1. A total of 6000 images (i.e., 3000 sunny and 3000 overcast) were sampled from the entire set of data for training. Additionally, 3000 more images (i.e., 1500 sunny and 1500 overcast) were sampled for validation purposes

2. The images were of high dimensionality; therefore, they were reduced to 128x128 pixels. Two feature extraction methods were used: histogram of gray-scale image and Histogram of Oriented Gradients.

3. Feature Extraction:

   a. Histogram of Grayscale Image – this method was used to determine the intensity of the image, which is useful since the intensity changes based off lighting conditions effected by overcast skies or sunny weather.
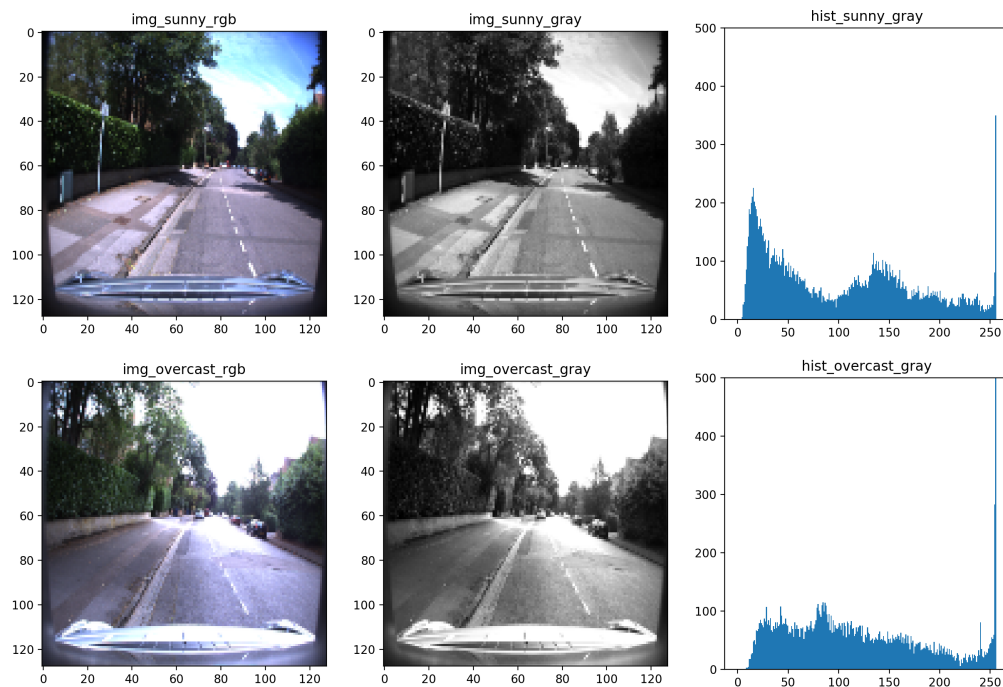
   b. Histogram of Oriented Gradients (HOG) – this was used to extract the local shape information from regions in the image, which changes with contrast differences due to weather conditions.

        i. Number of orientation bins = 6

        ii. Number of pixels per cell = [16,16]

iii. Number of cells per block = [1,1]
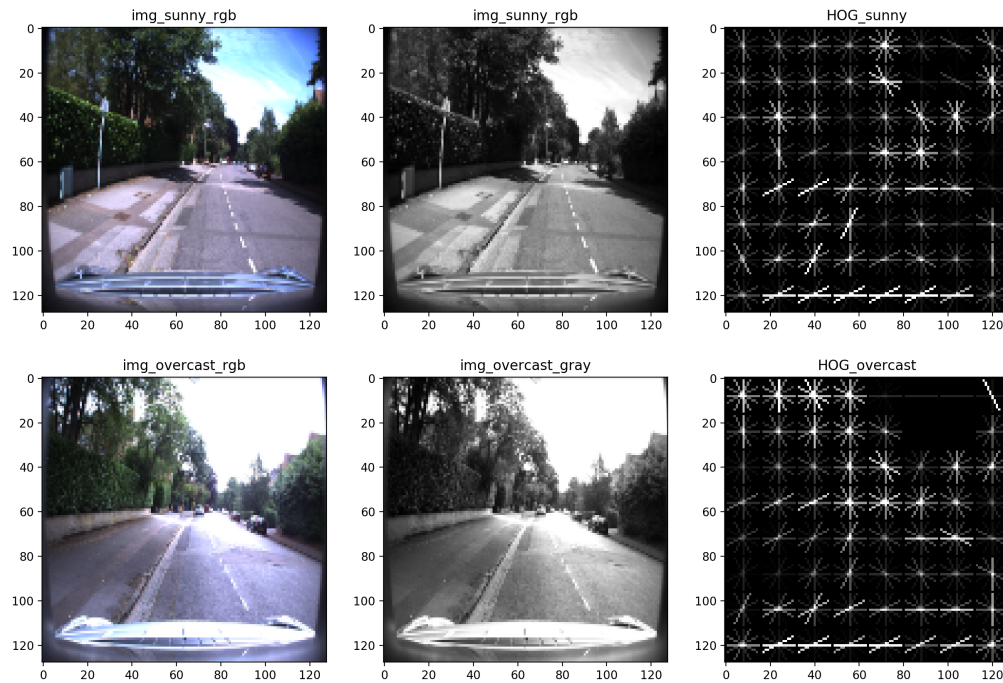
4. Bayesian Network:

    o A Naïve Bayes classifier was used to determine the weather condition of
       the image as sunny or overcast. Furthermore, the classifier was chosen
       because the input variables composed from the feature extraction
       methods were class-conditionally independent. Hence, the evidence
       generated from the HOG method doesn't provide information on the
       likelihood of evidence occurring from the greyscale histogram.

    • Confusion Matrix

		Actual Class	
		Sunny	Overcast
Predicted	Sunny	1388	112
Class	Overcast	349	1151

- Number of mislabeled points out of 3000: 461

- Naïve Bayes validation accuracy: 84.63%