International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Comprehensive Analysis of State-of-the-Art Transfer Learning Models for Remote Sensing Scene Classification

Divvela Srinivas Rao [1], S. Koteswari *[2], K. Bala Sindhuri [3], S. Ravi Chand[4], Mohan Appikonda[5], P. V. Sivarambabu[6]

Submitted: 26/08/2023 **Revised**: 15/10/2023 **Accepted**: 28/10/2023

Abstract: Remote Sensing classification plays a significantrole in numerous fields, such as Urban Planning, Environmental Monitoring, Land Management and Remote Sensing Analysis. The primary goal of this study is to compare the efficacy of DenseNet121, InceptionV3, and VGG16 as potential models for land use scene classification. To achieve this objective, a comprehensive experimental framework is constructed, encompassing data pre-processing, model training, and performance evaluation. The UC Merced datasetwas augmented four times and then was utilized in this study. The dataset consists of high-definition aerial photos that covera broad range of land use scenes, The models are refined through a process of fine-tuning, followed by a comprehensiveassessment of their performance using a wide array of evaluation metrics. These metrics encompass Accuracy, Precision, Recall, F1-score, Inference Time, and Model Size for all three models. DenseNet121 exhibited superior performance in capturing fine-grained features, achieving an accuracy of 91.94%. InceptionV3 excelled in handling variations in scale and rotation and achieved a relatively higher accuracy of 92.45%, while VGG16 demonstrated a balance between simplicity and accuracy achieving an accuracy of 88.89%.

Keywords*: Remote sensing, Land Use Scene Classification,Image Classification, Computer Vision, Transfer Learning*

1. Introduction

Remote Sensing refers to the process of acquiring data from aerial vehicles and unmanned space probes and satellites to assess the characteristics of objects on the Earth's surface. Satellite-based Remote Sensing systems offer a consistent and repetitive observation of the Earth, which is of great value in monitoring both short-term and long-term changes, as well as assessing the impact of human activities [1]. Remote sensing has undergone significant transformations in terms of data quality, spatial resolution, reduced revisit intervals, and expanded coverage area [2]. Over the course of the last few decades, the discipline of remote sensing has witnessed significant evolution, characterized by notable advancements in image spatial resolution and the velocity of data acquisition. These advancements have exerted a profound influence on the utilization and administration of remote-sensing images. The improved spatial granularity has opened up new avenues for making progressing analyses of images in remote sensing and interpretation, enabling the development of novel methodologies that were previously unattainable. [3].

The utilization of image classification to categorize land cover images constitutes a pivotal application in the domain of Remote Sensing. There are many practical real-world applications of Scene classification such as Land Cover[4], urban planning[5], land management[6], Environmental Risks Monitoring[7][8], vegetation survey[9], viticulture[10], hydrological modeling[11], a forensic investigation[12] etc. In this paper, we have focused on optimizing the classification of images related to covering and use of land in remote sensing. Scene categorization of remote-sensing photographs has been an active study topic, serving as a crucial and difficult Challenge for properly understanding remote-sensing photos [13]. Within the domain of remote sensing, scholars have categorized image classification into three distinct levels: (a) Pixel-level classification, which entails the assignment of class labels to individual pixels; (b) Object-level classification, which revolves around identifying objects within geospatial images; and (c) Scene-level classification, aiming to classify remote sensing image patches into meaningful classes. [14]. Our study focuses on the classification of images at the scene-level. Over the past few decades, there has been a lot of study into the

[1]Sr.Assistant Professor, Department of AI & DS, Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India. Email: srinivassowjanya2012@gmail.com
[2]Professor, Department of ECE, Pragati Engineering College (A), Surampalem, East Godavari District, AndhraPradesh,India.Email:eshwari.ngr@gmail.com
[3]Assistant Professor ,Department of ECE, S.R.K.R Engineering College, Bhimavaram, Andhra Pradesh, India. Email:k.b.sindhuri@gmail.com
[4]Professor & Head , Department of ECE, Nalla Narasimha Reddy Educational group of institutions –Integrated Campus, Hyderabad, Telangana, India.Email:ravichandsankuru1@gmail.com.
[5]Associate professor,Department of EEE, Pragati Engineering College, JNTUK Kakinada,Andhra Pradesh, India Email:appikonda.mohan@gmail.com
[6]Assistant Professor,Department of CSE, Koneru Lakshmaiah Educational Foundation, Vaddeswaram, Guntur Andhra Pradesh,India.Email:sivarambabu424@gmail.com
** Corresponding Author Email: eshwari.ngr@gmail.com*

capability of classifying supplied land-use scene pictures into specified meaningful class distinctions and of nuanced evaluation of images in remote sensing. The previous methods used for analyzing remote sensing images rely on extracting and representing features at fundamental-tier and intermediary-tier [15]. Various feature combinations and machine learning approaches have demonstrated promising results in terms of performance. Previous methods employed a limited-scale image dataset. However, current trends in the analysis of images in remote sensing have shifted towards the utilization of deep learning[16].Several research studies have been conducted to comprehensively examine various techniques of deep learning in the domain of image classification in remote sensing Xu etal.[17] used the methodology encompassed the utilization of principal component analysis (PCA) to address data duplication. Subsequently, a self-organizing network was trained for the purpose of effectively categorizing Landsat satellite images. Their approach surpassed the maximum likelihood method in terms of performance. Esam et al.[18]introduced an innovative methodology for classifying land use at the scene level, utilizing characteristics derived from convolutions and a sparse encoding autoencoder (AE). Marco et al.[19] implemented two modern architectures, namely CaffeNet and GoogleNet, using three distinct learning approaches. This combination resulted in a noteworthy enhancement in performance compared to existing cutting-edge references. Cheng et al.[20] scrutinized the use of deep CNNs - AlexNet, VGGNet and GoogleNet for scene classification and carried out the classification using linear SVM. Arijit et al. [21] implemented a transfer learning method incorporating Res2Net, which is capable of extracting multi- scale featured values from input and shows its capability of capturing relevant and complex features of the input at a granular level. In their research paper, Zhang et al.[22] presented a pre-trained Efficient Net model, a very fine- tuned methodology for image classification in remote sensing built on the basis of transfer learning method. Hung et al.[23] put forth a new CNN architecture - RSSCNet with high generalization capability and also applied the LIME(local interpretable model, agnostic explanation) algorithm to further improve results. Akhtar et al.[24] In their academic research paper, the authors undertook a process of fine-tuning the ResNet50 model. Their investigation revealed that the model demonstrated superior performance in terms of labeling precision metrics compared to a range of other models. Moreover, they emphasized that the presented outcomes of their implemented cross-domain transfer learning system exhibited precise and noteworthy performance improvements when compared to established benchmark evaluations. Bazi et al. [25] in their research paper proposed meticulous investigation of the Gaussian Process (GP) method for classification of multisource and spectral-rich remote sensing images. In addition, the investigation explored two separate analytical approaches for Gaussian Process classification, referred to as the methods of Laplace and expectation-propagation. These methodologies were employed in conjunction with two distinct functions of covariance: specifically, the covariance of squared exponential and neural networks. The focal objective of our comparative investigation is to assess the effectiveness of transfer learning [26], a deep learning method that utilizes pre-existing models to construct an effective model for categorizing remote sensing images at the scene level. Our image classification model incorporates a transfer learning methodology by leveraging pre-trained models for classification purposes. The augmented UC Merced dataset was used, divided into segments, and subsequently processed through pre-trained models such as VGG16[27], InceptionV3[28], and DenseNet[29]. The final layer of the model was linked to a flattened layer, which was then followed by a dense layer consisting of 21 neurons, each corresponding to one of the 21 distinct classes. The activation function used for this layer was softmax. A comprehensive comparative analysis was performed across all three models, exploring measures like Recall, Precision, F1-score, Training and Testing Accuracy, Model Size, and Inference Time.[30]

2. Methodology

Our methodology proposes a CNN-based framework that involves the preprocessing of data used as well as using pre- trained models for image classification. The dataset was augmented from its original size of 2100 images to 10,500 images by augmenting each class 4 times intending to produce results with efficacy and more precision. This led to data being more complex and bigger which required data to be preprocessed. Originally the size of each image in the dataset was 256x256x3, which was reshaped to a smaller shape of 224x224x3 as it prepares the input images as better and suitable input for Convolution neural network models. The smaller the size, less the quantity of parameters to be used making the model light and less complex and also resulting in better efficiency. Further Data Augmentation was carried out through sophisticated data augmentation techniques supported and provided by the Image Data Generator library of Tensor flow. The imagery underwent a sequence of augmentations facilitated by the Data Generator library. Notably, the Horizontal flip augmentation was activated, leading to stochastic horizontal image flips with a 50% probability. This augmentation methodology was employed to enrich the diversity and volume of the dataset, thereby enhancing model accuracy through exposure to diverse renditions of each image. Zoom augmentation was performed by setting the zoom range to a float value of 0.2 resulting in zooming in and out of each image each by 20%, yielding better results during training. Shear_range in Keras is used in transforming the image by slanting the image edges along a fixed axis and specifying themaximum shear angle to which the image is to be randomlysheared. The shear_range value can be set between 0 and 1, where 0 means no shear and 1 means shearing at an angle of 45 degrees. In our case, the shear_range was set to 0.2 which resulted in random shearing of images both horizontally and vertically in a range of [-0.2,+0.2] radians. The importance of the shear range is that it helps in improving the ability of themodel to generalize better. Three distinct CNN architectures are trained on the newly pre-processed data consisting of 21 distinct image classes - VGG16, DenseNet121 and InceptionV3, a flattening layer and an output layer in addition to it. The aforementioned trio of CNN architectures functions as adept feature extractors and holds a pivotal role in facilitating precise fine-tuning through the application of transfer learning. CNN, a prominent deep learning algorithm, proves instrumental in addressing computer vision challenges by eliminating the necessity for manual feature extraction. CNN requires a larger amount of data for training to provide efficient results and thus makes the training of models an arduous task. To surmount this challenge, the concept of transfer learning was introduced, offering a solution by harnessing the capabilities of pre-trained models derived from extensive research-driven datasets, notably the influential Image Net dataset. These pre- trained models inherit knowledge from the previous training, thus proving beneficial in training small scaled datasets more efficiently and helping it to generalize better on new and unseen data. Transfer learning has limited data requirements for which it provides fine tuning services for handling data. Fine tuning involves setting up various hyper parameters like Batch Size which we have set up to 32 in case of VGG16 and InceptionV3 while in case of DenseNet121, we allot a batchsize of 128. Batch Size allocation helps to control the accuracy of error gradient estimation during the process of training neural networks. The concluding layer of our methodology incorporated the utilization of the activation function named Softmax. The selection of the function for activation in the final layer is contingent upon the specific prediction needs, differentiating between binary classification and multiclass classification objectives. Softmax activation function was used in the last layer here as it is exclusively designed for multi-class classification problems (more than two classes).The dense layers in each model used the ReLU activation function. We used categorical_crossentropy as a loss function as we have a multi-class classification problem at hand. Numbers of epochs for all pre-trained models were set to 10. In addition, we used the adam optimizer in each of the VGG16, InceptionV3 and DenseNet121 pre-trained models with a learning rate set to 0.001 in each case. We have proposed this system of three pre-trained models - VGG16, InceptionV3 and Densenet121 among which Densenet121 turned out to be the most efficient than the rest. Image net forms the basis of all these pre-trained models as they are trained on it. It can be called as the source domain of the development chain as the models used for training have its origins traced to the ImageNet dataset. Conversely, the focus of the target domain pertains to the categorization of images into 21 distinct classes, encompassing designations such as agricultural, beach, baseball diamond, and others. Considering the parallels between the target domain and thesource, that is classifying images into multiple classes, we used freezing and training methods to counter that. It is used especially before compilation, where the layer. Trainable value is set to 'false', thus not allowing the pre-trained weights of the respective models to get updated duringtraining on the new data. It ensures that only the final layer trains on the data while ensuring the retain ability of the original pre-trained weights. The images after getting through pre-trained layers, were passed through the layer of Figure1. Conceptual representation of the process

Where the flattening of the array takes place, thus converting a multi-dimensional complex array into a single-dimensional linear feature vector. It would be further used as input to the next layer where the classification would take place. The final layer consists of fully connected nodes which contains a specialized activation function called Softmax Activation function which classifies the input image into 21 distinct categories [0,1…20]. The proposed flow of work is as per the below image.

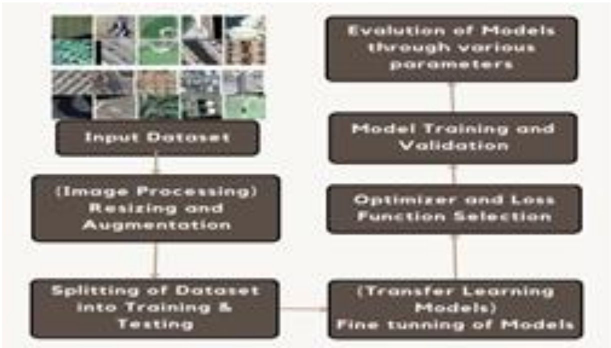

Fig 1: Conceptual representation of the process.

3. Dataset Description

The Original dataset belongs to UC Merced and was created on Oct 28, 2010. The USGS National Map Urban Area Imagery collection was used to meticulously compile the dataset, involving manual extraction from extensive images encompassing diverse urban regions across the United States. The image has a pixel resolution of 1 foot. There are 21 classes in the dataset, and each class has 100 photos of that type. Yi Yang and Shawn Neesham are the creators of the dataset and it was first used in their paper onLand Use Scene Classification. [31]

Table 1: Dataset Description

Collection	Original Image per Class	Augmented Image per Class
Agricultural	100	500
Airplane	100	500
Baseballdiamond	100	500
Beach	100	500
Buildings	100	500
Chaparral	100	500
Denseresidential	100	500
Forest	100	500
Freeway	100	500
Golfcourse	100	500
Harbor	100	500
Intersection	100	500
Mediumresidential	100	500
Mobilehomepark	100	500
Overpass	100	500
Parkinglot	100	500
River	100	500
Runway	100	500
Sparseresidential	100	500
Storagetanks	100	500
Tenniscourt	100	500
TOTAL	**2100**	**10,500**

4. Result Analysis

This section explains the findings made during the research project. The study was conducted employing a specific hardware configuration, which included an Intel 11th Gen 4-core 2.80GHz CPU, with a NVIDIA T500 GPU, 16GB Memory, and 1TB Solid State Drive. The development environment comprises Python version: 3.10.11, and various frameworks of deep learning like TensorFlow, Keras and Numpy have been employed to develop this project. The training dataset consists of 7350 images, a validation dataset of 2100 images and the test dataset of 1050 images. The training dataset constitutes 70% of the dataset while the latter ones - validation and test datasets constitute 30% in total. For training purposes, the images were reshaped from 256x256x3 to 224x224x3 and then were made to pass through pre- trained models. The batch size was opted to be 32 in VGG16 and InceptionV3 and in DenseNet121 it was 128. Epochs were set to 10 in all three model trainings. The output of the pre-trained layers was flattened with the help of the flattening layer where the complex multi-dimensional array was simplified to a linear single-dimensional feature vector which was further passed through the last layer containing the Softmax activation function which further classified the input images into one of the 21 distinct categories (harbor, intersection, storage tanks, beach, buildings, chaparral, forest, golf course, parking lot, runway, freeway, agricultural, tennis court airplane, mobile home park, medium residential, overpass, river, sparse residential, dense residential, baseball diamond,) of land cover. For the assessment of the results achieved, various parameters of evaluation were assessed - Training accuracy, Testing accuracy, Precision and F1 score, Recall, model size and time of inference. Among all the three pre-trained models in the research paper, the InceptionV3 model attained the highest testing accuracy of 92.45% while VGG16 and DenseNet121 achieved testing accuracies of 88.89% and 91.94% respectively. On the contrary, VGG16 had the highest training accuracy on the training data that is 95.83% while DenseNet121 and InceptionV3 achieved accuracies of 81.64% and 94.73% respectively. From close observation it can be assessed that DenseNet121 showed higher testing accuracy than its accuracy on training data; it is possible due to the regularization techniques it uses in order to generalize better on new unseen data. InceptionV3 had the highest Precision value of the other two models, i.e., 92% and for VGG16 and DenseNet121 it was 88.69% and 91.71% respectively. Similarly, in the case of F1 Score and Recall, InceptionV3 was the best with both values being 92.00% while VGG16 had 87.975 and 88.00%, and DenseNet121 with 90.65% and 90.95% respectively. The amount of time a mode takes to make predictions on new data, that is Inference time was found to be lowest in the case of VGG16 with 0.13s, and in the case of DenseNet121 and InceptionV3, it was 0.166s and 0.99s respectively. VGG16 model was the largest model with a size of 132.87 MB while InceptionV3 was 86.53 MB and DenseNet121 was 41.21 MB.

Table 2: Pretrained models and their various evaluation Parameters

Model Name	VGG16	DenseNet121	InceptionV3
Model Size	132.87MB	41.21MB	86.53MB
Inference Time	0.13s	0.166s	0.99s
Training Accuracy	95.83%	81.64%	94.73%
Testing Accuracy	88.89%	91.94%	92.45%
Precision	88.69%	91.71%	92.00%
F1 Score	87.97%	90.65%	92.00%
Recall	88.00%	90.95%	92.00%

5. Conclusion

In this paper, we carried out a comparative study with the help of 3 Transfer learning pre-trained models - VGG16, InceptionV3 and DenseNet121 for image classification of land cover in remote sensing. We used an augmented version of the UC Merced dataset that consists of 10,500 images distributed among 21 distinct land cover categories. Image preprocessing was carried out on the image data to make it ready for further training. Subsequent to this, the images underwent processing through the aforementioned three pre- trained models. Subsequently, they were directed through the final layer to accomplish precise image classification. Lastly, the proposed models were assessed thoroughly using 7 pre- defined evaluation parameters - Training accuracy, Testing accuracy, Model Size, Inference time, Precision, Recall and F1 Score. Among all three, InceptionV3 was the most efficient model, superficially demonstrating training accuracy of 94.73%, the highest testing accuracy of 92.45%, the highest Precision, F1 Score and Recall values of 92.00%,92.00% and 92.00% respectively and with a model size of 86.53 MB with inference time of 0.99s. In the future, the data could be accrued to a larger scaled dataset as the data is being available through satellites and drones with the research being proliferating in this domain. Different approaches as well as different proposals for developing deep learning models for classification with higher efficacy in the future could also help in better Land cover classification and assessment of land use.

Acknowledgements

None

Ethical approval

Not applicable

Conflicts of interest

The authors declare no conflicts of interest.

Funding

None declared

References

[1] Al-Doski, Jwan, Shattri B. Mansorl, and Helmi Zulhaidi Mohd Shafri. "Image classification in remote sensing." Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia 3.10(2013).

[2] R. Pires de Lima and K. Marfurt, “Convolutional Neural Network for Remote-Sensing Scene Classification: Transfer Learning Analysis,” Remote Sensing, vol. 12, no. 1, p. 86, Dec. 2019: http://dx.doi.org/10.3390/rs12010086

[3] Weixun Zhou, Shawn Newsam, Congmin Li, Zhenfeng Shao, PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 145, Part A, 2018, Pages 197-209, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2018.01.004.

[4] Dhanaraj, K., & Angadi, D. P. (2022). Land use land cover mapping and monitoring urban growth using remote sensing and GIS techniques in Mangaluru, India. GeoJournal, 87(2), 1133-1159.

[5] Netzband, M., Stefanov, W. L., & Redman, C. L. (2007). Remote sensing as a tool for urban planning and sustainability. Applied remote sensing for urban planning, governance and sustainability, 1- 23.

[6] Skidmore, A. K., Bijker, W., Schmidt, K., & Kumar, L. (1997). Use of remote sensing and GIS for sustainable land management. ITC journal, 3(4), 302-315.

[7] Ahmed, B., & Alam, A. (2022). Remote Sensing of Natural Hazards. MDPI, Basel.

[8] Li, J., Pei, Y., Zhao, S., Xiao, R., Sang, X., & Zhang, C. (2020). A review of remote sensing for environmental monitoring in China. Remote Sensing, 12(7), 1130.

[9] Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. Journal of plant ecology, 1(1), 9-23.

[10] Hall, A., Lamb, D. W., Holzapfel, B., & Louis, J. (2002). Optical remote sensing applications in viticulture-a review. Australian journal of grape and wine research, 8(1), 36-47.

[11] Thakur, P. K., Nikam, B. R., Garg, V., Aggarwal, S. P., Chouksey, A., Dhote, P. R., & Ghosh, S. (2017). Hydrological parameters estimation using remote sensing and GIS for Indian region: a review. Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, 87, 641-659.

[12] Davenport, G.. (2001). Remote sensing applications in forensic investigations. Historical Archaeology. 35. 87-100. 10.1007/BF03374530.

[13] R. Pires de Lima and K. Marfurt, "Convolutional Neural Network for Remote-Sensing Scene Classification: Transfer Learning Analysis," Remote Sensing, vol. 12, no. 1, p. 86, Dec. 2019, doi: 10.3390/rs12010086.

[14] G. Cheng, X. Xie, J. Han, L. Guo and G. -S. Xia, "Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 3735-3756, 2020, doi: 10.1109/JSTARS.2020.3005403.

[15] J. Zhang, T. Li, X. Lu and Z. Cheng, "Semantic Classification of High-Resolution Remote-Sensing Images Based on Mid-level Features," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 6, pp. 2343-2353, June 2016, doi: 10.1109/JSTARS.2016.2536943.

[16] Maryam Mehmood, Ahsan Shahzad, Bushra Zafar, Amsa Shabbir, Nouman Ali, "Remote Sensing Image Classification: A Comprehensive Review and Applications", Mathematical Problems in Engineering, vol. 2022, Article ID 5880959, 24 pages, 2022. https://doi.org/10.1155/2022/588095

[17] R. Naushad, T. Kaur, and E. Ghaderpour, "Deep Transfer Learning for Land Use and Land Cover Classification: A Comparative Study," Sensors, vol. 21, no. 23, p. 8083, Dec. 2021, doi: 10.3390/s21238083.

[18] Esam Othman, Yakoub Bazi, Naif Alajlan, Haikel Alhichri & Farid Melgani (2016) Using convolutional features and a sparse autoencoder for land-use scene classification, International Journal of Remote Sensing, 37:10, 2149-2167, DOI: 10.1080/01431161.2016.1171928

[19] Castelluccio, M., Poggi, G., Sansone, C., & Verdoliva, L. (2015). Land Use Classification in Remote Sensing Images by Convolutional Neural Networks. ArXiv, abs/1508.00092.

[20] G. Cheng, C. Ma, P. Zhou, X. Yao and J. Han, "Scene classification of high resolution remote sensing images using convolutional neural networks," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 2016, pp. 767-770, doi: 10.1109/IGARSS.2016.7729193.

[21] A. Das and S. Chandran, "Transfer Learning with Res2Net for Remote Sensing Scene Classification," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 796-801, doi: 10.1109/Confluence51648.2021.9377148.

[22] D. Zhang, Z. Liu and X. Shi, "Transfer Learning on EfficientNet for Remote Sensing image Classification," 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 2020, pp. 2255-2258, doi: 10.1109/ICMCCE51767.2020.00489.

[23] S.-C. Hung, H.-C. Wu, and M.-H. Tseng, "Remote Sensing Scene Classification and Explanation Using RSSCNet and LIME," Applied Sciences, vol. 10, no. 18, p. 6151, Sep. 2020, doi: 10.3390/app10186151.

[24] M. Akhtar, I. Murtza, M. Adnan, and A. Saadia, "Cross-Domain Transfer Learning for Natural Scene Classification of Remote- Sensing Imagery," Applied Sciences, vol. 13, no. 13, p. 7882, Jul. 2023, doi: 10.3390/app13137882.

[25] Y. Bazi and F. Melgani, "Gaussian Process Approach to Remote Sensing Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 1, pp. 186-197, Jan. 2010, doi: 10.1109/TGRS.2009.2023983.

[26] Bozinovski, Stevo. (2020). Reminder of the First Paper on Transfer Learning in Neural Networks, 1976. Informatica. 44. 10.31449/inf.v44i3.2828.

[27] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4.

[28] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer

vision. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 2818-2826).

[29] Jianming Zhang , Chaoquan Lu , Xudong Li , Hye-Jin Kim, Jin Wang. A full convolutional network based on DenseNet for remote sensing scene classification[J]. Mathematical Biosciences and Engineering, 2019, 16(5): 3345-3367. doi: 10.3934/mbe.2019167

[30] Ramoliya, D., Goel, P., Vaidya, J., & Vadaliya, P. (2022, September). University Indoor Scene Classification using Transfer Learning. In 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 793-797). IEEE

[31] Yi Yang and Shawn Newsam. 2010. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '10). Association for Computing Machinery, New York, NY, USA, 270–279. https://doi.org/10.1145/1869790.186982