International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Classification of Sentinel 2 Images using Customized Convolution Neural Networks

Geetha M*[1], Asha Gowda Karegowda[2], Nandeesha[3], Nagaraj B V[4]

Submitted: 22/10/2022 **Revised**: 25/12/2022 **Accepted**: 07/01/2023

Abstract: With the development of Convolutional Neural Networks and increased processing power in recent years, the discipline of deep learning and machine learning has made significant advancements. One of the most important networks in the deep learning space is the Convolutional Neural Network. In computer vision and natural language processing, convolutional neural networks have achieved remarkable successes. Based on the land usage and land cover of the specific area, satellite images are useful to constantly monitor. Classifying satellite images using cutting-edge deep learning is one of the potential and difficult tasks of remote sensing. Three of the most popular Convolutional Network models viz., Custom Architecture, VGG16, and Resnet34, were utilised for classification in order to assess and investigate deep learning convolutional models utilising satellite data. The multispectral Sentinel 2 image with its 13 spectral bands served as the training data for these three models. The dataset of the study area was created manually and the featured images were classified into six classes. The accuracy for VGG16 was found to be 90.70% and that using Resnet34 and custom architecture was respectively 91.50% and 93.73%, thus demonstrating the fact that Custom architecture produces more accurate results than the other two transfer learning techniques.

Keywords: *Convolution Neural Network, Deep Learning, Remote Sensing, Sentinel 2*

1. Introduction

In the modern world, remote sensing is essential [1]. It helps and supports society and the government in a variety of areas, including geology, agriculture, disaster management, and flood detection. Land use Land cover classification (LULC) is one such area that aids government decision-making for long-term planning. It is necessary to research and keep track of changes in land cover and land use because human land use has a significant impact on the environment [2]. Numerous applications, such as resource management, planning, and environmental monitoring, all make use of precise land cover maps [3]. The wide range land cover images needed to trace human activity and biophysical changes on earth's surface are evolved using satellite data [4, 5]. The accessibility to the development of novel methods to map land cover using remotely sensed data, such as Landsat data, has been driven by the availability of satellite data [6, 7]. Landsat photos are ideal for measuring the land cover because they have been available since 1972, and Sentinel 2 satellite data are frequently used to predict urban change [8, 9]. To create accurate land cover mapping, numerous automatic and semi-automated classification approaches have recently been developed [10]. Classifying land use and cover is a possible and important remote sensing task [3]. The speed and accuracy of categorization are increased when cutting-edge Deep Learning methods like Convolutional Neural Networks are used.

10 January 20231Department of MCA, BIET, Davanagere - 577004, INDIA
ORCID ID: 0000-0002-8762-980X
2Department of MCA, SIT, Tumakuru - 572103, INDIA
3Department of Civil Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, INDIA,
4Department of Management Studies, KIAMS, Harihara, Karnataka, INDIA
Corresponding Author Email: bmgeethamallad027@gmail.coma

Image classification, a key technique in the study of remote sensing images, is used to categorise pixels into groups based on how much of each category they contribute to the surface land cover. But classifying land cover photos is difficult, in large part because it is expensive and difficult to gather the training data needed to build reliable classifiers [4]. The classic image categorization techniques were used at the pixel level.But such algorithms are not suitable for land cover categorization based on Landsat photos.

The remote land cover's images are too complex to classify at the pixel level [11]. Deep learning has gained more attention recently and advanced a number of computer vision applications, including object identification [12], natural picture classification [13], categorization of remote sensing image [15, 16], and object categorization.

One deep learning structure that is frequently used in the classification of satellite images is the deep convolutional neural network. By adjusting the convolutional layer and pooling layer parameters, a CNN automatically extracts picture features. Although the training set for classic machine learning algorithms normally consists of a few hundred samples, CNN can learn from a very large dataset. These algorithms converge very slowly or don't converge at all as the training sets get bigger. CNNs can currently be used in many image categorization applications.

In this study, a Sentinel 2 image classification method based on Residual Neural Network (ResNet34), Custom architecture, and Visual Geometry Group (VGG-16) ConvNets is proposed. In order to train ResNet34 and VGG-16 ConvNets, more than 500 manually derived images from Sentinel 2 images were used.

2. Related Work

This section reviews about previous studies in land use and land cover classification. In this context, we present remotely sensed aerial and satellite image datasets. Furthermore, we present state-of-the-art image classification methods for land use and land cover classification.

Reham Gharbia et.al proposed a methodology for land use land cover classification using VGG-16 and AlexNet, the two deep CNN models on Landsat dataset. They have trained the models for approximately 500 image. They have derived the dataset manually using landsat5 and categorized into five different classes namely Agriculture, Desert, Roads, Urban and Water. The classification accuracy obtained was 74.8% using AlexNet and 90.2% using VGG16. Later by using Augmentation technique, they obtained 3500 images which results into 90.0% using AlexNet and 94.6% using VGG-16. The results shows that VGG-16 ConvNet achieved higher accuracy for both testing and augmented dataset [18].

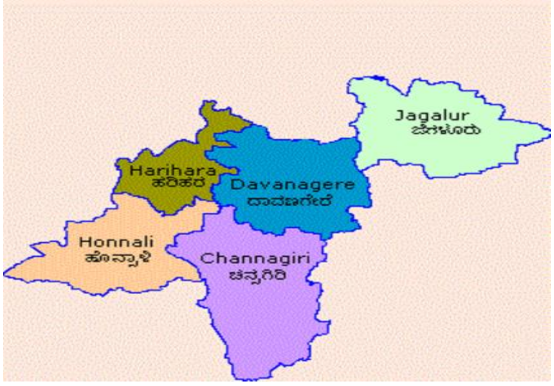
M. Voelsen et.al used remote sensing data for pixel based classification by combining Sentinel 2 images with data from geospatial database of LGLN. The data was trained by FCN to classify 4 land covers in Germany. It was found that the errors in trained data samples will not have larger influence on classifiers as the noise is randomly distributed [19].

Kavita Bhosle and Vijaya Musande have proposed a model to evaluate DL CNN models for classification of LULC and crop identification using EO-1 Hyperion hyperspectral images. They have proved that the accuracy of classification not only depends on the dataset but also depends on other factors like Optimizer, Activation function, Learning Rate, Filter Size and Batch Size also. It was found that deep learning CNN using optimized combination of parameters has provided 97.58% accuracy for the Indian Pines dataset, while 79.43% accuracy for the study area to justify that CNN is suitable in practice for both unstructured and small dataset [20].

Thus multiple studies on satellite images to accurately classify the land cover with the application of state of the art deep CNN have provided classification accuracy of above 90%. The present study is intended to apply three deep learning models like Customized CNN, VGG 16 and ResNet34 to get ascertain comparative accuracy of classification in the study area.

3. Proposed Modelling

Study area covers an area of 958 sq.kms of Davanagere, a central region of Karnataka State. It is located between north latitude 140 13’11.0” and 140 33’52.3” & east longitude 750 48’53.9” and 760 09’28.3”.



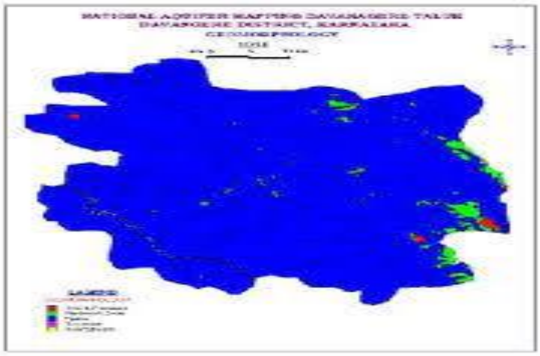


Fig 1. Study Area Davangere Taluk

3.1 Dataset Acquisition

In this study, the difficulty of classifying Land Use and Land Cover was addressed using Sentinel 2, a multispectral satellite image, which consists of 13 spectral bands with varying resolutions of 10 m, 20 m, and 60 m and a five-day return cycle, to record the entire surface of the earth's landmass. Sentinel 2's attributes are listed in Table 1 below. Four of the 13 bands—B2, B3, B4, and B8—with a 10m resolution were used in our study.

The classification of Land Cover will be affected by the voluminous satellite data when any machine learning or deep learning algorithms ate employed. Here, making use of the cutting-edge deep learning algorithms, Sentinel 2 data was collected to create the vast amount of data to increase the efficiency of the algorithm being used.

Two of the following steps were performed in order to construct the labelled dataset for image classification.

- Acquisition of Satellite Image: Acquiring cloud free, orthorectified, and level 1C sentinel 2 image for the study area.
- Dataset Creation: From the acquired data, dataset of 18000 labelled and geo-referenced image patches, each of size 64×64 pixel were generated and checked manually.

3.2 Satellite Image Acquisition

Sentinel 2 satellite image was acquired during April 2021 for the study area i.e., Davanagere region of Karnataka. The raw satellite images were collected from Copernicus which was orthorectified and atmospheric corrected. The acquired image was pre-processed to obtain False Colour Composite (FCC) image and then mosaicked to cover the entire area of interest (AOI) which mainly recorded under four tiles, finally clipped to obtain study area using GIS software and then cropped to small portions of 64 × 64 resolution and segregated to separate classes.

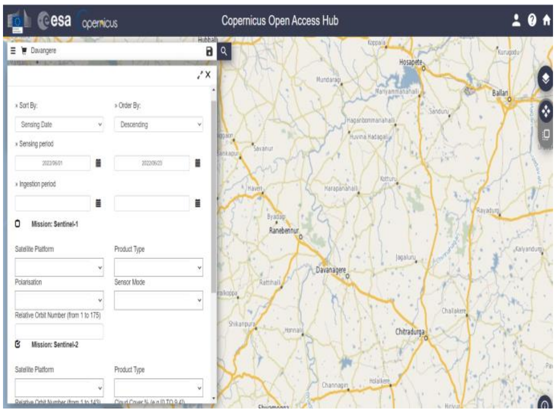


Fig 2. Acquisition of Sentinel 2

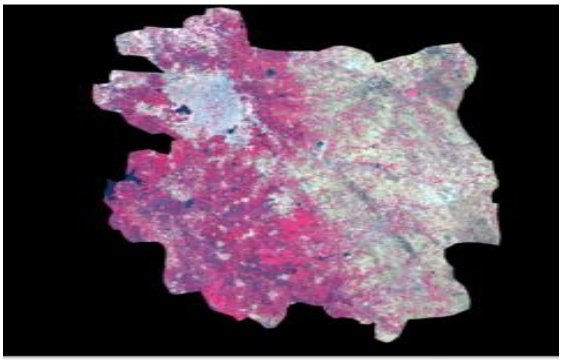


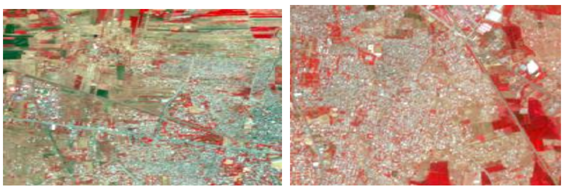
Fig 3. Study Area after Clipping

3.3 Creation of dataset

The dataset consists of 18000 labelled images of 6 classes. Each class contains 2000 images for training, 500 images for validation and 500 images for testing. Fig 2 and Fig 3 respectively shows the acquired satellite image and the obtained study area.

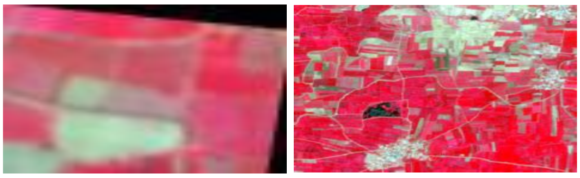
Further, data augmentation technique was used to increase the amount of datasets by adding slightly modified copies of existing images as it helps to increase the number of samples in the dataset as well as to get the best accuracy and to generalize the model for better test data. Fig.4 and Fig.5 show the sample image patches of six classes, each image is of 64 ×64 pixels.

(a)Waterbody



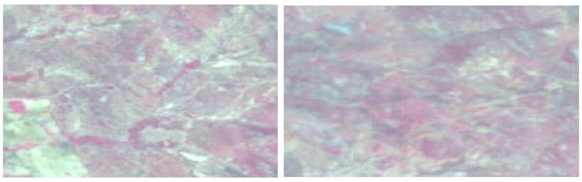
(b) BuiltUp

(c) Paddy



(d) Maize

(e) Horticulture



(f) Barren

Fig 4. Image Samples of Land Use Land Cover classes each having 64×64 pixels

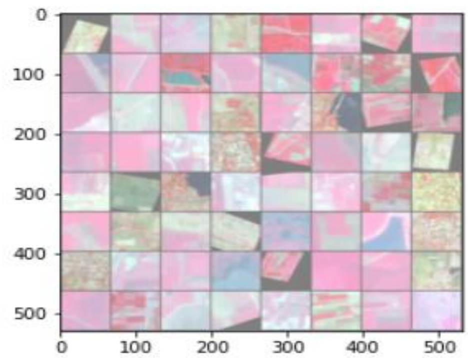


Fig 5. Sentinel 2 custom dataset patches of the study area

3.4 Processing Of Augmented Dataset Using Deep Learning Techniques

Augmented data set obtained is further processed for classification with high degree of accuracy using Convolutional Neural Network architecture involving (i) Custom neural network and transfer learning models such as (i) ResNet34 and (ii) VGG16.

4. Results and Discussion

4.1 Custom Neural Network Architecture

Custom CNN architecture comprises of 7 convolutional layers with kernel size 3 with increasing the depth of the network by doubling number of kernel every layer and stride 1 with valid padding so that input and output resolution of feature maps would same and 4 max pooling with kernel size 2 and stride2 for down sampling the feature maps. And one final fully connected layer with Softmax as a classification layer. In all the layers ReLu is used as an activation function and batch normalization. The Table 1 describes the model architecture related details and the figure 6 illustrates the convolutional neural network model with down sampling operations and workflow. For training the model we have experimented with different sets of hyper parameters, categorical cross entropy was used as a loss function, stochastic gradient descent (SGD) with learning rate 0.001, and momentum 0.9 with decaying the learning rate with every 7 steps. With these hyper parameters the model has trained with 50 epochs. Deep learning framework pytorch and google colab was used for experimenting purposes. The training accuracy was 99.692% and validation accuracy was 93.17% and each class level results are provided in the results section. The following Fig.6 shows the architecture diagram of Custom architecture.

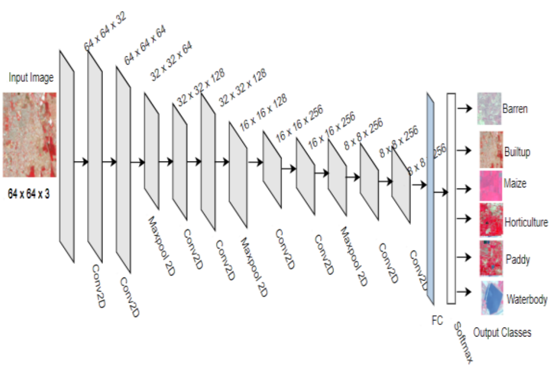


Fig 6. Custom Neural Network Architecture

4.2 Performance Evaluation of Models

Performance of classification is measured using parameters such as precision, recall and overall accuracy using the error matrix obtained for each of the model/technique used as follows.

Precision is measure of a positive predictive value. It is a quantity of relative classes among the retrieved classes. It evaluates what amount of positive identifications are certainly correct.

$$\text{Precision} = \frac{\textit{Number of positive predictions}}{\textit{Total number of positive predicts}} = \frac{TP}{TP+FP}$$

Recall is a measure of actual significant classes drawn from the dataset.

$$\text{Recall} = \frac{\textit{Number of correct actual positives}}{\textit{Total number of actual positives}} = \frac{TP}{TP+FN}$$

Accuracy is measure of correct predictions classified by the model

$$\text{Accuracy} = \frac{\textit{Number of correct Predictions}}{\textit{Total number of Predictions}} = \frac{TP+TN}{TP+TN+FP+FN}$$

Table 1. Model architecture of Custom Neural Network

Sl.	Layer type	Filter size	Channel size	Stride	Padding	Input size
1	Conv2D	3x3	32	1	1	64x64x3
2	Conv2D	3x3	64	1	1	64x64x32
3	MaxPool2D	2x2	-	2	-	32x32x64
4	Conv2D	3x3	64	1	1	32x32x64
5	Conv2D	3x3	128	1	1	32x32x64
6	MaxPool2D	2x2	-	2	-	32x32x128
7	Conv2D	3x3	128	1	1	16x16x128
8	Conv2D	3x3	256	1	2	16x16x128
9	MaxPool2D	2x2	-	2	1	16x16x256
10	Conv2D	3x3	256	1	1	8x8x256
11	Conv2D	3x3	256	1	1	8x8x512
12	Fully Connected	-	-	-	-	32768

Fig 7 shows the error matrix or confusion matrix and Table 2 depicts its precision, Recall and overall accuracy of all classified classes using custom architecture.

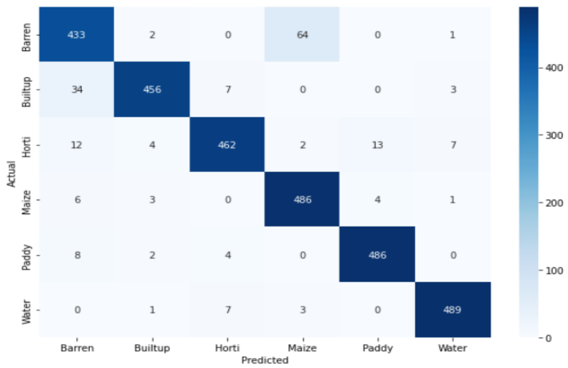


Fig 7. Confusion Matrix of Custom Architecture

Table 2. Accuracy Assessment using Custom Architecture

Sl. No	Class Name	Precision	Recall
1	Barren	0.868	0.866
2	Built Up	0.897	0.912
3	Horticulture	0.962	0.924
4	Maize	0.876	0.972
5	Paddy	0.966	0.972
6	Water	0.976	0.978

Over all Accuracy is 0.9733

4.3 Residual Neural Network - ResNet34

In this study, well-known designs to test a transfer learning method were used. ResNet34, a pre-trained architecture on the ImageNet dataset, is one of the architectures employed in the study. All of the ResNets versions share the same fundamental designs, with the exception that each pair of 33 filters has a shortcut connection added. Fig8 represents the architecture of ResNet34 and Fig 9 confusion matrix of the resultant model and Table 4 shows its Recall, Precision and accuracy of ResNet34.

Fig 8 represents the architecture of ResNet34 and Fig 9 confusion matrix of the resultant model and Table 3 shows its Recall, Precision and accuracy of ResNet34.

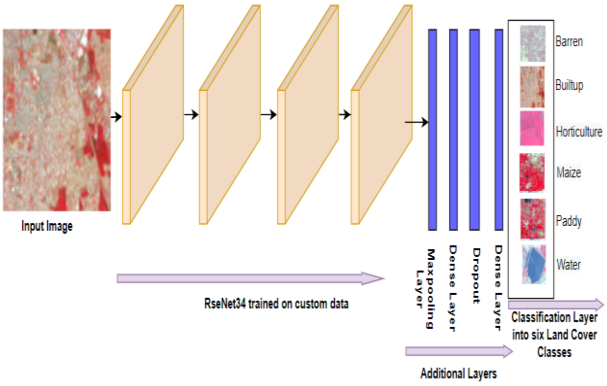


Fig 8. Custom Neural Network Architecture

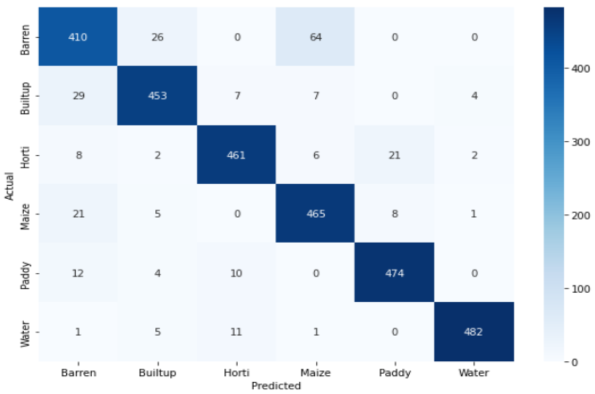


Fig 9. Confusion Matrix of ResNet34

Table 3. Precision & Recall of ResNet34

Sl. No	Class Name	Precision	Recall
1	Barren	0.852	0.820
2	Built Up	0.915	0.906
3	Horticulture	0.942	0.922
4	Maize	0.856	0.930
5	Paddy	0.942	0.948
6	Water	0.986	0.964

Over all Accuracy is 0.915

4.4 Visuaal Geometry Group - VGG16

Another transfer learning method with famous architecture is VGG16.Instead of large number of hyper parameters, it focuses on the convolutional (conv.) layers of 3 × 3 filters with stride 1 and padding, max pool layer of 2 × 2 filter with stride 2.Fig.10showsthe architecture of VGG16. The size of input image is 64 × 64 × 3 with RGB channels. The convolution and max pool layers are consistently arranged throughout the architecture.Conv-1 layer has 64 filters, layer 2 has 128 filters, layer 3 has 256 filters and layer 4 and 5 has 512 filters. Three fully connected layers have heap of convolutional layers. Softmax is the final layer consisting of six classes.

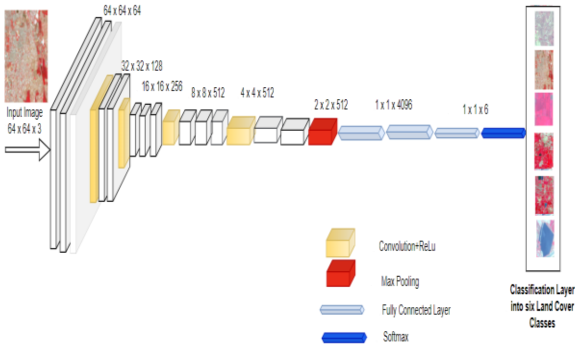


Fig 10. Architecture of VGG 16

The below Fig.11 and Table 4 represents confusion matrix and its Precision, Recall and its Accuracy obtained using VGG16 model.

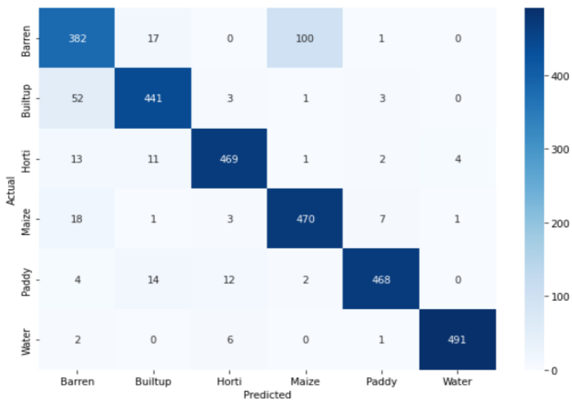


Fig 11. Confusion Matrix using VGG16

Table 4. Precision & Recall of VGG16

Sl. No	Class Name	Precision	Recall
1	Barren	0.811	0.764
2	Built Up	0.911	0.882
3	Horticulture	0.951	0.938
4	Maize	0.819	0.940
5	Paddy	0.971	0.936
6	Water	0.989	0.982

Overall Accuracy=0.907

Table 5. Comparison of three models using Precision and Recall

Algorithm	**Custom Architecture**		**ResNet34**		**VGG 16**	
	Precision	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**
Barren	0.868	0.866	0.852	0.820	0.811	0.764
Built Up	0.897	0.912	0.915	0.906	0.911	0.882
Horticulture	0.962	0.924	0.942	0.922	0.951	0.938
Maize	0.876	0.972	0.856	0.930	0.819	0.940
Paddy	0.966	0.972	0.942	0.948	0.971	0.936
Water	0.976	0.978	0.986	0.964	0.989	0.982
Overall Accuracy (%)	**93.73**		**91.50**		**90.70**	

The above Table 6 represents the comparison of all the three models which brief about the precision, Recall and Accuracy of all the six classes, which conclude that the overall accuracy of Custom Architecture is efficient compared to ResNet34 and VGG16.

5. Conclusion

In this research paper, the effort is made to apply and evaluate the different deep learning techniques like custom architecture and transfer learning models to address the accuracy issues of classifying land use and land cover studies using remote sensing (RS) images. As far as the approaches are concerned, the data activation and normalization is predefined in transfer learning models and that in custom architecture defined according to the accuracy needs of the researcher. Thus the overall accuracy was found to be 93.7%, 91.5%, and 90.7% respectively using Custom architecture, ResNet34 and VGG16. Though all three models seems to be consistent, classification using Custom architecture stands out with highest accuracy.

It may thus be concluded that deep learning tools classify the RS images to the best possible accuracy over machine learning models and custom architecture with necessary tools to increase the accuracy stands out to be the efficient model of classification when compared readily available transfer learning tools and hence used by decision makers in varied fields of application using RS images.

References

[1] H.Yassine1, K. Tout , and M. Jaber "Improving LULC Classification From Satellite Imagery Using Deep Learning - Eurosat Dataset", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLIII-B3-2021 XXIV ISPRS Congress (2021 edition).

[2] Basha, U.I., Suresh, U., Raju, G.S., Rajasekhar, M., Veeraswamy, G. and Balaji, E., "Land Use and Land Cover Analysis Using Remote Sensing and GIS: A Case Study in Somavathi River, Anantapur District, Andhra Pradesh, India", *Nature Environment & Pollution Technology*, 17(3), pp.1029-1033, 2018.

[3] Azzari, G. and Lobell, D.B., "Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring", *Remote Sensing of Environment*, 202, pp.64-74, 2017.

[4] Zhang, H.K. and Roy, D.P., "Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification", *Remote Sensing of Environment*, 197, pp.15-34.

[5] Turner, B.L., Lambin, E.F. and Reenberg, A., "The emergence of land change science for global environmental change and sustainability, *Proceedings of the National Academy of Sciences*, 104(52), pp.20666- 20671, 2017.

[6] Goldblatt, R., Stuhlmacher, M.F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C., Serrano- Candela, F.,

[7] Khandelwal, A.K., Cheng, W.H. and Balling, R.C" Using Landsat and night time lights for supervised pixel-based image classification of urban land cover", *Remote Sensing of Environment*, 205, pp.253-27, 2018.

[8] Gharbia, R., Hassanien, A.E., El-Baz, A.H., Elhoseny, M. and Gunasekaran, M., "Multi-spectral and panchromatic image fusion approach using stationary wavelet transform and swarm flower pollination optimization for remote sensing applications", *Future Generation Computer Systems*, 2018.

[9] Patel, N.N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F.R., Tatem, A.J. and Trianni, G.,"Multi temporal settlement and population mapping from Landsat using Google Earth Engine", *International Journal of Applied Earth Observation and Geo information*, 35, pp.199-208, 2015,

[10] Gharbia, R., El-Baz, A.H. and Hassanien, A.E., "*International Conference on Computing, Communication, and Automation (ICCCA)*, pp. 1080-1085. IEEE, 2016.

[11] Ban, Y., Jacob, A. and Gamba, P, "Space borne SAR data for global urban mapping at 30 m resolution using a robust urban extractor", *ISPRS Journal of Photogrammetry and Remote Sensing*, 103, pp.28-37, , 2015.

[12] Wu, S.S., Qiu, X., Usery, E.L. and Wang, L, "Using geometrical, textural, and contextual information of land parcels for the classification of detailed urban land use", *Annals of the Association of American Geographers*, 99(1), pp.76-98, 2009.

[13] Schmidhuber, J., "Deep learning in neural networks: An overview", Neural *networks*, 61, pp.85-117.

[14] Wu, X., Sahoo, D. and Hoi, S.C., "Recent advances in deep learning for object detection", *Neuro computing, 2020*. 2015

[15] He, K., Zhang, X., Ren, S. and Sun, J.,"Deep residual learning for image recognition", Proceedings of the *IEEE conference on computer vision and pattern recognition*, pp. 770-778. 10. 2016.

[16] Abdi, A.M., "Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data", *GIS Science& Remote Sensing*, 57(1), pp.1-20, 2020.

[17] Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. and Atkinson, P.M., "Joint Deep Learning for land cover and land use classification", *Remote sensing of environment*, 221, pp.173-187, 2019.

[18] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large scale image recognition",.*arXiv preprint* arXiv:1409.1556, 2014.

[19] Reham Gharbia Nour Eldeen M. Khalifa , and Aboul Ella Hassanien "Land Cover Classification using Deep Convolutional Neural Networks, Chapter · June 2021 DOI: 10.1007/978-3-030-71187-0_84.

[20] M. Voelsen1,∗ , J. Bostelmann , A. Maas , F. Rottensteiner , C. Heipke1 "Automatically Generated Training Data For Land Cover Classification With CNNs Using Sentinel-2 Images", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIII-B3-2020, 2020 XXIV ISPRS Congress (2020 edition).

[21] Kavita Bhosle and Vijaya Musande, "Evaluation of Deep Learning CNN Model for Land Use Land Cover Classification and Crop Identification Using Hyperspectral Remote Sensing Images", Journal of the Indian Society of Remote Sensing, 2019, https://doi.org/10.1007/s12524-019-01041-2

Author Contributions

Geetha M, DOB: 03-03-1978, Bellary, Karnataka, post graduated in MCA from Gulbarga University Gulbarga. Working as Assistant Professor, having teaching experience of 20 years. She has published two book chapters and three research papers in International Journals.

Dr. AshaGowda Karegowda DOB: 15-09-1971, Tumkur,Karnataka, received M.CA degree from Bangalore University, Karnataka, and M.Phil. Degree from Madurai Kamraj University, Tamil Nadu, Ph.D in Computer Science from Visvesvaraya Technological University, Belgaum, Karnataka. She is currently working as Associate Professor, in the Department of Master of Computer Applications, Siddaganga Institute of Technology, Tumkur, Karnataka since 1998. She is a board of study member and BOE member for few of the deemed and autonomous institutes in Karnataka. Her area of interest includes data mining, data analytics, data science, data structures, machine learning and bio-inspired computing. She has authored a few books, books chapters, published around 90+ papers in International Conferences & Journals and has one patent

Dr. Nandeesha DOB: 06-07-1960, Tumkur, having teaching experience of 37 years and post graduated in M.Sc Geology and Ph.D Hydrogeology from Bangalore University, Bangalore. He has published/presented 39 research papers in National/International Journals/Seminars. He is guiding 3 Ph.D students under Visvesvaraiah Technological University, Belaguam. His research funded projects are DST and VGST amounting to Rs. 65 lakhs on Groundwater Quality studies and Aquifer characterizations using Remote Sensing and GIS techniques.

Dr. Nagaraj B V DOB: 12-11-1973, Hospet, Karnataka, Awarded Ph.D from the University of Mysore. Presently working as Associate Professor at Kirloskar Institute of Advanced Management Studies, Harihar. He has a diverse background of both industry and academia. He was awarded doctorate in the area of agricultural value chains and supply chain systems. Basically being a statistician he has got experience in teaching Statistical Methods, Inferential Statistics, Operations Research & Bio Statistics. Besides teaching, his interest area is data analytics and research in the field of Agricultural marketing and innovation. He has published eight Research papers in reputed Journals of National and International repute.

Conflicts of interest

The authors declare no conflicts of interest.