International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-6, Issue-12, December 2020 Pages 122-128

A Survey on Indoor-Outdoor Scene Classification with Deep Learning Techniques

Deepika Bhardwaj, Vinod Todwal

***Abstract*—In the robotics and computer vision application, Scene classification is a fundamental challenge. A clear view or event with low& high-level featuresin Scene classification. Scene classification plays themain role in automated applications of surveillance comprising indoor orientation, pedestrian identification, semantic categorizations, etc. DL methods, especially CNN, which remove the features automatically and have no overhead for manual removal, have become a common solution for scene classification. DL, which is embedded in CNNs, improves its predecessors significantly. It uses graphic technologies for developing multilayered learning models of neuron transformations. A scene classification model was introduced dependsupon deep CNN features. The transfer learning method has been used to increase scene classification precision. Transfer learning is used to transfer control knowledge from a related domain to better the learner from one domain. We may learn about why transfer learning is feasible through real-world non-technical interactions. Transfer learning is aimed at using formerly trained models to transfer learning parameter values to new models. In this work, we outline Scene Classification its Pro & Cons of Diverse Scene Classification and after introduced Indoor-Outdoor Classification and various deep learning algorithms used in scene classification.**

***Index Terms*—Scene Classification, Indoor-Outdoor Classification, Transfer Learning, Deep Learning.**

I. Introduction

Computer vision has achieved a novel stage that enabled robots from labs to discover the real world. With progress in this area, robots face difficulties in understanding their environment. The classification of the scene is a key step for understanding the scene. In several applications such as the monitoring camera, automated driving, domestic robot, or database recovery, the data can be used for scene classification. Surveillance cameras are almost often installed. Both in public as well as in private areas, the installation of a camera is essential as the rate of crime increases every day. The need for operators would increase to control all activity on the camera. Yet human mistakes can still be made, and information cannot be monitored in all minutes. So, the classification may be useful for certain activities in order to solve those issues. [1].

The scene is a real-time view that includes numerous useful viewpoints such as a Hospital Room, railway station, a waiting room for the airport, a kitchen, and so on. When someone from the human brain talks about the kitchen at that period, the kitchen scene, loaded with utensils as well as cellars, is automatically created. Identify the proper object as a scene recollection in respect of its environment. Scene Recognition is a complex issue requiring accurate methods of obtaining adequate results. It is an important phase for different applications, such as robotic navigation, map construction, etc. In different computer vision activities like image retrieval, visual surveillance, and more, Scenes Recognition helps label an image according to a set of definitional categories. While different recognition approaches in the past few decades have been suggested, it is still a challenge due to internal diversity and interclass similarity among scene images. [2].

Scene classification is an interesting field of study in computer vision as it may be utilized in many applications, such as, self-conduct, surveillance, robots, etc. It is to describe the scene as predefined classes as the coast, kitchen, salon, forest, etc. Scene classification aims to categorize scene image1 into predefined scene categories (such as beach, kitchen, and bakery) based on the image's ambient content, objects, and layout. Visual scene understanding requires reasoning about the diverse and complicated environments we encounter in our daily lives [3]. The fundamental issues with the classification of the scene from Indoor/Outdoor. Since indoor-outdoor scene classification is a fundamental issue of classification, the effects of indoor-outdoor scene classifications contribute to classification. Characterization of indoor-outdoor scenes also draws considerable attention from researchers involved in content-based image retrieval. In addition to assuming that outdoor and indoor images are generally obtained under different lighting conditions, more imaging applications like image orientation recognition, map generation, color constancies enhancement, and robot usage can also be used to determine.

Machine learning (ML) has become increasingly popular with researchers in recent years. It has been used in numerous applications like image classification, multimedia concept retrieval, social network analysis, text mining, video recommendation, etc. DL, also called representation learning, is commonly utilized in these applications by different machine learning algorithms. DL was recently the fastest-growing development in big data analysis as well as, for its outstanding performance relative to conventional learning algorithms, has been extensively and effectively implemented in numerous fields, for instance, natural language processing, image recognition, and speech enhancement. A different DL design of CNNs has successfully obtained effective facts in computer vision, which is attributable to the deep structure that enables the capturing & generalization of filtering processes by image-domain convolutions leading to very abstract &

Deepika Bhardwaj, M.Tech Student, Department of CSE , RCEW
Vinod Todwal, Assistant Professor, Department of CSE , RCEW

efficient features [4].

Transfer learning (TL) is utilized by transferring information from a deep network to enhance the learner from one domain. We will discover that transfer learning is feasible from nontechnical interactions in the real world. Take a look at two people who want to learn how to play the piano. One person has no prior musical background, and the other person has advanced knowledge of music from guitar performance. The music context person can learn the piano more effectively by moving learned material music to the learning process to play the piano. [5].

II. Scene Classification

The scene is a real-time view that includes numerous useful viewpoints such as a Hospital Room, railway station, a waiting room for the airport, a kitchen, and so on. When someone from the human brain talks about the kitchen at that period, the kitchen scene, loaded with utensils as well as cellars, is automatically created. Identify the proper object as a scene recollection in respect of its environment. Two different kinds of scenes: - 1^{st} is real-time,and2^{nd} is an artificial environment—the real-time natural environment, as well as the simulated indoor scene. Scene recollection centered primarily on object identification and low-level scene characteristics. Look-alike picture has a wide variety of real-world uses as scene recognition and object detection. To obtain meaning and meaningful features of an image, image matching is also a significant/previous operation [6-9].

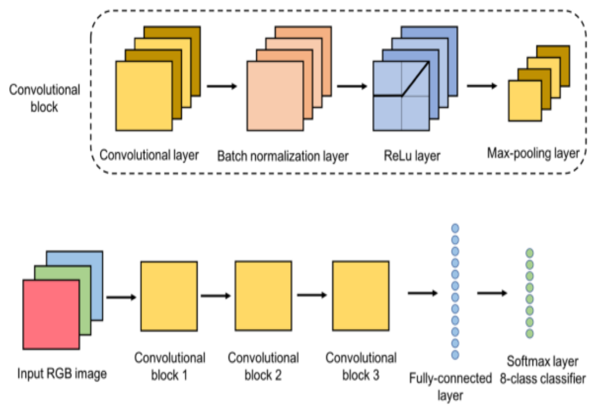

Fig. 1. Scene Classification

Scene Classification is a key image perception problem. In the past, various automated approaches have been utilized to label scenes with semantic labels that can enhance computer vision applications such as navigation, retrieval, and object recognition. [10,11]. Object recognition & description are specific computer vision that has a variety of real applications, from computer games to self-driving vehicles. But this method is not effective because of our limitation in learning various scenes and classifications. [12].

Table I. Pro & Cons of Diverse Scene Classification Methods [1]

Method	**Pro**	**Cons**	**Category**
Feature-based algorithms such as Color histogram, Harris Corner detector, FAST	Performs when the scene has clear colors & details or distinguishable corners or boundaries.	Never work properly as images differ in size, lighting, rotation, etc.	Beach, Forest, Mountains generally outdoor
Texture based algorithms such as SIFT, Principal Component Analysis – SIFT (PCA)	Transform for illumination, affinity as well as robust image deformation.	Weak in performance.	Outdoor
SURF	Easier to perform.	Some features are observed.	Outdoor
BoVW	Identify pixel-level spatial characteristics.	Weak in performance.	Outdoor
Spatial Pyramid Matching	It contains spatial image information.	In outdoor scenes but not indoors, the solution fits well.	Outdoor
GIST + ROI	Hybrid local & global indoor scene characteristics	The method for complicated scenes is not good because of the lack of semanthetical information.	Indoor & Outdoor
Object Bank	Using the object detector output, high-level information is collected.	The structural knowledge on the scene is ignored, and their accuracy depends on the object detector output. Thus, the effects of classification are sharply diminished if the performance is contradictory.	Indoor
Semantic Segmentation	Useful for correctly positioning and	Una ble to obtain exact pixel	Indoor

	shaping objects in the scene.	employed in each object	
DeCAF	Strong to give semanticized information on the point.	Failure to include structural data. For complicated scene classification, it does not fit well.	Indoor & Outdoor
DBRC	Performs well on the scene with a standard global architecture & some traditional artifacts (e.g., classroom, waiting room). More specific findings can be achieved in a dynamic scene in combination with DeCAF.	It takes more time for new datasets to be learned.	Indoor & Outdoor
SSCN	In deep architecture, fewer no. Training images are included. For accurate results, semantic segmentation has been applied.	Require pixel-wise gr ound truth values. It may be a constraint if the new task is implemented.	Indoor & Outdoor
Feature matching algorithm	Overfitting is not feasible even though training data is limited	When extended to a large dataset of large category numbers, it is difficult to obtain accuracy	Indoor

A. *Various Methods in Scene Classification*

The existing scène classification algorithms are categorized in different ways. The grouping may be based on the classifying techniques, on the parameters of the data being used, or on the pixel information was using, or the knowledge of the ancillary data or the image attributes used. Scene classification, as well as unsupervised classification, may be monitored predicated on or analyst's role. Due to using data parameters, parametric and non-parametric scenes are being categorized. Depending on image pixels, per-pixel, subpixel, per-field & contextual classification may be used in scene classification. Images can be classified as non-knowledge as well as knowledge classification based on the accessibility of information. In this study, scene classification algorithms are explained in the following paragraph predicated on per-pixel classifications.

1. *Pre-pixel Classification*

Each pixel is classified into a class in per-pixel classification in light of the spectral similarities with individual classes. Parametric or non-parametric per-pixel classification can be. The probability distribution for each class is assumed to be known in the parametric classification. The training data are generally used to make parameters such as a medium vector as well as a covariance matrix. For complicated landscapes, every class's assertion will distribute its normal probability is often violated. Also, insufficient samples of training can lead to a single matrix of covariance.

- **Nearest Neighbor (NN) Classification:** The NN-based algorithms in statistical classification are simple and efficient methods. The classification of unlabeled samples is predicated in the training data on their distance from tests.
- **Support Vector Machine Classification:** SVM is an effective supervised binary classification method. SVM techniques of classification have often shown a higher level of accuracy than other techniques, like MLC and ANN. When the optimality issue is convex, SVM classifiers always bring different solutions. Several significant aspects of SVM classification givea classification of remote-sensing images based on cluster assumption.
- **Decision Tree-based Classification:** In comparisonwith the ANN, a supervised classification system requiring less difficult training is premised on a decision tree. A decision tree splits a complex decision into several easier decisions so that the method is similar to the required one. The decision tree consists of nodes as well as directed edges—hierarchical structure. So, every node is an observer attribute, categorized, while each edge is an attribute value. The root node is the most common attribute, so although each leaf node is given a class label. Hunt algorithm is the part of developing a decision tree most frequently utilized.

- **Artificial NN-based Classification:** ANN is a biological neural network-inspired computational model. It may be classified as a weighted directed graph where neurons are nodes and neurons are connected to edges with weight. Each artificial neuron calculates and creates a weighted total amount of its input signals dependent on positive activation functions, like linear, sigmoid, Gaussian, etc. It has one input layer, one output layer, which may or may not contain hidden layers, depending on the specific application.[13]

III. Indoor-Outdoor Scene Classification

The classification of indoor-outdoor scenes is a major issue in the scene classification, and the findings in the classification of indoor-outdoor scenes help to generalize the classification of the scenes [14][15][16][17]. The classification of indoor-outdoor scenes also draws substantial interest from scientists interested in content-based image retrieval [18][19].In addition to assuming that pictures are normally taken indoors and outdoors in various illumination environments, more imaging applications, including image processing orientation detection[20], map depth creation[21], and color constancy improvement[22] and robot applications, can also be decided[23]. As the classification of the indoor-outdoor scenes is clear, we conclude that it is very useful to study the classification of indoor-outdoor scenes suggested by different researchers in the last 20 years.

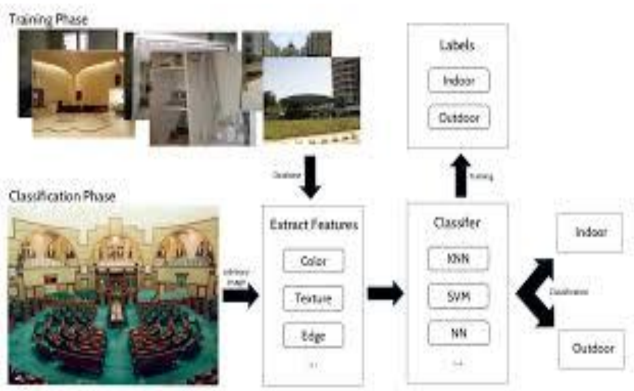

Fig. 2. indoor/Outdoor Scene Classification

Indoors, outdoors, or unspecified, researchers want to identify each graphic. An unexplained photo is where a camera observing a scene is not specifically seen indoors or outside by the viewer. An instance of this scenario is a close-up image with very little view of the background & several figures or objects in the foreground. Any image is divided into a grid of 4 to 4. Each grid cell is separately determined for the color, border, line & texture features to generate a series of feature vectors. These feature vectors are being applied to train the support vector machine by radial function kernel along with the indoor, outdoor or undetermined label of the image. Classifier predicts whether area belongs to an indoor, an outdoor, or an unspecified image is given feature vector for a rectangular portion of an image. On each image grid cell, outcomes from this classifier are normalized & then utilized as a feature vector for an SVM classifier, with a linear kernel estimating the whole class of objects. The effects are between zero & one for three groups. The two-stage classification is used on each frame in the video being processed. If the frame is undefined or indoor, the frame would be the result. Further processing is carried out outside to limit the class. [24,25].

IV. DEEP LEARNING (DL)

DL is a solution that enables the machine to know specific vision tasks as accurately as possible. Deep learning is also defined as deep organized and hierarchical learning that comprises multi-layered, non-linear, conversion, and efficient extract processing units [26]. Deep belief networks (DBNs) have been suggested by using the restricted Boltzmann machines (RBMs) to represent a significant and initial advancement in DL—implemented by the advancement of auto-encoder-based work that trains several intermediate levels of interpretation locally at every step. In recent times a different DL architecture from CNN's obtained significant outcomes in computer vision, attributing to deep structure, which enables model, by image-domain convolutions, to capture as well as generalize filtering processes leading to abstract and efficient properties. [4].

Low-level traditional approaches are also difficult to manage when database capacity approaches one million, whereas the DL technique to learning is effective. DNN, in particular, introduced a remarkable breakthrough in the classification challenge of the scene. This CNN can learn common image attributes from many image data. The deep network response characteristics increasingly have become a universal image recognition representation [26]. In the world of computer vision, CNN also has huge promise. It is also important that CNN will play a significant role in the future creation of classification of the scene.

A. Convolutional Neural Network for Scene Recognition

CNN has become a multilayered neural network constructed on animals' visual cortex. LeCun et al. also created the first CNN. CNN's areas of use specifically include the retrieval of images and handwriting recognition of characters, including such postal analysis. About the architecture, earlier layers are used to describe characteristics like edges and later layers to recombine features to form high-level input attributes and the classification to be taken into consideration. This ensures pooling, which decreases the dimension of the derived characteristics.

Often CNN consists of three layers: Convolutional, fully connected, and Pooling. Convolution layers operate the function extractor. However, they're not produced. In the course of the training, Convolution kernel filter weights are determined. Convolutional layers can extract local characteristics since they limit the hidden layers' receptive fields to local ones. Pooling Layer performs downsampling. It reduces the size of each item that contains the most important information. [27, 29].

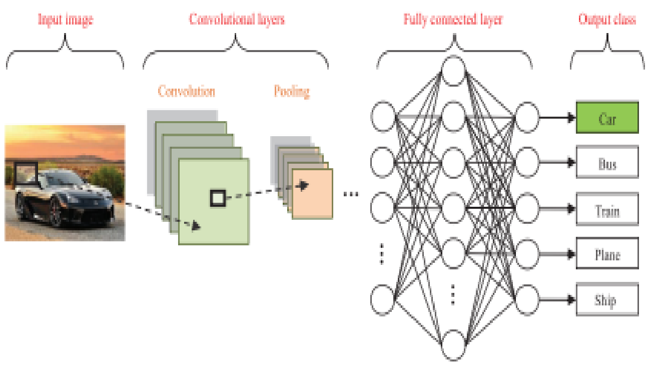

Fig. 4. CNN image classification pipeline

Diverse layers of CNN utilized are:

- **Input Layer:** The first CNN layer used here is an 'input layer' that takes images and resizes them for features extracted into more layers.
- **Convolution Layer:**Next,some layers are 'Convolution layers' that serve as image filters, then

detect image functions and also calculate the corresponding functional points throughout testing.

- **Pooling Layer:**Feature sets extracted are then transferred to the pooling layer. This layer takes big pictures also reduces them while retaining the most valuable detail. It maintains the optimal utility in each window and retains the right match for each element in the window.
- **Rectified Linear Unit Layer:**Next,the "Rectified Linear Unit" or ReLU layer swaps all negative pooling layer numbers with 0. This allows CNN to remain stable by preserving learned values that stay near to 0 or that expand into infinity.
- **Fully Connected Layer:**The last layer is fully connected layers that take high-level filtered images into labels. [30].

V. Transfer Learning

Since 1990, literature has addressed transfer learning with various names. These include learning, life-long learning, transfer of knowledge, multi-task learning, inductive transfer, context-sensitive learning,knowledge consolidation. All these techniques are in effect designed to acquire data from one or more tasks as well as extend to another new job, with the exception of multitask learning. As normally described by various authors, modeling based on learning needs labeled training data and a supervised learning approach with only a few labeled data is almost difficult to study in the target domain. In semi-supervised research, a large number of unlabeled data and a limited number of named data are available to create a classifier provided the source and target tasks are equitably represented. However, in some situations, the model is no longer valid as the feature space is changed, and new data has to be re-trained. This study is informed by transfer learning & analysis techniques that permit information to be transferred between various areas or tasks.

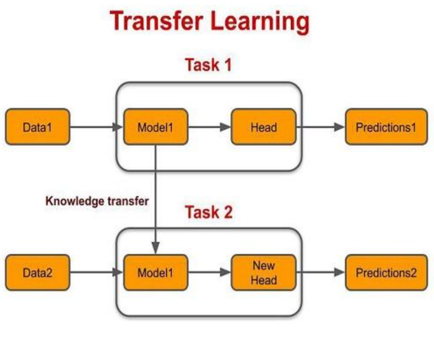

Fig. 5. Transfer Learning

Three main steps are in place to improve learning by transfer. First of all, the initial output can only be achieved with translated information compared to the performance requirements of an ignorant actor, prior to any further learning. Second is the time it takes to properly understand the desired task in comparison with time from scratch, provided the transferred information. Thirdly, the final level of output can be achieved in comparison with the final level without transition.

A. *Transfer Learning Strategies*

NMany authors of various studies examine three critical issues relating to transfer learning techniques:

1. *Inductive Transfer Learning*

One typical inductive method of learning is to use a source domain as well as a task for a trained deep learning model and to various fine-tuning layers ofthe target task. For instance, a pre-trained model is provided by the ImageNet dataset, then modified to recognize various object classes.

2. *Transductive*

Input and output tasks are also identical in this context, but domains vary. No marked information is available in the target only the source that resembles semi-supervised learning. The functional spaces among domains are equal, μS = ·T, but the distribution of input data's marginal probability is specific, P (XS) = P(XT), transductive transfer learning has to do with domain adaption, sampling bias, or covariate changes. The use of synthetic images to increase accuracy for real-world images' objective task is an example of transductive translation.

3. *Unsupervised Transfer Learning*

This environment is the same as inductive TL, and tasks vary from goal to source. However, unsupervised TL concentrates on solving unsupervised learning tasks in the target domain, including such clustering & dimensional reducing. The information on both domains is not accessible [31].

B. *Applications of Transfer Learning*

The five primary fields are mostly covered:

- NLP
- Computer vision & image processing.
- Biology
- Finance
- Business management [32].

VI. Literature Review

W. Tahir et al. (2015) propose an easy but efficient pipeline for the initial measurement of the GIST vector of an image. A feed-forward neural network with a comprehensive training data set is trained for the classification task. The analysis indicates that their method improves numerous state-of-the-art classification algorithms. The final classification pipeline has been built on the Amazon EC2 server for live smartphone uses, owing to computational constraints on mobile devices [33].

E. Rezende et al. (2017) In this study, they introduce a malware classification method focused on ResNet-50 architecture, using a DNN. Malware samples are seen as grayscale byteplot images, while DNN freezes fully convolutional ResNet-50 layers pre-trained on ImageNet data set to apply very last layer to the classification of the malware family. An empirical result from 9,339 samples from 25 families has shown that their methodology may be utilized successfully to categorize families of malware with 98.62 percent accuracy [34].

Y. Jiang et al. (2017) Transfer learning is used to minimize the disparity among data distribution as well as testing data; semi-supervised learning is used forutilizing unlabeled testing data in order to correct the shortage of training information; also, TSK fuzzy method to improve model interpretability is adopted. The method should be trained using two learning algorithms. Their research findings suggest that the methods presented may be faster than many state-of-the-art seizure classification algorithms. [35].

L. Zhang et al. (2019) They suggest a set of indoor/outdoor classification learning scheme of the typical urban area, deepened on cellular data collected within the enterprise LTE network. The self-validation outcome indicates that the ensemble model performs a highly reliable classification for indoor & outdoor conditions (with an out-of-bag error below 1 percent). Also, prominent variables dependent on the different metrics in the initial training are identified. Compared to other classic machine learning approaches, the system reconfigured to fewer variables, and fewer poor students also achieve maximum accuracy and relative fast computing speed [36].

E. Cengıl and A. Çinar (2019) The study concentrates on the use of pre-trained models to solve problems. For the classification of images, pre-trained DL models like Alexnet, Googlenet, VGG16, DenseNet & ResNet are also used with the method called TR. The findings indicate that the models used produce reasonable rates while the VGG16 model achieves the best efficiency [37].

Y. Wang et al. (2019) Intended to address timeliness & lack of partial image information in existence. An approach is applied for transfer learning, based on aCNN, incorporating the feature-extraction process image histogram as the procedure for pre-classification of the SVM system. The simulation results show that, as against the conventional classification algorithm & CNN algorithm, the classification accuracy of 5 categories of elephants or dinosaurs in this work can be efficiently increased & total classification accuracy can exceed 95 percent. The accuracy of classification was increased by about 5 percent [38].

T. Fukumura et a. (2020) In this research, they built a test device to collect and analyze sound using a microphone. First, to identify cracks in concrete structures, they identified the data found by CNN 90.2percent precision. In their past study, the K-mean system of classification accuracy was about 80 percent, so DL results better. Even so, the concrete construction conditions are very different in the real situation, and they must change the classification function to satisfy these criteria so that they have tried to use TL. Tried by using 40 training data, they evaluated TL's ability and created a learning model. The specificity of the description was 90.0%. This outcome was almost the same as CNN [39].

A. A. Rafique et al. (2020) The techniques of recognition suggested is a new frame of segmentation that uses multi-object statistical segmentation to learn robust scene models and separate objects in the scene. Such segregated objects then remove specific features for further processing for linear SVM recognition. Ultimately MLP is provided with scene recognition features as well as weights. In the estimation of state-of-the-art programs, their method showed a major increase. The system provided is efficient in autonomous vision systems, including GPS location detection, robotic vision, protection& sports [40].

S. C. Cetindag et al. (2020) Deep models are implemented at low computing costs so that models can be trained rapidly. In this subject, transfer learning technology has been used to realize CNN models very common in the image recognition of the DL world. A deep learning algorithm known as CAT-Net compares and evaluates the transfer learning system's performance compared to those. With each model, the outcomes are contrasted with the total accuracy, accuracy, recall, or F1 values. [41].

VII. Conclusion

Scene classification plays the main role in computer vision. One of the difficult tasks in computer vision is the scene classification problem. Classification problems indoor/outdoor scenes have been suggested for almost 20 years and are commonly used for general scene classification, image processing, image retrieval, & robot applications. However, a particular technique of classification for the scene is not known to solve the issue of classification of the indoor-outdoor scene perfectly. In recent times, another CNN technology DL has obtained significant computer vision performance due to the deep structure that makes it easier to capture & generalize the process of filtering by image-domain convolutions, resulting in highly abstract and efficient features. In the scene classification work, the deep-CNN has made a breakthrough in particular. Traditional machine-learning algorithms forecast upcoming data by mathematical models trained on labeled or unlabeled data that are similar to future information that has already been collected. The areas, functions including distributions of training and testing used for Transfer learning, by contrast, vary. We see several cases of transfer learning in the real world.

References

[1] T. A.Patel, et al., "Survey on Scene Classification techniques," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020.

[2] S. Masood, et al., "Scene Recognition from Image Using Convolutional Neural Network," Procedia Comput. Sci., vol. 167, no. 2019, 2020, pp. 1005–1012.

[3] M. R. Greene and A. Oliva, "The briefest of glances: The time course of natural scene understanding," Psychological Science, vol. 20, no. 4, 2009, pp. 464–472.

[4] Y.Li, et al.,"Deep learning for remote sensing image classification: A survey," WIREs Data Mining and Knowledge Discovery, vol. 8, no. 6, 2018.

[5] W. Karl, "A survey of transfer learning," DOI 10.1186/s40537-016-0043-6

[6] Z. Bolei,"Places: A 10 million Image Database for Scene Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.40, 2017, pp.1452-1464.

[7] B. Xiaoyong, C. Chen, "Fusing Local and Global Features for High-Resolution Scene Classification," IEEE Geoscience & Remote Sensing Society, vol. 10, 2017, pp. 2889-2901.

[8] Z. Jinyi, L.Wei, "Scene Classification using Local and Global Features with Collaborative representation Fusion," Information Sciences, vol.348,2016, pp. 209-226.

[9] Z. Bolei, "Learning Deep Features for Scene Recognition Using Places Database," In Advances in Neural Information Processing Systems, vol.11, 2014, pp. 1-9.

[10] S.Warren, et al., “A logical calculus of the ideas immanent in nervous activity,” The bulletin of mathematical biophysics 5.4, 1943, pp.115-133.
[11] S.Kuppuswamy, and B.Panchanathan, “Similar Object Detection and Tracking in H,” 264 Compressed Video Using Modified Local Self Similarity Descriptor and Particle Filtering.
[12] Ross G., et al, “Rich feature hierarchies for accurate object detection and semantic segmentation,” Proceedings of the IEEE conference on computer vision and pattern recognition, 2014.
[13] S. Arunadevi, S. Daniel Madan Raja, “A Survey on Image Classification Algorithm Based on Per-pixel,” Int. J. Eng. Res. Gen. Sci., 2014.
[14] R B. Matthew, et al.,“Learning multi-label scene classification,” Pattern recognition, vol. 37, no. 9, 2004, pp. 1757–1771.
[15] L. Li-Jia, et al., “Object bank: A high-level image representation for scene classification & semantic feature sparsification,” In Advances in neural information processing systems, 2010, pp. 1378–1386.
[16] Q. Ariadna and T. Antonio,“Recognizing indoor scenes,” In Computer Vision and Pattern Recognition, CVPR 2009 IEEE Conference, 2009, pp. 413–420.
[17] W. Jianxin and M R. Jim.,“Centrist: A visual descriptor for scene categorization,” IEEE transactions on pattern analysis and machine intelligence, vol. 33, no. 8, 2011, pp. 1489–1501.
[18] V. Aditya, et al.,“Image classification for content-based indexing,” IEEE transactions on image processing, vol. 10, no. 1, 2001, pp. 117–130.
[19] D. Ritendra, et al., Image retrieval: Ideas, influences, and trends of the new age,” ACM Computing Surveys (Csur), vol. 40, no. 2, 2008.
[20] Z. Lei,et al., “Boosting image orientation detection with indoor vs. outdoor classification,” In Applications of Computer Vision, (WACV 2002), Proceedings. Sixth IEEE Workshop on pages,2002, pp. 95–99.
[21] B. Sebastiano, et al.,“Depth map generation by image classification,” In Electronic Imaging, International Society for Optics and Photonics, 2004,pp. 95–104.
[22] C. Gianluigi, et al., “Improving color constancy using indoor–outdoor image classification,” IEEE Transactions on image processing, vol. 17, no. 12, 2008, pp. 2381– 2392.
[23] R. S. Alejandro &C. Jack, “Environment classification for indoor/outdoor robotic mapping,” In Computer and Robot Vision, CRV’09, Canadian Conference on IEEE, 2009, pp. 276–283.
[24] H. Dunlop, “Scene classification of images and video via semantic segmentation,” 2010.
[25] S. Arunadevi, S. Daniel Madan Raja, “A Survey on Image Classification Algorithm Based on Per-pixel,” Int. J. Eng. Res. Gen. Sci., 2014.
[26] S. Pouyanfar, et al., “A Survey on Deep Learning: Algorithms, Techniques, and Applications,” ACM Comput. Surv., [Online]. Available: https://doi.org/10.1145/3234150.
[27] T. Akilan, et al., “A late fusion approach for harnessing multi-CNN model high-level features,” In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 566-571.
[28] Z. Jinyi, L. Wei, “Scene Classification using Local and Global Features with Collaborative representation Fusion,” Information Sciences, vol.348, 2016, pp. 209-226.
[29] L.Herranz, S.Jiang, & X. Li, “Scene recognition with CNNs: objects, scales, and dataset bias,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 571- 579.
[30] S. Neha, “An Analysis of Convolutional Neural Networks for Image Classification,” International Conference on Computational Intelligence and Data Science (ICCIDS), 2018, pp.1-8.
[31] R. Ricardo, “A Survey of Transfer Learning for Convolutional Neural Networks,” 2019 32nd SIBGRAPI Conference on Graphics, Patterns & Images Tutorials (SIBGRAPI-T), 2019.
[32] J. Lu, et al., “Transfer learning using computational intelligence: A survey,” Knowledge-Based Syst., 2015.
[33] W. Tahir, et al.“Indoor/outdoor image classification using GIST image features and neural network classifiers,” 12th International Conference on High-capacity Optical Networks & Enabling/Emerging Technologies (HONET), Islamabad, Pakistan, 2015, pp. 1-5.
[34] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. de Geus, “Malicious Software Classification Using Transfer Learning of ResNet-50 Deep Neural Network,” 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 2017, pp. 1011-1014.
[35] Y. Jiang et al., “Seizure Classification from EEG Signals Using Transfer Learning, Semi-Supervised Learning, and TSK Fuzzy System,” in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 12, 2017, pp. 2270-2284.
[36] L. Zhang, et al., “An Ensemble Learning Scheme for Indoor-Outdoor Classification Based on KPIs of LTE Network,” IEEE Access, vol. 7, 2019, pp. 63057-63065.
[37] E. Cengıl and A. Çinar, “Multiple Classification of Flower Images Using Transfer Learning,” 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 2019, pp. 1-6.
[38] Y. Wang, C. Wang, L. Luo, and Z. Zhou, “Image Classification Based on Transfer Learning of Convolutional neural network,” 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 7506-7510.
[39] T. Fukumura, et al.,“An Efficient Learning Method for Sound Classification using Transfer Learning for Hammering Test,” 2020 IEEE SENSORS, Rotterdam, Netherlands, 2020, pp. 1-4.
[40] A. A. Rafique, A. Jalal, and K. Kim, “Statistical Multi-Objects Segmentation for Indoor/Outdoor Scene Detection and Classification via Depth Images,” 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 2020, pp. 271-276.
[41] S. C. Cetindag, K. Guran, and G. Bilgin, “Transfer Learning Methods for Using Textural Features in Histopathological Image Classification,” 2020 Medical Technologies Congress (TIPTEKNO), Antalya, Turkey, 2020, pp. 1-4.