Multi-class Weather Classification using Single image via feature fusion and selection

Ganesh Deepak[1], Venkat Siddish Gudla[2]
And Faheema AGJ[3]
[1] Indian Institute of Technology Kharagpur, India
v.ganeshdeepak@iitkgp.ac.in, durgajp94@iitkgp.ac.in
[2]Centre for AI & Robotics, Bangalore, India
faheema.cair@gov.in

Abstract. Weather classification using multiple classes is a most sought technique which has many potential applications. Due to diverse nature of weather, it is extremely difficult to get discriminative features from weather images. In this paper we have tried to capture the discriminative feature by using feature fusion and feature selection. The proposed method uses combination of Histogram of Gradient (HOG) [1] & deep features, feature selection, feature reduction and classification to classify the weather images. Extensive experiments on the benchmark datasets were carried out using various features extraction and selection/reduction methods in conjunction with various classifiers. The extensive experimental evaluation demonstrates fusion of Histogram of Gradient (HOG) & DenseNet-161 features with linear SVM classifier achieves the best classification accuracy of 99.65% & 95.2% for MCWRD and MWI dataset respectively. The proposed method uses fusion of Histogram of Gradient (HOG) & DenseNet-161 features and linear SVM classifier to achieve the classification accuracy of 99.65% and 95.2% for MCWRD and MWI dataset respectively.

Keywords: ResNet, DenseNet, Vision Transformer (ViT), Histogram of Gradient (HOG), SVM, PCA, Mutual Information

1 Introduction

However, especially for fast-moving objects such as cars, or vulnerable participants such as pedestrians, real-time perception is crucial to avoid accidents. Multi-class weather classification systems plays critical & an indispensable role in supporting the decisions of self driving cars, especially in severe and adverse circumstances. Development of technologies related to self driving cars is very active research area. A self driving car requires development of robust and accurate localization, environment perception and behavioral planning. Environment perception is one of the important technologies which sense the environment around the vehicle to generate the obstacle map for vehicle navigation. Environment perception enables the vehicle to discern stationary and moving objects. In order to develop a robust perception models for adverse weather conditions, multimodal sensors are employed to capture complementary features. Existing literature in the field of computer vision is mostly based on the assumption that the weather condition in outdoor images or videos is clear. However, in reality different weather conditions such as rain, snow or haze, fog will decrease the quality of images or videos, which will result in degraded performance. In order to develop accurate and robust environment perception algorithm, it is essential to sense the weather conditions and accordingly choose the right set of sensors for obstacle perception. In order to improve the detection of obstacle in adverse weather conditions, a reliable weather detection/classification system is essential. With deep learning techniques, self driving cars can effectively identify outdoor weather conditions using single image and thus make appropriate decisions to easily adapt to new conditions and environments. Several researchers have reported good accuracy for weather classification using images. In this paper, we have proposed fusion of deep and conventional features for multi-class weather classification. Adverse weather conditions, such as heavy fog, sleeting rain, snowstorms, dusty blasts, and low light conditions, have a significant impact on the image quality, which in turn effects the obstacle detection and scene segmentations for self driving cars.

Therefore, in this paper, we make use of deep CNN based deep feature and Histogram of Oriented Gradient (HOG) features for weather detection. The proposed model tends to be accurate, sensitive and precise for efficient deployment on autonomous vehicles for making proper decisions, especially in adverse weather.
The main contribution of this paper are listed as under

1. Fusion of deep and Histogram of Gradient (HOG) features for characterization of images
2. Evaluation of statistical, entropy and transform based feature selection methods such as variance, entropy and mutual information for feature selection on fused feature vectors.
3. Weather classification using fusion, selection and reduction of feature vector with various classifiers.
4. Comprehensive experimental results and analysis to provide more insight into solution, which is supported by various metrics like classification accuracy, confusion matrix, precision, recall, F1-score and support.
5. Comparison of our results on weather classification benchmark datasets with existing related research to show the advantage of our findings

The paper is structured as follows: the next section, section 2 describes and discusses the related work. Section 3 describes proposed method and discusses in detail various feature extraction, selection and reduction techniques. Section 4 provides details about experimental environment, dataset, evaluation metrics, and discussion. Finally, Section 5 captures the conclusion drawn and future work.

2 Related Work

Weather classification is an active research area in the domain of computer vision. Despite advances in deep CNN, the existing works in this domain still suffer from challenges. Few authors [2, 3] have focused on weather recognition from vehicle camera images for driver assistance system for recognition of rainy weather. These methods focus only on fixed target scene images which are not a practical for self driving cars. The authors of [4] proposed a method to label images of the same scene with three

weather conditions including sunny, cloudy, and overcast. In [5], authors has attempted for any scenario multi-class weather classification using multiple features and multiple kernels learning. They extract multiple features and combine them into high dimensional vector and use multiple kernels learning to learn adaptive classifier.
C. Zheng [6] et al. tried weather classification by extracting more comprehensive features for sky and non-sky regions. In addition they used dictionary learning in conjunction with active learning for labeling images.

Zhang et al. tackled the weather image classification using Histogram of Oriented Gradients for their feature extraction method [7, 8]. Then combining that with a combination of dictionary learning and multiple kernel learning which attempts to reduce the feature space using feature selection methods and improves the learning capacity of models like SVM. They went on to create their own dataset, the Multi-class Weather Image (MWI) [9] containing 20K images and reported classification accuracy of 71.39% on the four classes in the dataset. Delving into the CNN approach, Al-Haija et al. proposed a deep CNN classification model using ResNet-18 architecture [10]. They employed the use of transfer learning for feature extraction and SoftMax for classification. Fully connected neural network was run through the network which outputs the numerical probabilities of each class. Their model was tested on the same dataset and achieved a classification accuracy of 98.22%. Another Deep CNN approach was proposed by Xia et al. in which they develop a model called MeteCNN [11]. They created their own dataset, which has 11 classes. The CNN was developed as an optimized version of the VGG-16 which is easy to train and occupies small memory footprint. Their network consisted of 13 convolutional layers, 6 pooling layers and a SoftMax classifier which was trained on the 6,877 images belonging to 11 classes. The accuracy achieved by the MeteCNN was 92.68%. Developing their own framework Ibrahim et al. proposed WeatherNet which consisted of parallel Deep CNN models which were used to analyses street-level images of urban scenes [12]. The WeatherNet comprised of four models which were used to detect dawn/dusk, day, night-time, glare, rain, snow, and fog, respectively. These models were NightNet which detected the differences between dawn/dusk, day and night-time, GlareNet which detected images with glare, PrecipitationNet detected clear, rainy, or snowy weather and FogNet detected the occurrence of fog. Training this parallel model on their curation of images from various sources, they achieved class accuracies ranging from 91.6-95.6%. Al-Haija et al. proposed new improved model which characterized the performance of 3 different CNNs namely: SqueezeNet, ResNet-50 and EfficientNet-b0 [13]. For evaluation they combined two popular datasets which gave them a dataset of 6 classes containing 1656 images. Using transfer learning and fine tuning of the fully connected of the three deep CNNs and SoftMax classifier was used. They reported accuracies of the 3 models ranged from 95.68-98.44% in which the ResNet-50 had the highest accuracy. We have tried to capture most of state-of-art literature which has leveraged MCWRD and MWI dataset. As per our knowledge none of the methods in the tried fusion of deep and HOG features for weather classification.

3 Proposed Method

In this paper we propose a model which uses a fusion of deep and conventional features, feature selection, feature reduction techniques in conjunction with classifier to achieve high accuracy on the benchmark datasets. The proposed method uses transfer learning using pre-trained models to extract features and fuse deep features with Histogram of Oriented Gradients (HOG) features. The proposed method is described in detail in Figure 1.

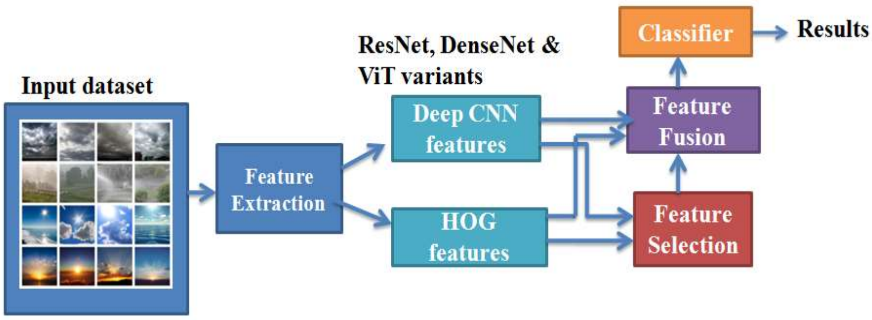

Fig. 1. Proposed Method

Fig 1. describes in detail proposed method. It consists of blocks for feature extraction, feature selection, feature fusion and classification. Input to feature extraction is dataset. Feature extraction module deal with feature extraction from images. In this case we extracted deep features from various variants ResNet, DenseNet, latest ViT and HOG features. Feature selection module select only relevant features and thereby enabling in handling curse of dimensionality. As part of this module we tried statistical, entropy, transform and manifold based feature selection methods. Finally the feature fusion model concatenates the feature vectors from both deep and HOG to generate fused feature vector. The feature selection module takes the input feature vector applies relevant feature selection method to generate low dimension feature vector. The classification module uses various classifiers ranging from; K-Nearest Neighbor (KNN), Linear SVM, Decision trees (DT), Random Forest (RF) and Gaussian Naïve Bayes (GNB). We carried out exhaustive experiments to evaluate various feature vectors, reduced feature vectors, fused feature vector and reduced fused feature vectors in conjunction with various classifiers. It was found out fused feature vector of DenseNet-161 and HOG in conjunction with linear SVM gave the best classification accuracy for MCWRD [14] and MWI [9] dataset respectively.

3.1 Feature extraction

Accurate weather classification system is hugely dependent on strong environmental features. In the following section various feature extraction methods are described in detail.

Histogram of Oriented Gradients (HOG).
The HOG [1] descriptor focuses on the structure or the shape of an object. In case of edge features, we only identify if a pixel is an edge or not. HOG is able to provide the edge direction as well. This is done by extracting gradients and orientation of the edges. To obtain features using HOG, the images are first resized into the shape of where the ratio of height is to width is 1:2. Before extracting features, all the images undergo preprocessing process. The dimensionality of HOG feature vector is 3780. One draw-back with HOG features is that it is very sensitive to image rotation. To improve the results, deep learning methods are used. The introduction of very deep CNN features helped the models to achieve state-of-art results on the tasks like image recognition and image classification.

Deep Features.
Transfer learning is leveraged to extract deep features from pre-trained model. We have only the chosen top performing deep architectures which include DenseNet, ResNet and ViT. For non-deep learning we have only used HOG features due to its good per-formance on weather images.

Residual Networks (ResNet).

In order to solve the problem of vanishing gradient, this architecture introduced the concept called residual blocks. Skip connections a technique was introduced in this method. The skip connection connects activations of a layer to further layers by skip-ping some layers in between. This forms a residual block. ResNets are made by stacking these residual blocks together. The approach behind this network is instead of layers learning the underlying mapping, we allow the network to fit the residual mapping.
So, instead of say H(x), initial mapping, let the network fit,

$$F(x) := H(x) - x \text{ which gives } H(x) := F(x) + x. \tag{1}$$

To extract features from ResNet models, we imported the pretrained ResNet models from Pytorch. After importing the models, we replaced *model.fc* with *nn.Identity*. Iden-tity() will just return the input without any clone usage or manipulation of the input and since the features are its input, the output of the entire model will be the features. We have leveraged pre-trained models of various variants of ResNet model such as ResNet-18, ResNet-34, ResNet-50, ResNet-101 and ResNet-152 for extracting deep features for our experiments.

Dense Networks (DenseNets).

Each layer in DenseNet receives inputs from all preceding levels, implying that each layer has access to the cumulative knowledge of all preceding layers. They solve the vanishing-gradient problem, improve feature propagation, promote feature reuse, and cut the number of parameters. DenseNet has better accuracies over SOTA methods and requires less memory and computation to attain good performance. To extract features from DenseNet models, we imported the four pre-trained DenseNet models from Pytorch. After importing the models, we replaced *model.classifier* with *nn.Identity*. We have leveraged DenseNet-121, DenseNet-161, DenseNet-169 and DenseNet-201 for extracting deep features

Vision Image Transformers (ViT).

The Vision Transformer, or ViT, is a model for image classification that employs a Transformer-like architecture over patches of the image. An image is split into fixed-size patches, each of them are then linearly embedded, position embeddings are added, and the resulting sequence of vectors is fed to a standard Transformer encoder. In order to perform classification, the standard approach of adding an extra learnable “classification token” to the sequence is used.
Again, using the similar approach, we extracted the feature vectors from ViT models. The pretrained ViT models which were leveraged are: Vit_b_16, Vit_b_32, Vit_l_16 and Vit_l_32.

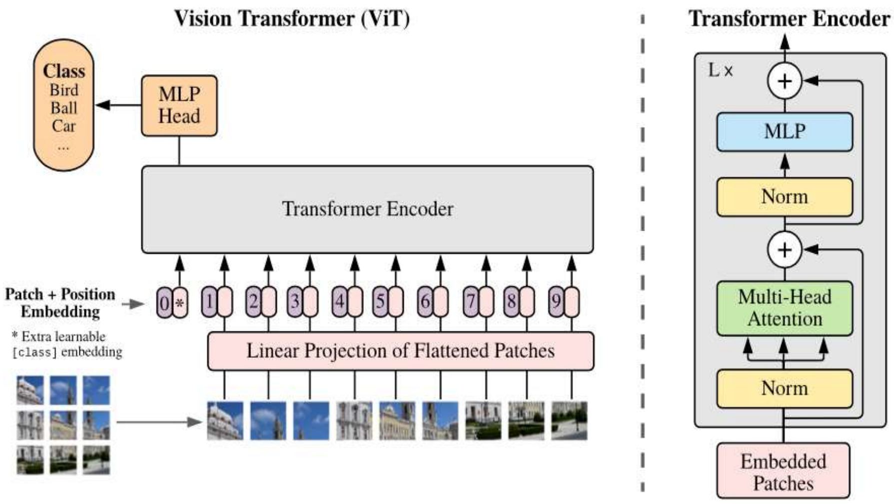

Fig. 2. Overview of the ViT architecture [15]

3.2 Feature Selection

Selection of important features is extremely crucial for weather recognition. Feature selection aims at choosing a subset of relevant features for effective classification of data. In high dimensional data classification, the performance of a classifier often depends on the feature subset used for classification. We evaluated statistical, entropy based and transform based method to reduce the dimensionality.

Statistical & Entropy based methods

As part of statistical based feature selection variance was computed for each feature element. The features whose variance is above some threshold were retained and remaining features were dropped to reduce the dimensionality of feature vector. This selection method was tried for both dataset for different combination of feature vectors. Variance feature selection did not give not much improvement in the accuracy. In entropy based methods, both entropy and mutual information were used to select the relevant features. Entropy based feature selection did help in selection of discriminative features as it did not improve the accuracy; hence it was dropped from further experiments. Mutual information based features selection was employed to select the relevant features. It is based on feature–feature mutual information to determine an optimal subset of features to minimize redundancy and to maximize relevance among features. The effectiveness of the selected feature subset is evaluated using multiple classifiers on both datasets. We call MI based feature selection as feature selection (F.S) in the remaining sections of paper.

The mutual information between two random variables X and Y can be stated formally as follows:

$$I\ (X{:}\ Y) = H(X) - H\ (X \mid Y) \tag{2}$$

Where $I\ (X{:}\ Y)$ is the mutual information for X and Y, $H(X)$ is the entropy of X and $H\ (X|Y)$ is the conditional entropy for X given Y. Mutual information is a measure of dependence between two random variables

Dimensionality reduction using transform-based methods

The curse of dimensionality basically means that the error increases with the increase in the number of features. It refers to the fact that algorithms are harder to design in high dimensions and often have a running time exponential in the dimensions. Principal Component Analysis (PCA) helps us to identify patterns in data based on the correlation between features. In a nutshell, PCA aims to find the directions of maximum variance in high-dimensional data and projects it onto a new subspace with equal or fewer dimensions than the original one. We also tried manifold based methods like UMAP and t-SNE for selection of relevant features. These methods also did not help in improving the accuracy.

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 down 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 (MI) 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 feature vector.

3.3 Classification

Feature vectors are fed to the classifier for categorizing weather images into relevant classes. Combining the features of the base DL models, base DL models and HOG along with the four aforementioned feature spaces, we trained our classifier models on these six feature spaces. We used Random forest, K-Nearest Neighbors, Decision trees, Support Vector Machines and Gaussian Naïve Bayes classifier to classify the weather dataset images. Experimental evaluation of various features and classifiers is discussed in detail in experiment section.

All the different feature spaces obtained through the mentioned methods were shuffled randomly to reduce the model bias towards a particular class. For the features obtained through fusion of two different kinds of features, shuffling was performed after the fusion of features. We split the data into training and testing set in a 75-25% train-to-test fashion for both the benchmark dataset discussed in next section. We did this so that our model could be tested first on unseen data before being deployed into real-world scenarios.

3.4 Data set details

We have utilized two benchmark datasets for evaluation of proposed method: Multi-class Weather Image (MWI) [7, 8] and Multi-class Weather Recognition Dataset (MCWRD) [14].

The Multi-class Weather Image dataset called MWI [7, 8]. MWI dataset consists of 20K images obtained from many web albums and films such as Flicker, Picasa, Moji Weather, PoCo, Fengniao. The main motivation of authors of this dataset was to generate an extensive test bed for the evaluation. We have used around 4K images of this dataset for the experiments. This dataset comprises of four categories; Sunny, rainy, snowy and Haze.

Multi-Class Weather Recognition Dataset (MCWRD) [14] comprises of multi-classification of the outdoor weather condition images into four categories: cloudy, rain, shine, and sunrise. This dataset is one of the benchmark dataset which is publicly available. It is a comprehensive dataset composed of 1125 colored images with different image dimensions and bit-depths. The dataset consists of Cloudy (300 images), Rainy (215 images), Shine (253 images), and Sunrise (356 images).

3.5 Implementation Details

The experiments were conducted using Pytorch open-source framework. The experimental platform uses High end workstation 64GB memory & 4TB hard disk with NVIDIA GeForce RTX 1080 Ti graphics card. Pre-trained deep models & vision transformers are used for feature extraction from the weather images. The algorithms are implemented using Pytorch, OpenCV and Scikit-learn libraries.

4 Experiments Results

In this section we will discuss various experiments which we carried out. We carried out extensive experiments in order to evaluate the performance of the multi-class weather classification using various feature extraction, selection and fusion in conjunction with variety of classifiers. We have also applied the feature selection method such as variance to eliminate features with various below threshold. We found that it was not improving the accuracy. Similarly we also tried using entropy for feature selection. Combining the features of the base DL models, base DL models and HOG along with the four aforementioned feature spaces, we trained our classifier models on these six feature spaces. We trained classical Machine learning models and obtained the results. In our experiments, we found out that the classifier with the highest accuracy was a Support Vector (Machines) classifier which outperformed Random Forest, K-Nearest Neighbors, Decision Tree and Naive Bayes Classifiers.

4.1 Evaluation of classifier

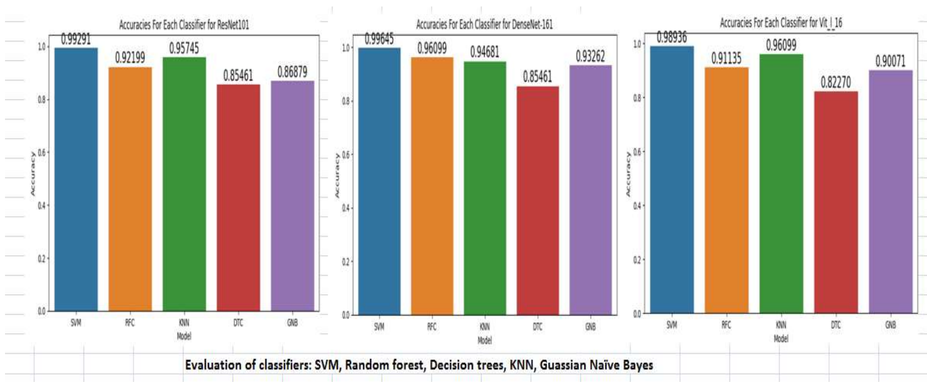

Fig. 3. Evaluation of classifiers using ResNet-101, DenseNet-161 & Vit_l_16

We carried out experiment using various classifiers with different feature extraction techniques as discussed in section 3.1. We can draw conclusion from Fig 3. that Support Vector Classifier was giving the best results as compared to Random Forest (RF), Decision trees Classifier (DTC), K-nearest neighbor (KNN) and Gaussian Naïve Bayes (GNB). In our subsequent experiments we have selected Support Vector Machine classifier.

4.2 Classification using various features

In this experiment we have used transfer learnt feature from various variant of ResNet, DenseNet and ViT. We chose only state-of-the-art network architecture like ResNet and DenseNet and ViT for features extraction using pre-trained models. We tried weather classification techniques only with various deep model features, fusion of deep model and HOG, application of PCA on fused feature vector of deep model and HOG, PCA followed by feature selection using fused feature vector of deep model and HOG, feature selection on fused deep model and HOG feature vector and lastly feature selection followed by PCA on fused feature vector of deep model and HOG. The results of our experiments are depicted by Fig. 4 and table1.

From Fig. 4 & table 1 it is clear that for MCWRD dataset DenseNet161+HOG (fv=5988) gave 99.65% accuracy, this is followed by ResNet101+HOG+PCA (fv=757) which has recorded 99.29% accuracy. Both these methods have outperformed the state-of-the-art method. Among Vision transformer variants ViT-l-32+HOG+PCA+FS (fv=875) & ViT-l-16+HOG+PCA+FS (fv=909) and Vit-b-32+HOG+PCA+FS (fv=853) gave accuracy of 98.93% & 98.58% respectively. ViT has shown promising accuracy.

For MWI dataset DenseNet-161+HOG (feature vector size=5988), DenseNet-161+ HOG+PCA+FS (feature vector size=853) and DenseNet-169+HOG+PCA (feature vector size=838) gave accuracy 95.2%, 94.44% and 94.4% respectively. Among the vision transformer variants ViT-l-32+HOG+PCA +F. S (fv=875) gave the best accuracy of 92.4% and among ResNet variants ResNet-101+HOG+PCA +FS (fv=757) gave 94.2% accuracy.

From the above analysis it is very clear that DenseNet-161+HOG model has outperformed for both the datasets. This shows that fusion of deep and HOG features captures the distinguishing features which enable to achieve higher accuracy. One more insight we can draw from above analysis is that PCA followed by feature selection on fused feature vectors of various deep model and HOG has resulted in promising accuracy. It can be noted that dimensionality of feature vector with plane fusion of deep model and HOG is very high as compared to the deep model+HOG+PCA+FS. Section 4.3 captures the details of various feature vector size for various variants of deep model &HOG along with PCA and feature selection. Our claim of using feature tapping from pre-trained model is validated by [16]. It is evident from the extensive experiments that the tapped features using various variants of pre-trained deep models in conjunction with HOG has captured robust and discriminative features which has resulted in getting good accuracy. Our detail experiments enable us to select right set of feature vector dimension based on available compute. Tradeoff between accuracy and feature vector dimensions is depicted in table 1 & table 2.

4.3 Feature vector size

Table 2 shows the number of features classifier model is trained with in different approaches as discussed above. F.S stands for Feature Selection in below table.

Resuts with MCWRD dataset

	Model	Model + HOG	Model+ HOG+PCA	Model+HOG + PCA + F.S	Model + HOG + F.S	Model + HOG + F.S + PCA
ResNet models	ResNet-18	98.22	98.22	98.22	98.22	98.22
	ResNet-34	96.81	97.16	97.52	96.45	96.81
	ResNet-50	98.22	97.87	98.22	97.52	97.16
	ResNet-101	98.58	**99.29**	98.93	98.22	98.22
	ResNet-152	98.22	97.87	98.22	97.52	97.16
DenseNet models	DenseNet-121	98.22	96.45	97.16	96.81	96.45
	DenseNet-161	**99.65**	98.93	98.58	98.58	98.22
	DenseNet-169	98.58	98.58	98.58	98.22	98.22
	DenseNet-201	98.93	98.58	98.93	98.22	97.87
ViT models	ViT_b_16	97.51	97.16	97.51	97.87	97.87
	ViT_b_32	97.87	98.22	98.58	97.87	98.22
	ViT_l_16	98.58	98.58	98.93	98.22	98.22
	ViT_l_32	98.22	98.58	98.93	98.22	98.22
	Max	99.65	99.29	98.93	98.58	98.22
	Avg	98.27769231	98.11461538	98.33153846	97.84153846	97.75846154

Resuts with MWI dataset

	Model	Model + HOG	Model + HOG + PCA	Model + HOG + PCA + F.S	Model + HOG + F.S	Model + HOG + F.S + PCA
ResNet models	ResNet-18	91.4	90	90.4	90.2	90.2
	ResNet-34	91.4	90.4	90.2	89.6	89
	ResNet-50	93.4	93.4	93	93.6	93.6
	ResNet-101	94.2	93.4	94.2	93.4	93.6
	ResNet-152	92.4	92.2	92.8	92.8	92.4
DenseNet models	DenseNet-121	92.6	91	92.6	90.8	91.2
	DenseNet-161	**95.2**	94	94.44	92	91.8
	DenseNet-169	94	94.4	93.8	92.22	92.22
	DenseNet-201	93.2	92.2	93.4	91	91.2
ViT models	ViT_b_16	91.4	90	90.4	89.6	89.6
	ViT_b_32	91.2	91.6	91.8	90.6	90.6
	ViT_l_16	90.6	91	90.6	89.6	90
	ViT_l_32	90.8	91.6	92.4	90.8	90.2
	Max	95.2	94.4	94.44	93.6	93.6
	Avg	92.44615385	91.93846154	92.31076923	91.24769231	91.20153846

Table 1. Accuracy of various variants of ResNet, DenseNet and ViT with feature fusion, PCA and F.S

Model	Model + HOG	Model + HOG + PCA	Model + HOG + PCA + F. S	Model + HOG + F. S	Model + HOG + F.S + PCA
ResNet-18	512+3780= 4292	181+556= 737	181+556=737	256+1890=2146	114+397=511
ResNet-34	512+3780= 4292	184+556=740	184+556=740	256+1890=2146	116+397=513
ResNet-50	2048+3780=5828	204+556=760	204+556=760	1024+1890=2914	153+397=550
ResNet-101	2048+3780=5828	201+556=757	201+556=757	1024+1890=2914	151+397=548
ResNet-152	2048+3780= 5828	192+556=748	192+556=748	1024+1890=2914	145+397=542
DenseNet-121	1024+3780=4804	216+556=772	216+556=772	512+1890=2402	79+397=476
DenseNet-161	2208+3780=5988	297+556=853	297+556=853	1104+1890=2994	103+397=500
DenseNet-169	1664+3780=5444	282+556=838	282+556=838	832+1890=2722	105+397=502
DenseNet-201	1920+3780=5700	298+556=854	298+556=854	960+1890=2850	103+397=500
ViT_b_16	768+3780=4548	274+556=830	274+556=830	384+1890=2274	185+397=582
ViT_b_32	768+3780=4548	273+556=820	273+556=829	384+1890=2274	185+397=582
ViT_l_16	1024+3780=4804	353+556=909	353+556=909	512+1890=2402	247+397=644
ViT_l_32	1024+3780=4804	319+556=875	319+556=875	512+1890=2402	222+397=619

Table 2. Feature vector size for various variants of ResNet, DenseNet & ViT

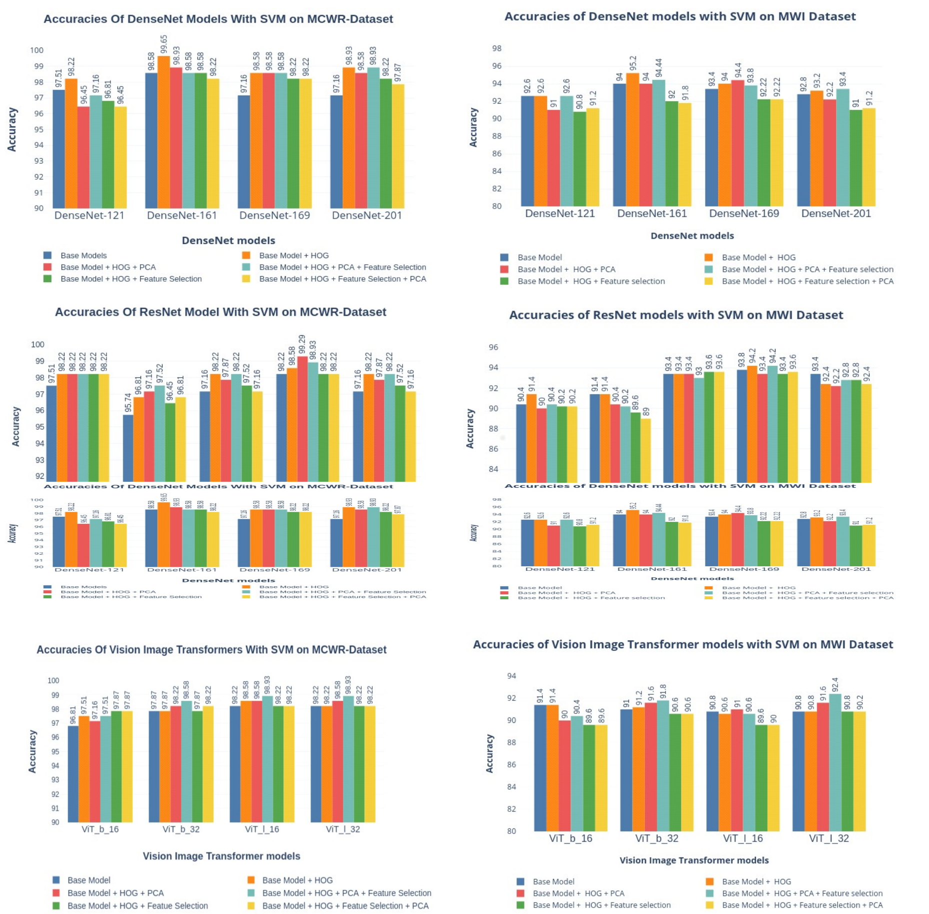

Fig. 4. Classification results on MCWRD dataset & classification results on MWI dataset

4.4 Comparison with other methods

We carried out comparison with related methods on MCWRD [14] and MWI [7,8] datasets. Table 3(a) & (b). depicts the classification accuracies reported by various papers as discussed in the related work section & our proposed method. It is clearly seen that for both the datasets our method has outperformed the other proposed methods.

MCWR dataset	
Method/Year	Accuracy
C. Zheng et. al. [10] [2016]	94.00%
W. Chu, et. Al[11][2017]	96.30%
Z. Zhu et. al.[12][2017]	95.46%
Y. Shi et. al.[13][2018]	94.71%
L. Kang et. al.[14][2018]	92.00%
O. Luwafemi et. al.[15][2019]	86.00%
M. Ibrahim et. al.[16][2019]	97.69%
Y. Wang et. al.[17][2020]	81.25%
J. Xia et. al.[18][2020]	96.03%
Qasem Abu Al-Haija[19] [2020]	98. 22%
Proposed Method	99.65%

MWI dataset	
Method/Year	Accuracy
M. Roser, F. Moos ann[2]/2008	0.2267
X. Yan, Y. Luo, X. Zheng[3]/2009	0.1889
Z. Chen, F. Yang, A. Lindner et al. [4]/2012	0.4158
Z. Zhang, H.D. Ma[5]/2015	0.7139
Proposed Method	0.952

Table 3(a) and 3(b). Comparison with other methods for both MCWRD (left) and MWI (right) datasets

4.5 Evaluation Metrics

The well-known metrics for evaluating classification algorithms comprises of computation of confusion matrix, precision, recall, F1-score and support. Precision is the number of relevant images retrieved with respect to total number of retrieved images & Recall is number of relevant images retrieved with respect to total relevant images.

		Predicted Class			
		Cloudy	Rain	Shine	Sunrise
Actual Class	Cloudy	79	1	0	0
	Rain	0	46	0	0
	Shine	0	0	65	0
	Sunrise	0	0	0	91

Classes	Precision	Recall	F1-score	Support
Cloudy	1	0.99	0.99	80
Rain	0.98	1	0.99	46
Shine	1	1	1	65
Sunrise	1	1	1	91
Average	0.995	0.9975	0.995	282

		Predicted Class			
		Cloudy	Rain	Shine	Sunrise
Actual Class	Cloudy	79	0	1	0
	Rain	0	46	0	0
	Shine	1	0	64	0
	Sunrise	0	0	0	91

Classes	Precision	Recall	F1-score	Support
Cloudy	0.99	0.99	0.99	80
Rain	1	1	1	46
Shine	0.98	0.98	0.98	65
Sunrise	1	1	1	91
Average	0.9925	0.9925	0.9925	282

Fig. 5(a). Confusion matrices & classification report for DenseNet-161+HOG & ReseNet101+HOG+PCA for MCWRD dataset

Actual Class \ Predicted Class	Haze	Rainy	Sunny	Snowy
Haze	120	1	4	2
Rainy	4	110	0	0
Sunny	1	2	131	0
Snowy	5	5	0	115

Classes	Precision	Recall	F1-score	Support
Haze	0.92	0.94	0.93	127
Rainy	0.93	0.96	0.95	114
Sunny	0.97	0.98	0.97	134
Snowy	0.98	0.92	0.95	125
Average	0.95	0.95	0.95	500

Actual Class \ Predicted Class	Haze	Rainy	Sunny	Snowy
Haze	120	1	3	3
Rainy	3	108	0	3
Sunny	1	1	131	1
Snowy	5	7	0	113

Classes	Precision	Recall	F1-score	Support
Haze	0.93	0.94	0.94	127
Rainy	0.92	0.95	0.94	114
Sunny	0.98	0.98	0.98	134
Snowy	0.94	0.9	0.92	125
Average	0.942	0.942	0.942	500

Fig. 5(b). Confusion matrices and classification report DenseNet-161+HOG & DenseNe-161+HOG+PCA for MWI dataset

Fig 5(a) and (b) captured the confusion matrix and classification support report for top two performing model; 1). DenseNet+16+HOG and 2) ResNet101+HOG+PCA and Fig 5(c) and (d) has depicted the top two performing confusion matrix; 1) Dense-Net+161+HOG and 2) DenseNet-161+HOG+PCA.

It is evident from the results that the DenseNet-161 variant captures most robust, accurate and discriminative features in conjunction with HOG as compared to other deep variants. We will discuss in detail confusion matrix and classification report of DenseNet161+HOG.

Fig 5(a) demonstrates the four-class confusion matrix and classification report for top 2 best performing methods for MCWRD dataset. The confusion matrix analysis for the testing dataset for DenseNet-161+ HOG based method indicates its robustness which is visible through the large number of correctly predicted samples represented in the diagonal of the matrix as compared to only one incorrectly predicted samples represented in the upper diagonal. ResNet-101+HOG+PCA is second best performing method with 99.25% accuracy. Fig5 (b) depicts the classification summary report for DenseNet-161+ HOG based & DenseNet-161+HOG+PCA methods. It shows Precision, recall, F1-score and support for each class for MWI dataset. The classification report of DenseNet-161+HOG shows the precision for 3 classes out 4 is 1, which indicates the ability of classifier not to label a positive instance as negative. We can also notice that Recall of 3 classes out of 4 is 1, which indicates the ability of classifier to find all positive instances, which is very good. F1-score is a weighted harmonic mean of precision and recall. F1-score of 1 indicates best score whereas 0 indicates the worst score.

5 Conclusion

This study investigates the performance of various deep learning based features and Histogram of Oriented Gradient (HOG) features for the task of weather classification problem. We carried out exhaustive experiments to demonstrate that the fusion of transfer learned deep feature and HOG yields the best classification over the latest reported literature on weather classification using single image. We have evaluated various deep-learned features, HOG features, fusion of features, feature selection using various statistical, entropy and transformation based feature selection methods. We also evaluated deep and HOG features with different type of classification methods and obtained promising results from our experiments. It was evident from our study that PCA followed by feature selection on fused features vector of various deep model and HOG has resulted in promising accuracy with a very low dimension feature vector. It can clearly evident with minor reduction in accuracy fused feature, PCA and feature selection combination has given very good performance. Another insight we can draw conclusion from the above study that the deep models have outperformed the state-of-the-art Vision transformer (ViT) variants. In future, we focus on deployment and evaluation of models for weather detection on self driving cars.

References

1. N. Dalal, B.Triggs: Histogram of oriented gradients for human detection,2005 IEEE Computer Society Conference on Computer Vision and Pattern recognition (CVPR'05)
2. M. Roser, F. Moosmann, Classification of weather situations on single color images, in: Proceedings of IEEE Intelligent Vehicles Symposium, 2008, pp. 798–803.
3. X. Yan, Y. Luo, X. Zheng, Weather recognition based on images captured by vision system in vehicle, in: Advances in Neural Networks–ISNN 2009, 2009,
4. Z. Chen, F. Yang, A. Lindner, G. Barrenetxea, M. Vetterli, How is the weather: automatic inference from images, in: Proceedings of IEEE International Conference on Image Processing, 2012, pp. 1853–1856.
5. Z. Zhang, H.D. Ma, Multi-class weather classification on single images, in Proceedings of IEEE International Conference on Image Processing, 2015, pp. 4396–4400
6. Zhang, Zheng & Ma, Huadong. (2015). Multi-class weather classification on single images. 4396-4400. 10.1109/ICIP.2015.7351637.
7. Zhang, Zheng & Ma, Huadong. (2015). Multi-class weather classification on single images. 4396-4400. 10.1109/ICIP.2015.7351637.
8. Zhang, Zheng & Ma, Huadong. (2015). Multi-class weather classification on single images. 4396-4400. 10.1109/ICIP.2015.7351637.
9. [1] Zheng Zhang, Huadong Ma. Multi-class weather classification on single images. Proc. of IEEE International Conference on Image Processing, 2015. [2] Zheng Zhang, Huadong Ma, et al. Scene-free multi-class weather classification on single images. Neurocomputing, 2016, 207: 365-373.
10. Q. A. Al-Haija, M. A. Smadi and S. Zein-Sabatto, "Multi-Class Weather Classification Using ResNet-18 CNN for Autonomous IoT and CPS Applications," 2020 International Conference on Computational Science and Computational Intelligence (CSCI), 2020, pp. 1586-

1591, doi: 10.1109/CSCI51800.2020.00293.Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010).

11. Xiao, Haixia & Zhang, Feng & Shen, Zhongping & Wu, Kun & Zhang, Jinglin. (2021). Classification of Weather Phenomenon from Images by Using Deep Convolutional Neural Network. Earth and Space Science. 8. 10.1029/2020EA001604.
12. Ibrahim, Mohamed & Haworth, James & Cheng, T.. (2019). WeatherNet: Recognising Weather and Visual Conditions from Street-Level Images Using Deep Residual Learning. International Journal of Geo-Information. 8. 549. 10.3390/ijgi8120549.
13. Al-Haija, Q.A.; Gharaibeh, M.; Odeh, A. Detection in Adverse Weather Conditions for Autonomous Vehicles via Deep Learning. *AI* 2022, *3*, 303-317. https://doi.org/10.3390/ai3020019
14. Ajayi, Gbeminiyi (2018), "Multi-class Weather Dataset for Image Classification", Mendeley Data, V1, doi: 10.17632/4drtyfjtfy.1
15. Alexander Kolesnikov, Alexey Dosovitskiy, Dirk Weissenborn, Georg Heigold, Jakob Uszkoreit, Lucas Beyer, Matthias Minderer, Mostafa Dehghani, Neil Houlsby, Sylvain Gelly, Thomas Unterthiner, & Xiaohua Zhai (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale