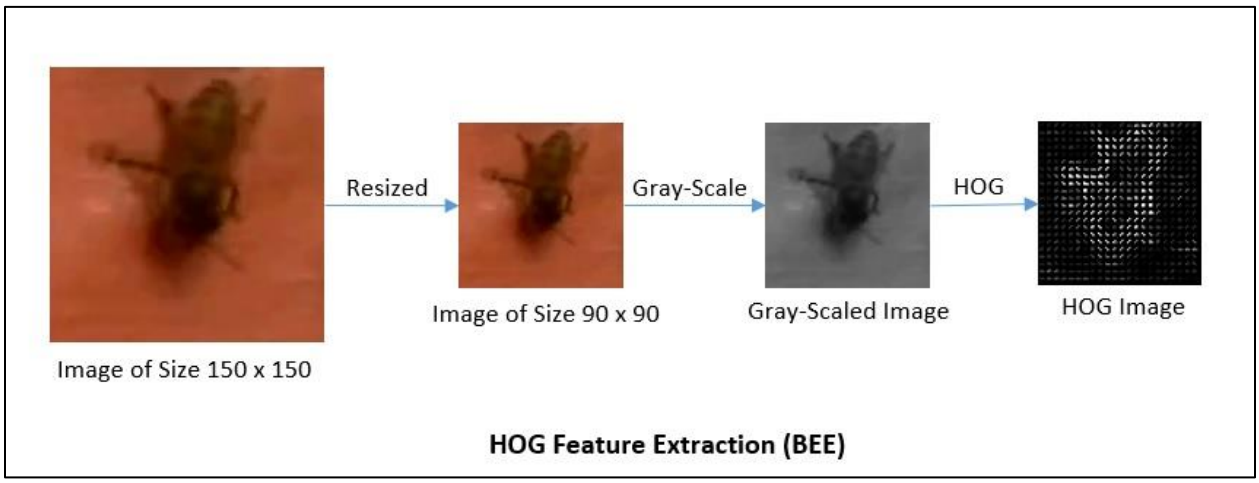
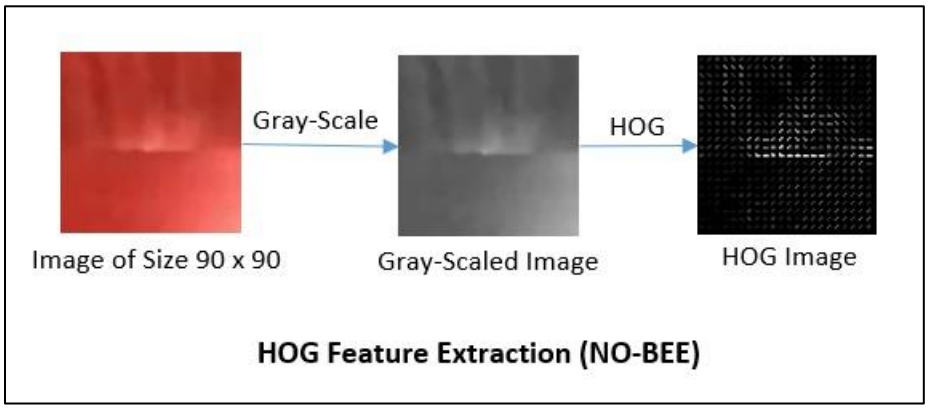
BEE Classification Using Feature Extraction and Random Forests

All the images are preprocessed through the below process and then used for the training of the classifiers which gives us the feature images for both the BEE and NO-BEE.





Built the code with all the data processing, data reshaping, random forest classifier training in the bee_rf.py file.

The function named **bee_image_processing()** function is used for processing the image into the HOG image as shown above and save the files into a separate directory to avoid image processing every time.

The function named **bee_process()** function is used for processing and labelling the data into the required form. The reshaping and the ground truth formatting is done in the **load_data_wrapper()** function.

The function named **fit_image_rf()** is designed for the testing the classifier on a particular image. It takes two parameters as the input which are the random forest classifier's pickle file and the image path which the classifier is tested against. **This function returns [0] for a bee and [1] for a no-bee.**

BEE Classification Using Feature Extraction and Random Forests

Experimental Results

I have trained the random forest classifier with number of decision trees as [5,20,40,50,60,80,100,150,200,250,300] and recorded all the classification reports and the confusion matrices. The below are the summary of my classification results.

For the BEE2 One Super Dataset:

	BEE 2 One Super					
	Validation			Testing		
	BEE Acc.	NO-BEE Acc.	Overall Acc.	BEE Acc.	NO-BEE Acc.	Overall Acc.
RF(5)	93.34	33.85	49.83	62.55	84.64	80.67
RF(20)	96.15	34.27	49.98	71.67	86.52	83.87
RF(40)	96.65	33.68	48.49	76.29	86.26	84.65
RF(50)	96.36	33.01	46.93	79.29	86.31	85.22
RF(60)	96.68	33.12	47.13	78.83	86.43	85.23
RF(80)	96.77	33.13	47.13	78.11	86.07	84.84
RF(100)	96.71	32.95	46.71	79.4	86.23	85.18
RF(150)	96.96	32.72	46.08	80.76	86.12	85.32
RF(200)	97.24	32.69	45.95	80.39	86.18	85.31
RF(250)	96.72	32.72	46.15	80.67	86.19	85.37
RF(300)	96.94	32.68	45.97	82.04	86.33	85.7

For the BEE2 Two Super Dataset:

	BEE 2 Two Super					
	Validation			Testing		
	BEE Acc.	NO-BEE Acc.	Overall Acc.	BEE Acc.	NO-BEE Acc.	Overall Acc.
RF(5)	50.33	61.32	58.06	83.16	84.9	84.16
RF(20)	53.85	62.47	60.04	87.72	89.38	88.66
RF(40)	55.56	61.92	60.44	90.41	89.44	89.84
RF(50)	56.3	61.86	60.63	90.97	89.26	89.97
RF(60)	57.43	62.2	61.15	90.98	89.64	90.2
RF(80)	58.09	61.97	61.18	91.48	89.67	90.42
RF(100)	59.1	62.09	61.49	91.77	89.6	90.49
RF(150)	59.53	61.96	61.5	91.99	89.5	90.53
RF(200)	60.1	62.05	61.68	92.3	89.53	90.66
RF(250)	59.81	61.88	61.5	92.3	89.73	90.79
RF(300)	59.68	61.97	61.54	92.37	89.6	90.74