. Olsen, Marcus K. Nielsen & Anders L. Bæk.
Case 2 in 02582 Computational Data Analysis

Figure 1: Examples of the data we are working on. The left image is classified as clear and the right is classified as foggy.

Introduction

DMI have asked the participants of this course to attempt to classify a number of images of foggy and clear roads and scenes. Our take on the problem is to test two different methods for classifying the images: A manual extraction of features and a convolutions deep artificial neural network (ANN).
The data sets we have available contains both clear and foggy images, but it is very skewed with about 1.500 foggy images and 150.000 clear images. Additionally there exists duplicates and the images are not independent since they are taken within a short time frame.

Feature extraction approach

The main goal of the hand crafted feature extractions is to see how well they are to generalize across locations. The model has been trained on images from Skive and the test of generalization are measured on a subset of the Billund images.

Features: The extracted features were; the mean value of the Dark channel, the variance and the squared sum of the Sobel filter values, the absolute sum and variance of the Laplace filter and the pct. of overexposed pixels[1]. All features have been adjusted to the size of the image to create a standard of reference.
Figure 2 illustrates the properties of the features by the first and second PCA.

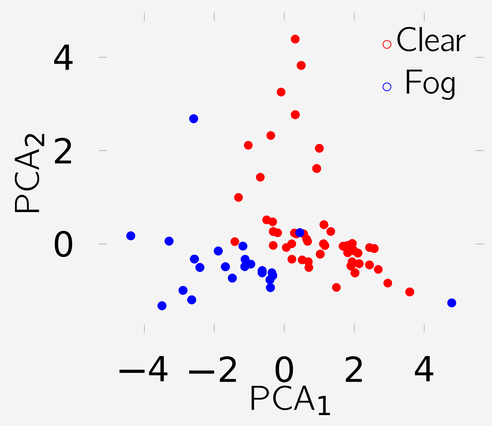

Figure 2: The first and the second PCA represents 87% of the total variation.

The model: A Random Forrest Classifier (RFC) has been trained on images from Skive. The selection of the hyper-parameters was by a randomized search-grid followed by a thorough search-grid both of which was done by 5-fold CV.

		Pred.		Class metrics			
		Fog	Clear	Pre.	Sens.	F_1	Acc.
Actual	Fog	97	3	93	97	95	91
	Clear	0	100	98	100	99	99

Table 1: Performance metrics on the independent test partition. All values are in percentage.

Deep learning approach

Binary image classification is particularly well suited for *Deep Convolutional Artificial Neural Networks (ConvNets)*.

Data preparation: The general concern with this specific data set, was that there was around 100 times more clear images than foggy images. A random subset of the clear images was sampled, equal in size to the foggy.

Network architecture: Considerations were the limited amount of data. Making a too large network prone to over-fitting.

Figures/conv_net_fig-eps-converted-to.pdf

Figure 3: Architecture of the convolutional neural network.

The loss function: we are using in the *ConvNet* is *cross-entropy*.

$$C(p,q) = -\sum_i p_i \log(q_i) \qquad (1)$$

$$\Longleftrightarrow$$

$$-y\log(\hat{y}) - (1-y)\log(1-\hat{y}) \qquad (2)$$

Training: was done using a *stochastic gradient descent* method known as the *Adam-optimiser*.

Results

Feature extraction approach: Table 2 reports the classification metrics from the Billund images.

		Pred.		Class metrics			
		Fog	Clear	Pre.	Sens.	F_1	Acc.
Actual	Fog	97	3	93	97	95	91
	Clear	80	20	35	20	26	91

Table 2: Performance metrics of predicting the appearance of fog in an unseen location. All values are in percentage.

The RFC does a poor job of detecting the appearance of clear weather in the unseen location. Therefore the generalization of the features is poor and location specific.

Deep learning approach:

		Pred.		Class metrics			
		Fog	Clear	Pre.	Sens.	F_1	Acc.
Actual	Fog	87	13	91	87	89	89
	Clear	9	91	88	91	89	89

Table 3: Performance metrics of predicting the appearance of fog. All values are in percentage.

Conclusion

The handcrafted features are shown to be location specific and does not generalize to unseen locations. The RFC has an average accuracy of 98.5% on the test set from Skive and an average accuracy of 58.5% on unseen location from Billund.

The ConvNet proved to be quite effective, at an accuracy of 89.0%. It should be mentioned that given the sampling, Billund will very likely be over represented in the training, validation and test data sets.

So In all it seem to be the ANN which is performing best on the data but in order to be certain a bigger dataset and more tests are required.

[1] The selected features are inspired by the presented features in "European Study Group with Industry 121" by M. Lyksborg et. al.