Analysis of deep learning architecture for patch-based land cover classification

Hisyam Fahmi
Mathemathics Department, Faculty of Science and Technology
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Malang, Indonesia
hisyam.fahmi@uin-malang.ac.id

Wina Permana Sari
Computer Science Department, School of Computer Science
Bina Nusantara University
Jakarta, Indonesia
wina.sari001@binus.ac.id

***Abstract*— In recent years, the usage of computer vision methods for mapping land cover and land use has increased. Patch-based change detection may produce much superior results than pixel-based change detection and generate accurate change maps. Convolutional Neural Network (CNN) is an excellent option for remote sensing applications using hyperspectral data. The training of models using deep learning techniques takes a lot of time. This research aims to determine the most efficient deep learning model for patch-based land-cover classification by examining and comparing three CNN architectural models for land cover classification: LeNet-5, VGG-16, and ResNet-50. EuroSAT data derived from Sentinel-2A remote sensing imagery are utilized in this work. Comparing the three CNN architectures indicates that ResNet-50 has the highest validation accuracy, with a testing accuracy of 0.877, and a training time that is neither too quick nor too slow. The LeNet-5 model has the quickest training time but the lowest accuracy. VGG-16 has the longest training period yet has the highest test score of 0.878.**

Keywords*— *Convolutional neural network, EuroSAT, land cover classification, LeNet-5, ResNet-50, VGG-16

I. Introduction

Land-cover data are essential for land-use planning, urban management, and environmental monitoring. Recent years have seen an increase in the use of computer vision techniques, namely machine learning algorithms, for land cover and land use mapping, which can make these procedures more efficient and cost-effective [1]. It was discovered that the approach combining a sufficient number of larger scales and multi-type features is suggested for improved performance [2]. In comparison to pixel-based change detection, patch-based change detection can produce much better results and give accurate detection to support manual land cover update [3].

Convolutional Neural Networks (CNN) have made significant advancements in image classification, object recognition, and image semantic segmentation in recent years, as a result of its superior feature learning and feature expression capabilities. The current method of choice for addressing computer vision problems like image classification is deep learning approach. This is the case because deep learning architecture is able to extract intricate statistical features from high dimensional data [4]. Also, for hyperspectral remote sensing applications, CNN is a good option, due to its internal products' excellent classification accuracy and interpretability [5]. Nevertheless, conventional CNN models have drawbacks, such as a large number of layers, which results in a greater computational cost [6]. The accuracy of the CNN model is influenced by the optimizer, activation function, filter size, learning rate, and batch size. The evaluation of deep learning CNN can be done by modifying these parameters. The most important factors in enhancing HSR picture land-cover classification included effective super pixel segmentation, networks with high learning capabilities, an optimized multi-scale and multi-feature solution, and adaptive attention-based feature fusion [7].

Different CNN architectures have been developed over time, resulting in incredible deep learning advancements. Error rates in contests like the ILSVRC ImageNet challenge are an effective gauge of this development. In order to classify the land cover data, this research examines and compares multiple CNN architectural models. Three models— ResNet-50, VGG-16, and LeNet-5—represent the architectures that are being compared. Model accuracy and training and testing execution times for the four architectures are compared. The VGG-16 network was used to identify the trees using data from the WorldView-3 remote sensing system's eight visible-near infrared bands [8].

The land cover classification was evaluated using a model developed using EuroSAT land cover data comprised of land cover data from the Europe area. Land cover classification training and test data are generated using Sentinel-2 satellite images. The dataset can be used to track changes in land use and land cover, which can help to enhance geographic maps. Helber et al. [9] makes this geo-referenced dataset of EuroSAT available to the public. According to that research, the RGB band combination outperformed all other band combinations and produced classification results that were more accurate than those of all single-band tests.

So that the contribution of this research is, first, we aim to do a land-cover classification procedure based on image patches using a deep learning approach with three alternative models. After comparing the three models' trial outcomes, we want to determine which one is the most effective for classifying land cover image.