

A Comparative Analysis of Remote Sensing Land Use Land Cover Image Classification using Deep Convolutional Neural Networks based on SuperPixels

Namdeo Baban Badhe¹, Dr. Vinayak Ashok Bharadi² and Dr. Nupur Giri³

¹Research Scholar, Information Technology department, FAMT, Mumbai University, Ratnagiri, India

²Professor, Information Technology department, FAMT, Mumbai University, Ratnagiri, India

³Professor, Computer Engineering Department, VESIT, Mumbai University, Mumbai, India

Abstract.

The categorization of human activities and natural elements on the landscape over time using scientific and statistical methodologies is referred to as land use or land cover (LULC). Satellite imagery is a great tool for studying terrestrial resources, and land cover classification is critical in this analysis. A benchmark dataset “EuroSAT dataset”, based on Sentinel-2 satellite is used for study purpose. Transfer learning was utilized rather than building training models from scratch. Deep learning approaches, particularly those trained on the ImageNet dataset, have grown in popularity for identifying land use and land cover. In this paper, the pretrained Deep Convolutional Neural Network models, like VGG16, ResNet50, and InceptionV3 have been studied. For comparison, parameters like accuracy, precision, recall and F1-Score have been used. The InceptionV3 model achieved an overall classification accuracy(88.18%),Precision(0.9711), Recall(0.9800) and F1-Score(0.9887) outperforming both VGG16 and ResNet50 models. In this paper, a Modified Mean Shift Algorithm(MMSA) is used to generate valuable features known as Super pixels from the input image to enhance accuracy and reduce execution time. Those super pixels were subsequently fed to the pretrained models as input. A combination of pretrained models and the MQSA is trained and evaluated on the same dataset. The Experimental results shows the combination slightly improves the performance.

Keywords: convolutional neural networks; high spatial resolution; pretrained CNN models; remote sensing scene classification.

I. Introduction

In recent years, it has become possible to obtain remote sensing images with great spatial and spectral resolution. Remote sensing imagery has a wide range of applications due to its high resolution for fine spectra, including the environment, military, mining, and medicine. Unlike normal images, hyperspectral photographs contain a wealth of spectral information that might reflect the physical characteristics and chemical composition of the object of interest, making image classification easier. Hyperspectral image classification is the most prominent area of

remote sensing field. There are numerous applications for remote sensing very high-resolution image classification. It is used in urban planning, environmental monitoring, agriculture, disaster management, land use land cover mapping, transportation planning, geological investigations, archaeology, and cultural heritage management, among other things. It aids in the analysis of land use patterns, monitoring vegetation, assessing crop health, mapping disasters, planning infrastructure, investigating geology, and protecting cultural heritage. Its uses include forestry, water resource management, climate change research, and more.

Land cover refers to physical features on the Earth's surface, whereas land use refers to the human activities linked with such features. Deep learning approaches, particularly deep convolution neural networks (DCNNs), are rapidly becoming popular for accurate multispectral and hyperspectral data classification. CNNs outperform prior approaches in the processing of complex contextual pictures.

In this paper, we review the use of deep learning in land use and land cover classification based on multispectral and hyperspectral images. We investigate the state-of-the-art in deep learning approaches in order to develop whole or partial solutions to the problems. One of the most important aspects of this research is how the LULC classification accuracy varies when different Convolutional Neural Networks have been applied. The following is a breakdown of the structure of the paper. The related works are introduced and examined in the second part. Section III then introduces the proposed approach. In Section IV, we'll look at and discuss an experimental implementation. Finally, Section V concludes this investigation and gives some reflections on the work.

II. Related Work

Here LULC Classification based on machine learning and deep learning have been studied and addressed

Paper[1] investigates the effect of spatial resolution unification approaches on Land Use and Land Cover (LULC) categorization accuracy for Sentinel-2 imagery. It compares upscaling and downscaling techniques and suggests that downscaling outperforms upscaling. The study includes additional spatial characteristics such as GLCM and EAPs, which increase classification accuracy greatly. Overall, results show that downscaling and incorporating spatial characteristics improves LULC classification for Sentinel-2 images.

Paper[2] presents ASPP-Unet and ResASPP-Unet are novel deep learning models for urban land cover classification using high-resolution satellite images. ASPP-Unet contains multi-scale features, whereas ResASPP-Unet adds residual units to the architecture. The models outperform previous models in terms of accuracy on WorldView-2 and WorldView-3 imagery of Beijing. These models provide accurate and effective results for urban land cover classification.

Paper[3] suggests a change to the mean shift algorithm for guaranteed convergence without a stopping requirement. The improved version restricts the amount of iterations and shows the convergence of the resulting sequence. Simulations show that it improves previous technique in estimating cluster centers.

Paper[4] uses remote sensing data to present a system based on deep learning for land cover and land use classification. In terms of accuracy, the proposed convolutional neural network (CNN) surpasses conventional techniques such as SVM, random forests, and k-NN. The CNN is faster for training and testing, and its efficacy has been proven by cross-validation studies and a GRSS competition. The proposed design does not address the issue of deep learning model interpretability, which is critical for many applications.

Paper[8] uses multitemporal Sentinel-2 data to evaluate the accuracy of four machine learning algorithms in categorizing land cover and land use over a diverse boreal region. The algorithms obtained equivalent overall accuracies that ranged from 0.733 to 0.758. According to the study, the Sentinel-2 Multispectral Instrument is unique among presently operating Earth-observation satellites due to its three red edge bands capable of capturing plant chlorophyll content and its medium-high spatial resolution of 10 m.

III. Methodology

A pretrained CNN based feature extraction framework with machine learning SoftMax classifier is proposed to extract the discriminating information in hyperspectral images to improve classification performance. The pretrained VGG16, InceptionV3, Resnet50 Models captures features with/without feature reduction techniques using Modified Mean Shift Algorithm(MMSA)[3]. Furthermore, SoftMax machine learning classifier is employed to get an effective prediction. The block diagram of the proposed framework is shown in Figure 1.

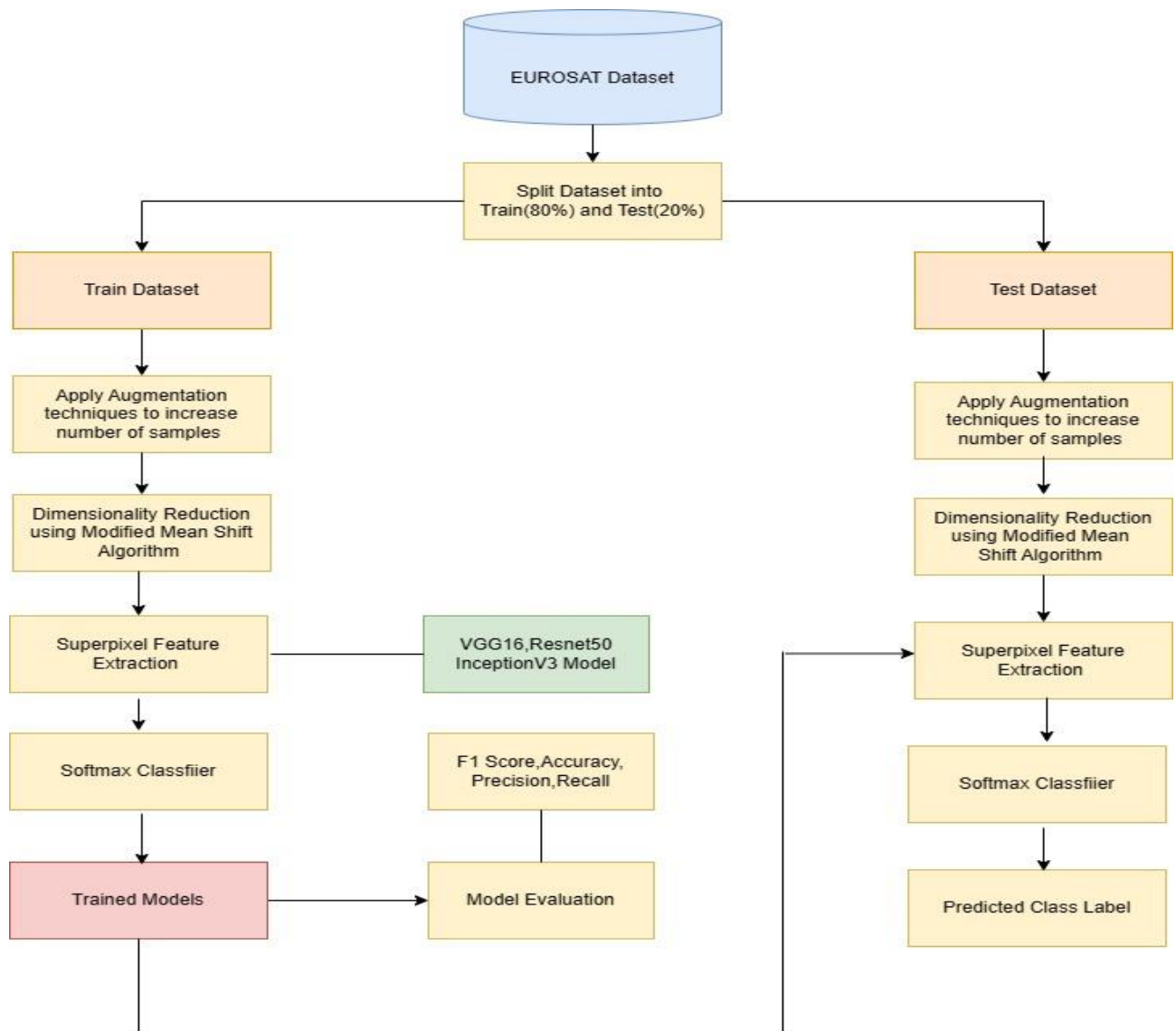


Fig 1: proposed framework for LULC classification

The hyperspectral dataset in the Live Aerial Image Hyperspectral Dataset is composed of spectral pictures, which are two-dimensional vectors of pixels in picture format. The framework's implementation is divided into two phases: training and testing, with the dataset requiring an 80-20% sample split for this process. Preprocessing for the spectral band includes extracting pixels and ground truth data from raw data, as well as scaling or shrinking the spectral image. The MMSA Algorithm is implemented to select superpixel features from preprocessed images. Following are the steps for MMSA Algorithm.

Steps for Modified Mean Shift Algorithm(MMSA):

1. Input: Preprocessed EuroSAT image (RGB).
- 2 Initialization:
 - Choose a window size or radius.
 - Set the convergence threshold and maximum number of iterations.

- Define the spatial bandwidth and color bandwidth parameters.
 - Select an initial set of seed points within the image using regular grid technique.
3. Mean Shift Iteration:
- For each seed point:
 - Calculate the mean shift vector by considering the color and/or spatial information within the defined window.
 - Update the seed point by shifting it in the direction of the mean shift vector.
 - Repeat the mean shift calculation and update until convergence or the maximum number of iterations is reached.
4. Super pixel Generation:
- Assign each pixel in the image to the nearest seed point or cluster center based on color or spatial similarity and Build a super pixel map, where each pixel is labeled with the corresponding super pixel identifier.
5. Output: Super pixel map, where each pixel is labeled with a super pixel identifier.

After super pixel map generation, VGG16, InceptionV3, and Resnet50 models are trained individually by following setting hyperparameters to get best model weights .

Loss Function	Categorical cross-entropy loss function
Optimizer	Adam optimizer
EarlyStopping mechanism	monitor='val_categorical_accuracy', patience=10, restore_best_weights=True, mode='max'
ReduceLROnPlateau Mechanism	monitor='val_categorical_accuracy', factor=0.5, patience=3, min_lr=0.00001

All Models are evaluated on test dataset by using performance metrics like accuracy, precision, recall and F1-Score.

IV. Experimental Result And Discussion

Experimental Setup

All of the trials were carried out using a Google Colab with 16 GB of RAM, GPU Tesla T4. Keras, Tensorflow, and the Scikit-Learn toolkit are used to achieve diverse features and classifications.

Dataset Description

The EuroSAT dataset is based on Sentinel-2 satellite images spanning 13 spectral bands and consisting of 10 categories with 27000 recognized and geo-referenced samples. There are two datasets provided: - rgb: only the optical R, G, and B frequency bands are represented as JPEG images. - all: All 13 bands in the original value range (float32) are included. The rgb version dataset is used in this paper. Annual Crop, Forest, Herbaceous Vegetation, Highway, Industrial, Pasture, Permanent Crop, Residential, River, Sea Lake are the 10 distinct categories. Sample Images are shown in Figure 2.

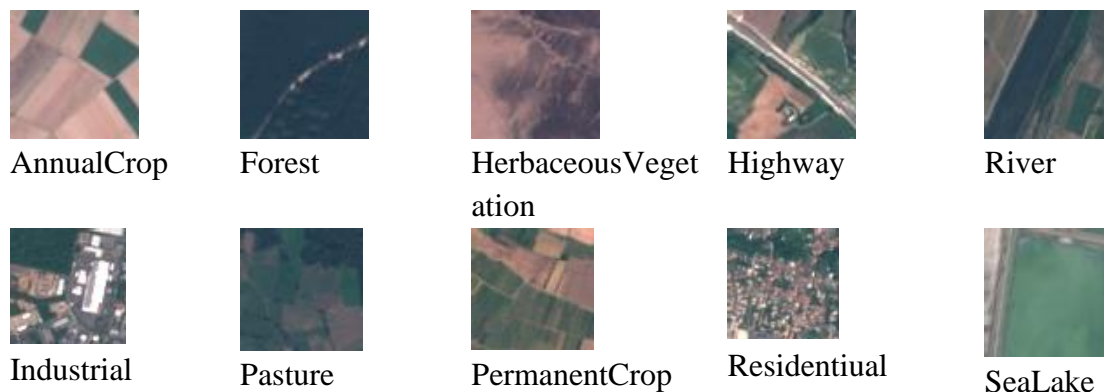


Figure 2 EuroSAT Sample Images

Model

In this study ,the pretrained VGG16,Inception v3 and Resnet50 Models pretrained on ImageNet dataset have been used by only weights are updating during training, the architecture of the model remains the same.

Performance Analysis

The performance of with and without MMSA combination with pretrained models is evaluated on testing dataset and shown in in table 2-8. Table 1 shows class numbers and its abbreviation

Table 1 Class Labels

Class	Abbreviation
	AnnualCrop
	Forest
	HerbaceousVegetation
	Highway
	Industrial
	Pasture
	PermanentCrop
	Residential
	River
	SeaLake

Table 2 Results of Training and Testing of Remote Sensing images on VGG16 pretrained network:

Confusion Matrix											
	Class	1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	95.0	0.0	0.0	0.2	0.0	0.3	5.2	0.0	0.2	0.2
	2	0.0	99.7	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.0
	3	0.3	0.5	95.8	0.2	0.0	1.0	1.0	1.2	0.0	0.0

	4	0.5	0.0	0.0	80.2	0.8	0.0	0.2	0.3	1.3	0.0
	5	0.0	0.0	0.0	0.0	82.7	0.0	0.0	0.7	0.0	0.0
	6	0.2	0.8	0.8	0.0	0.0	64.3	0.5	0.0	0.0	0.0
	7	0.0	0.0	1.2	0.2	0.5	0.0	81.3	0.2	0.0	0.0
	8	0.0	0.0	0.0	0.0	0.2	0.0	0.0	99.8	0.0	0.0
	9	0.3	0.2	0.0	1.3	0.2	0.0	0.0	0.0	81.3	0.0
	10	0.3	0.5	0.3	0.0	0.0	0.0	0.0	0.0	0.5	98.3
		True_classes									

Table 2 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2, 3, 8, and 10 is higher than 95%. The overall right classification rate is 87.75%, with an incorrect classification rate of 12.2%.

Table 3 Results of Training and Testing of Remote Sensing images on Resnet50 pretrained network:

Confusion Matrix											
		1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	95.0	0.2	0.0	0.5	0.0	0.5	2.7	0.0	1.2	0.0
	2	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.2	2.0	93.3	0.7	0.0	0.3	2.3	1.2	0.0	0.0
	4	0.3	0.0	0.5	79.0	0.3	0.0	0.3	1.2	1.7	0.0
	5	0.2	0.0	0.0	0.2	78.0	0.0	0.2	4.8	0.0	0.0
	6	1.2	2.2	1.3	0.3	0.0	60.7	0.5	0.0	0.5	0.0
	7	1.3	0.0	1.8	0.3	0.2	0.0	78.5	1.2	0.0	0.0
	8	0.2	1.2	0.0	0.0	0.0	0.0	0.0	98.7	0.0	0.0
	9	0.8	0.2	0.2	2.3	0.0	0.2	0.0	0.0	79.7	0.0
	10	0.0	2.5	0.0	0.2	0.0	0.0	0.0	0.0	0.8	96.5
		True_classes									

Table 3 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2 and 3 is higher than 95%. The overall right classification rate is 85.9%, with an incorrect classification rate of 14.14%.

Table 4 Results of Training and Testing of hyperspectral images on InceptionV3 pretrained network:

Confusion Matrix											
	Class	1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	93.3	0.2	0.0	0.2	0.0	0.5	5.3	0.0	0.3	0.2
	2	0.0	99.5	0.0	0.0	0.0	0.3	0.0	0.0	0.2	0.0
	3	0.5	1.5	94.2	0.2	0.0	0.8	1.7	1.2	0.0	0.0
	4	0.8	0.0	0.2	84.2	0.3	0.0	0.0	0.3	1.2	0.0
	5	0.0	0.0	0.0	0.0	82.0	0.0	0.0	1.3	0.0	0.0
	6	0.3	0.7	1.0	0.0	0.0	67.2	0.5	0.0	0.0	0.0
	7	0.5	0.0	1.3	0.3	0.2	0.0	80.7	0.2	0.2	0.0
	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
	9	0.0	0.2	0.0	1.0	0.0	0.0	0.2	0.0	82.0	0.0
	10	0.5	0.3	0.3	0.0	0.0	0.0	0.0	0.0	1.2	98.3
	True_classes										

Table 4 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2, 3, 8 and 10 is higher than 93%. The overall right classification rate is 88.18%, with an incorrect classification rate of 11.82%.

Table 5 Results of Training and Testing of Remote sensing images on Modified Mean shift Algorithm + VGG16 pretrained network:

Confusion Matrix											
		1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	95.0	0.0	0.0	0.2	0.0	0.3	5.2	0.0	0.2	0.2
	2	0.0	99.7	0.0	0.0	0.0	0.2	0.0	0.0	0.2	0.0
	3	0.3	0.5	95.8	0.2	0.0	1.0	1.0	1.2	0.0	0.0
	4	0.5	0.0	0.0	83.0	0.8	0.0	0.2	0.3	1.3	0.0
	5	0.0	0.0	0.0	0.0	84.0	0.0	0.0	0.7	0.0	0.0
	6	0.2	0.8	0.8	0.0	0.0	66.0	0.5	0.0	0.0	0.0
	7	0.0	0.0	1.2	0.2	0.5	0.0	81.3	0.2	0.0	0.0
	8	0.0	0.0	0.0	0.0	0.2	0.0	0.0	99.8	0.0	0.0
	9	0.3	0.2	0.0	1.3	0.2	0.0	0.0	0.0	81.3	0.0
	10	0.3	0.5	0.3	0.0	0.0	0.0	0.0	0.0	0.5	98.3
	True_classes										

Table 5 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2, 3, 8 and 10 is higher than 95%. The overall right classification rate is 88.63%, with an incorrect classification rate of 11.36%.

Table 6 Results of Training and Testing of hyperspectral images on Modified Mean shift Algorithm + Resnet50 pretrained network:

Confusion Matrix											
		1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	95.0	0.2	0.0	0.5	0.0	0.5	2.7	0.0	1.2	0.0
	2	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.2	2.0	93.3	0.7	0.0	0.3	2.3	1.2	0.0	0.0
	4	0.3	0.0	0.5	85.0	0.3	0.0	0.3	1.2	1.7	0.0
	5	0.2	0.0	0.0	0.2	78.0	0.0	0.2	4.8	0.0	0.0
	6	1.2	2.2	1.3	0.3	0.0	65.0	0.5	0.0	0.5	0.0
	7	1.3	0.0	1.8	0.3	0.2	0.0	78.5	1.2	0.0	0.0
	8	0.2	1.2	0.0	0.0	0.0	0.0	0.0	98.7	0.0	0.0
	9	0.8	0.2	0.2	2.3	0.0	0.2	0.0	0.0	79.7	0.0
	10	0.0	2.5	0.0	0.2	0.0	0.0	0.0	0.0	0.8	96.5
		True_classes									

Table 6 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2, 3, 8 and 10 is higher than 93%. The overall right classification rate is 86.96%, with an incorrect classification rate of 13.03%.

Table 7 Results of Training and Testing of Remote Sensing images on Modified Mean shift Algorithm + InceptionV3 pretrained network:

Confusion Matrix											
		1	2	3	4	5	6	7	8	9	10
Predicted_classes	1	93.3	0.2	0.0	0.2	0.0	0.5	5.3	0.0	0.3	0.2
	2	0.0	99.5	0.0	0.0	0.0	0.3	0.0	0.0	0.2	0.0
	3	0.5	1.5	94.2	0.2	0.0	0.8	1.7	1.2	0.0	0.0
	4	0.8	0.0	0.2	88.0	0.3	0.0	0.0	0.3	1.2	0.0
	5	0.0	0.0	0.0	0.0	82.0	0.0	0.0	1.3	0.0	0.0
	6	0.3	0.7	1.0	0.0	0.0	73.0	0.5	0.0	0.0	0.0
	7	0.5	0.0	1.3	0.3	0.2	0.0	80.7	0.2	0.2	0.0
	8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
	9	0.0	0.2	0.0	1.0	0.0	0.0	0.2	0.0	82.0	0.0
	10	0.5	0.3	0.3	0.0	0.0	0.0	0.0	0.0	1.2	98.3
		True_classes									

Table 7 shows the accuracy of classification for each distinct class. We can easily see that the classification accuracy for class 1, 2, 3, 8 and 10 is higher than 93%. The overall right classification rate is 89.1%, with an incorrect classification rate of 10.9%.

Table 8 Comparative Analysis of All Results

Methods	Accuracy	Precision	Recall	F1-Score
Resnet50 Model	85.93	0.9687	0.9670	0.9675
VGG16 Model	87.85	0.9746	0.9748	0.9745
Inception3 Model	88.18	0.9711	0.9800	0.9887
Modified Mean shift+Resnet50 Model	86.96	0.9789	0.9710	0.9790
Modified Mean shift +VGG16 Model	88.63	0.9866	0.9868	0.9865
Modified Mean shift +InceptionV3 Model	89.10	0.9821	0.9910	0.9997

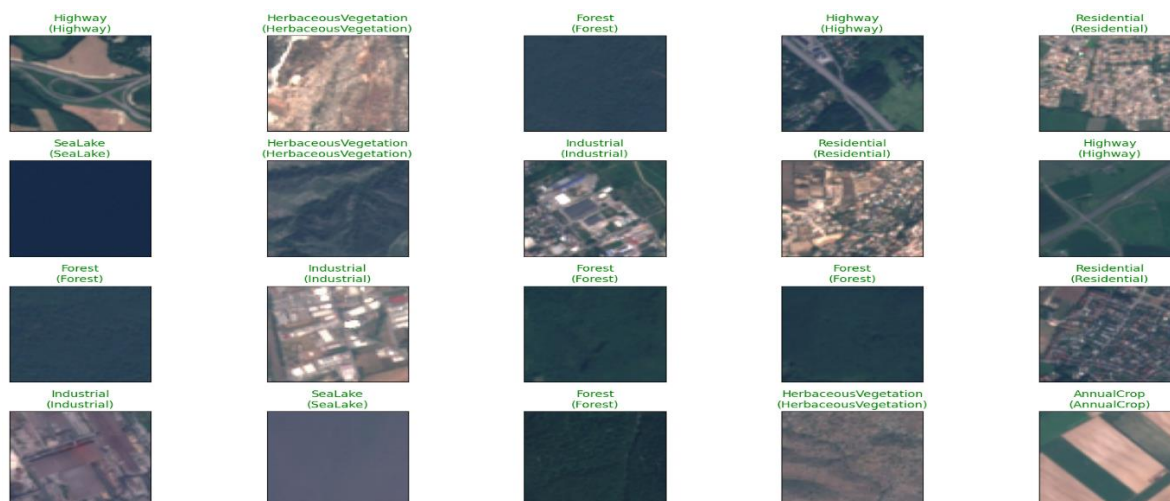


Figure 3 EuroSAT Classified Images

V. Conclusion

The proposed framework compares various Pretrained Models to fully use the discriminant feature extraction power of a pre-trained deep neural network model as well as the classification capabilities of machine learning softmax classifier models. Features are extracted from multilayer convolutional and fully connected layer, the feature extractor in this framework are VGG16 Resnet50 and InceptionV3 pretrained models. Experiments on benchmark hyperspectral EuroSAT dataset show that Inception v3 Model outperforms in term of overall accuracy, with the highest OAs as 87.88%, which is 0.13% more than VGG16 Model(87.75%) and 1.95% more than Resnet50 Model(85.93%).

To improve the classification performance, a Modified Mean Shift Algorithm is applied on every image and important features- superpixels are generated and store in NumPy array. Then VGG16, Resnet50 and InceptionV3 pretrained models have been applied on it. This strategy improves accuracy of individual models as shown in table 8. To further increase classification performance, Models can fine-tuned with different strategies . There is also a

more in-depth analysis of alternative network topologies, performance evaluation using various training techniques, and minibatch training adjustments.

References

1. Hongrui Zheng, Peijun Du and N. Yokoya, "Performance Evaluation of Downscaling Sentinel-2 Imagery for Land Use and Land Cover Classification by Spectral-Spatial Features", *Environmental Science Remote Sensing* 7 December 2017
2. Pengbin Zhang, Y. Ke, Zhenxin Zhang, Mingli Wang, Peng Li, Shuangyue Zhang "Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery", *Environmental Science, Mathematics Italian National Conference on Sensors* 1 November 2018
3. Youness Aliyari Ghassabeh, Frank Rudzicz, "Modified Mean shift algorithm", *IET Image Process.* 2018, Vol. 12 Iss. 12, pp. 2172-2177
4. Manuel Carranza-García, Jorge García-Gutiérrez, José Cristóbal Riquelme Santos "A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks", *Computer Science, Environmental Science Remote Sensing* 30 January 2019
5. Nagesh Kumar Uba, "Land Use and Land Cover Classification Using Deep Learning Techniques" *Environmental Science, Mathematics, arXiv.org* 1 May 2019
6. Agrawal, S. A., Umbarkar, A. M., Sherie, N. P., Dharme, A. M., & Dhabliya, D. (2021). Statistical study of mechanical properties for corn fiber with reinforced of polypropylene fiber matrix composite. *Materials Today: Proceedings*, doi:10.1016/j.matpr.2020.12.1072
7. Anupong, W., Azhagumurugan, R., Sahay, K. B., Dhabliya, D., Kumar, R., & Vijendra Babu, D. (2022). Towards a high precision in AMI-based smart meters and new technologies in the smart grid. *Sustainable Computing: Informatics and Systems*, 35 doi:10.1016/j.suscom.2022.100690
8. Anupong, W., Yi-Chia, L., Jagdish, M., Kumar, R., Selvam, P. D., Saravanakumar, R., & Dhabliya, D. (2022). Hybrid distributed energy sources providing climate security to the agriculture environment and enhancing the yield. *Sustainable Energy Technologies and Assessments*, 52 doi:10.1016/j.seta.2022.102142
9. Eleni Kroupi, Maria Kesa, Victor Diego Navarro-Sánchez, Salman Saeed, Camille Pelloquin, Bahaa Alhaddad, Laura Moreno, Aureli Soria-Frisch, and Giulio Ruffini "Deep convolutional neural networks for land-cover classification with Sentinel-2 images," *Journal of Applied Remote Sensing* 13(2), 024525 (20 June 2019). <https://doi.org/10.1117/1.JRS.13.024525>
10. Patrick Helber, Benjamin Bischke, Andreas Dengel, and Damian Borth EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification, *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, VOL. 12, NO. 7, JULY 2019
11. A. Abdi, "Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data", *Environmental Science GIScience & Remote Sensing* 22 August 2019

12. A. Vali, S. Comai, M. Matteucci "Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review", Environmental Science Remote Sensing 3 August 2020
13. M. Salwa Thasveen, S. Suresh, "Land - Use and Land - Cover Classification Methods: A Review" Environmental Science, Mathematics International Conference on Management 18 November 2021