**Report on Assignment 8**

Mohammed Jasam

**Summary**

In this project we have to implement different Regression models on the dataset provided. The dataset consists of training and testing images each of which has two folders the background images and the face images. We have to train different regression models using the training data supplied and have to classify the testing images into either face or background.

The different regression models to be implemented:

1. Logistic Regression
2. Bayesian Logistic Regression
3. Dual Logistic Regression
4. Dual Bayesian Logistic Regression
5. Kernel Logistic Regression
6. Relevance Vector Logistic Regression

**Algorithmic Approach**

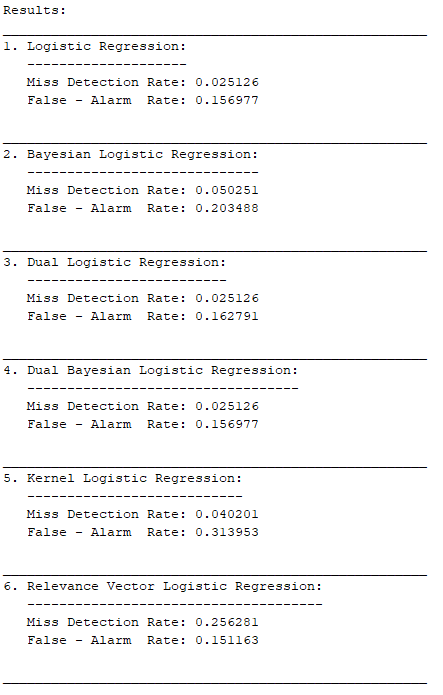
**Pre-processing:**

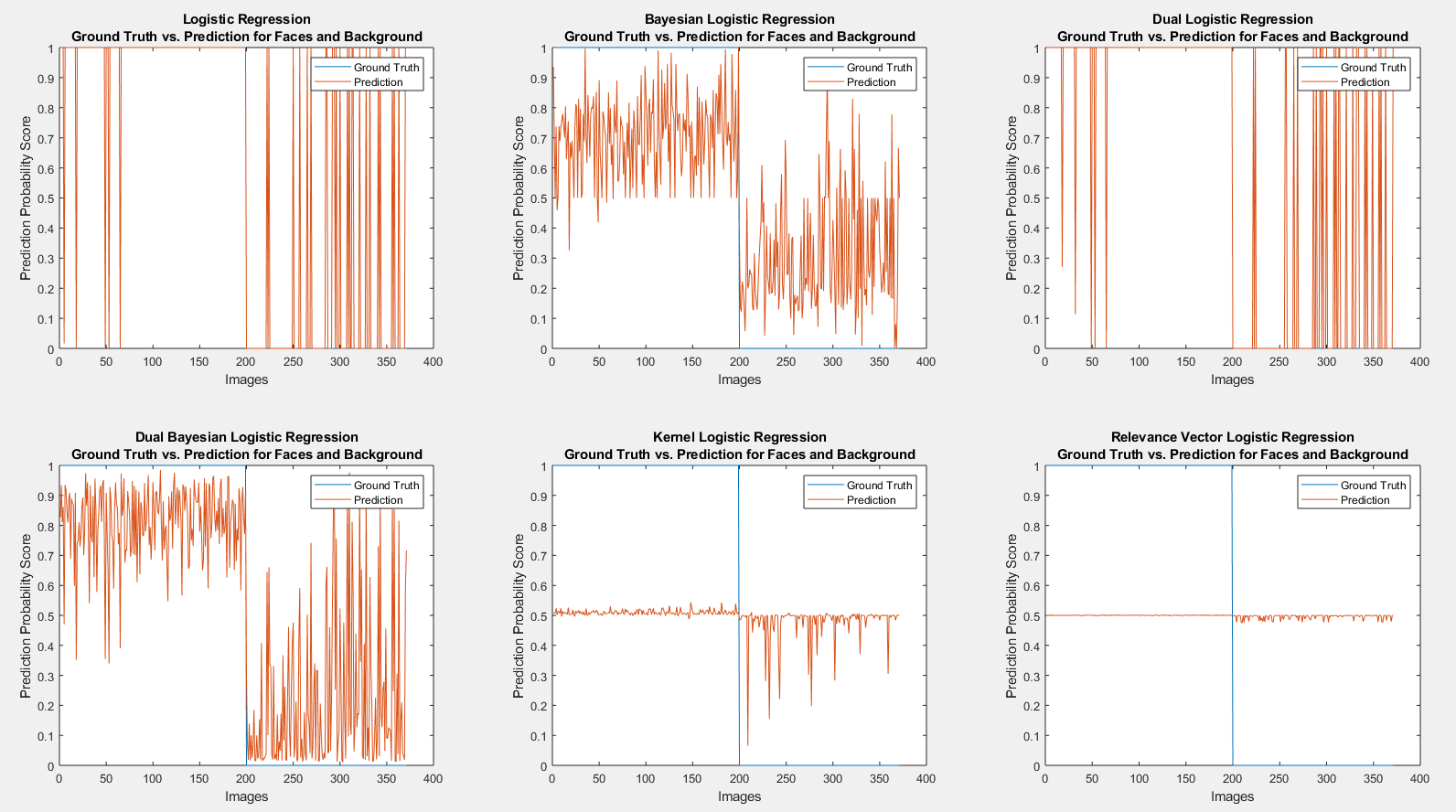
1. All the images are read into matrices, then the face and the background data are combined into a single matrix. Then the WTrain and WTest matrices are generated, these are initialized values (0 or 1) in the order in which the above face and background matrices are combined.
2. Different methods and their corresponding cost functions are provided so we have to transform the data to fit the requirements.
3. Different parameters like VarPrior, InitialPhi, InitialPsi, Lambda, Nu are set accordingly and supplied to the fit function of the corresponding method.
4. Here the fminunc function will utilize the parameters supplied and compute the gradient vector and hessian matrix using the cost function.
5. This gives the prediction values on the test images.
6. We then convert these values using a step operation where the values greater than 0.5 and equated to one and those less than 0.5 are equated to 0.
7. These values are now compared with the ground truth WTest.
8. During this process the Miss Detection Rate and the False Alarm Rate are calculated for Face and Background respectively.

**Results**

The models were run in different colorspaces:

**RGB**

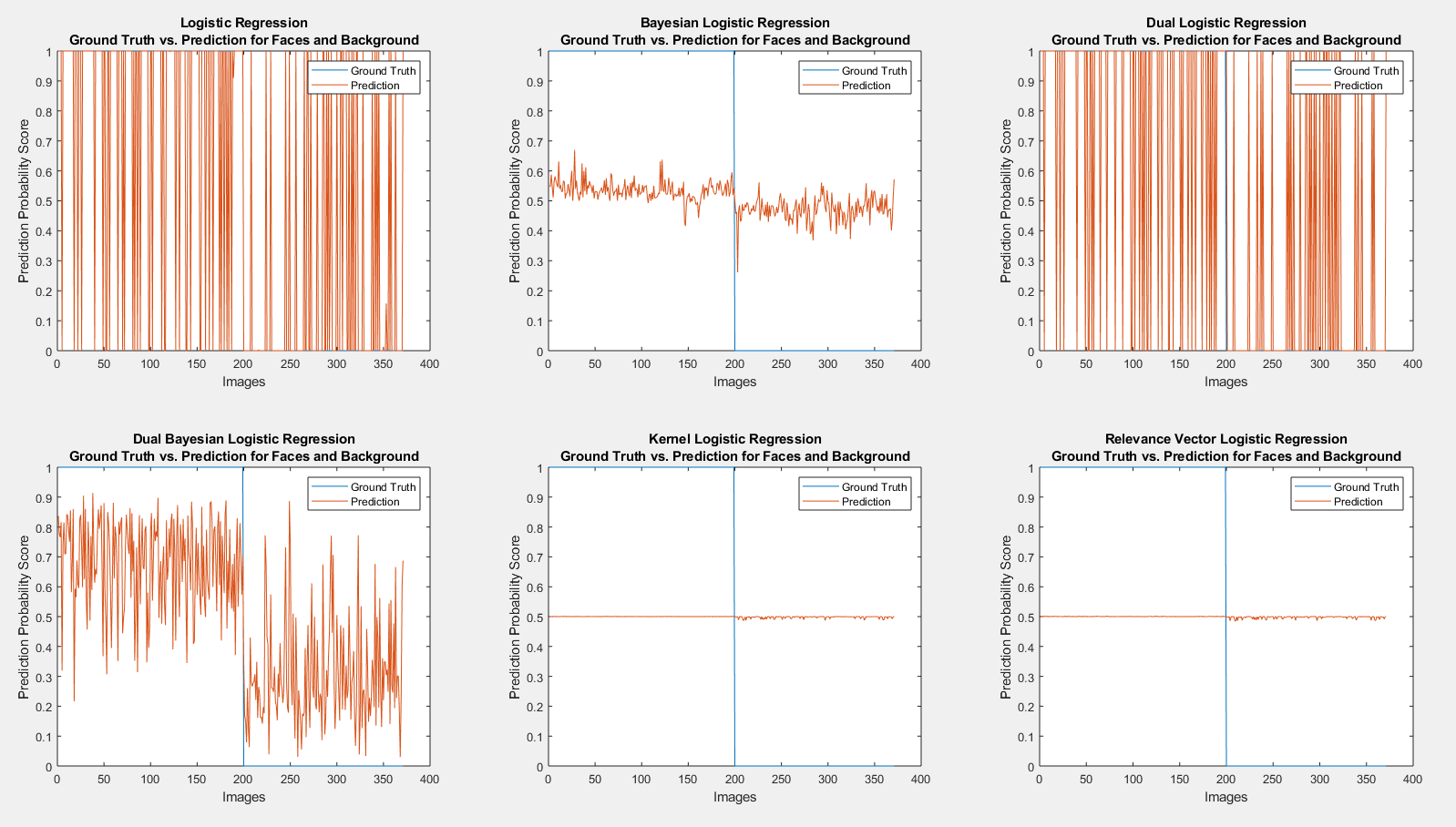
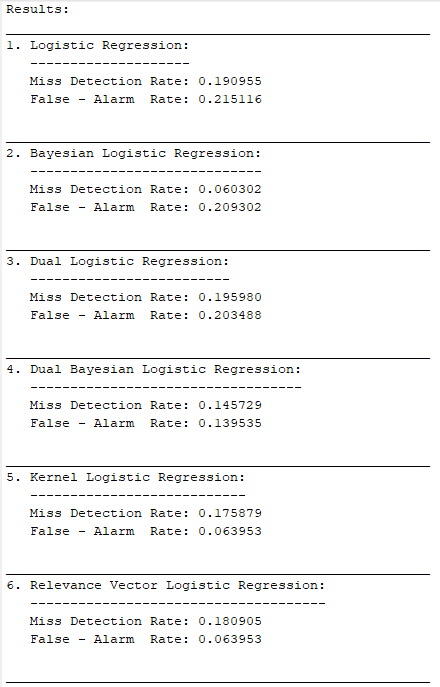




***Observation:***

We can observe that the Logistic Regression method classifies the face data better than the background data. The Bayesian does a good job but has less accuracy compared to the previous method. The Dual LR Works similar to the LR method and there is an increase in accuracy in the Dual BLR. Finally, the Kernel and the RVLR perform way worse than its predecessors. All their face and background predictions lie close to 0.5 with high degree of ambiguity

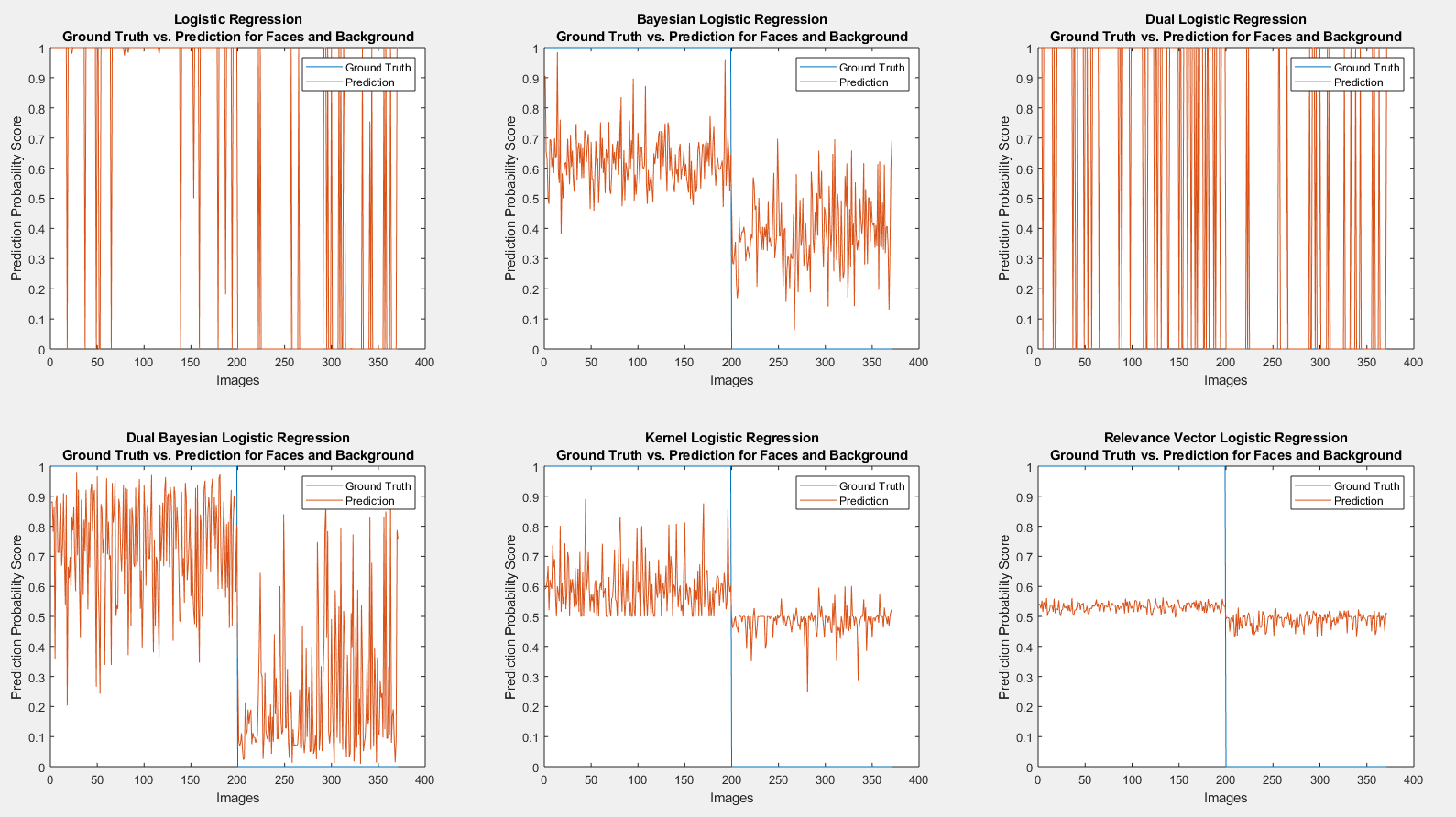
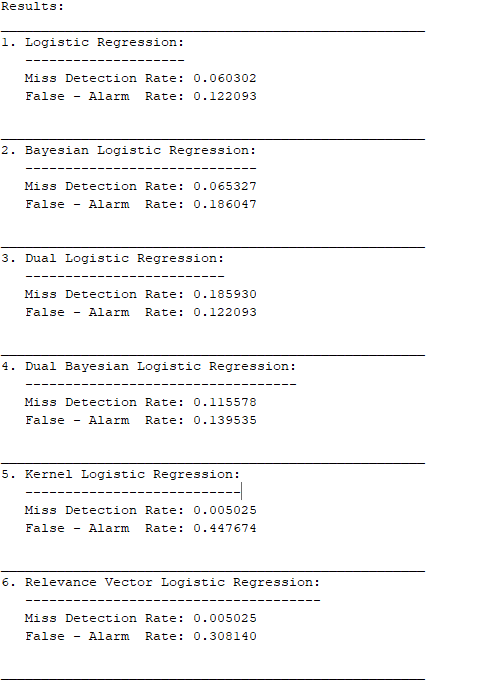
**HSV**



***Observations:***

There is a high degree of ambiguity in the logistic regression model, we can see that the model classifies the data as both face and background in both datasets. The Bayesian LR method performs better, we can see that it classifies max face data as face and max background data as background. The dual LR performs similar to the LR method classifying the data as both in face and background sets. The Dual BLR works similar to the BLR but with variable prediction accuracy. Finally, the Kernel and RLVR show negligent change in the face data and very small change in the background data. These are by far the worst among the methods again.

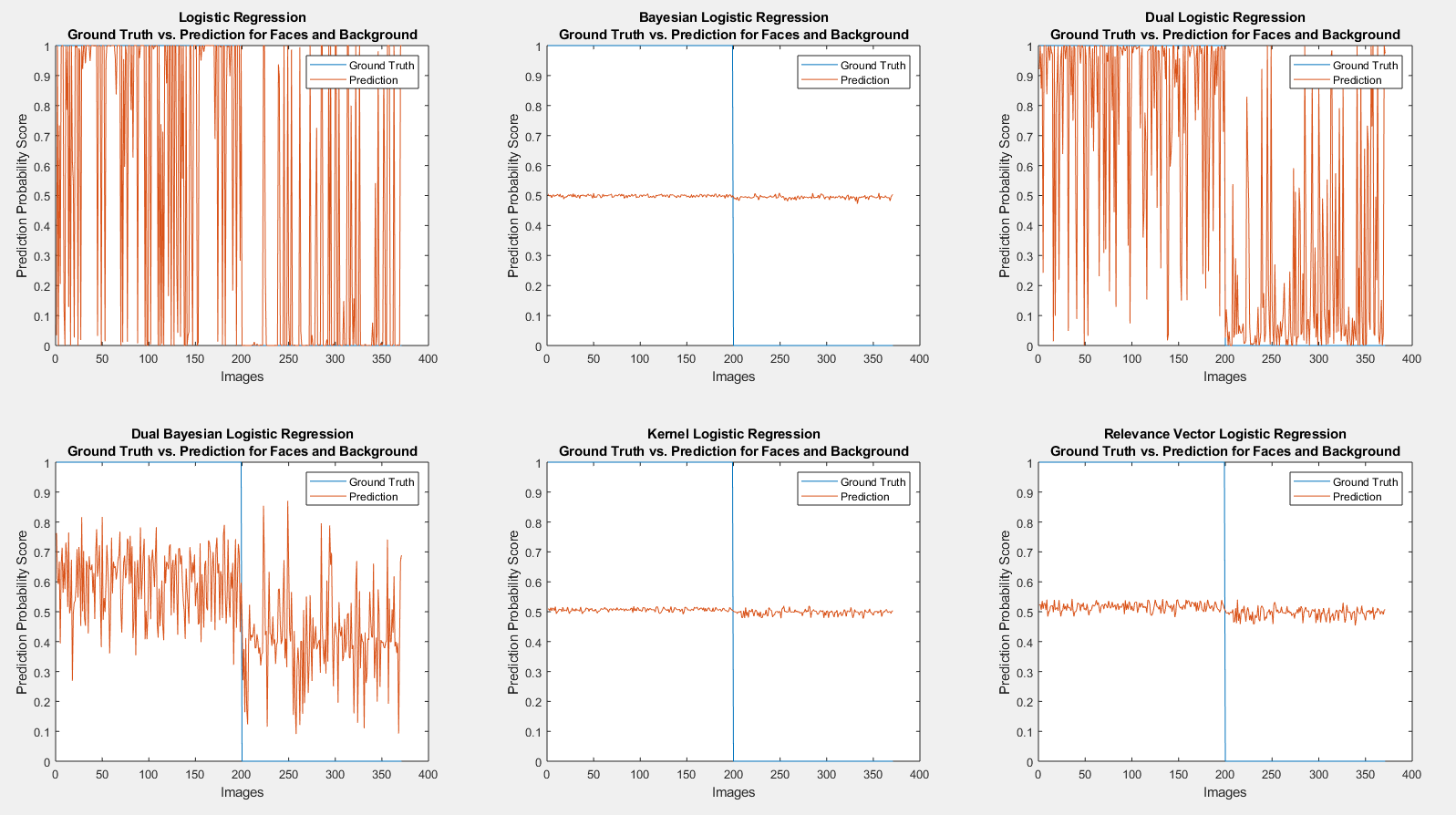
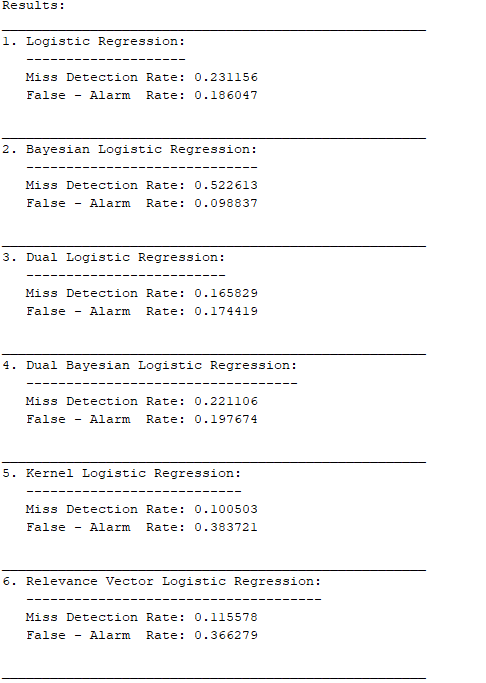
**YCbCr**



***Observations:***

In this colorspaces the first four methods show high degree of ambiguity in the prediction result classifying the data as face and background in both the datasets. While the last two methods, the kernel and RVLR work efficiently classifying the data as face and background in both the datasets.

**Gray**

****

***Observations:***

In this final colorspaces we can observe that LR and Dual LR work worst showing very high degree of ambiguity in classifying the data as face and background in both the datasets. The BLR and Kernel and RVLR do a fairly decent job at predicting the face and background in the respective datasets and the Dual BLR although predicts most of the face images correctly, it shows high degree of ambiguity in the background prediction!