

A Deep Learning-based Classification Method for Land Cover Monitoring Using UAV Images

Abstract

A land cover map is a crucial tool for urban planning as it provides essential insights into the landscape's composition and distribution. However, traditional techniques for creating and maintaining maps often require significant temporal and financial investments. Embracing deep-learning-based approaches offers a promising way to revolutionize aerial map generation, providing efficiencies that were previously unattainable. The present study aims to harness the power of neural networks rooted in deep learning to craft a comprehensive land cover map. Focusing on Shiraz City, this study aimed to delineate urban land uses into four categories: Almond, Pistachio, Bare soil, and Shadow of trees. Leveraging imagery captured by a Phantom DJI 4 drone, the research scrutinizes ground features to facilitate accurate classification. The convolutional neural network (CNN) emerges as a pivotal component of the methodology, serving as the bedrock for the automated classification process. Preliminary findings underscore the efficacy of the CNN approach, yielding an impressive overall accuracy rate of approximately 86.56%. Such results underscore the viability of deep-learning-based methodologies in land cover mapping. They also demonstrate its potential for scalability and applicability in various urban landscapes. By reducing the resource-intensive nature of traditional mapping techniques, this study paves the way for more agile and cost-effective urban planning attempts, poised to accommodate the dynamic nature of modern cities.

Keywords: Deep Learning, Land Cover, CNN, UAV Image, Classification.

I. INTRODUCTION

Earth Observation (EO) is an invaluable tool for acquiring insights into the dynamics of our planet through remote sensing techniques. Satellites Positioned in space constitute pivotal platforms for capturing a wealth of data about Earth's surface, atmosphere, and oceans. This vantage point enables the collection of a diverse array of information crucial for monitoring environmental changes, managing natural resources, and informing various fields ranging from agriculture to disaster management. Deep learning, a subset of machine learning, has emerged as a potent framework for analyzing EO data with unprecedented accuracy and efficiency. At its core, deep learning harnesses neural networks and their variants to process and interpret vast volumes of Earth observation imagery. Unlike traditional machine learning approaches, where manual feature extraction and classification are commonplace, deep learning techniques automate these processes, endowing the system with the capability to discern complex patterns and features directly from the raw data [1-4].

The integration of deep learning neural networks into EO analysis holds immense promise for enhancing the identification and characterization of features within satellite and aerial imagery. By enabling automated feature extraction and classification, deep learning algorithms streamline the analysis pipeline, reducing the need for laborious manual intervention and accelerating decision-making processes. Moreover, the adaptability of deep learning models allows them to continuously improve their performance through iterative learning, ensuring their efficacy in handling diverse and evolving Earth observation datasets [5-8].

Furthermore, the synergy between EO and deep learning extends beyond mere data analysis, encompassing wide applications such as land cover classification, object detection, and change detection. From monitoring deforestation patterns in the Amazon rainforest to tracking urban expansion

in rapidly growing cities, the marriage of EO and deep learning technologies empowers researchers, policymakers, and stakeholders with actionable insights for addressing pressing environmental and societal challenges on a global scale. As advancements in both fields continue to unfold, the potential for unlocking new frontiers in Earth observation and environmental monitoring remains boundless, promising a future where our understanding of the planet is more profound and comprehensive than before.

In the contemporary landscape, the relentless march of technological progress has ushered in an unprecedented era of data collection. At the forefront of this data revolution stands the utilization of uncrewed aerial vehicles (UAVs) to meticulously survey and map vast expanses of land, with a particular emphasis on agricultural areas. Indeed, the production of high-resolution maps has swiftly emerged as an indispensable requirement of our times, driven by the burgeoning needs of industries, enterprises, and governmental bodies alike. As companies proliferate and organizations diversify, the demand for accurate and up-to-date spatial information continues to surge, amplifying the imperative for advanced data acquisition methodologies. This surge signals not merely a phase but the culmination of a data collection revolution, underscoring the pivotal role that spatial data infrastructure plays in shaping contemporary decision-making processes. However, as we stand on the cusp of this data-intensive era, it becomes increasingly apparent that the sheer volume of collected data necessitates a commensurate shift towards the era of data processing. In essence, the forthcoming epoch will be defined by our capacity to extract meaningful insights from the deluge of raw information amassed through remote sensing and aerial imaging endeavors.

Remote sensing techniques, augmented by cutting-edge aerial image processing methodologies, will serve as the linchpin of this transformative phase, facilitating the parsing, analysis, and interpretation of vast datasets with unprecedented precision and efficiency. From identifying crop health trends to monitoring environmental changes, the amalgamation of remote sensing and data processing technologies promises to unlock new frontiers in our knowledge about the world. Indeed, as we navigate the complexities of the data-driven age, the ability to harness the power of spatial data analytics will emerge as a defining factor in driving innovation, fostering sustainability, and addressing pressing societal challenges. By leveraging the insights gleaned from remote sensing and aerial image processing, we stand poised to usher in a future where data isn't merely collected but transformed into actionable intelligence, shaping a world more informed, interconnected, and resilient than before. Text:

Numerous studies have been conducted on land classification, each with its own set of challenges. These challenges include the high cost of acquiring images, particularly SAR images [9], the unavailability of ground control points, and the increased cost and time required to obtain ground data for training purposes [10]. Overall, one of the major obstacles is accessing the necessary data. In 2016 a study was conducted on car extraction from UAV aerial images. The results revealed the importance of pre-processing and processing [11]. Another study in China focused on the classification of UAV images and used the AlexNet network [12]. However, as mentioned earlier, we are in the late era of data collection. The selected images, which serve as the primary data for our issue, are collected by the photogrammetry team, while the control points are managed by the land surveyor team. UAV images have a high spatial resolution, and if we have georeferenced UAV images, we can easily design neural networks and achieve optimal accuracy [13]. The area studied in this research is an area on the border between Shiraz and Firozabad in Fars province. The deep neural network used in this study is Convolutional Neural Network (CNN), which the Keras library implements.

II. METHODOLOGY

Keras is a powerful and user-friendly open-source library for developing and evaluating deep learning models. Keras covers two numerical machine learning libraries, Theano and TensorFlow, and lets you determine and train neural network models in just a few lines of code. Keras is written based on different libraries, and because the TensorFlow library is newer, stronger, and more universal, we will install the Keras based on this library [6]. In Machine Learning, models are created to predict the results of certain events. The accuracy of these models can be measured using a train/test method to measure if the model is well enough. Train/Test is a method for measuring the accuracy of a model. The model is trained using the train set, and the model's accuracy will be calculated using the test set. This study used a Convolutional Neural Network (CNN) for the abovementioned purpose [7]. Figure 1 illustrates The algorithm of methodology.

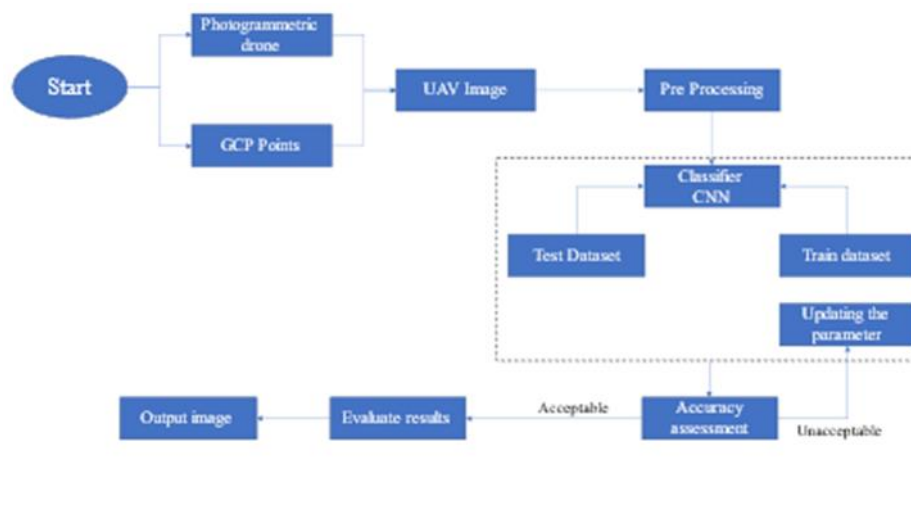


Figure 1: Flowchart of the proposed method

Initially, the original image and the test and training data should be normalized. For this purpose, the available data is divided by 256. Then, the model's parameter should be defined using two dense layers. The model has ten nodes in the first layer and, in the second layer, four nodes that indicate the number of our classes. The compilation is the final step in creating a model. Once the collection is done, we can move on to the training phase. Finally, the model was trained by using 60% of the data and was tested by using 40% of the test data in 150 epochs. To evaluate the model's accuracy, we can form the confusion matrix.

A confusion matrix (or the error matrix) [9] is usually used to characterize image classification accuracy quantitatively. A table shows the similarity between the classification output and a reference image. Also, to create the confusion matrix, we require the ground truth data, such as cartographic information, manually digitizing an image, and fieldwork/ground survey results recorded with a GPS receiver. The structure of the confusion matrix is shown in Table 1.

Table 1: The structure of the confusion matrix.

Classified image	Reference image			Total
	A	B	C	
A				Σa

B				Σb
C				Σc
Total	ΣA	ΣB	ΣC	$N = \Sigma$

We can get two sorts of information from the confusion matrix. Firstly, we can get the overall accuracy by using the diagonal elements of the matrix. Diagonal cells include the number of correctly identified pixels. If we divide the sum of these pixels by the total number, we will get the classification's overall accuracy. This index will be equal to the following:

$$p_o = \frac{aA + bB + cC}{N} \quad (1)$$

In this formula, p_o indicates the overall accuracy.

Another accuracy indicator is the kappa coefficient [10]. It measures how the classification results compare to values assigned by chance. It can take matters from 0 to 1. If the kappa coefficient equals 0, there is no agreement between the classified image and the reference image. If the kappa coefficient equals 1, then the classified and ground truth images are identical. So, the higher the kappa coefficient, the more accurate the classification is. The formula for estimation of the kappa coefficient is:

$$p_e = \frac{\frac{\sum a \times \sum A}{N} + \frac{\sum b \times \sum B}{N} + \frac{\sum c \times \sum C}{N}}{N} \quad (2)$$

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (3)$$

III. CASE STUDY AND DATA

The case study in our study is an area on the border between Shiraz and Firozabad in Fars province in Iran. The investigated image is a georeferenced image shown in Fig.2.

This area is located in zone 39 and the northern hemisphere of the UTM map projection. General information on this area is shown in Table 4.

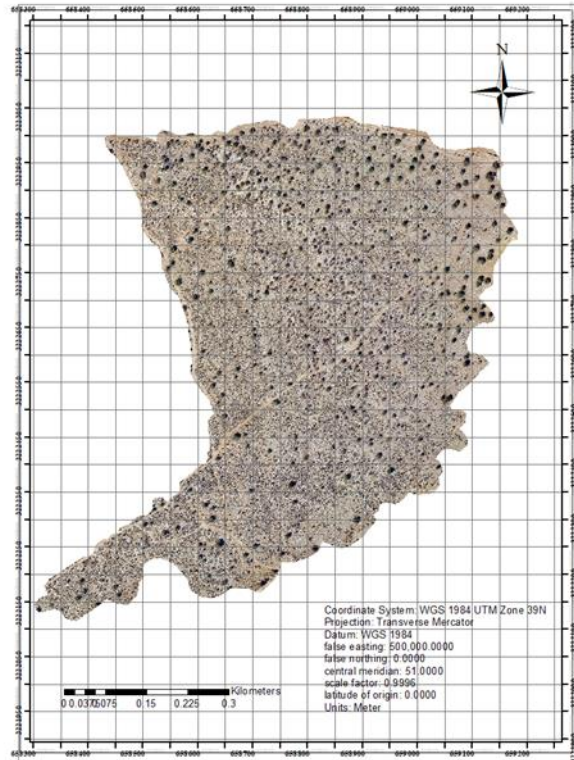


Figure 2: Image of area.

Table 2: general information on the study area.

Map projection	Pixel size	Datum	UL-Geo	UL-Map
UTM, Zone 39	0.25 m	WGS-84	52°37'37.69"E, 29°7'33.43"N	658301.682, 3223034.397

IV. RESULTS AND DISCUSSION

After the implementation, the confusion matrix of the model will be as follows:

Table 3: Confusion matrix of the model.

Classified image	Reference image				
	Bare soil	Pistachio	Almond	Shadow of trees	Total
Bare soil	23	157	0	1926	2106
Pistachio	0	2349	0	236	2585
Almond	2	1421	76	3	1502
Shadow of trees	6	216	1	23097	23320
Total	31	4143	77	25262	29513

Overall accuracy and kappa coefficient can be extracted from the above matrix.

Table 4: Final Result.

Model	Architecture	Number of Class	Epoch	Kappa Coefficient	Overall Accuracy
CNN	DenseNets	4	150	59.80%	86.56%

The model's accuracy was increased during the training, which means the model's movement was successful. Figure 3 illustrates the model's accuracy and loss trend.

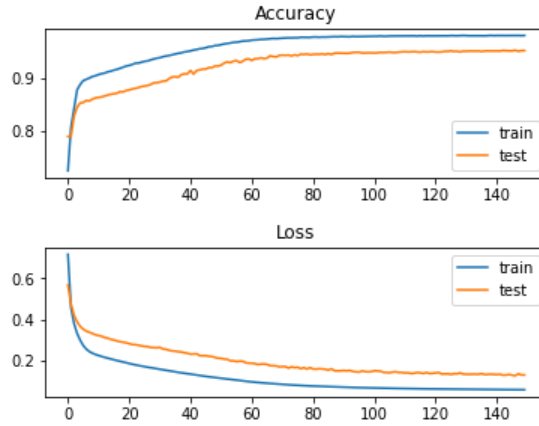


Figure 3: Accuracy and Loss for the train and test set.

As Figure 3 shows, the model's accuracy increases and the loss decreases for each train and test set.

Figure 4 illustrates the original image and the classified image.

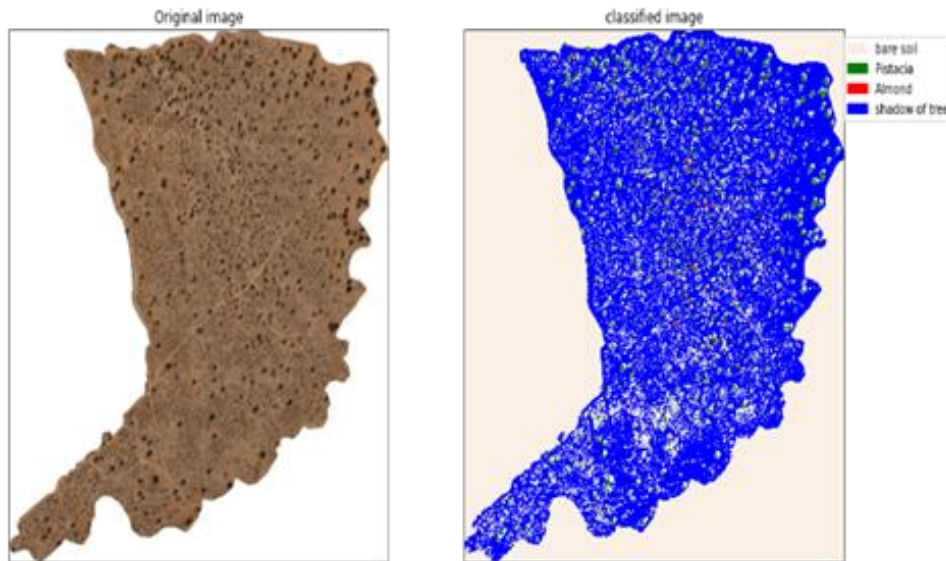


Figure 4: Original image (left) and the classified image (right).

v. CONCLUSION

In today's data-driven landscape, the emphasis on data collection reigns supreme, with aerial imagery assuming a pivotal role in creating accurate and comprehensive maps. Unmanned Aerial Vehicles (UAVs) have emerged as indispensable tools in this endeavor, offering a cost-effective alternative to ground surveying while delivering superior resolution compared to satellite images. Consequently, the processing of UAV-derived imagery has garnered heightened significance, presenting a gateway to unlocking actionable insights from vast swathes of aerial data.

This study delves into the world of UAV image classification, leveraging the capabilities of deep learning-based methodologies to extract meaningful information from aerial datasets. Focusing on a targeted area straddling Shiraz and Firozabad in the Fars province of Iran, the study meticulously

categorizes the imagery into four distinct classes: pistachio tree, almond tree, bare ground, and tree shade, achieving an impressive accuracy rate of 86.55%. Central to the methodology employed is the utilization of convolutional neural networks (CNNs), a cornerstone of deep learning frameworks renowned for their ability to discern intricate patterns and features within complex datasets. Through iterative training and optimization, the CNN model attains peak performance, enabling precise classification of aerial imagery with remarkable efficiency and accuracy.

The findings of this study underscore the efficacy of CNN-based approaches in aerial image classification, affirming the suitability of deep learning techniques for extracting actionable intelligence from UAV-derived datasets. Through the use of advanced algorithms, researchers and practitioners can streamline the analysis pipeline, expedite decision-making processes, and discover valuable insights for a variety of applications, such as agriculture, environmental monitoring, and urban planning. As data collection continues to evolve, incorporating deep learning techniques into UAV image processing is a significant step toward fully leveraging aerial data analytics.

With the continuous advancement in technology and methodology, we are on the verge of a future where the seamless integration of UAV imagery and deep learning is key to understanding the complexities of our ever-changing world.

REFERENCES

- [1] Z. Huang, M. Datcu, Z. Pan, and B. Lei, "A Hybrid and Explainable Deep Learning Framework for SAR Images," *Int. Geosci. Remote Sens. Symp.*, pp. 1727–1730, Sep. 2020, doi: 10.1109/IGARSS39084.2020.9323845.
- [2] B. Zheng, S. W. Myint, P. S. Thenkabail, and R. M. Aggarwal, "A support vector machine to identify irrigated crop types using time-series Landsat NDVI data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 34, no. 1, pp. 103–112, 2015, doi: 10.1016/J.JAG.2014.07.002.
- [3] N. Ammour, H. Alhichri, Y. Bazi, B. Benjdira, N. Alajlan, and M. Zuair, "Deep Learning Approach for Car Detection in UAV Imagery," *Remote Sens. 2017, Vol. 9, Page 312*, vol. 9, no. 4, p. 312, Mar. 2017, doi: 10.3390/RS9040312.
- [4] C. Fan and R. Lu, "UAV image crop classification based on deep learning with spatial and spectral features," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 783, no. 1, p. 012080, May 2021, doi: 10.1088/1755-1315/783/1/012080.
- [5] M. Der Yang, K. S. Huang, Y. H. Kuo, H. P. Tsai, and L. M. Lin, "Spatial and Spectral Hybrid Image Classification for Rice Lodging Assessment through UAV Imagery," *Remote Sens. 2017, Vol. 9, Page 583*, vol. 9, no. 6, p. 583, Jun. 2017, doi: 10.3390/RS9060583.
- [6] M. Si, "Development of Predictive Emissions Monitoring System Using Open Source Machine Learning Library – Keras: A Case Study on a Cogeneration Unit," vol. 7, no. x, 2019, doi: 10.1109/ACCESS.2019.2930555.
- [7] G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for high resolution remote sensing imagery using a fully convolutional network," *Remote Sens.*, vol. 9, no. 5, May 2017, doi: 10.3390/rs9050498.
- [8] N. Pourhasan, R. Shah-Hosseini, and S. T. Seydi, "Deep Learning-based Classification Method for Crop Mapping Using Time Series Satellite Images," *J. Geomatics Sci. Technol.*, vol. 11, no. 1, pp. 129–142, 2021, Accessed: Oct. 14, 2022. [Online]. Available: <http://jgst.issge.ir/article-1-938-en.html>

- [9] M. Ohsaki, P. Wang, K. Matsuda, S. Katagiri, H. Watanabe, and A. Ralescu, "Confusion-matrix-based kernel logistic regression for imbalanced data classification," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 9, pp. 1806–1819, 2017, doi: 10.1109/TKDE.2017.2682249.
- [10] J. Wang, Y. Yang, and B. Xia, "A Simplified Cohen's Kappa for Use in Binary Classification Data Annotation Tasks," *IEEE Access*, vol. 7, pp. 164386–164397, 2019, doi: 10.1109/ACCESS.2019.2953104.
- [11] Chasmer, L.; Hopkinson, C.; Veness, T.; Quinton, W.; Baltzer, J. A decision-tree classification for low-lying complex land cover types within the zone of discontinuous permafrost. *Remote Sens. Environ.* **2014**, *143*, 73–84.
- [12] Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258.
- [13] Zhang, L.; Li, X.; Yuan, Q.; Liu, Y. Object-based approach to national land cover mapping using HJ satellite imagery. *J. Appl. Remote Sens.* **2014**, *8*, 083686.