ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.6, 3797–3802
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.58 Spec Issiue on ICAML-2020

THE INDOOR-OUTDOOR IMAGE CLASSIFICATION AND COMPARISON OF MACHINE LEARNING METHODS USING THE MPEG-7 DESCRIPTORS

A. GOEL[1] AND N. SINGHAL

ABSTRACT. This work presents the use of MPEG-7 descriptors to extract the low-level image attributes like color, texture, and edge features in the indoor-outdoor scenes. The extraction of these low-level features became highly efficient using the MPEG-7 tools. The obtained features are then supplied for the training of the various machine learning classifiers for the categorization of indoor/outdoor scenes. A comparative study is also made among different machine learning models, that depicts the accuracy levels up to 93%. The results ascertain the effectiveness of MPEG-7 descriptors for accurate results in scene classification while maintaining low computation costs and efficient time utilization.

1. INTRODUCTION

The indoor/outdoor image classification has wide application importance. Many previous studies are concerned for outdoor-indoor scene classification from photographs. We presented the use of MPEG-7 descriptors for features extraction while maintaining the good classification accuracy and less computations. We obtained low level-attributes (color, texture and edge features) from the images using the MPEG-7 descriptors and integrated them for the training of classifiers.

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 62H35, 65D18.

Key words and phrases. Indoor-outdoor classification, Machine learning, Image processing, MPEG-7 descriptors.

2. Literature and Overview

Some previous studies, one given by Payne and Singh in [1] makes the use of edge features, but the issue here arrises to describe the images where similar objects are encountered in both outdoor-indoor pictures. A study by Vailaya and Jain in [2] proposed the importance of visual features. Their work depicts the importance of color and shape attributes of pictures for the efficient output of required images. This shows that the low-level features of images can be used for the successful classification of image scenes. Our study shows the usefulness of various low-level image features like color, edges, and texture, that are extracted using MPEG-7 descriptors while maintaining high efficiency and less computational load, and also the integration of all these features proved useful to obtain high accuracy of classification models.

3. Methodology

3.1. **System Overview.** The high definition indoor and outdoor images are acquired from multiple sources. The various low-level image features like color, texture and edges are extracted from images using MPEG-7 descriptors, and the obtained features matrices are used as the input for classification models. The trained classification models are used on the test set to classify the scene of input images. The system overview is shown in Figure 1.

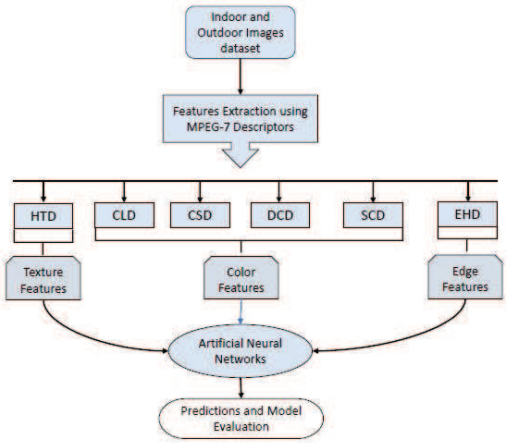

Figure 1. System Overview

3.2. **MPEG-7 Feature Descriptors.** The MPEG–7 is a low-level feature extraction tool for multimedia content developed by the Moving Picture Experts Group, [3]. The primary goal is to obtain attributes from the multimedia content and facilitate the classification process. Feature extraction is done by computing descriptors from each pixel of a target multimedia file (images in our case), and the output from a descriptor is a vector representation of an image, [4]. The feature descriptors used in our study are shown in Table 1. A full description of these descriptors is elaborated in MPEG-7 overview, [5], and a brief overview of MPEG-7 descriptors is presented below.

TABLE 1. No. of attributes associated with each descriptor

Descriptors	**CLD**	**CSD**	**EHD**	**HTD**	**SCD**	**DCD**
No of features	120	64	80	62	64	58

(a) Color Layout Descriptor (CLD): This descriptor gives the spatial distribution of the color, that is very useful for to get the color differences on indoor-outdoor scenes.

(b) Color Structure Descriptor (CSD): This descriptor captures both structured contents of the image and color. It is utilized for static image retrieval system,

$$p = \max\{0, \text{round}(0.5 \times \log_2 WH - 8)\},$$

where, H is the height of image; W is the width of image

(c) Edge Histogram Descriptor (EHD): This descriptor do the image-to-image matching by edges present in them. The edge features can significantly improve the results even in the case of natural images, where edge distribution is non-uniform.

(d) Homogenous Texture Descriptor (HTD): This descriptor is used for image-to-image matching using texture-based similarities.

(e) Scalable Color Descriptor (SCD): It is a compound descriptor that contains color space, color quantization, and histogram descriptor. It is a color histogram represented in HSV color space that is uniformly quantized into 256 bins.

3.3. **Artificial Neural Networks.** An ANN is a deeply interconnected network of numerous computing components, known as nodes. The nodes in the ANN

mimics the work of neurons in brain. The ANN consists of densely interconnected layers of nodes, and the signals are transferred through these layers. The signals are the weights corresponding to each input feature, and the model is trained by adjusting these weights, [6]. The nodes are able to store and process the input features for the experimental learning, [7]. In our study, we have used a multi-layer perceptron feed-forward artificial neural network model.

4. Experimentation and Results

4.1. **Data Set distribution.** The data set contains a total number of 326 high-resolution images- 181 and 145 for outdoor and indoor scenes, respectively. Both types of scene are composed of a wide variety of views, such as mountains, lakes, beach, bedroom, conference room, auditorium, etc. Some of these images are shown in Figure 2. The data-set has a 0.7/0.3 train-test split, with 229 training images and 97 test images.

Figure 2. Samples of Outdoor and Indoor Images from the data-set

4.2. **Implementation Framework.** Features are extracted from all images using MPEG-7 descriptors, [8], that took approximately 5 hours to complete. Machine specification- Intel i-7 processor, 8 GB main memory, and 2 GB graphics memory. Model training is done using Python 3.x based 'Orange' package, which is GUI based Python module for data analysis.

4.3. **Results and Discussion.** A comparison is made among various ML classification algorithms: artificial neural networks (ANN), K- nearest neighbour (KNN), and Support vector machines (SVM) classifiers. The MPEG-7 features are used to train these classifiers, and their corresponding accuracy are tabulated. The ANN predictions are most accurate and reliable among other models.

This model was able to classify indoor/outdoor images with an accuracy of approximately 93%- the confusion matrix is presented in Table 2.

TABLE 2. Confusion Matrix of ANN classifier

Actual/Predicted	**Indoor**	**Outdoor**	**Total**
Indoor	36	3	39
Outdoor	4	54	58
Total	40	57	97

$$\text{accuracy} = \frac{36+54}{97} = \frac{90}{97} \times 100\% = 92.8\%$$

The confusion matrix shows that only three indoor pictures and four outdoor pictures are incorrect. Our data set consists of various types of mixed photos, with this, the 93% accuracy of the classifier depicts the good information extraction by MPEG-7 descriptors and the applicability of these descriptors for indoor/outdoor scene classification. Also, the coalesce of low level features proved highly significant for the scene discrimination. We also carried out analysis and comparison using other classifiers like the KNN and SVM. However, the predictions made by these models are could not outperform the neural-network model performance. A comparison chart to show the classifiers performance is given in the Table 3.

TABLE 3. Classification results summary on different ML classifiers

Name of the Classifier	**Model Parameters**	**Pred. accur.**	**ROC Curve Area**	**Recall**	**Precision**
Support Vector Machines(SVM)	Polynomial Kernel	87.6%	0.937	0.876	0.881
K-nearest neighbours (KNN)	Number of neighbours- 5	86.6%	0.946	0.866	0.877
Artificial Neural Network (ANN)	ReLU Activation function	92.8%	0.986	0.928	0.928

5. Conclusion

This study presents a generic framework based on MPEG-7 descriptors, that confirms the importance of integration of low-level image features for the classification of indoor-outdoor scenes. This study also ascertains the substantial applicability of MPEG-7 feature descriptors and their incorporation with machine learning algorithms for the accurate image classification. The time efficient method and low computational cost while maintaining good accuracy are indeed distinctive in this work. The applicability of this work would make the way of real-world classification problems a lot easier and more approachable.

References

[1] Payne, S. Singh: *Indoor vs. outdoor scene classification in digital photographs*, Pattern Recognition, **38** (2005), 1533–1545.

[2] A. K. Jain, A. Vailaya: *Image retrieval using color and shape*, Pattern Recognition, **29** (1996), 1233–1244.

[3] B. S. Manjunath, P. Salembier, T. Sikora: *Introduction to MPEG-7*, Multimedia content description interface, First edition, John Wiley & Sons, 2002.

[4] A. Mittal, L. H. Cheong: *Addressing the problems of Bayesian network classification of video using high-dimensional features*, IEEE Transactions on knowledge and data engineering, **16**(2) (2004), 230–244.

[5] J. Martínez: *Mpeg-7 overview (version 10)*, 70th MPEG meeting Palma de Mallorca, (2004).

[6] N. Karunanithi, W. J. Grenney, D. Whitley, K. Bovee: *Neural Networks for river flow prediction*, Journal of Computing in Civil Engineering, **8**(2) (1994).

[7] J. M. Zurada: *Introduction to Artificial Neural Systems*, First Edition edition, West Group, 1992.

[8] M. Bastan, H. Cam, U. Gudukbay, O. Ulusoy: *BilVideo-7, An MPEG-7compatible video indexing and retrieval System*, IEEE Multimedia, **17**(3) (2010), 62–73.

Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology
Patiala, Punjab, India
Email address: abhishek94goel@gmail.com

Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology
Patiala, Punjab, India
Email address: nitikasinghal566@gmail.com