International Journal of Engineering and Techniques-Volume11Issue3,May - June - 2025

PERFORMANCE EVALUATION OF DEEP LEARNING MODELS USED FOR REMOTE SENSING IMAGE CLASSIFICATION

[1]Dr. P. Nagendra Kumar, [2]Y.Dakshayani, [3]T.Nirmitha, [4] T.Sushmitha, [5] P.Mahima
[1]Professor & HOD, [2,3,4,5]UG Student, [1,2,3,4,5]Department of Computer Science & Engineering (AI&ML),
Geethanjali Institute 0f Science And Technology, Nellore, India

Abstract

Remote sensing is the technology used for extracting information about the earth surface with the help of sensors installed on the satellites. It is mainly used in the production of land cover and land use (LCLU) maps that helps to classify the land cover types like forests, urban areas, water bodies and more. The major problems that occur while analyzing the remote sensing images are atmospheric effects, geometric errors, weather conditions. These problems are essential for environmental monitoring, agricultural decision-making, and urban planning and can be overcome by using deep learning models. The deep learning methods could be designed starting from scratch or using pre-trained networks. We proposed evaluating and comparing the deep learning models convolutional neural network feature extractor (CNN-FE) by developing it from scratch, transfer learning, and fine-tuning it for the LCLU classification system using remote sensed images. We used UCM (University of California, merced) public dataset to train the deep learning models and compared their performances using the performance measurement metrics like accuracy, precision, recall, F1-score, and confusion matrix. The proposed deep learning algorithms can adapt and learn the features of the remote sensing images, and the Transfer Learning and Fine-tuning classification performances are significantly improved. As a result the Fine-tuning deep learning model achieved more accuracy in the UCM dataset.

Keywords: Deep learning, Remote sensing, image classification, CNN-FE, LCLU,

Introduction

Remote sensing is a technology for acquiring information about the earth's surface without actually being in contact with it. This is done by sensing and recording reflection or emitted energy & processing, analyzing, and applying that information. In this technology, special cameras collect images from the objectives and then sense the images accordingly to provide information and sense things about the earth.It is the science & art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in direct contact with the object, area, or phenomena under the investigation. It helps in essential global matters such as climate change, global warming, etc.

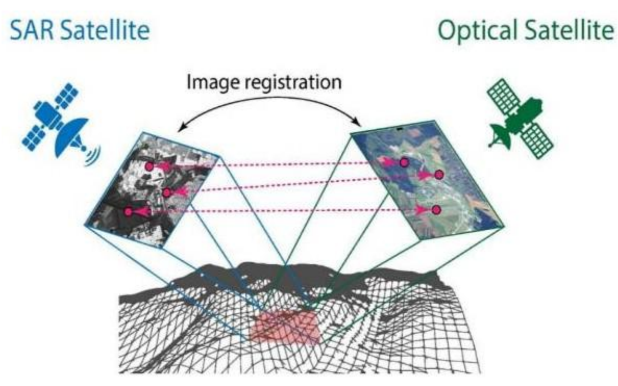


Fig.1. Remote Sensing

Challenges Faced In Remote Sensing Images

Analyzing remote sensing images presents several challenges due to the complexity and variability of the data, the nature of the sensing technology, and the task at hand. Below are some of the key challenges encountered when working with remote sensing images.

Literature Review

[1] In 2022, B. Yang, S. Hu, Q. Guo and D. Hong, "Multisource domain transfer learning based on spectral projections for hyperspectral image classification", IEEE access on may 2022. The experiments show that the proposed method can effectively preserve the source domain features, especially for the scenarios with very few samples in the target domain, which can significantly improve the classification accuracy and reduce the risk of model overfitting. Meanwhile, this strategy greatly reduces the requirement of source domain data, using multisensor data to jointly train a more robust general feature model. The proposed method can achieve high accuracies even with few training samples compared to currently many state-ofthe-art classification methods.

[2] In 2022, A. Alem and S. Kumar, "Transfer learning models for land cover and land use classification in remote sensing image", This paper is aimed to apply one of the DL methods called transfer learning (TL). TL is the recent research problem in machine learning and DL approaches for image classification. DL consumes much time for training when starting from scratch. This problem could be overcome in the TL modeling technique, which uses pre-trained models to build deep TL models efficiently. This work applied the TL model using bottleneck feature extraction from the pre-trained models: InceptionV3, Resnet50V2, and VGG19 to LCLU classification in the UC Merced dataset.

[3] In 2021, R. Naushad, T. Kaur and E. Ghaderpour, "Deep transfer learning for land use land cover classification: A comparative study". Efficiently implementing remote sensing image classification with high spatial resolution imagery can provide significant value in land use and land cover (LULC) classification. The new advances in remote sensing and deep learning technologies have facilitated the extraction of spatiotemporal information for LULC classification. In this study, instead of training CNNs from scratch, the transfer learning was applied to fine-tune pre-trained networks Visual Geometry Group (VGG16) and Wide Residual Networks (WRNs), by replacing the final layers with transfer learning. In this study, instead of training CNN from additional layers for LULC classification using the red–green–blue version of the EuroSAT dataset.

[4] In 2021, A. Shabbir, N. Ali, J. Ahmed, B. Zafar, A. Rasheed, M. Sajid, et al., "Satellite and scene image classification based on transfer learning and fine tuning of ResNet50". This paper aim to fine-tune ResNet50 by using network surgery and creation of network head along with the fine-tuning of hyperparameters. The learning of hyperparameters is tuned by using a linear decay learning rate scheduler known as piecewise scheduler. To tune the optimizer hyperparameter, Stochastic Gradient Descent with Momentum (SGDM) is used with the usage of weight learn and bias learn rate factor.

[5] In 2020, M. Rashid, M. A. Khan, M. Alhaisoni, S.-H. Wang, S. R. Naqvi, A. Rehman, et al., "A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection", Sustainability, vol. 12, no. 12, pp. 5037, Jun. 2020. The proposed approach comprises three steps: (1) By utilizing two deep learning architectures, Very Deep Convolutional Networks for Large-Scale Image Recognition and Inception V3, it extracts features based on transfer learning, (2) Fusion of all the extracted feature vectors is performed by means of a parallel maximum covariance approach, and (3) The best features are selected using Multi Logistic Regression controlled Entropy-Variances method. For verification of the robust selected features, the Ensemble Learning method named Subspace Discriminant Analysis is utilized as a fitness function. The experimental process is conducted using four publicly available datasets, including Caltech-101, Birds database, Butterflies database and CIFAR-100, and a ten-fold validation process which yields the best accuracies of 95.5%, 100%, 98%, and 68.80% for the datasets respectively.

[6] In 2020, D. Zhang, Z. Liu and X. Shi, "Transfer learning on EfficientNet for remote sensing image classification", Proc. 5th Int. Conf. Mech. Control Comput. Eng. (ICMCCE), pp. 2255- 2258, Dec. 2020. In this

paper a transfer learning method based on pre-trained EfficientNet models with fine tuning strategy for remote sensing image classification. EfficientNet achieves the current state-of-the-art performance using significantly less parameters than other latest models in this field of image classification. But nowadays the network is still lacking used in remote sensing tasks. This method is validated on five remote sensing data sets, the experimental results show the effectiveness and superiority of the proposed methods for scene classification in remote sensing imagery.

Proposed Model

System architecture is the foundation upon which the entire system is built. It defines the system's boundaries, components, data flow, and communication channels, making it a crucial aspect of any technology-driven endeavor. It is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. It consists of system components and the sub-systems developed, that will work together to implement the overall system. The System Architecture Diagram of our system is presented below:

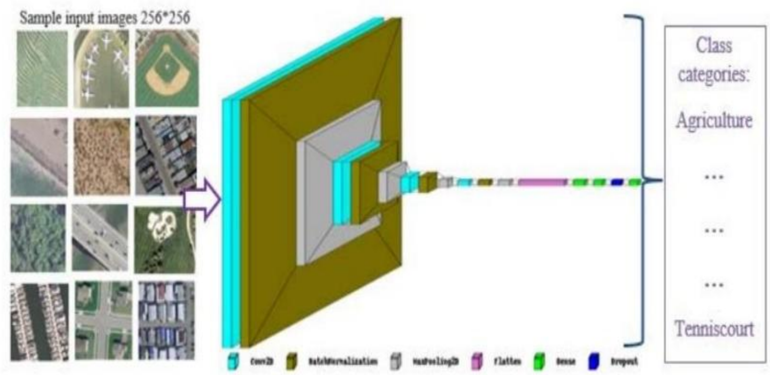


Fig.2. Architecture Diagram

Workflow of the system

1. Collect and preprocess the UCM dataset, including resizing, normalization, and data augmentation, followed by dataset splitting into training, validation, and testing sets.
2. Develop and train three deep learning models: CNN from scratch (CNN-FE), Transfer Learning (TL and Fine-Tuning using VGG16, ResNet50, InceptionV3 and DenseNet121, optimizing hyperparameters using TensorFlow.
3. Evaluate model performance using Accuracy, Precision, Recall, F1-score, and Confusion Matrix, comparing CNN-FE, TL, and Fine-Tuning to determine the most efficient approach.
4. Select Fine-Tuned as the best-performing model and deploy it for environmental monitoring, agricultural decision-making, and urban planning.
5. Continuously improve the model by retraining with larger datasets, fine-tuning hyperparameters, and optimizing for cloud or edge computing deployment.

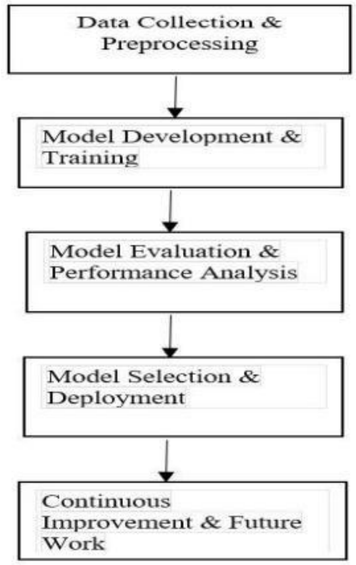


Fig.3. Workflow of the system

Dataset Deep

learning heavily depends on data and dataset makes the training feasible. A dataset is used to train the model for performing various actions, to work automatically. The dataset in which machine learning and deep learning algorithms have been trained and the dataset we use to validate the accuracy of our model is called testing dataset. We used the publicly available University of California Merced (UCM) dataset for modelling the CNN, CNN-based TL, and fine-tuning. The UCM dataset is an LCLU data set collected from the earth, labeled manually, and introduced by at the University of California Merced. It contains twenty-one classes. Each class includes 100 images with a resolution of 256 × 256 pixels and a spatial resolution of about 30 centimeters per pixel. However, the UCM dataset is inconsistent, as about 44 images have different pixel shapes

Fig.4. Dataset collected from kaggle

Models Architectures evaluated

1. VGG16 - VGG16 is a deep convolutional neural network architecture introduced by the Visual Geometry Group (VGG) at the University of Oxford. It is widely used for image classification and feature extraction
2. ResNet50 - ResNet50 is a deep convolutional neural network (CNN) that introduced the concept of residual learning to solve the vanishing gradient problem in deep networks. It was developed by Microsoft Research and won the ILSVRC 2015 competition.

3. InceptionV3 - InceptionV3 is a deep learning convolutional neural network (CNN) architecture designed for image recognition. It was developed by Google and is an improved version of the Inception (GoogLeNet) architecture. InceptionV3 is widely used for image classification and transfer learning due to its efficiency and accuracy.
4. DenseNet121 - DenseNet-121 is a type of Dense Convolutional Network (DenseNet), designed to improve feature reuse and reduce computational cost. It was introduced in the paper "Densely Connected Convolutional Networks" (2017) by Huang et al.

Convolutional Neural Network(CNN)

A Convolutional Neural Network (CNN) is a type of deep learning algorithm primarily used for processing and analyzing structured grid data, such as images, video, and sometimes even audio or time-series data. CNNs are widely used in computer vision tasks, like image classification, object detection, segmentation, and recognition. They are particularly effective in identifying spatial hierarchies in images, making them ideal for analyzing visual data like remote sensing images

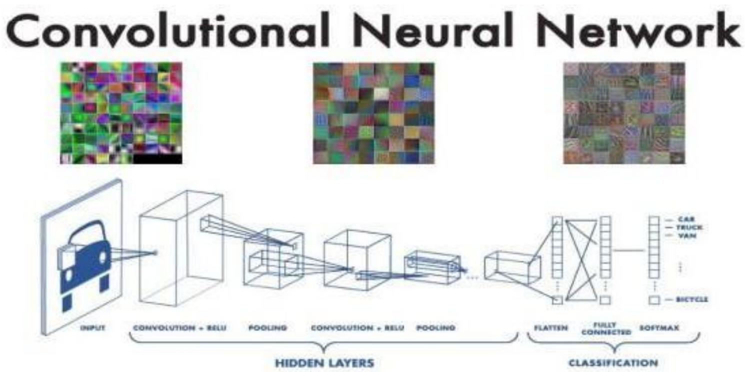


Fig.5. Convolutional Neural Network

A typical CNN architecture is made up of three types of layers, each playing a specific role. Convolutional layer - The convolutional layer, which is like the feature detector of the network. Imagine sliding a small window—called a filter or kernel—across an image. This filter picks up specific patterns like edges, corners, or textures. Each filter creates a different feature map, capturing different aspects of the input. Pooling layer - the pooling layer, which is responsible for downsampling the feature maps. This means reducing the spatial size (width and height) while keeping the most important information. The most common type is max pooling, where the maximum value in a small patch is kept, and the rest are discarded. This step not only makes the computation more efficient but also adds a bit of translational invariance, meaning the network becomes more robust to small shifts or changes in the input image.

Fully connected layer - The fully connected layer in a CNN comes toward the end of the architecture and acts as the decision-making part of the network. Unlike convolutional or pooling layers that deal with spatial data, the fully connected layer takes all the features extracted from previous layers and flattens them into a one-dimensional vector. The final layer often includes an activation function like softmax, which turns the output into probabilities, helping the network choose the most likely class. In essence, the fully connected layer ties everything together and produces the final result.

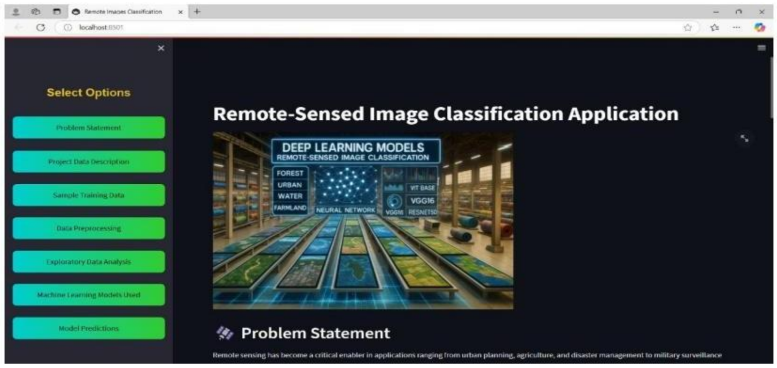
Results & Analysis

COMPARISION TABLE

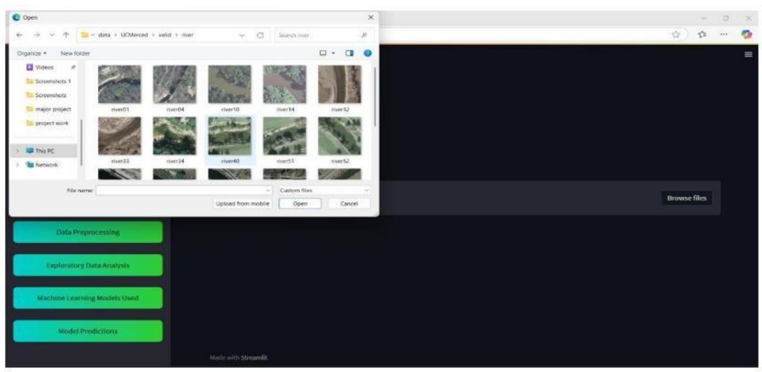
The below table consists of evaluation metric values for the models used, based on which we are finding the best approach.

Models	Accuracy	Precision	Recall	F1 Score
CNN-FE	75.8	74.9	75.2	74.8
CNN-TL	85.6	86.4	84.9	85.6
CNN-FT	91.3	91.3	90.8	91.4

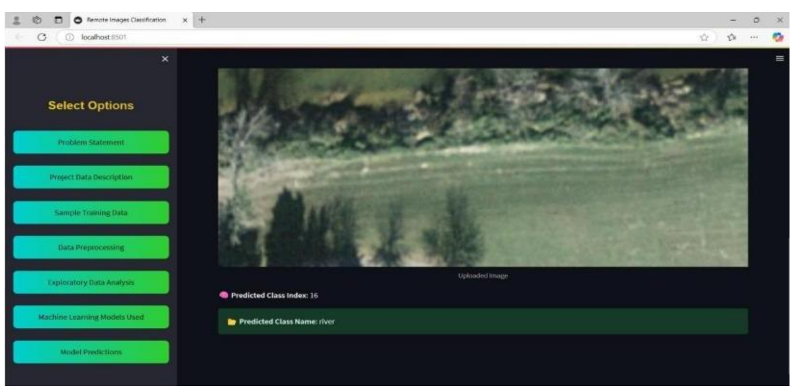
Table.1. Comparison table



HOME PAGE



INPUT TO PREDICT THE CLASS



PREDICTED OUTPUT

Conclusion

This project presents three deep learning models—CNN-FE, TL, and fine-tuning—for Land Cover and Land Use (LCLU) classification using remote sensing (RS) images. The TL and finetuning models were trained on the VGG16, ResNet50, InceptionV3, DenseNet121 pre-trained networks, while CNN-FE was trained from scratch using the UCM dataset. The model's performances were evaluated using accuracy, precision, recall, f1-score, and CM metrics. Results showed that finetuning performed best with accuracy 91.3%, precision 91.3%, recall

90.8% and F1 score 91.4%. The study highlights the efficiency of TL and fine-tuning models in reducing training time, recommending further DL optimization for larger datasets.

FUTURE SCOPE

For further enhancements in remote sensing image classification include the integration of advanced deep learning models, such as transformer-based architectures, to improve accuracy and generalization across diverse landscapes. Incorporating multi-source data fusion—combining optical, radar, and LiDAR data can offer richer contextual understanding. Additionally, real-time classification using edge computing and cloud-based platforms could accelerate decision-making in applications like disaster response or precision agriculture. Enhancements in explainable AI will also be critical to interpret classification results more transparently, aiding in trust and adoption across various domains

References

1. S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. Benediktsson, ‘‘Deep learning for hyperspectral image classification: An overview,’’ IEEE Trans. Geosci. Remote Sens., vol. 57, no. 9, pp. 6690–6709, Sep. 2019.
2. B. Liu, X. Yu, P. Zhang, A. Yu, Q. Fu, and X. Wei, ‘‘Supervised deep feature extraction for hyperspectral image classification,’’ IEEE Trans. Geosci. Remote Sens., vol. 56, no. 4, pp. 1909– 1921, Apr. 2018.
3. G. Cheng, Z. Li, J. Han, X. Yao, and L. Guo, ‘‘Exploring hierarchical con-volutional features for hyperspectral image classification,’’ IEEE Trans. Geosci. Remote Sens., vol. 56, no. 11, pp. 6712–6722, Nov. 2018.
4. M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang, ‘‘Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery,’’ Remote Sens., vol. 10, no. 7, p. 1119, 2018.
5. M. A. Shafaey, M. A. Salem, H. M. Ebied, M. N. Al-Berry, and M. F. Tolba, ‘‘Deep learning for satellite image classification,’’ in Proc. Int. Conf. Adv. Intell. Syst. Inform., vol. 845, 2019, pp. 383– 391.
6. G. J. Scott, M. R. England, W. A. Starms, R. A. Marcum, and C. H. Davis, ‘‘Training deep convolutional neural networks for land–cover classification of high-resolution imagery,’’ IEEE Geosci. Remote Sens. Lett., vol. 14, no. 4, pp. 549–553, Apr. 2017.
7. A. B. Hamida, A. Benoit, P. Lambert, and C. B. Amar, ‘‘3-D deep learning approach for remote sensing image classification,’’ IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4420–4434, Aug. 2018.
8. C. Deng, Y. Xue, X. Liu, C. Li, and D. Tao, ‘‘Active transfer learning network: A unified deep joint spectral–spatial feature learning model for hyperspectral image classification,’’ IEEE Trans. Geosci. Remote Sens.,vol. 57, no. 3, pp. 1741–1754, Mar. 2019.
9. F. Özyurt, ‘‘Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures,’’ J. Supercomput., vol. 76, no. 11, pp. 8413–8431, Dec. 2019