Internatonal Journal of Science And Engineering
ISSN (Online): 2208-2204
Volume 11 Issue 01 May 2025
DOI: 10.53555/ephijse.v11i1.309

AUTOMATED LAND COVER MAPPING FROM LISS-III MULTISPECTRAL IMAGES USING SUPERVISED CLASSIFICATION WITH DEEP LEARNING

Nirav Desai[1*], Akruti Naik[1]

[1]D-UIAS AND D-SIM&C, Valsad, INDIA

***Corresponding author**
**Email: niravdesai.research@gmail.com*

Abstract
Remote sensing (RS) is the technique of finding and understanding information from a long distance or remote location using sensors. Land use and land cover mapping are fundamental tasks for planning and management. The deep neural network was used to perform the study. This study shows the semantic segmentation of LISS-III multispectral image using a fully convolutional network (FCN): U-net and Tiramisu. We present an innovative dataset, based on these LISS-III images that contained 4 different spectral bands (Band – 2 (Blue), Band-3 (Green), Band-4(Red), and Band-5 (Nearly Infrared), the false color composite (FCC) images and the ground truth mask images to classify total 4 classes (Surface water , Flora , Idle Land, and Housing zones) which holds 1470 labeled images where 1,255 images are used to train the model while 215 images are reserved for validation and evaluation. An FCN-based U-Net model with skip connections was trained on 256 × 256 × 3 input images to produce 256 × 256 × 4 one-hot encoded segmentation masks, representing four distinct land cover classes. Results confirmed that the U-Net architecture performs exceptionally well in semantic segmentation of LISS-III multispectral imagery for land use/land cover classification. The U-Net model successfully classified four land cover classes with an overall accuracy of 84%, significantly outperforming the Tiramisu architecture, which achieved a comparatively moderate accuracy of 52%.

Keywords: *remote sensing, LISS- III multispectral image, land use land cover classification, deep learning, Fully Convolutional Network (FCN), U-Net, Tiramisu*

1. INTRODUCTION

Remote sensing (RS) is the art of finding and understanding information from a long distance, using sensors without communication with the object being observed [19]. Land use land cover classification intends to form space-born images into a particular class, which was dependent on the distribution of recognized land use land cover classes. Land use and land cover mapping are fundamental tasks for planning and management. Artificial intelligence methods for pattern recognition and computer vision, for example, neural networks, K-means, Random Forests, and Support Vector Machines are an active way for remote sensing image classification. In 2006, deep learning theory was proposed by Hinton et al. [3]. Semantic Segmentation is defined as a pixel-level classification of images where a class is allotted to each pixel of the image.

The deep learning algorithms are emphasized by their importance on automatic feature knowledge from the datasets. In deep learning, mainly Convolutional Neural Networks (CNNs), have been effectively applied for image classification, target detection, and scene understanding [1][4][7][10][22]. In 2015, a fully convolutional neural network (FCN) was proposed by Long [11]. FCN is fully associated with layers in CNN with up-convolutional layers and concatenates with a shallow, finer layer to produce end-to-end labels. FCN is more appropriate for pixel-based image classification, i.e., labeling every pixel to a respective class. The FCN framework has also revealed great potential in remote sensing image classification. The U-Net model projected by [18] is an enhanced FCN model categorized by balanced U-shaped architecture containing a symmetric contracting path and expansive path. It combines low-level features with complete spatial information with high-level features with semantic information to expand segmentation accuracy. U-Net achieved good results in one-class segmentation tasks like road network mining or biomedical image segmentation and for multi-class labeling tasks such as land use land cover classification.

This study was performed in the South Gujarat region, Gujarat state in India. A total of 4 classes have been classified - Surface water , Flora , Idle Land, and Housing zones areas successfully in the study. A pixel-level classification of an image where every individual pixel is allotted to the respective class is called semantic segmentation of an image. In this study, the deep neural network was applied to handle this task. A fully-convolutional network (FCN) with skip connections is trained to take an input image of size 256 x 256 x 3 and outputs a matrix of shape 256 x 256 x 4 i.e., a one-hot encoded version of the mask. U-net and Tiramisu applied for the land use land cover classification of LISS-III multispectral images on a large dataset. The U-Net model achieved 84% of accuracy while Tiramisu achieved 51 % accuracy.

2. Survey of Literature

Several studies have explored deep learning approaches for aerial and satellite image analysis. For instance, the work in [16] introduced a CNN architecture tailored for aerial image labeling. In [6], a pre-trained CNN was employed to differentiate and organize distinct regions in high-resolution remote sensing imagery. The approach in [15] focused on extracting residential and building areas from ortho-imagery using a specialized deep learning model. A land cover classification method based on GaoFen-2 four-band satellite imagery in rural regions was proposed in [17]. The study in [21] presented a Fully Convolutional Network (FCN) framework for semantic labeling using the ISPRS Vaihingen and Potsdam benchmark datasets, which consist of publicly available false-color aerial images with a spatial resolution of 9 cm and digital surface models (DSMs) generated via LiDAR data. The comparative study conducted in [12–14] evaluated the performance of CNN and FCN architectures for aerial image labeling and also introduced a multilayer perceptron-based framework. Additionally, Fu et al. [2] developed a model for post-processing land cover classifications using imagery from both GF-2 and IKONOS satellites.

3. Approach and Techniques

3.1 Information repository

The study was performed in the South Gujarat Region, State of Gujarat, country INDIA. Multispectral space-born remote sensing images (IRS LISS – III) were used to perform the study. These multispectral remote sensing images have a total of 4 different bands in separate .tiff files and the number of bands is Band – 2,3,4 and 5(Blue, Green, Red, and near Infrared). These images have more than 100 nm resolution and less the 10 bands. These images contain a total of 4 bands and the spatial resolution is 30 meters. Quadrats of 30m X 30m size were placed across the study area. Data was acquired from the website https://bhuvan-app3.nrsc.gov.in which is provided by the ISRO. A field study was accomplished to collect environmental landscapes of different land use and land cover. The Latitude, Longitude for the location of the respective class are being verified and recorded which are called Ground Control Points (GCPs). GPS device used to collect GCPs. GPS Garmin – eTrex 30 is used for the study. The GCPs reserved for the respective category were reliant on the allocation of recognized land use classes within the study area.

3.2 Preliminary processing

Indian Remote Sensing (IRS) LISS - III multispectral images contain 4 bands, and these bands were combined and created a False Colour Composite (FCC) image. FCCs are designed using heaping these multiband.TIFF image files by the grouping of Band – 4 (Red), Band – 3 (Green), and Band – 2 (Blue). The ground truth masks were formed to train the model after the creation of FCCs for each image. These Masks were created using the maximum likelihood (ML) algorithm on the region of interest for each class. Figure I shows the FCC image and Figure II represents the ground truth mask.

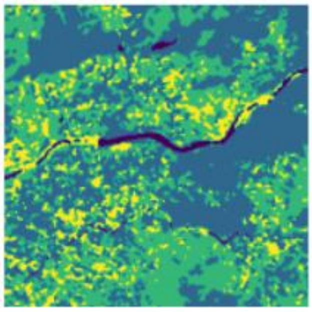


Figure I. FCC Image **Figure II. Ground Truth Mask**

The study was performed on a novel dataset that contains the FCC image of different seasons and Ground Truth Masks. FCCs and their corresponding masks are resized to 1024 * 1024 pixels. The size of dataset was 1470 images, where 1,255 images are used to train the model while 215 images are reserved for validation and evaluation.

3.3 Implementation Details

In the Maximum Likelihood (ML) method, each pixel is assigned to the class it most likely belongs to, based on statistical probability. The mean vector and covariance matrix, derived from the training data, are essential for this classification approach [20]. Fully Convolutional Networks (FCNs) are widely used for tasks like character recognition, image analysis, and medical image segmentation, especially for pixel-wise classification. FCNs replace fully connected layers with convolutional ones, preserving spatial information throughout the network [8]. In a typical FCN, the encoder extracts features while the decoder reconstructs the segmented output. For instance, the U-Net architecture takes a 256 × 256 × 3 image and produces a 256 × 256 × 4 output—representing a one-hot encoded mask with four classes.

The encoder has five blocks, each with two layers (convolution + batch normalization + ReLU), followed by max-pooling (except the last). The decoder includes four upsampling blocks, each beginning with upsampling, a 1×1 convolution, and skip connections that merge encoder features. Finally, a 1×1 convolution with four filters generates the final segmented output.

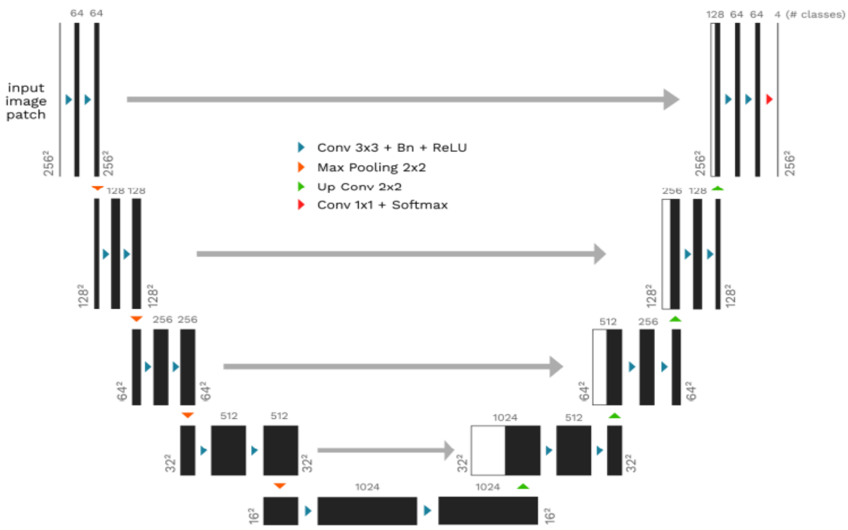


Figure III. U-Net Architecture

Figure.III shows the U-Net architecture [9]. The arrows symbolize the various processes, the black containers denote the feature map and the gray containers denote the cropped feature maps from the contracting path. Tiramisu is a polyhedral compiler for dense and sparse deep learning and data-parallel algorithms and directs a huge set of loop optimizations and data design alterations. It is the only open-source DNN compiler that optimizes sparse DNNs and marks distributed architectures. It can perform complex loop transformations and uses dependence investigation to assure the accuracy of optimizations. Tiramisu has also demonstrated its performance on various standards like deep learning operations (Convolution, ReLu, MaxPool, Sparse Neural Networks, etc.) and linear algebra. However, the Tiramisu network, which itself is a modified UNet, is much larger and took longer to train. Figure IV shows the tiramisu architecture. DenseNet input concatenates all previous feature outputs in a feedforward fashion for convolution.

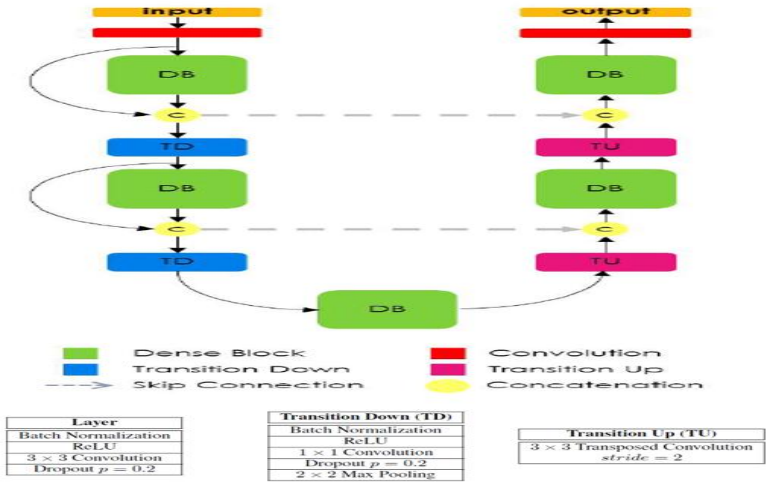


Figure IV. Tiramisu Architecture [23]

3.4 Model Training Setup

Before passing the remote sensing images and ground truth masks to the model for training, normalize the input images by clipping them while the masks are one-hot encoded according to the total number of classes i.e. 4. Random augmentations were also applied to the batch of images and masks before the model training. This expands the dataset and makes the model robust enough to encounter different orientations than just the training data. A custom image data generator is created to fulfill the requirements of this data ingestion pipeline. Table I represents the hyperparameters and other configurations used for training the model.

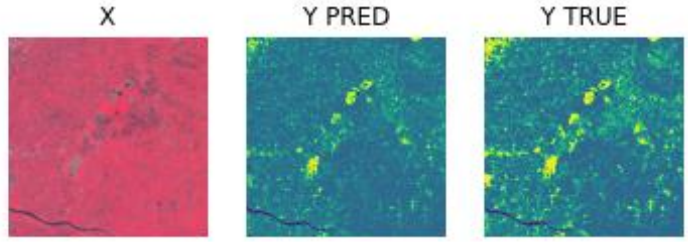
Hyperparameters & Configurations	Values
Train Batch Size	16
Validation Batch Size	16
Input Image Shape	(256,256,3)
Number of classes	4
Epochs	50
Loss	Categorical Focal Loss*
Optimizer	Adam
Metrics	Dice Coefficient*
Class Weights	[1.69941,0.53043,1.23977, 1.38949]

Table I. Hyperparameters and other configurations used for training the model.

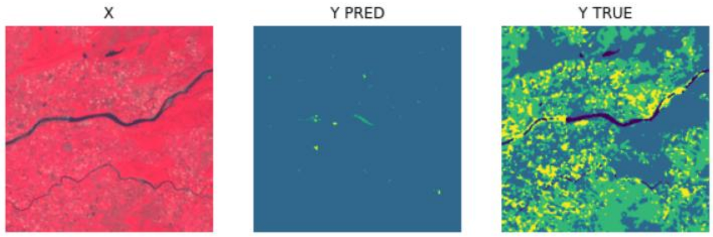
4. Experimental Findings and Discussion

The environment for the experiment is as follows: Windows 11 operating system and Envi 4.7 is used to process the remote sensing information, such as the selection of training sample, generation of masking, etc. The algorithm for classification developed in U-Net and Tiramisu was coded with Python + OpenCV.Figure 5 shows the classification of land cover land of the south Gujarat region, India. Water Surface are black , Flora are light green , Idle Land are light blue , and Housing zones are in yellow.

FCC image predicted by the models. In Figure V(a), X represents the FCC image, and Y PRED represents the result predicted by the U-net model. And Y TRUE represents the ground truth mask. In Figure V(b) X represents the FCC image, and Y PRED represents the result predicted by the Tiramisu model. And Y TRUE represents the ground truth mask. Table -II represents the experiment outcomes and accuracy with different epochs and different models. Experiment figured out the best fine-tuning parameters for U-Net and Tiramisu with RGB bands of the dataset of LISS- III multispectral remote sensing images. The parameters which give the best performance are used to develop the final model. The model also used data augmentation. And from the experiment results, it was detected that U-net gives better results in classifying land use land cover classes.



(a)



(b)

Figure V. Experiment result

MODEL	Optimizer	EPOCH Trained	Accuracy
U-Net	adam	50	84
Tiramisu	adam	50	52

Table II. Experiment Result

Figure VI shows the resultant image predicted by the model with the quantification for each class. Class 0 represents the water surface, 1 represents the Flora, 2 represents the Idle Land and 3 represents the Housing zones.

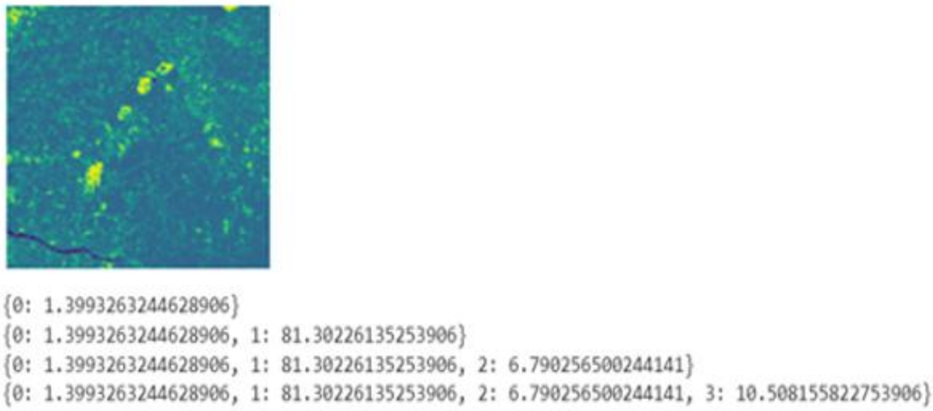


Figure VI: quantification for each class

5 Conclusion

With the proposed model, we achieved very good accuracy in the land use land cover classification using a deep learning approach. Deep learning for LULC classification becoming more evident. It will deliver a cost-effective and time management resolution than the visual understanding In the present study, we proposed a land use land cover classification model built on U-Net and tiramisu algorithms. The models were trained and tested on LISS-III multispectral space-born image dataset. Experiments show that model detected a total of 4 land use land cover classes i.e., Water body, vegetation, uncultivated land, and residential with very good accuracy. Outcomes confirmed that the U-Net classifier holds massive potential for accurate detection of land use land cover classes than tiramisu. The U-Net model achieves a very good accuracy of 84 % while Tiramisu achieved an accuracy of 52 %.

References:

[1]. Codts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke,W.; Roth, S.; Schiele, B.The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 3213–3223.

[2]. Fu, G.; Liu, C.; Zhou, R.; Sun, T.; Zhang, Q. Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional Network. Remote Sens. **2017**, 9, 498.

[3]. Fu, G.; Zhao, H.; Li, C.; Shi, L. Segmentation for High-Resolution Optical Remote Sensing Imagery Using Improved Quadtree and Region Adjacency Graph Technique. Remote Sens. **2013**, 5, 3259–3279.
[4]. Hinton, G.; Osindero, S.; Welling, M.; Teh, Y.-W. Unsupervised Discovery of Nonlinear Structure Using Contrastive Backpropagation. Science **2006**, 30, 725–732.
[5]. Hosseiny, B.; Mahdianpari, M.; Brisco, B.; Mohammadimanesh, F.; Salehi, B. WetNet: A Spatial-Temporal Ensemble Deep Learning Model for Wetland Classification Using Sentinel-1 and Sentinel-2. IEEE Trans. Geosci. Remote Sens. 2021, 1–14,
[6]. Hu, F.; Xia, G.S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sens. **2015**, 7, 14680–14707.
[7]. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the Neural Information Processing Systems (NIPS) Conference, La Jolla, CA, USA,3–8 December 2012.
[8]. Kumar, S., Kumar, N., Rishabh, I. K., & Keshari, V. (2021). Automated brain tumour detection using deep learning via convolution neural networks (CNN). Int. J. Cur. Res. Rev, 13(02), 148.
[9]. Längkvist, M.; Kiselev, A.; Alirezaie, M.; Loutfi, A. Classification and segmentation of satellite orthoimagery using convolutional neural networks. Remote Sens. **2016**, 8, 329.
[10]. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature **2015**, 521, 436–444.
[11]. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 5–7 June 2015; pp. 3431–3440.
[12]. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Fully Convolutional Neural Networks for Remote Sensing Image Classification. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 5071–5074.
[13]. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification. IEEE Trans. Geosci. Remote Sens. **2017**, 55, 645–657.
[14]. Maggiori, E.; Tarabalka, Y.; Charpiat, G.; Alliez, P. High-resolution aerial image labeling with convolutional neural networks. IEEE Trans. Geosci. Remote Sens. **2017**, 55, 7092–7103.
[15]. Maltezos, E. Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds. J. Appl. Remote Sens. **2017**, 11, 1–22.
[16]. Mnih, V.; Hinton, G.E. Learning to detect roads in high-resolution aerial images. In Computer Vision—ECCV 2010, Lecture Notes in Computer Science; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6316, pp. 210–223.
[17]. Pan, X.; Zhao, J. A central-point-enhanced convolutional neural network for high-resolution remote-sensing image classification. Int. J. Remote Sens. **2017**, 38, 6554–6581.
[18]. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
[19]. Sarah C. Goslee" Analyzing Remote Sensing Data in R: The Landsat Package", Journal of Statistical Software, July 20i I, Volume 43, Issue http://www.jstatsoft.orgl.
[20]. Schowengerdt, R.A. "Remote Sensing Models and Methods for image processing". 3.ed, Elsevier, 2007.
[21]. Sherrah, J. Fully convolutional networks for dense semantic labeling of high-resolution aerial imagery. arXiv,2016; arXiv:1606.02585.
[22]. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geosci. Remote Sens. Mag. **2016**, 4, 22–40.
[23]. https://towardsdatascience.com/review-fc-densenet-one-hundred-layer-tiramisu-semantic-segmentation-22ee3be434d5