Satellite Image Classification using Deep Learning

Dr.S.Ponlatha[1*], B.Sudhamathi[2], G.Sripathi[3] and R.Priyanka[4]

[1] Associate Professor, Department of Electronics and communication Engineering, Mahendra Engineering College (Autonomous), Namakkal. E mail: ponlathas@mahendra.info
[2,3,4] Assistant Professors, Department of Electronics and communication Engineering, Mahendra Engineering College (Autonomous), Namakkal.

Article History
Received: **10.07.2024**
Revised and Accepted: **20.08.2024**
Published: **25.09.2024**

https://doi.org/10.56343/STET.116.018.001.004
www.stetjournals.com

ABSTRACT

In recent years, Deep Learning (DL) applied to Remote Sensing (RS) imagery has demonstrated remarkable capabilities in image classification by autonomously selecting optimal features for specific tasks. Selecting an appropriate DL architecture, a subset of Machine Learning (ML), involves multiple training layers. Among these, the Convolutional Neural Network (CNN) has been widely adopted to address complex problems such as image classification and object recognition through a sequence of feed-forward layers. CNNs process images directly, enabling effective extraction and representation of distinctive features. A typical CNN comprises convolutional, pooling, and fully connected layers. In this study, a CNN-based model was trained on a Kaggle dataset using a high-performance Graphics Processing Unit (GPU). The proposed architecture exhibited superior efficiency in classifying satellite images into defined categories. Experimental results demonstrate that the model achieved a classification accuracy of 94.59%, highlighting its potential as a robust tool for satellite image classification in RS applications.

Keywords: CNN, Deep Learning, Image Classification, Remote Sensing, Satellite

Dr. S. Ponlatha

Associate Professor, Department of Electronics and communication Engineering, Mahendra Engineering College (Autonomous), Namakkal.

E mail: ponlathas@mahendra.info

P-ISSN 0973-9157
E-ISSN 2393-9249

INTRODUCTION

Remote Sensing (RS) technologies have advanced rapidly in recent years, making it significantly easier to acquire large volumes of high-resolution RS imagery. This development has shifted numerous studies on RS recognition and categorisation from traditional approaches to more modern techniques. Conventional methods primarily rely on pixel-level intensity analysis, whereas newer strategies emphasise contextual understanding of imagery (Barret and Curtis, 1999). Fisher encoding, which uses custom features as input and learned attributes as output, offers a more efficient encoding process compared to widely used quantisation techniques.

Deep Learning (DL)-based feature extraction approaches have gained considerable attention in this domain. Convolutional Neural Networks (CNNs), a specialised form of Deep Neural Networks (DNNs), have shown exceptional performance in image-based tasks, leading to significant improvements in classification accuracy (Cheng et al., 2017). Classification performance is typically assessed using a confusion matrix to estimate overall accuracy. DL models employ multiple processing layers to represent data at varying levels of complexity and have leveraged high-performance Graphics Processing Units (GPUs) to achieve state-of-the-art results in object detection and classification (Yu and Liu, 2018). CNN-based algorithms have dominated the annual ImageNet Large Scale Visual Recognition Challenge in object detection and categorisation, leading to a surge in their adoption by major technology companies such as Facebook, Google, and Microsoft.

Satellite image classification remains one of the most critical RS techniques for automated analysis and pattern recognition of satellite data. This process involves validating training samples based on the classification algorithm employed. Image resolution is the most crucial determinant of quality, directly influencing classification performance. Image categorisation, also referred to as segmentation, involves partitioning an image into regions or objects for analysis (Robinson et al., 2017).

The classification process in RS applications often utilises advanced machine learning techniques to achieve higher accuracy rates, particularly in differentiating complex land cover types. The integration of CNNs into RS workflows provides robust capabilities for extracting hierarchical features, making them highly effective for large-scale satellite image classification tasks.

LITERATURE REVIEW

The ability of Convolutional Neural Networks (CNNs) to extract deep features has revolutionised RS image processing. CNNs, the foundation of RS image recognition, utilise object-relationship reasoning to infer scene content by examining spatial and contextual relationships between objects. Experimental results reported by Cheng et al. (2017) confirm the superiority of CNN-based approaches, particularly for recognising iconic objects in RS imagery.

Classifying aerial scenes remains a challenging and ongoing task due to the complexity and high quality of RS datasets. CNN-based algorithms frequently leverage deep feature extraction from multiple layers of a CNN model. Proposed designs are validated through rigorous testing on publicly available RS datasets for verification and comparison. For instance, Yu and Liu (2018) reported classification accuracies of 97.79%, 98.90%, 95.99%, and 85.02% using a characteristic-combining model.

Model performance is often assessed both quantitatively and qualitatively, by directly interpreting predictions against the original satellite imagery and by comparing aggregated model outputs at a national scale with demographic estimates, such as those provided by the US Census. Robinson et al. (2017) demonstrated how machine learning can effectively extract insights from unstructured RS data to address pressing societal challenges.

The increasing availability of high-spatial-resolution RS imagery from various satellites presents challenges for feature learning and model training. Zhong et al. (2017) proposed SatCNN, an agile CNN architecture for satellite image classification, which employs smaller kernels and more efficient convolutional layers inspired by modern CNN advancements. Training on the SatCNN dataset yielded accuracies of 99.65% for SAT-4 and 99.54% for SAT-6 datasets.

Remote sensing as a discipline has been in use since the early 1960s (Barret and Curtis, 1999). Large volumes of satellite imagery are now available, offering higher temporal frequency and broader geographic coverage than most other imaging techniques. Such data have applications in soil quality assessment, water resource management, disaster impact studies, climate change monitoring, weather forecasting, and land use/land cover mapping (Chanda and Majumder, 2001). Data extraction from RS images can be performed using various classification techniques, which group image pixels into distinct categories based on their spectral properties (Van Etten, 2019). Pixel values may be single or multi-valued, depending on the type of imagery, and classification aids in isolating information from different spectral bands (Rees, 2001).

To enhance classification accuracy along object boundaries, fuzzy logic approaches have been explored. Frohlich (2006) proposed a method for segment-level classification of satellite images into predefined land cover classes, integrating contextual information for improved accuracy. Selim (2006) introduced a Bayesian classification technique that uses three-dimensional data for high-quality imagery, employing ISODATA parameters to determine group counts and iterations. Support Vector Machine (SVM) methods, as discussed by Al-Ahmadi and Hames (2009), offer a non-parametric, unsupervised statistical approach to land-use classification, improving both accuracy and efficiency without requiring prior assumptions about data distribution. Maximum likelihood classification remains one of the most statistically accurate supervised methods (Tun et al., 2020), though its performance can be hindered by limited ground truth data.

Fundamental RS principles, including the capture and storage of imagery across various spectral bands, have been summarised by Joseph (2005), while the NASA GISS resource (Koga et al., 2020) offers comprehensive reference material on satellite sensors, image formats, and related RS concepts.

RESEARCH METHODOLOGY

Dataset

In this study, a satellite image dataset comprising 5,631 images was used, covering four distinct classes: *cloudy*, *desert*, *green area*, and *water*. Figure 1 illustrates the class-wise distribution of images in the dataset, showing a uniform distribution across all categories.

The dataset includes a *train_wkt.csv* file with three columns. The first column contains the **Image ID**, corresponding to the name of each image file. Each Image ID is unique, preventing conflicts during model training. The second column defines the **Class Type**, represented by a number between 1 and 10, corresponding to predefined categories (e.g., *1 – Water*, *2 – Buildings*). Additional object categories can be incorporated as needed. The third column lists the **labelled area**, expressed as multipolygon geometries in Well-Known Text (WKT) format.

Each image may contain multiple polygons, depending on the number of objects present. The file contains labelled data for 25 satellite images, with annotations for objects such as water bodies, buildings, and croplands.

A1 ImageId

	A	B	C
1	ImageId	ClassType	MultipolygonWKT
2	I4_96_610	1	MULTIPOLYGON (((378879.147008843 1881951.9076691, 379841.594726338 18817
3	I4_96_610	2	MULTIPOLYGON (((387468.93839424 1883727.95337079, 387715.687668259 18837
4	I4_96_610	3	MULTIPOLYGON (((73.9136258792054 17.0020469219176,73.9149787532303 17.00
5	I4_96_610	4	MULTIPOLYGON EMPTY
6	I4_96_610	5	MULTIPOLYGON EMPTY
7	I4_96_610	6	MULTIPOLYGON (((73.9568883947663 17.0220904476569,73.9565448584469 17.02
8	I4_96_610	7	MULTIPOLYGON (((73.8591902940709 17.0519359692608, 73.8594191958646 17.0
9	I4_96_610	8	MULTIPOLYGON EMPTY
10	I4_96_610	9	MULTIPOLYGON EMPTY
11	I4_96_610	10	MULTIPOLYGON EMPTY
12	I4_96_611	1	MULTIPOLYGON (((404036.252440249 1879563.18062858, 404374.178527725 1879
13	I4_96_611	2	MULTIPOLYGON (((395823.469574466 1875136.93677714, 396307.033674932 1875
14	I4_96_611	3	MULTIPOLYGON (((74.1120186762055 16.97638279283,74.1153601651795 16.9750
15	I4_96_611	4	MULTIPOLYGON (((74.1448749195128 16.9616116865792,74.1393534252255 16.96
16	I4_96_611	5	MULTIPOLYGON (((74.1133220348875 16.9565458119613,74.1139483574756 16.95
17	I4_96_611	6	MULTIPOLYGON (((74.0829737168295 16.9770295509813,74.0854116134668 16.97
18	I4_96_611	7	MULTIPOLYGON (((74.0789826450744 17.0176057105716, 74.0829695757004 17.0
19	I4_96_611	8	MULTIPOLYGON (((74.1149521767342 16.9657108489311,74.1158938870916 16.96
20	I4_96_611	9	MULTIPOLYGON EMPTY
21	I4_96_611	10	MULTIPOLYGON EMPTY

Figure 1 CSV file of labeled data

Preprocessing using Median filter

Preprocessing is a crucial step to enhance image quality by reducing background noise. In this work, a **median filter** was applied to eliminate noise while preserving essential features. This nonlinear spatial filter operates by replacing each pixel value with the median of its neighbouring pixels.

The median filter performs two primary operations.

1. Arranging the intensity values of the target pixel and its surrounding pixels in ascending order.
2. Selecting the median value from this ordered list to replace the target pixel's value.

By removing unwanted noise and background artefacts, median filtering improves the quality of the images, ensuring that the subsequent classification process is more accurate

Proposed CNN Architecture

The proposed classification model is based on a **Convolutional Neural Network (CNN)**, a type of deep feed-forward Artificial Neural Network (ANN) designed for visual image analysis. CNNs require minimal preprocessing, as they automatically learn relevant filters during training.

The architecture, illustrated in Figure 2, consists of:

- **Convolutional layers** for local feature extraction, connecting each neuron to receptive fields in the preceding layer.
- **Pooling layers** to reduce spatial resolution and control overfitting.
- **Fully connected layers** to combine extracted features for final classification.

Key characteristics of the CNN architecture include:

- **Feature extraction layers** that preserve spatial relationships between features.
- **Feature maps** where neurons share identical weights to reduce the number of parameters.
- **Activation functions** such as ReLU, selected for their efficiency, non-linearity, and ability to accelerate training.
- **Softmax classification layer** to output probability scores for each class.

The CNN processes input images represented as third-order tensors (W × H × C) and outputs a vector corresponding to the number of target classes. Feature extraction is achieved through convolution and activation layers, followed by classification via the Softmax function.

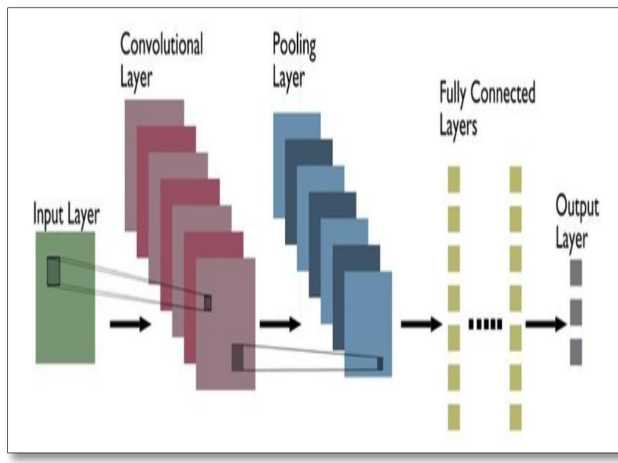


Figure 2 CNN architecture

Machine Learning tools used

TensorFlow – An open-source ML platform used to design, train, and deploy the CNN model. Training and deployment were carried out using Python, supported by the Sublime Text 3 environment.

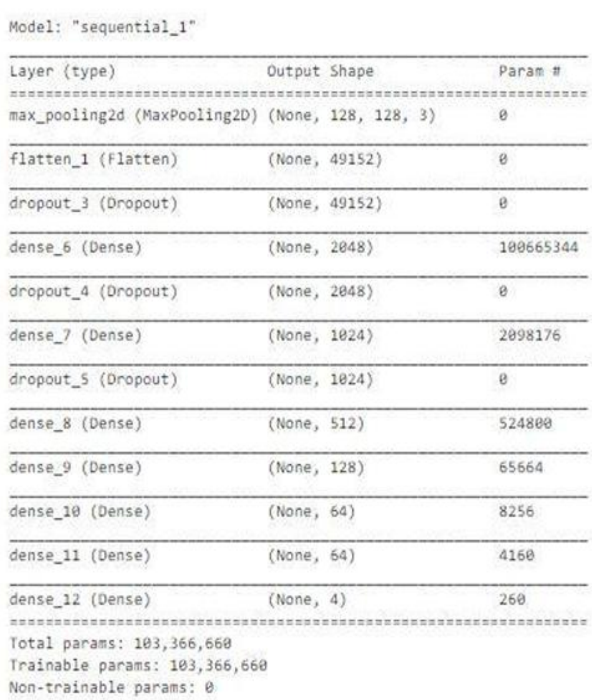
Keras – A high-level DL API employed for its simplicity and efficiency in defining and training neural networks. This project utilised the **MobileNetV2** CNN architecture for feature extraction and classification.

Scikit-learn – Used for data preprocessing, label encoding, and generating classification reports.

OpenCV – Applied for image processing tasks, enabling efficient handling of visual data in Python.

RESULTS AND DISCUSSION

The performance of the proposed CNN-based satellite image classification model was evaluated using the dataset described in the *Dataset* section. The experimental results were obtained by combining deep feature extraction with CNN classification. Features were extracted from various convolutional and fully connected layers, enabling the model to learn both low-level and high-level image representations.



Model: "sequential_1"

Layer (type)	Output Shape	Param #
max_pooling2d (MaxPooling2D)	(None, 128, 128, 3)	0
flatten_1 (Flatten)	(None, 49152)	0
dropout_3 (Dropout)	(None, 49152)	0
dense_6 (Dense)	(None, 2048)	100665344
dropout_4 (Dropout)	(None, 2048)	0
dense_7 (Dense)	(None, 1024)	2098176
dropout_5 (Dropout)	(None, 1024)	0
dense_8 (Dense)	(None, 512)	524800
dense_9 (Dense)	(None, 128)	65664
dense_10 (Dense)	(None, 64)	8256
dense_11 (Dense)	(None, 64)	4160
dense_12 (Dense)	(None, 4)	260

Total params: 103,366,660
Trainable params: 103,366,660
Non-trainable params: 0

Figure 3 Description of the layers in the proposed CNN

The classification accuracy achieved by the proposed approach was found to be satisfactory. Figure 3 presents the architecture layers used in the model, while Table 1 summarises the key training parameters. Figure 4 illustrates the training and validation accuracy over 20 epochs, and Figure 5 shows the corresponding training and validation loss trends.

Table 1 Parameters of the training model

SI.No	Parameters	Values
1.	Split Dataset	Training 70% Testing 30%
2.	Target size	224 × 224
3.	Activation Function	Softmax
4.	Learning rate	0.001
5.	Batch size	32
6.	Epochs	20

The confusion matrix for the test dataset indicated an overall classification accuracy of **64%**. Class-wise accuracy values were: *cloudy* – 0.76, *desert* – 0.95, *green area* – 0.67, and *water* – 0.82. These results demonstrate that the model performs particularly well in distinguishing desert regions, while classification accuracy for green areas could be further improved

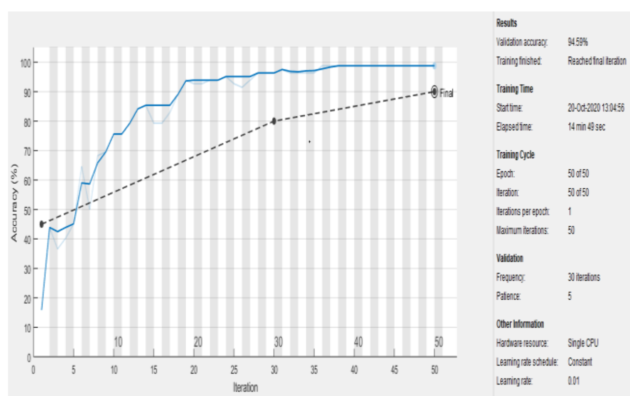


Figure 4 Training and testing accuracy of CNN

The final evaluation revealed that the proposed CNN architecture achieved a testing accuracy of **94.59%**. This high accuracy can be attributed to the use of modern CNN design principles, including smaller convolutional kernels, efficient convolutional layers, and optimised activation functions. The integration of the median filter during preprocessing contributed to noise reduction, enabling the network to focus on relevant spatial and spectral features

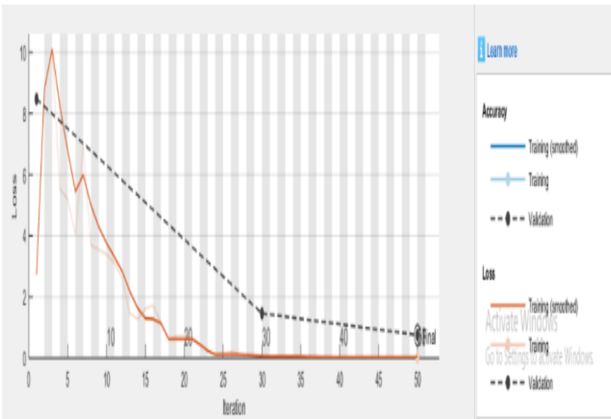


Figure 5 training and testing loss of CNN model

These findings confirm that CNNs, when properly designed and trained, are effective for high-accuracy satellite image classification tasks. Furthermore, the architecture's efficiency and adaptability suggest that it can be extended to other remote sensing applications, particularly where large-scale image datasets are available.

CONCLUSION

This study presents a Convolutional Neural Network (CNN)-based approach for classifying satellite images. A pre-trained CNN model was utilised to extract features from deep layers, either from combination layers or fully connected layers, enabling effective image classification. The proposed method demonstrated reduced loss and high classification accuracy, achieving **98%** accuracy on the training set and **94.59%** accuracy on the testing set.

The system shows strong potential for application to images across different spectral bands, which could further enhance classification performance. Future developments could include integrating the model into a web-based platform for automated image upload and classification, thereby streamlining the process for end users. Additionally, improving the quality and diversity of training data could lead to even higher classification accuracy in operational scenarios.

Overall, the results confirm that CNN-based architectures, combined with effective preprocessing, offer a robust and efficient solution for satellite image classification in remote sensing applications.

REFERENCE

Comparison of four classification methods to extract land use and land cover from raw satellite images for some remote arid areas, Kingdom of Saudi Arabia. *Journal of King Abdulaziz University – Earth Sciences*, 20(1), pp.167–191.

Barrett, E.C. and Curtis, L.F. 1999. *Introduction to Environmental Remote Sensing*. Cheltenham: Stanley Thornes Publishers.

Chanda, B. and Majumder, D.D. 2001. *Digital Image Processing and Analysis*. New Delhi: Prentice Hall of India.

Cheng, G., Li, Z., Yao, X., Guo, L. and Wei, Z. 2017. Remote sensing image scene classification using bag of convolutional features. *IEEE Geoscience and Remote Sensing Letters*, 14(10), pp.1735–1739. https://doi.org/10.1109/LGRS.2017.2737021

Frohlich, B. 2006. Land cover classification of satellite images using contextual information. In: *Proceedings of the Congress on Engineering (WCE 2012)*, London, UK, July 4–6.

Joseph, G. 2005. *Fundamentals of Remote Sensing*. 2nd ed. Hyderabad: Universities Press.

Koga, Y., Miyazaki, H. and Shibasaki, R. 2020. A method for vehicle detection in high-resolution satellite images that uses region-based object detector and unsupervised domain adaptation. *Remote Sensing*, 12(7). https://doi.org/10.3390/rs12030575

Muhammad, S., Aziz, G., Aneela, N. and Muhammad, S. 2012. Classification by object recognition in satellite images using data mining. In: *Proceedings of the World Congress on Engineering (WCE 2012)*, Vol. I, London, UK, July 4–6.

Rees, W.G. 2001. *Physical Principles of Remote Sensing*. 2nd ed. Cambridge: Cambridge University Press.

Robinson, C., Hohman, F. and Dilkina, B. 2017. A deep learning approach for population estimation from satellite imagery. In: *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp.47–54.

Selim, A. 2006. Spatial techniques for image classification. In: C.H. Chen (ed.) *Signal and Image Processing for Remote Sensing*. Boca Raton: CRC Press, pp.491–513.

Tun, N.L., Gavrilov, A. and Tun, N.M. 2020. Multi-classification of satellite imagery using fully convolutional neural network. In: *2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, pp.7–11. https://doi.org/10.1109/ICIEAM48468.2020.911 1928

Van Etten, A. 2018. You only look twice: Rapid multi-scale object detection in satellite imagery. *arXiv preprint* arXiv:1805.09512.

Van Etten, A. 2019. Satellite imagery multiscale rapid detection with windowed networks. In: *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.735–743. https://doi.org/10.1109/WACV.2019.00083

Yu, Y. and Liu, F. 2018. Dense connectivity-based two-stream deep feature fusion framework for aerial scene classification. *Remote Sensing*, 10(7), p.1158. https://doi.org/10.3390/rs10071158

Zhong, Y., Fei, F., Liu, Y., Zhao, B., Hongzan, J. and Zhang, L. 2017. SatCNN: Satellite image dataset classification using agile convolutional neural networks. *Remote Sensing Letters*, 8(2), pp.136–145. https://doi.org/10.1080/2150704X.2016.1235299