**Internship Report**

**Weather classification using single image by fusion of features**

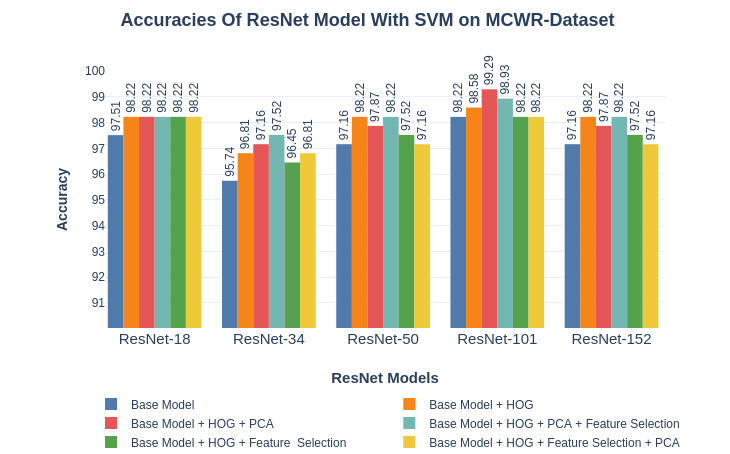
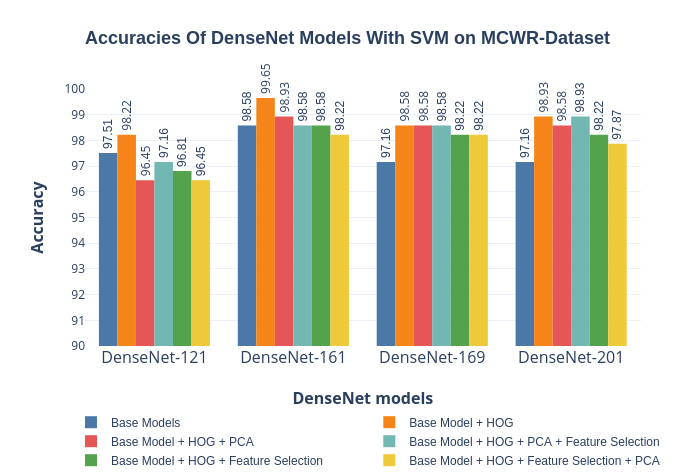
**-Venkat Siddish Gudla**

**-Valmiki Ganesh Deepak**

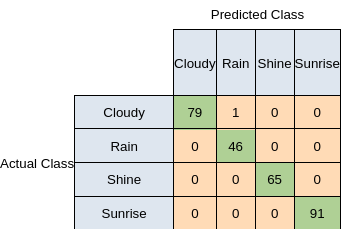
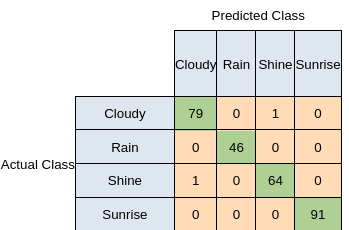
* In this internship we created a model which uses a fusion of features and subsequent feature transformation techniques to achieve a very high accuracy on the dataset. We employ the use of transfer learning to import models pretrained on the ImageNet dataset and fuse those features with features obtained through Histogram of Oriented Gradients (HOG). This proved to give higher accuracies than if the model classified using only the features obtained from the pretrained models.
* Extensive experiments were done on standard datasets like MCWRD and MWI
* In our extensive experiments, we found that the dimensions of the raw features were quite high owing to the fact that the features obtained through HOG were of the dimension 3780. This contributed to a high dimension feature space which slows the execution of the classification model.
* Over the course of our research, we found out that the best results were obtained when we fused the features obtained through a Deep learning method and a non-DL method. The non-DL features obtained through HOG were able to enrich the information of the feature space which previously contained the feature vectors from the DL models.
* We used different approaches to reduce dimensionality and apply feature selection. First was to apply PCA to the individual feature vectors from DL models and HOG and thereby concatenating them to obtain the final features.
* The second approach was to apply feature selection to the individual features after applying PCA and then concatenating them.
* A different approach was to first apply feature selection and concatenating them to obtain the final features. Also, a variation of this would be to apply PCA after feature selection and finally concatenating them to obtain the fourth feature space.
* Various model evaluation metrics were calculated and plotted for ensuring model performance.

**Plots and Results:**

* Accuracies of two best models with both the datasets: -



* Confusion matrix analysis for the two best models: -



* Precision-recall curve and ROC curve for evaluating model confidence: -
* Area Under the Graph values of specifying how well the best model is classifying different classes.

|  |  |
| --- | --- |
| One vs Rest | ROC-AUC Value |
| Cloudy | 0.9996 |
| Rain | 1 |
| Shine | 0.9999 |
| Sunrise | 1 |
| Average AUC-ROC for one vs rest | 0.9999 |

* Additional evaluation metrics like precision-recall curves, roc curves and histograms were plotted to test model performance. These metrics are calculated for the two best models as defining earlier.

