International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-2, July 2021

Techniques of Indoor-Outdoor Scene Classification using the VGG-16 CNN Model

Kajal Gupta, RK Sharma

Abstrac: In the world of today, computers have begun to rule the people as the machines carry out practically every work that people can accomplish. Scene classification is one such concept that becomes increasingly important when robots replicate the actions of a human being Scene categorization may be done on interior or exterior scenes using various extraction techniques, as well as categorization of indoor and outdoor scenes in these two categories is more difficult. The methodology for the indoor/outdoor classification scene has the drawback of inadequate accuracy. This research aims to enhance the accuracy by using the Convolution Neural Network Model in VGG-16. This paper proposes a new approach to VGG-16 to classify images into their classes. The algorithm results are tested using SUN397- indoor-outdoor dataset & the tentative data reveal that the methodology proposed is superior to the existing technology for the scene classification of indoor-outdoor (I/U).

Keywords: Scene Classification, Indoor-Outdoor Classification, Deep Learning, Neural Network Model VGG 16, CCN, Data Augmentation, Imagedatagenerator, Optimizers.

I. INTRODUCTION

As technology progresses, humans in many areas are overpowered. There are various considerations where technology is employed for environmental analysis rather than human analysis. The classification of the scene is one such use in which the scene analysis may be performed by the machine. The classification of scenes is a method in which a computer visions a stage and that machine then tries to describe the scene according to deep learning (DL). The scene to be classified may be an indoor scene, such as the bakery, the airport, the garage, the bedroom, and so on. Comparing to the outside scene classification, indoor scene classification is more complex due to indoor scene variability [1]. In recent years, numerous indoor classification systems have been devised, and they face multiple issues, as well as accuracy is the main difficulty.

In a variety of applications dealing with consumer photography, scene classification is significant. Understanding the scene category is useful in the case of categorization as an essential component of automatic methods for albumin. [2].

Manuscript received on July 19, 2021.
Revised Manuscript received on July 26, 2021.
Manuscript published on 30 July, 2021.

* Correspondence Author

Kajal Gupta*, M.Tech degree, Department of Computer Science and Engineering, Faculty Of Engineering and Technology, Agra College, Agra (U.P.), India.

R K Sharma, Associate Professor, Faculty Of Engineering and Technology, Agra College, Agra (U.P.), India.

The study looks forward to computer vision to master machines in the most complex cognitive capacity of humankind: vision. However, when you look, you don't only aim to capture and comprehend pixels. In this regard, considerable study has been conducted to generate visual sentences. These models utilize either sentence templates & graphics or rely on deep networking capabilities [3-5].

The presentation of scene classes based on their kind is generally taken since most scientists consider that the classification is significantly reliant on the scene type [6], [7], [8]. The scenes are organized into several categories appropriately; some of the examples are the recreation and recreational sceneries vs cultural & ancient, industrial & commercial against open-country, mountain against the forest, etc. [6], [9], [10]. However, the representation is not constrained to the classifications above as well as the most common presentation is provided by the human beings indoor vs outdoor or natural. [6], [8].

All inside scenes are artificially produced, whereas outdoor images are either natural or man-made. Indoor spatial areas generally contain vast entity collections (e.g. items), while outdoor settings feature limited kinds of items[6]. In contrast, the generally horizontal orientation of natural sceneries, whereas man-made outdoor ones tend to dominate both vertically and horizontally [9] overall. In addition,[8][6],[11] has been demonstrated to have low performance in indoor environments most scene recognition methods which function well for outdoor scenes.

In this study, we examine the fundamental challenge in the classification of the indoor-outdoor scene. As the classification of the IndoorOutdoor scene is a classification issue, the outcomes of the classification of the indoor-outdoor scene contribute to the general classification of a scene. Also, the categorization of indoor-outdoor scenes receives significant interest from the scientific population concerned with content picture recovery. Furthermore, you may also decide on extra image processing applications orientation detection, map depth development, color consistency improvement, and robot application if indoor and outdoor pictures are regularly taken under various lighting situations. As the issue of categorization of the outdoor and indoor scene has a better understanding & wider application, we related to the severity important to evaluate the ways to the classification of the scene in IndoorOutdoor which different academics have suggested during the last 20 years. By analyzing alternative great techniques, we determine that several difficult problems remain unresolved and provide possible answers in this study. [12].

As the core of scene classification, scene representation is the process of transforming a scene image into its concise descriptors (i.e., features), and still attracts tremendous and increasing attention. The recent revival of interest in Artificial Neural Networks (ANNs), particularly deep learning, has revolutionized computer vision and been ubiquitously used in various tasks like object classification & detection, semantic segmentation, & scene classification. Recent deep learning advances have opened a scenario description to large-scale and wild-scene datasets and have been suggested for multiple scene representations. Deep learning has significantly improved the classification of the scene [13]. During the past years, in several classification activities, DL architectures like CNNs have outperformed conventional approaches. These models have shown high classification efficiency where extensive and varied data sets for training are available. [14]. In DL, deep belief networks (DBNs) were suggested by the operation of the restricted Boltzmann (RBMs)engines as an initial breakthrough. Preceded by the production focusing on auto-encoder work, which forms the multiple intermediate layers of participation at each stage locally. A new DL design of fully convolutional neural networks (CNNs, for its initials) has achieved notable important computer vision outcomes, attributed to a deep framework that facilitates the model to grab and generalize filters via Image domain convolutions which lead to abstract and efficient characteristics.

Section II includes the corresponding studies from several investigators. Discuss the methods, problem identification, and proposed algorithm presented in Section III, Section IV offerings the findings & evaluation of the suggested model categorization. Lastly, Section V concludes with future work.

II. REVIEW OF LITERATURE

In the last few years, researchers on applied Indoor/ outdoor scene classification have developed various practical methods.

Shahriari, M., & Bergevin, R. (2016) This research developed an indoor-to-outdoor hierarchical two-stage classification framework. The technique proposed is a straightforward yet very effective global scene-representation model. We validate the methodology presented by adding two benchmark datasets to the algorithm: 15-Scene and SUN397. The efficacy of the suggested technique is shown by experimental tests. The model they proposed achieves an average accuracy of 97.84% & 55.1% in outdoor scenes & 93.09% as well as 92.4% correspondingly in 15- & SUN397 indoor scene classes. [15].

K. Pujar et al. (2017) This paper's approach utilizes a neural network that has become more important with progress in the methodology of machine learning. The technique is more efficient than existing models because they aim to categorize background as an entire environment instead of utilizing item identification for that. They evaluate this methodology using our data set that includes RGBs and deep images of frequent places in academic surroundings like classrooms, laboratories, etc. The methodology described is up to 98 percent more accurate than earlier approaches [16].

L. Zhang et al. (2019) In this research we offer an ensemble learning approach, based upon cellular data acquired in a commercial LTÉ network, for indoor-outdoor classification for a representative urban area. The findings of self-validation reveal that the ensemble model has a categorization of the interior and outdoor environment extremely precisely (with an out-of-bag error of less than 1 percent). The important variables are also chosen depending on the variable significance of the first training. In contrast to other conventional machine learning approaches, the reconfigured model founded on fewer variables & less weak learners also gets maximum accuracy & comparatively short time. [17].

Gill, J. S., & Brar, A. S. (2019) this paper aims to improve precision by integrating the SIFT feature, SURF (Speed-Up Robust Feature) & Tamura features to remove functionalities (SVM) and then by use of them for feature match. In this paper, the innovative scene categorization is addressed. The algorithm outcomes are assessed using MIT-Indoor data, and the findings from the experiments reveal that the methodology provided exceeds several current indoor classification algorithms. [18].

A. A. Rafique et al. (2020) presented a scene model to analyze and recognize depth data to allow robots to perceive situations like humans in real-time. To learn the robust scene model & discrete items in a scene, the suggested recognition technic is the new segmentation framework. These separated items then extract the unique characteristics for further identification utilizing linear SVM. Finally, the characteristics and weights for the identification of the scene are given via an MLP. The enhancement of their system associated with a state of the art schemes was substantial. In automated driving such as robotic viewing, GPS-focuses, sports & safety the suggested approach is efficient. [19].

K. Abdullah et al. (2020) The ML-based IO user categorization technique is proposed in this study for 3G network applications in cellular systems. They take many circumstances into account. The testing findings suggest that the best IO classification machine learning technique is the boost method to 88.9 percent with accuracy. [20].

Carbonneau et al. (2020) evolve a novel supervised learning process predicated on NASNet CNN named 'CNN-Supervised Classification' (CSC). First, they try comparing classification accuracy of posterior probability, an MLP, a random forest, as well as CSC. Outcomes show median F1 scores (usually used ML quality metrics) of 71%, 78%, 72% as well as 95%, respectively. Secondly, they use data for 5 of 11 rivers to train their classification. The test set for all eleven rivers is then predicted. Median F1 scores reach 98 percent for five rivers that are used in model training. Median F1 scores are 90 percent for the 6 rivers not used in model training. Investigations must now concentrate on further developing a new generation of DL classification algorithms that will encode human image interpretation abilities as well as enable fully automated, possibly in real-time, river landscape imaging interpretations. [21].

III. PROPOSED WORK

A. Problems Identification

In this work, we analyze the basic problem of scene classification about scene classification for indoor-outdoor. In previous work, they have used AlexNet DCNN for the scene classification of indoor-outdoor. The depth of this model is much smaller and thus it is hard to learn from the collection of images. More time is required to produce better results. AlexNet stacks fewer layers and maximum size filters. To deal with such an issue we have used VGG16 deep CNN model scene understanding and classification that intends to understand the activations from images of various public scene environments with the help of CNN.

B. Proposed Methodology

The work proposes a VGG16 deep CNN model for indoor-outdoor scene classification from images of various public scene environments. A well-known pretrained network, 'VGG16' has been chosen for our proposed learning model. In this first, we have to perform preprocessing. Preprocessing is the overall term for all the transformation of the data, including centering, normalization, rotation, shifting, shear, etc., before being transformed into the model.

The Data pre-processing is one stage in the solution of each problem of ML. For data representation, training, and testing to be performed efficiently, it must be processed, cleansed, and transformed in ML. Preprocessing is the overall term for all the transformation of the data, including centering, normalization, rotation, shifting, shear, etc., before being transformed into the model. The pre-processing objective is to enhance image data that eliminates unwanted distortions or optimizes some image features that are needed for more processing, while geometric image transformation is classed among pre-processing methods here, given that similar technologies are utilized for the pre-processing process.

1. Data augmentation (DA)

The data augmentation plays an important role in enhancing the model's performance in the proposed process. The data augmentation concerns the modification of image data and a sequence of operations so that the changed image remains in the same class. Data augmentation helps to generalize the data input, reflecting a better test accuracy. There are three key strategies for augmenting training data: expanding data set, in-place or on the fly augmentation, incorporating data set, and on-site augmentation. Training DL-NN models in more data will cause skilled models & Improving technology can cause visual changes that strengthen fit models' capacity to generalize what they learn into new pictures. Kera's DNN learning library gives the ability to fit image data increase models through the ImageDataGenerator class.

- ***ImageDataGenerator***

The validation dataset and the evaluation dataset can also be defined by a data generator. A separate instance of ImageDataGenerator is also used that has the same configuration for pixel scaling (not covered) as that utilized for ImageGenerator's training dataset. The reason is that the DA is only employed to artificially expand the training data set to boost model efficiency on an unaugmented dataset.

We will concentrate on 5 major forms of image DA strategies; in particular:

- Image shifts through *height_shift_range* & *width_shift_range* arguments.
- The image flips through *vertical_flip* & *horizontal_flip* arguments.
- Image rotations through *rotation_range* argument
- Image brightness through *brightness_range* argument.
- Image zoom through *zoom_range* argument.

To increase our training data, we employed Keras ImageDataGenerator. It offers different transformations to increase image data such as scale, rotating, shear, brightness, zoom, channel shift, width and height changes, and horizontal and vertical shifts. We applied geometric transforms such as Scaling, zoom, Horizontal flip, Image size, Batch size, Images, Classes, Color channel, Test data image, and Validation image in the proposed method.

2. Neural network model VGG16

Deep learning has strong performance in image classification and several deep learning models like AlexNet, VGGNet and InceptionNet have been used in recent years. In this work, we have used VGG-16 in CNN for this purpose. In the paper, VGG16 is a CNN model presented by A. Zisserman and K. Simonyan from Oxford University "Very Deep Convolutional Networks of Large-Scale Image Recognition". In ImageNet, a dataset of over 14 million images belonging to 1000 classes, this model achieves 92.7% of the highest test accuracy. This was one of the ILSVRC-2014's popular models. It increases on AlexNet by replacing broad kernel-sized filters (5, & 11, respectively, on 1st & 2nd layers of convolution) by multiple 3×3 kernel-sized filters one by one.

VGG 16 is a CNN 16-layer, pre-trained 100-class model. The Sun397 image network dataset is trained in this VGG 16 model. We utilized this model as a feature extraction tool & extracted from each image 4132 features & stored them in hdf5 file format. To resize all the images to mentioned dimensions VGG 16 models need images of 224 x 224 dimensions. We have such a good result. VGG-16 has an excellent ability to extract the image to get a strong image classification effect.

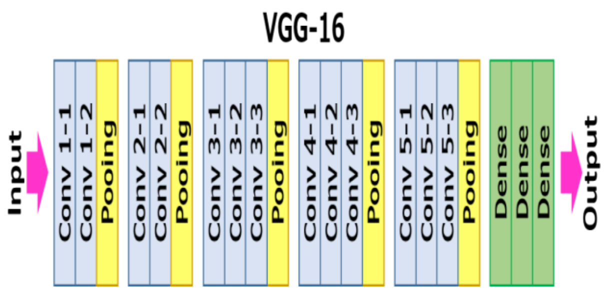

Figure 1: VGG16 **Model**

- ***Loss Function***

CategoricalCrossEntropy: For categorical classification, cross-entropy loss contributed by training data point $i, (x_i, y_i)$, is simply the "negative log-likelihood (NLL)":

$$L_i = -\log (p_{y_i})$$

since the ground truth probability is one for the correct label y_i and zero for every other label.

- ***Adam Optimizer***

The estimation of the adaptive moment is an approach for gradient descent optimization techniques. For dealing with big problems containing a lot of data or variables, the approach is highly effective. It needs less memory and therefore is effective. Instinctively, it is a mix of the algorithm 'gradient descent as well as the algorithm 'RMSP.'

Mathematical Aspect of Adam Optimizer

If we take the formulae used in the different methods above, we obtain:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\left[\frac{\delta L}{\delta \omega_t}\right] v_t = \beta_1 v_{t-1} + (1-\beta_1)\left[\frac{\delta L}{\delta \omega_t}\right]$$

... (1)

m_t and v_t are estimates of 1st moment (mean) & 2nd moment (uncentered variance) of gradients correspondingly, henceforth method name. As m_t & v_t are initialized as vectors of 0's, researchers of Adam detect that they are biased towards 0, especially throughout primary time steps, & especially when decay rates are small (such that β_1 & β_2 are nearly 1).

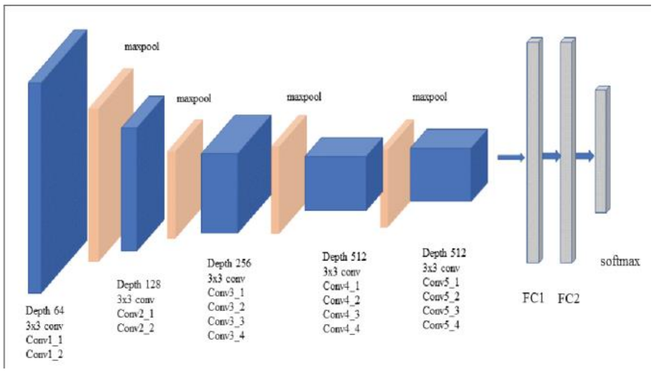

Figure 2: Overview of Proposed model

We modified VGG16 with different layers like the Batch normalization *layer* after every convolution layer and the dropout layer after every dense layer.

C. Proposed Algorithm

step1 Collect the input images from SUN397 Dataset.
step2 Preprocess the image
step3 Training with CNN model
step4 Generate the Gwalior and Kota traffic with the use of sumo.
step5 Modify the VGG 16 with different layers.
step6 Test the images.
step7 Predicted Results.

IV. RESULTS AND DISCUSSION

This work has implemented using python programming language and the platform used is Jupiter notebook and done for the result at the proposed approach. After training the networks separately for indoor and outdoor classes, it is observed that the accuracy of indoor classes is better than that of outdoor classes from the chosen dataset. The dataset used for the purpose is publicly available SUN397dataset.

A. Data Set

Scene UNderstanding 397 (SUN397) dataset consists of 397 scene categories, in which each category has more than 100 images. The dataset contains 108,754 images with an image size of about 500×300 pixels. SUN397 spans over 175 indoor, 220 outdoor scene classes, and 2 classes by mixed indoor & outdoor images, e.g., a promenade deck with a ticket booth. There are several train/test split settings with 50 images per category in the testing. SUN397 is a wider scene benchmark of 397 categories like indoor, man-made & natural groups (at the least before places). The SUN397dataset is used for public use. This dataset is very demanding not only for large no. of groups however also as a result of the smaller no. of trained data and a much wider variability of object and layout properties (50 images per category). It is generally accepted as the scene classification reference benchmark. Our experiments consider seven scales which are 227x227 by scale images.

Table 1: Parameters information

Parameters	value
Dataset	sun397
Scaling	1./255
zoom	20%
Horizontal flip	True
Image size	227*227
Batch size	64
Images	4132
Classes	2
Color channel	3 (RGB)
Test data image	518
Validation image	515
Neural network model	VGG16
Epoch	100

Table 2: Model Summary

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 224, 224, 64)	1792
conv2d_1 (Conv2D)	(None, 224, 224, 64)	36928
max_pooling2d (MaxPooling2D)	(None, 112, 112, 64)	0
Batch_normalization Batch No	(None, 112, 112, 64)	256
conv2d_2 (Conv2D)	(None, 112, 112, 128)	73856
conv2d_3 (Conv2D)	(None, 112, 112, 128)	147584
max_pooling2d_1 (MaxPooling2	(None, 56, 56, 128)	0
batch_normalization_1 (Batch	(None, 56, 56, 128)	512
conv2d_4 (Conv2D)	(None, 56, 56, 256)	295168
conv2d_5 (Conv2D)	(None, 56, 56, 256)	590080
conv2d_6 (Conv2D)	(None, 56, 56, 256)	590080
max_pooling2d_2 (MaxPooling2	(None, 28, 28, 256)	0
batch_normalization_2 (Batch	(None, 28, 28, 256)	1024
conv2d_7 (Conv2D)	(None, 28, 28, 512)	1180160
conv2d_8 (Conv2D)	(None, 28, 28, 512)	2359808

Retrieval Number: 100.1/ijrte.B62970710221
DOI: 10.35940/ijrte.B6297.0710221
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-2, July 2021

```
conv2d_9 (Conv2D)            (None, 28, 28, 512)       2359808
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 14, 14, 512)       0
_________________________________________________________________
batch_normalization_3 (Batch (None, 14, 14, 512)       2048
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 14, 14, 512)       2359808
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 14, 14, 512)       2359808
_________________________________________________________________
conv2d_12 (Conv2D)           (None, 14, 14, 512)       2359808
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 7, 7, 512)         0
_________________________________________________________________
batch_normalization_4 (Batch (None, 7, 7, 512)         2048
_________________________________________________________________
flatten (Flatten)            (None, 25088)             0
_________________________________________________________________
dense (Dense)                (None, 4096)              102764544
_________________________________________________________________
dropout (Dropout)            (None, 4096)              0
_________________________________________________________________
dense_1 (Dense)              (None, 4096)              16781312
_________________________________________________________________
dropout_1 (Dropout)          (None, 4096)              0
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 8194
=================================================================
Total params: 134,274,626
Trainable params: 134,271,682
Non-trainable params: 2,944
```

B. Results Analysis

This subsection represents the outcomes study obtained by the proposed model.

Table 3: Comparison of accuracy, Loss, Val_loss and Val_Accuracy for Base and Proposed Model

Model	Loss	Accuracy	Val_loss	Val_Accuracy
Base model	0.2623	0.9113	1.7929	0.5615
Proposed model	0.1029	0.9637	1.2143	0.8913

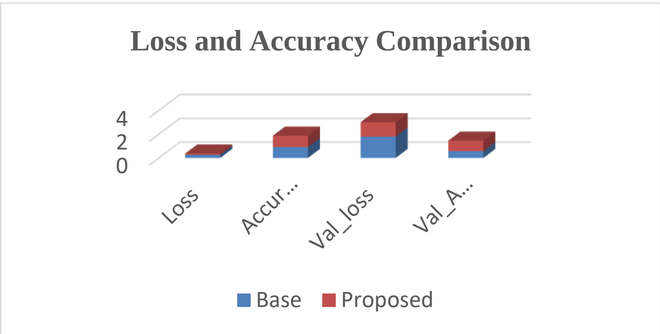

Figure 3: Graph comparison of accuracy, Loss, Val loss & Val Accuracy for Base and Proposed Model

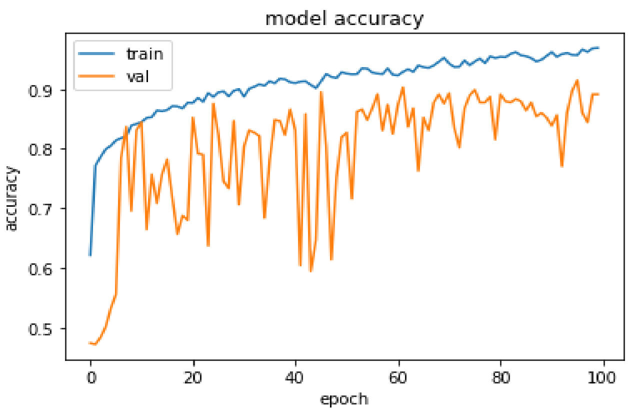

Figure 4: Line Graph for Model Training Accuracy

Figure 4 represents a line graph for model training accuracy. This process continues up to 100 epochs. It shows training and validation accuracy. Initially, it starts training accuracy from 62%, which is gradually increased up to 96% accuracy at 100 epochs. It also shows validation accuracy. Initially, it starts validation accuracy at approximately 23%, which has a variable increment inaccuracy. It constantly increases approximately 89% accuracy.

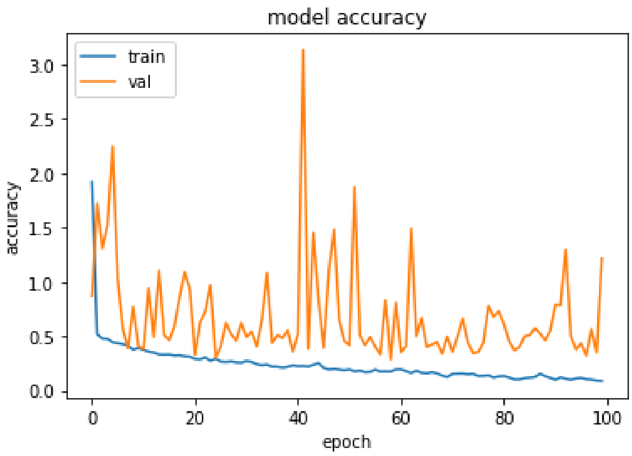

Figure 5: Line Graph for Model Loss Accuracy

Figure 5 represents a line graph for model loss accuracy. This process continues up to 100 epochs. It shows training and validation accuracy. Initially, it starts loss accuracy from 62%, which is gradually increased up to 96% accuracy at 100 epochs. It also shows validation accuracy. Initially, it starts validation accuracy at approximately 23%, which has a variable increment inaccuracy. It constantly increases approximately 89% accuracy.

V. CONCLUSION

In this work, the two-class scene classification model is clear and efficient, based on an indoor-outdoor classification proposed. The concept of indoor-outdoor scenes has been studied over the years, to our understanding, all study was dedicated to a VGG-16 where both indoor and outdoor scenes are categorized. To explain images and important issues in computer graphics for geometry comprehension, scene understanding is still an extremely difficult issue in computer view. In the previous years, several systems for indoor categorization have been established, with distinct exacting requirements. We also introduced a methodology that classifies all indoor/outdoor scenes and provides an indoor/outdoor scene label, Compared to earlier works. In this study, we have established deep CNNs. In ImageNet, a collection of over 14 million 1000 class pictures, 92.7% achieved top-5 test accuracy. It improves on AlexNet by substituting big filters (11 & 5 in 1st & 2nd convolutional layers), with multiple 3×3 kernel size filters one by one. We attain Training loss = 10% and Training Accuracy = 96% in our proposed work.

FUTURE SCOPE

The deep learning method has recently worked well in several computer vision tasks. But how to classify the outdoor-indoor scene perfectly is still rarely explained. Therefore, multidisciplinary researchers do need to focus on the IndoorOutdoor scene classification problem, in particular from researchers, like neuro-biology & machine learning, for more analysis. We will create an IndoorOutdoor dataset based on current existing databases in future work. We would also test with a deep learning model and associate them with previous studies.

REFERENCES

1. Quattoni and A. Torralba, "Recognizing Indoor Scenes", in IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 413- 420.
2. Serrano, N., Savakis, A., & Luo, A. (n.d.). A computationally efficient approach to indoor/outdoor scene classification. Object Recognition Supported by User Interaction for Service Robots. doi:10.1109/icpr.2002.1047420
3. G. Kulkarni, V. Premraj, V. Ordonez, S. Dhar, S. Li, Y. Choi, A. Berg, and T. Berg, "Babytalk: Understanding and generating simple image descriptions," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, no. 12, pp. 2891 – 2903, Dec 2013.
4. K. Xu, J. Ba, R. Kiros, A. Courville, R. Salakhutdinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," 2015, vol. abs/1502.03044. [Online]. Available: http://arxiv.org/pdf/1502.03044v2.pdf.
5. A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on, June 2015, pp. 3128 – 3137.
6. J. Xiao, K. Ehinger, J. Hays, A. Torralba, and A. Oliva, "Sun database: Exploring a large collection of scene categories," International Journal of Computer Vision, pp. 1 – 20, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11263-014-0748-y
7. A. Quattoni and A. Torralba, "Recognizing indoor scenes," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, June 2009, pp. 413 – 420.
8. A. Oliva and A. Torralba, "Chapter 2 building the gist of a scene: the role of global image features in recognition," in Visual PerceptionFundamentals of Awareness: Multi-Sensory Integration and High-Order Perception, ser. Progress in Brain Research. Elsevier, 2006, vol. 155, Part B, pp. 23 – 36. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0079612306550022
9. M. M. Gorkani and R. W. Picard, "Texture orientation for sorting photos "at a glance"," in Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision amp; Image Processing., Proceedings of the 12th IAPR International Conference on, vol. 1, Oct 1994, pp. 459 – 464 vol.1.
10. S. Lazebnik, C. Schmid, and J. 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, vol. 2, 2006, pp. 2169 – 2178.
11. A. Quattoni and A. Torralba, "Recognizing indoor scenes," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, June 2009, pp. 413 – 420.
12. Zhehang Tong* , Dianxi Shi, Bingzheng Yan and Jing Wei," A Review of Indoor-Outdoor Scene Classification", 2nd International Conference on Control, Automation, and Artificial Intelligence (CAAI 2017),
13. Delu Zeng, Minyu Liao, Mohammad Tavakolian, Yulan Guo Bolei Zhou, Dewen Hu, Matti Pietikainen, and Li Liu," Deep Learning for Scene Classification: A Survey", arXiv:2101.10531v2 [cs.CV] 20 Feb 2021
14. Alina Matei et al.," Deep learning for scene recognition from visual data: a survey", arXiv:2007.01806v1 [cs.CV] 3 Jul 2020
15. Shahriari, M., & Bergevin, R. (2016). *A Two-Stage Outdoor - Indoor Scene Classification Framework: Experimental Study for the Outdoor Stage. 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA).* doi:10.1109/dicta.2016.7796998
16. K. Pujar, S. Chickerur and M. S. Patil, "Combining RGB and Depth Images for Indoor Scene Classification Using Deep Learning," 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2017, pp. 1-8, doi: 10.1109/ICCIC.2017.8524231.
17. L. Zhang, Q. Ni, M. Zhai, J. Moreno and C. Briso, "An Ensemble Learning Scheme for Indoor-Outdoor Classification Based on KPIs of LTE Network," in IEEE Access, vol. 7, pp. 63057-63065, 2019, doi: 10.1109/ACCESS.2019.2914451.
18. Gill, J. S., & Brar, A. S. (2019). *Support Vector based Indoor Scene Classification Technique using Different Features. 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA).* doi:10.1109/iceca.2019.8822153
19. 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), 2020, pp. 271-276, doi: 10.1109/IBCAST47879.2020.9044576.
20. K. Abdullah et al., "A Machine Learning-Based Technique for the Classification of Indoor/Outdoor Cellular Network Clients," 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2020, pp. 1-2, doi: 10.1109/CCNC46108.2020.9045473.
21. Carbonneau, P. E., Dugdale, S. J., Breckon, T. P., Dietrich, J. T., Fonstad, M. A., Miyamoto, H., & Woodget, A. S. (2020). *Adopting deep learning methods for airborne RGB fluvial scene classification. Remote Sensing of Environment, 251, 112107.* doi:10.1016/j.rse.2020.112107

AUTHORS PROFILE

Kajal Gupta, received her B.Tech degree in Computer Science and Engineering from Faculty Of Engineering and Technology, Agra College, Agra, Dr APJ Abdul Kalam Technical University, Lucknow, India and pursuing M.Tech degree in Computer Science and Engineering from Faculty Of Engineering and Technology, Agra College, Agra, Dr. APJ Abdul Kalam Technical University, India. Her major research is in the field of Deep Learning

R. K Sharma, received his M.Tech degree in Engineering System in Computer Science from Dayalbagh Educational Institute, Agra and PhD form Xavier University from Bhubaneshwar, India. Presently working as Associate Professor at FET Agra college, Agra and also holding many administrative posts. His major researches are in the field of Network Security, Image Processing, Software cost Estimation and Optimization. He published Around 30 papers in international journals, conferences, workshops. He is also the member of IEEE and more than 6 societies and also served as academic experts in different govt. organizations DRDO, BSNL etc. and Industries

Retrieval Number: 100.1/ijrte.B62970710221
DOI: 10.35940/ijrte.B6297.0710221
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication