ATLANTIS PRESS
*Advances in Intelligent Systems Research, volume 134*
2nd International Conference on Control, Automation, and Artificial Intelligence (CAAI 2017)

# A Review of Indoor-Outdoor Scene Classification

Zehang Tong*, Dianxi Shi, Bingzheng Yan and Jing Wei
National Laboratory for Parallel and Distributed Processing, School of Computer, National University of Defense Technology, Changsha, Hunan, China 410073
*Corresponding author

***Abstract*—Indoor-Outdoor scene classification problem have been proposed for almost 20 years and widely applied to general scene classification, image retrieval, image processing and robot application. But there is no consensus on one particular scene classification technique that can solve the Indoor-Outdoor scene classification problem perfectly. As larger image dataset has been developed and machine learning technology especially deep learning based methods achieve remarkable performance in computer vision, we aim to provide guidance and direction for researchers to tackle the Indoor-Outdoor scene classification problem with more powerful and robust solution through concluding the Indoor-Outdoor scene classification approaches which have been proposed in last 20 years. In this paper, we review the Indoor-Outdoor scene classification including feature extraction, classifier and related dataset. Their advantages and disadvantages are discussed. At last we conclude some challenging problems remain unsolved and propose some potential solutions.**

***Keywords-indoor-outdoor; scene classification; computer vision***

## I. Introduction

The scene classification problem is one of the challenging task in computer vision. Given any arbitrary image, scene classification problem is that the computer can associate it with a particular scene category properly such as indoor scene, urban scene and nature scene etc. The problem of scene classification has been explored from a variety of angles in the literature for many years. Various methods have been proposed and achieve good performance in specific image dataset. But there is no consensus on one particular scene classification technique that can solve the scene classification problem perfectly.

In this paper, we review the basic scene classification problem about Indoor-Outdoor scene classification. As Indoor-Outdoor scene classification is one of the basic scene classification problems, the results of Indoor-Outdoor scene classification contribute to general scene classification[1][2][3][4]. Indoor-outdoor scene classification also attracts considerable attention of scientific population involved in content based image retrieval[5][6]. Besides assumption that indoor and outdoor images are usually taken under different illumination conditions can be used for decision about further image processing applications such as image orientation detection[7], depth map generation[8], improving color constancy[9] and robot application[10]. As the Indoor-Outdoor scene classification problem has a clearer definition and broad application prospects, we think it is very meaningful to review the Indoor-Outdoor scene classification approaches which have been proposed by various researchers in last 20 years. By comparing some excellent approaches, we conclude some challenging problems remain unsolved and propose some potential solutions in this paper.

The research on Indoor-Outdoor scene classification can be traced back to the work of Szummer and Picard[11] in 1998. They applied a two-stage classification approach on features that combine Ohta color space histogram and multiresolution simultaneous autoregressive model (MSAR)[12]. At the first stage, they used K-Nearest Neighbors(KNN) to classify subblocks of the image, while the final decision was based on the majority rule. The accuracy of 90.3% is achieved on a set of over 1300 consumer images. Several approaches for Indoor-Outdoor scene classification have been proposed after that. Especially a lot of features of color, texture and edge with high variance and good distribution over category samples were designed to distinguish the indoor images and outdoor images. On the other hand, various machine learning classifiers have been considered and applied. These classifiers include K-Nearest Neighbors(KNN), Support Vector Machine(SVM), Hidden Markov Models(HMM), Neural Networks(NN), Random Forest(RF), Bayesian methods, etc. Last but not least, larger image dataset has been developed as millions of images have been created every day due to the popularity of smart phones. With the image dataset is getting bigger and bigger, more challenges will be arised but it also means more opportunity to build powerful and robust Indoor-Outdoor scene classification technology.

Almost all the Indoor-Outdoor scene classification approaches can be summarized as the Figure 1 shows. There are normally two phases called training phase and classification phase. For both training phase and classification phase, extract features from image is the first step. Many kinds of features have been designed by researchers to depict the difference between the Indoor image and Outdoor image. It is generally believed that feature extraction is crucial for Indoor-Outdoor scene classification[13]. Once features have been extracted, method for automatic image classification should be applied. In training phase, we extract features from a image dataset labelled by Indoor or Outdoor. And then features and labels will tune the parameter of the classifier to achieve good classification performance with specific optimization algorithm. In classification phase, features extraction is the same but classifier is just to judge Indoor or Outdoor based on the features extracted from the image rather than training. Another

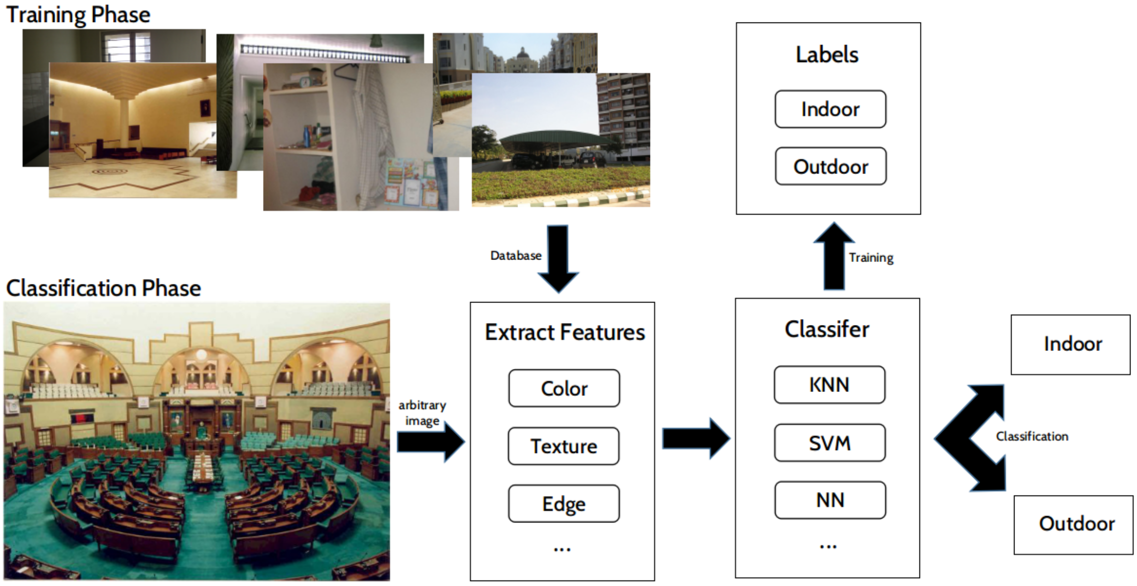


FIGURE I. INDOOR-OUTDOOR SCENE CLASSIFICATION APPROACHES CAN BE SUMMARIZED AS TWO PHASES CALLED TRAINING PHASE AND CLASSIFICATION PHASE.

labelled image dataset would be used to test performance in general which is called test dataset.

The main contribution of this paper is providing a detailed review of every important components of Indoor-Outdoor scene classification including features extraction, classifier and dataset. And then we conclude some challenging problems remain unsolved and propose some potential solutions to provide guidance and direction for researchers to tackle the Indoor-Outdoor scene classification problem with more powerful and robust solution. All the contribution of this paper can be shown in the following list:

- We review various features and classifiers proposed to solve Indoor-Outdoor scene classification problem in last 20 years.

- We summary and compare all the image datasets used to Indoor-Outdoor scene classification.

- We conclude some challenging problems remain unsolved and propose some potential solutions aiming to promote the development of Indoor-Outdoor scene classification.

To the best of our knowledge, such a detailed review of Indoor-Outdoor scene classification problem has not been proposed before. In the following we discuss the feature extraction methods in Section II and introducing applied classifiers in Section III. Section IV presents the datasets related to Indoor-Outdoor scene classification problem. Section V we conclude some challenging problems remain unsolved and propose some potential solutions. Finally we draw conclusion and discuss future work in Section VI.

## II. FEATURE

It is generally believed that feature extraction is crucial for Indoor-Outdoor scene classification. Many researchers focus on the design and selection of discriminant features. Low-level features such as color, texture and edge have been widely examined. To enhance the performance, some other semantic features or camera info also have been studied. Recently, with the introduction of convolutional neural network (CNN), it is believed that the best solution to achieve a system with high performance on scene classification is to learn deep scene features using CNN[14]. In these section, we review all the feature extraction methods proposed to solve the Indoor-Outdoor scene classification problem.

### *A. Color*

Color features have been widely studied to distinguish Indoor images and Outdoor images. As Outdoor images probably consist of green grass and blue sky, color features are very discriminative to classify Indoor images and Outdoor images. On the one hand, various color space have been experimented. Ohta color space[15] is that the color channels are approximately decorrelated. [11] and [16] extracted color features in Ohta color space. They also proved that color features in Ohta color space can achieve better performance than RGB color space. [17] and [18] used LUV color space and [19] pointed out that features in the LUV color space yielded better results during image retrieval than in other color spaces. LST color space[20] was used in [21] and [22] as they reduced feature dimensionality reduced by half by comparing with [11] and [16]. Recently, HSV color space was widely used such as [23][24][25] as it has been reported that HSV color space is invariant to scale and illumination and also more close to human perception[26]. Besides extracting features in single color space, some researchers tried multiple color space such as [27][28] and [13].

On the other hand, different representations of color have been studied. Color histograms and color moments are two popular features in early stage. Besides [23] defined the color orientation histogram as color feature. [24] proposed color correlated temperature feature. [25] used NBHS(Normalized Bins of Hue and Saturation)[29] as color feature. Although so many color features have been proposed and experimented, still no color feature is effective and valid to distinguish all Indoor images and Outdoor images.

### *B. Texture*

Texture aids in identifying objects of interest or region of interest irrespective of the source of the image. Along with color feature, the texture feature is considered to improve the classification accuracy. Several texture feature extraction methods have been introduced to solve the problem of texture analysis and classification. The texture features are computed using the multiresolution, simultaneous auto regressive model (MSAR) in [11][17][16]. These are among the best texture features on benchmarked which was proved by[12]. However MSAR texture features are computationally intensive and thus, [21][22][30][27][31] considered the more efficient wavelet texture representation[32]. Homogeneous Texture Descriptor (HTD) characterizes the region texture by the mean energy and the energy deviation from a set of 30 frequency channels which was proposed by [13]. By making use of Gabor filter[33][34] have extracted WHGO[35] texture feature from Gabor convolved images. Many researchers reported that classification accuracy would be improved by adding texture features.

### *C. Edge*

Edge features also have been widely studied. As organic objects have a larger amount of small erratic edges due to their fractal nature and the synthetic objects, in comparison, have edges that are straighter with less erratic, [36] proposed edge straightness feature and proved the effectiveness. Another widely used edge feature is edge orientation as described in [17][23][18][27]. But when the outdoor scene contains plenty of synthetic objects rather than organic objects, edge features make less contribution.

### *D. Others*

In addition to color, texture and edge feature, many researchers have used additional information about scene or camera to improve performance. [16] and [22] proved that information of sky and grass would improve the accuracy. [37] enhanced performance by using camera info. However this kind of information such as exposure time or flash fired info is commonly unavailable. In [38], they first learned mixture models for 20 basic classes of local image content and then produced 20 probability density response maps (PDRM) indicating the likelihood that each image region was produced by each class. Those PDRMs can be seen features. In [27], they used Normalized Cuts(NCuts)[39] as mid-level cues. Recently, the computer vision community has shown inclination towards global feature based image classification[3][4][40]. A global feature provides a holistic representation of an image by treating it as one single entity, instead of segmenting the image into various subblocks and calculating local features for each sub block. [34] have used a global GIST[41] feature based approach for the indoor-outdoor classification task[22].

With the introduction of CNN, it is believed that the best solution to scene classification is to learn deep scene features using CNN[42][43][44][45][46]. These studies add that deep features can be learned through neural networks where they provide more promising results than complicated handengineered features. But there is no report that deep scene features have been used in Indoor-Outdoor scene classification.

## III. CLASSIFIER

Once features have been extracted, method for automatic image classification should be applied. Approaches for image classification can be roughly grouped into two categories: machine learning method and Bag of Word model.

### *A. Machine Learning*

There are two kinds of machine learning methods for automatic classification: non-parametric methods and learning-based methods. Non-parametric methods perform classification directly on the data, without learning the parameters. The most widely used non-parametric method is KNN which determines image class based on the class of its most similar images [11][47][36][24]. Some Bayesian methods also can be treated as non-parametric methods[17][16][22]. Learning-based methods are able to learn optimal parameters based on input training samples. These methods include SVM[21][23][13], Neural Networks[18][34], Random Forest[27] and etc. Although non-parametric methods require no learning steps and are able to naturally handle a large number of classes, they often suffer from high variation along the decision boundary caused by finite sampling in terms of bias-variance decomposition[48]. As a consequence, their accuracy could be inferior compared to learning-based methods[49]. In addition, processing time of nonparametric methods is considerably larger than the learning-based methods, which makes them inconvenient for large scale classification systems. In [28], they achieved better performance in large-scale dataset by using bagging method[50].

Although CNN can reach state-of-the-art performance based on millions of training images in scene classification, method based on CNN to solve Indoor-Outdoor scene classification problem has not been reported yet.

### *B. Bag of Word*

In computer vision, the bag of word model(BoW model) can be applied to object image classification by treating image features as words[51]. A bag of visual words is a vector of occurrence counts of a vocabulary of local image features. [52][53] proved that BoW model can be applied to scene classification. As deep scene features do not come easily and they require an intensive learning/training stage and large-scale training image sets, [31] proposed a simple and efficient approach based on BoW model and reached a highly accurate performance on Indoor-Outdoor scene classification.

## IV. Dataset

In the Indoor-Outdoor scene image classification literature, datasets used by researchers are varied. The Kodak consumer image dataset tested by [11][16][21][22] contains 1343 images. A benchmark of 1000 images was proposed by [54]. Coral was used by [5]. But to our best knowledge, all these datasets are not available to the public. [17][38][34]collected images from the Internet but they were not released to public either. IITM-SCID2 is a public Indoor-Outdoor classification image dataset containing 902 images and can be accessible from the website. [30][18][13] used this dataset and all achieved accuracy exceeding 90%. But in our opinion IITM-SCID2 is not diverse and large enough. Fifteen Scene dataset[52] contains nearly 4500 images with 15 scene categories. It was used by [25]. SUN was published by [55] for the general scene classification benchmark. It consists of 397 well-sampled scene category indexes and 108,754 images. [28] labelled the whole SUN dataset into 47260 indoor images and 61494 outdoor images. Their experiments were conducted with respect to this dataset. [31] also used SUN dataset. Such very large datasets are meaningful and challenging for scene image classification. By gathering and labelling the datasets into Indoor scene and Outdoor scene, it will be very helpful to promote and verify Indoor-Outdoor scene classification methods.

## V. Open Challenge

We conclude the primary approaches in last 20 years in the Indoor-Outdoor scene classification literature as table I shows. Although all the approaches with varying degrees of success reported classification accuracy could exceed 88%, there is no consensus on one particular scene classification technique that can solve the Indoor-Outdoor scene classification problem perfectly. With the image dataset is getting bigger and bigger, more challenges will be arised.

### *A. Benchmark*

Many novel approaches have been developed to tackle Indoor-Outdoor scene classification problem, but several approaches rely on their own database of images thus reducing the confidence in the success of the approach. As we discussed in Section IV, Kodak dataset was widely used in early stage of Indoor-Outdoor scene classification, but is not available to the public to our best knowledge. Fifteen Scene and SUN dataset are both general scene classification benchmark rather than Indoor-Outdoor scene classification benchmark. [54] presented a Indoor-Outdoor scene classification benchmark only with 1000 images but it is not available to the public either. To promote the development of Indoor-Outdoor scene classification and evaluating all the approaches, a complete and Indoor-Outdoor scene specific benchmark is essential. We define the following constraints on a benchmark for this purpose:

- The images should be diverse. Classification systems can only be well verified if the ground truth data is well placed into real-world types.
- There should be a sufficient number of images as we are facing an increasingly data-driven future.
- The dimensions of the images should be suitable for most image processing techniques with consideration taken for storage size. Images smaller than 640x480 pixels tend to loose the quality in detail that is required by higher level semantic analysis.
- Besides accuracy, some other task related performance criteria such as computation cost must be considered.

### *B. High-Level Feature*

As we discuss in section II, it is usually applied to low-level feature in Indoor-Outdoor scene classification. But low-level features ignore the spatial and structural information of the background and objects in the image. Effective scene image features not only characterize the whole image, but also capture the background and objects information in the image. Therefore, the research on the new feature extraction technology of scene images is a trend of the future development of scene classification. [56] is on the way.

### *C. Dynamic Scene*

In some specific tasks such as in robotics, scene classification sometimes is dynamic. Dynamic means images are continuous and time correlated. The majority of studies have limited their scope to scenes from single image. The major difference between single image classification and dynamic scene classification is that single image classification ignores potentially informative temporal and spatial cues. [57][58] have tackled dynamic scene classification in different way. Dynamic Indoor-Outdoor scene classification is still needed to explore.

### *D. Deep Learning*

Traditional methods based on low-level feature are often difficult to deal with these massive amounts of data when the database capacity exceeds one million, while the deep learning based approach has a good performance. Especially the deep convolution neural network, has achieved a new breakthrough in the scene classification task. This convolution neural network can learn common attributes of the image from a large number of image data. The response characteristics of the deep network have gradually become a universal representation of image recognition[14]. Convolution neural network still has great potential in the field of computer vision. So there is no doubt that convolution neural network will play an important role in the future development of the scene classification.

## VI. Conclusions

In this paper, we review the major Indoor-Outdoor scene classification approaches which have been proposed in last 20 years in the aspect of the feature, classifier and dataset. Although all the approaches with varying degrees of success reported classification accuracy could exceed 88%, there are still open challenges as the dataset becomes larger and larger and application requirements change. Recently the deep learning based approach has a good performance in many computer vision tasks. But it is still difficult to explain how to classify Indoor-Outdoor scene perfectly. So it still needs multidisciplinary scholars to concentrate on Indoor-Outdoor

TABLE I. APPROACHES HAVE BEEN PROPOSED TO SOLVE INDOOR-OUTDOOR SCENE CLASSIFICATION PROBLEM

| Year | Method | Feature | | | | Classifier | Dataset | | Accuracy |
|---|---|---|---|---|---|---|---|---|---|
| | | Color | Texture | Edge | Other | | Name | Size | |
| 1998 | [11] | Ohta Histograms | MSAR | × | DCT | KNN | Kodak | 1343 | 90.3% |
| 1999 | [17] | LUV Moments | MSAR | Orientation | × | Bayesian | Unknown | 6931 | 90.8% |
| 2000 | [47] | × | × | Orientation | × | KNN | Unknown | 470 | 88.7% |
| 2001 | [16] | Ohta Histograms | MSAR | × | Sky,Grass | Bayesian | Kodak | 1179 | 90.1% |
| 2002 | [21] | LST Histograms | Wavelet | × | × | SVM | Kodak | 1200 | 90.2% |
| 2004 | [22] | LST Histograms | Wavelet | × | Sky,Grass | Bayesian | Kodak | 1200 | 90.7% |
| 2004 | [38] | × | × | × | PDRM | LDA | Unknown | 1500 | 93.8% |
| 2005 | [36] | × | × | Straightness | × | KNN | Unknown | 872 | 90.7% |
| 2007 | [30] | RGB Mean | Wavelet | Straightness | × | PNN | IITM-SCID2 | 902 | 92.4% |
| 2010 | [23] | HSV Orientation | × | Orientation | × | SVM | Unknown | 626 | 90.3% |
| 2010 | [18] | LUV Histogram | × | Orientation | DCT | NN | IITM-SCID2 | 902 | 93.1% |
| 2012 | [24] | CCT | × | × | × | KNN | Unknown | 800 | 88.3% |
| 2013 | [27] | C1,C2,C3 | Wavelet | Orientation | NCuts | RF | Gehler | 568 | 88.4% |
| 2014 | [28] | TN,HSH,HDH | × | × | × | EDF | SUN | 108754 | 91.2% |
| 2014 | [13] | SCD,CLD,CSD | HTD | EHD | × | SVM | IITM-SCID2 | 902 | 93.7% |
| 2015 | [25] | NBHS | WHGO | × | × | SRC | 15-Scene | 4485 | 88.6% |
| 2015 | [34] | × | × | × | GIST | NN | Unknown | 2420 | 90.8% |
| 2016 | [31] | × | Wavelet | × | × | ANN | SUN | 108754 | 92.4% |

scene classification problem, especially from academics such as neurobiology and machine learning to carry out cross-research to obtain further break. In future work, we will build an Indoor-Outdoor dataset based on current existing datasets. We will also try deep learning based approach and compare with existing approaches.

## ACKNOWLEDGMENT

The authors would like to express their gratitude towards National Laboratory for Parallel and Distributed Processing for providing with the research resources.

## REFERENCES

[1] Matthew R Boutell, Jiebo Luo, Xipeng Shen, and Christopher M Brown. Learning multi-label scene classification. Pattern recognition, 37(9):1757–1771, 2004.

[2] Li-Jia Li, Hao Su, Li Fei-Fei, and Eric P Xing. Object bank: A highlevel image representation for scene classification & semantic feature sparsification. In Advances in neural information processing systems, pages 1378–1386, 2010.

[3] Ariadna Quattoni and Antonio Torralba. Recognizing indoor scenes. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 413–420. IEEE, 2009.

[4] Jianxin Wu and Jim M Rehg. Centrist: A visual descriptor for scene categorization. IEEE transactions on pattern analysis and machine intelligence, 33(8):1489–1501, 2011.

[5] Aditya Vailaya, M´ario AT Figueiredo, Anil K Jain, and Hong-Jiang Zhang. Image classification for content-based indexing. IEEE transactions on image processing, 10(1):117–130, 2001.

[6] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z Wang. Image retrieval: Ideas, influences, and trends of the new age. ACM Computing Surveys (Csur), 40(2):5, 2008.

[7] Lei Zhang, Mingjing Li, and Hong-Jiang Zhang. Boosting image orientation detection with indoor vs. outdoor classification. In Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on, pages 95–99. IEEE, 2002.

[8] Sebastiano Battiato, Salvatore Curti, Marco La Cascia, Marcello Tortora, and Emiliano Scordato. Depth map generation by image classification. In Electronic Imaging 2004, pages 95–104. International Society for Optics and Photonics, 2004.

[9] Simone Bianco, Gianluigi Ciocca, Claudio Cusano, and Raimondo Schettini. Improving color constancy using indoor–outdoor image classification. IEEE Transactions on image processing, 17(12):2381–2392, 2008.

[10] Jack Collier and Alejandro Ramirez-Serrano. Environment classification for indoor/outdoor robotic mapping. In Computer and Robot Vision, 2009. CRV'09. Canadian Conference on, pages 276–283. IEEE, 2009.

[11] Martin Szummer and Rosalind W. Picard. Indoor-outdoor image classification. In Content-Based Access of Image and Video Database, 1998. Proceedings., 1998 IEEE International Workshop on, pages 42–51. IEEE, 1998.

[12] Jianchang Mao and Anil K Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. Pattern recognition, 25(2):173–188, 1992.

[13] Stevica S. Cvetkovic, Saa V. Nikoli, and Slobodan Ilic. Effective combining of color and texture descriptors for indoor-outdoor image classification. Facta Universitatis, Series: Electronics and Energetics, 27(3):399–410, 2014.

[14] Bolei Zhou, Agata Lapedriza, Jianxiong Xiao, Antonio Torralba, and Aude Oliva. Learning deep features for scene recognition using places database. In Advances in neural information processing systems, pages 487–495, 2014.

[15] Yu-Ichi Ohta, Takeo Kanade, and Toshiyuki Sakai. Color information for region segmentation. Computer graphics and image processing, 13(3):222–241, 1980.

[16] Jiebo Luo and Andreas Savakis. Indoor vs outdoor classification of consumer photographs using low-level and semantic features. In Image Processing, 2001. Proceedings. 2001 International Conference on, volume 2, pages 745–748. IEEE, 2001.

[17] Aditya Vailaya, Mrio Figueiredo, Anil Jain, and Hong Jiang Zhang. Content-based hierarchical classification of vacation images. In Multimedia Computing and Systems, 1999. IEEE International Conference on, volume 1, pages 518–523. IEEE, 1999.

[18] Li Tao, Yeong-Hwa Kim, and Yeong-Taeg Kim. An efficient neural network based indoor-outdoor scene classification algorithm. In Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on, pages 317–318. IEEE, 2010.

[19] Wei-Ying Ma and HongJiang Zhang. Content-based image indexing and retrieval. Handbook of multimedia computing, pages 227–254, 1999.

[20] Jiebo Luo, Robert T Gray, and Hsien-Che Lee. Towards physics-based segmentation of photographic color images. In Image Processing, 1997. Proceedings., International Conference on, volume 3, pages 58–61. IEEE, 1997.

[21] Navid Serrano, Andreas Savakis, and A. Luo. A computationally efficient approach to indoor/outdoor scene classification. In Pattern Recognition, 2002. Proceedings. 16th International Conference on, volume 4, pages 146–149. IEEE, 2002.

[22] Navid Serrano, Andreas E. Savakis, and Jiebo Luo. Improved scene classification using efficient low-level features and semantic cues. Pattern Recognition, 37(9):1773–1784, 2004.

[23] Wonjun Kim, Jimin Park, and Changick Kim. A novel method for efficient indooroutdoor image classification. Journal of Signal Processing Systems, 61(3):251–258, 2010.

[24] A. Nadian Ghomsheh and Alireza Talebpour. A new method for indoor-outdoor image classification using color correlated temperature. Int. J. Image Process, 6(3):167–181, 2012.

[25] R. Raja, S. Md Mansoor Roomi, and D. Dharmalakshmi. Robust indoor/outdoor scene classification. In Advances in Pattern Recognition(ICAPR), 2015 Eighth International Conference on, pages 1–5. IEEE,2015.

[26] Koen Van De Sande, Theo Gevers, and Cees Snoek. Evaluating color descriptors for object and scene recognition. IEEE transactions on pattern analysis and machine intelligence, 32(9):1582–1596, 2010.

[27] Yang Liu and Xueqing Li. Indoor-outdoor image classification using mid-level cues. In Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific, pages 1–5. IEEE, 2013.

[28] Chen Chen, Yuzhuo Ren, and C.-C. Jay Kuo. Large-Scale Indoor/Outdoor Image Classification via Expert Decision Fusion (EDF).In Asian Conference on Computer Vision, pages 426–442. Springer, 2014.

[29] R Raja, S Md Mansoor Roomi, and D Dharmalakshmi. Classification and retrieval of natural scenes. In Computing, Communications and Networking Technologies (ICCCNT), 2013 Fourth International Conference on, pages 1–8. IEEE, 2013.

[30] Indoor versus Outdoor Scene Classification Using Probabilistic Neural Network. In ResearchGate, 2007.

[31] Mana Shahriari and Robert Bergevin. A Two-Stage Outdoor-Indoor Scene Classification Framework: Experimental Study for the Outdoor Stage. In Digital Image Computing: Techniques and Applications (DICTA), 2016 International Conference on, pages 1–8. IEEE, 2016.

[32] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. IEEE transactions on pattern analysis and machine intelligence, 11(7):674–693, 1989.

[33] Ondrej Drbohlav and Mike Chantler. Illumination-invariant texture classification using single training images. In Texture 2005: Proceedings of the 4th international workshop on texture analysis and synthesis, pages 31–36, 2005.

[34] Waleed Tahir, Aamir Majeed, and Tauseef Rehman. Indoor/outdoor image classification using GIST image features and neural network classifiers. In High-Capacity Optical Networks and Enabling/Emerging Technologies (HONET), 2015 12th International Conference on, pages 1–5. IEEE, 2015.

[35] Li Zhou, Zongtan Zhou, and Dewen Hu. Scene classification using multi-resolution low-level feature combination. Neurocomputing, 122:284–297, 2013.

[36] [36]Andrew Payne and Sameer Singh. Indoor vs. outdoor scene classification in digital photographs. Pattern Recognition, 38(10):1533–1545, 2005.

[37] Matthew Boutell and Jiebo Luo. Beyond pixels: Exploiting camera metadata for photo classification. Pattern recognition, 38(6):935–946, 2005.

[38] Le Lu, Kentaro Toyama, and Gregory D Hager. A two level approach for scene recognition. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 688–695. IEEE, 2005.

[39] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. IEEE Transactions on pattern analysis and machine intelligence, 22(8):888–905, 2000.

[40] Congcong Li, Adarsh Kowdle, Ashutosh Saxena, and Tsuhan Chen. Toward holistic scene understanding: Feedback enabled cascaded classification models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(7):1394–1408, 2012.

[41] Aude Oliva and Antonio Torralba. Building the gist of a scene: The role of global image features in recognition. Progress in brain research, 155:23–36, 2006.

[42] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

[43] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In Icml, volume 32, pages 647–655, 2014.

[44] Yunchao Gong, Liwei Wang, Ruiqi Guo, and Svetlana Lazebnik. Multi-scale orderless pooling of deep convolutional activation features. In European conference on computer vision, pages 392–407. Springer, 2014.

[45] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Object detectors emerge in deep scene cnns. arXiv preprint arXiv:1412.6856, 2014.

[46] Mandar Dixit, Si Chen, Dashan Gao, Nikhil Rasiwasia, and Nuno Vasconcelos. Scene classification with semantic fisher vectors. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2974–2983, 2015.

[47] Anne Gurin-Dugu and Aude Oliva. Classification of scene photographs from local orientations features. Pattern Recognition Letters, 21(13):1135–1140, 2000.

[48] Hao Zhang, Alexander C Berg, Michael Maire, and Jitendra Malik. Svm-knn: Discriminative nearest neighbor classification for visual category recognition. In Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, volume 2, pages 2126–2136. IEEE, 2006.

[49] Manik Varma and Debajyoti Ray. Learning the discriminative powerinvariance trade-off. In Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pages 1–8. IEEE, 2007.

[50] Leo Breiman. Bagging predictors. Machine learning, 24(2):123–140, 1996.

[51] Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and C´edric Bray. Visual categorization with bags of keypoints. In Workshop on statistical learning in computer vision, ECCV, volume 1, pages 1–2. Prague, 2004.

[52] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In Computer vision and pattern recognition, 2006 IEEE computer society conference on, volume 2, pages 2169–2178. IEEE, 2006.

[53] Li Fei-Fei and Pietro Perona. A bayesian hierarchical model for learning natural scene categories. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 524–531. IEEE, 2005.

[54] Andrew Payne and Sameer Singh. A benchmark for indoor/outdoor scene classification. In International Conference on Pattern Recognition and Image Analysis, pages 711–718. Springer, 2005.

[55] Jianxiong Xiao, James Hays, Krista A Ehinger, Aude Oliva, and Antonio Torralba. Sun database: Large-scale scene recognition from abbey to zoo. In Computer vision and pattern recognition (CVPR), 2010 IEEE conference on, pages 3485–3492. IEEE, 2010.

[56] Mayank Juneja, Andrea Vedaldi, CV Jawahar, and Andrew Zisserman. Blocks that shout: Distinctive parts for scene classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 923–930, 2013.

[57] Konstantinos G Derpanis, Matthieu Lecce, Kostas Daniilidis, and Richard P Wildes. Dynamic scene understanding: The role of orientation features in space and time in scene classification. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pages 1306–1313. IEEE, 2012.

[58] Christian Th´eriault, Nicolas Thome, and Matthieu Cord. Dynamic scene classification: Learning motion descriptors with slow features analysis. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pages 2603–2610, 2013.