Classification of Remotely Sensed Images Using Transfer Learning

Pooja Vijay Mestri [1], Aditya Sanjay Mali [2], Tanvi Vishnu Gawas[3]

[1,2,3]*Students, Dept. of Computer Engineering, Pillai HOC College of Engineering and Technology, Mumbai, India*

---***---

Abstract - *Hyperspectral images are images captured in thousands of bands of the electromagnetic spectrum. Hyperspectral remote sensors are widely used in remote sensing to monitor the earth's surface with high spectral resolution. Recently, hyperspectral image (HSI) classification proposals have been proposed and demonstrated assuring performance based on deep learning models. However, because of very limited available training samples and huge model parameters, Deep Learning methods may encounter overfitting. Due to a complex network structure and large parameters, it is challenging to achieve satisfying classification accuracy with only a small number of training samples. In this project, we are using a 3D Convolutional Neural Network, to encode pixels spectral and spatial data and a Multi-Layer Perceptron to conduct the classification task.*

Key Words: **3D-CNN, Transfer Learning, Hyperspectral images, Remote Sensing**.

1. INTRODUCTION

The goal of hyperspectral imaging is to obtain the spectrum for each pixel in a captured images in order to find objects, identify materials, or detect processes. Hyperspectral images are spectrally overdetermined, which means that they present adequate spectral information to recognize and characterize spectrally unique materials. Hyperspectral representation provides the potential for more well-defined and detailed information extraction than possible with any other type of remotely sensed data. Hyperspectral sensors simultaneously measure hundreds of continuous spectral bands with a fine resolution to form a three-dimensional hyperspectral image data cube. The data obtained are highly correlated and comprise a significant amount of redundant data. All the image bands are not equally important for the specific application. Because hyperspectral images have a high dimensionality, they generate a large amount of data; thus, processing such a large volume of images necessitates a large amount of computational power. Various high performance computing approaches have recently been used to accelerate the computations required for Hyperspectral data processing.

There is a very wide range of thematic applications in modern society, such as Ecological science, Geological science, Hydrological science, Precision agriculture. Transfer learning focuses on transferring knowledge gained while solving one problem to a different but related problem. Broadly speaking, the purpose of transfer learning is to use training data from related tasks to aid learning on a future problem of interest. Some of the currently available methods of image classification require significant user interaction to produce satisfactory results. In this project, we aim to produce an image classification model with better accuracy. We also aim to make the model more resource-efficient to reduce the amount of training data required.

2. PROPOSED SYSTEM

In the field of hyperspectral image processing, researchers usually use a 1D CNN to obtain spectral features in the spectral domain separately. It is important to capture joint features in both spatial and spectral dimensions while using HSI classification problems. We investigate 3D kernel for joint action recognition using 3D convolution to extract spatial-temporal features.

2.1 Transfer Learning

DL models have already shown promise in a number of fields, including classification, identification, and tracking, to name a few. Many models, on the other hand, only perform well with a large number of training samples. In particular, advanced models and a large number of training samples are used to achieve success in classification and recognition. Lack of sufficient training samples may lead to a poor performance. In such cases, it would be helpful if transfer learning is adopted. Pretraining a model on one data set with a large number of labelled samples, such as ImageNet, and then transferring lightweight the pretrained model to the target data set to finetune is a common transfer learning strategy. Transfer learning is critical for data sets with a small number of training samples, particularly when the model used is deep CNN, which has a large number of parameters. Since HSI data sets often have a small number of training samples, transfer learning can be useful.

DL models have been used in HSI classification and have shown to be efficient. However, there is still room for progress in HSI classification approaches based on DL. ResNet's excellent success demonstrates the importance of depth in DL-based image processing methods. We used this as inspiration to build a very deep 3-D convolutional network for HSI classification. The specifics of the proposed network are presented in this section.

2.2 CNN Architecture

CNN is one of the most effective tools for classifying large amounts of data. Two-dimensional (2D) CNNs primarily capture features from the spatial domain, but three-dimensional (3D) CNNs could aid in the extraction of spatial-spectral tensor features. A Convolutional Network is a series of layers in which each layer uses a differentiable function to transform one volume of activations into another. An input layer, a convolution layer, a pooling layer, a fully-connected layer, ReLU and an output layer make up a typical 3D-CNN.

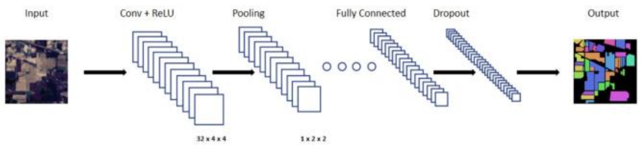


Fig -1: CNN Architecture

Convolutional layer: The most critical aspect of the CNN structure is the convolutional layer. Activation functions are often used in convolution operations to remove features and add certain non-linear factors to the network. The spatial-spectral feature mapping can be achieved using 3D convolutional kernels with the input data of the HSI tensor containing spatial and spectral dimensions. By convolution with a set of filters, each convolutional layer transforms one set of feature maps into another set of feature maps.

Pool layer: A ReLU often precedes the max-pooling layer, which is followed by convolution or completely linked. Calculate the maximum number of r × c pixels for each batch, which will be multiplied by an adjustable weight and then applied to a bias expression. To produce an output for the batch r × c, the result is passed through an activation function. It performs down sampling operation along the spatial dimensions.

ReLU: This layer serves as a correction layer, so all negative values obtained as inputs are replaced with zeros. It performs the purpose of activation. It is often followed by the pooling layer and preceded by the convolution layer or totally connected.

Fully Connected layer: At the end of the network, the FC layer is often mentioned. A completely connected layer, as the name implies, is one in which each neuron is connected to each neuron of the previous layer and each connection has a weight. This is a very broad relation model that makes no assumptions about the data's properties. It also costs a lot of money in terms of memory (weight) and computation. Every activations in the previous layer are completely connected to neurons in a fully connected layer. As a result, their activation can be calculated using a matrix multiplication.

2.3 Methodology

For feature extraction, the proposed deep network is primarily made up of three convolutional layers. Instead of a very deep network structure, we used only three convolution layers to increase the performance of the proposed model, it uses ReLU as an activation function. The 3D-CNN model processed spectral features by fusing them with neighbor band information, and the pooling layer implemented an activator in each convolution component to reduce data variance and boost the feature's nonlinearity. The 3D convolution operation is applied to a 3D block for the capture of the integration of a spatial-spectral patch with a 3D kernel, and the convolution calculation is done in a sequential fashion, for example, from top to bottom, left to right. The pooling layer comes after the convolution layer, and it is used as an activator to create nonlinear features in the 3D-CNN model. In this paper, we use a zero-padding policy to keep the size of the next input feature for the following convolution layer the same. Fully connected layer is used separately to reassemble the obtained local features for the output layer. We select a neuron from the final set of spatial neighbors of the active pixel for classification. Hold the knowledge base as a vector of both columns during the classification stage, the first column indicates values obtained during the partial learning, and the second column indicates the groups of each of them. The spectral feature is used as channel information in the development of feature extraction to strengthen the feature representation capacity. This module captures the spatial feature while also incorporating intra-band information, which is expressed in the feature quantification not only in the spatial domain but also in the band domain.

Convolutional Network layer by layer convert the original image from the original pixel values to the final class scores in this manner. It's worth noting that certain layers have parameters and others don't. The Convolution and Fully Connected layers, in particular, perform transformations that are a function of both the activations in the input volume and the parameters (the weights and biases of the neurons). The ReLU and Pool layers, on the other hand, will use a fixed feature. Gradient descent will be used to train the parameters in the Convolution layer and Fully Connected layer so that the class scores computed by the Convolutional Network are compatible with the labels in the training set for each picture. Finally, we use dropout layer to discard values below 0.5.

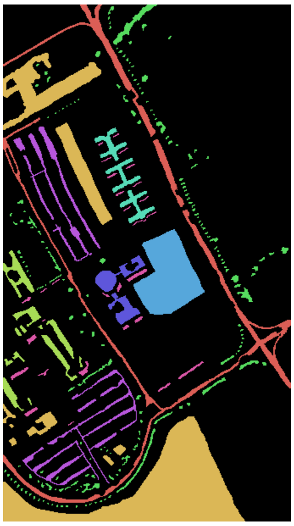
3. RESULTS

Fig -2: Classification results of Indian Pines dataset.

(a) Input (b) Output

Fig -3: Classification results of PaviaU dataset.

(a) Input



(b) Output

4. CONCLUSIONS

The number of training samples available in HSI classification is limited. This paper proposes an improved 3D-CNN classification system based on transfer learning to solve the problem of insufficient samples and improve HSI classification. Transfer learning helps us to pretrain a classification model with enough training samples on the source data set, then transfer the model to the target HSI data set with much smaller training samples for fine-tuning to complete the classification task. The proposed 3D-CNN HSI classification system utilizes both spectral and spatial information contained in HSI data. It can be demonstrated that, when compared to the standard approaches tested, the proposed method is capable of detecting a better subspace that provides the best classification accuracy.

ACKNOWLEDGEMENT

This project required a significant amount of work, research, and dedication. Nonetheless, implementation would not have been possible without the support of Prof. Rajashree Gadhave our guide, and Head of Department, Dr. Ashok Kanthe. First and foremost, we appreciate their financial and logistical support, as well as their advice on project execution. We are also thankful for the expertise and technical assistance given during the implementation. Without their superior expertise and experience, the Project's consistency of results will suffer, so their assistance has been critical. We would like to express our heartfelt gratitude to the volunteer researchers who volunteered their time and expertise to help us complete this project. Nevertheless, we want to express our gratitude to our families and co-worker's for their support and motivation in helping us finish this project.

REFERENCES

[1] Chunyan Yu, Rui Han, Meiping Song, Caiyu Liu, Chein-I Chang, Life Fellow, IEEE, "A Simplified 2D-3D CNN Architecture for Hyperspectral Image Classification Based on Spatial-Spectral Fusion", 2019.

[2] Yanan Luo, Jie Zou, Chengfei Yao, Tao Li, Gang Bai, College of Computer and Control Engineering, Nankai University Tianjin, China, "HSI-CNN: A Novel Convolution Neural Network for Hyperspectral Image", 2018.

[3] Md. Tanvir Ahmed, Md. Ali Hossain and Md. Al Mamun, Department of Computer Science & Engineering Rajshahi University of Engineering & Technology Rajshahi, Bangladesh, "Target class oriented subspace detection for effective hyperspectral image classification", 2018.

[4] Xiangrong Zhang , Senior Member, IEEE, Yujia Sun, Kai Jiang, Chen Li, Licheng Jiao, Fellow, IEEE, and Huiyu Zhou, "Spatial Sequential Recurrent Neural Network for Hyperspectral Image Classification", 2018.

[5] Weisheng Dong, Member, IEEE, Huan Wang, Fangfang Wu, Guangming Shi, Senior member, IEEE, and Xin Li, Fellow, IEEE, "Deep Spatial-spectral Representation Learning for Hyperspectral Image Denoising", 2018.

[6] Hyungtae Lee, Sungmin Eum, Heesung Kwon, Booz Allen Hamilton Inc., McLean, Virginia, U.S.A. , U.S. Army Research Laboratory, Adelphi, Maryland, U.S.A., "CROSS-DOMAIN CNN FOR HYPERSPECTRAL IMAGE CLASSIFICATION", 2018.

[7] Mingyi He, Bo Li, Huahui Chen, Northwestern Polytechnical University, School of Electronics and Information International Centre for Information Acquisition & Processing, Xi'an, Shaanxi, China, "MULTI-SCALE 3D DEEP CONVOLUTIONAL NEURAL NETWORK FOR HYPERSPECTRAL IMAGE CLASSIFICATION", 2017.

[8] Yushi Chen, Member, IEEE, Hanlu Jiang, Chunyang Li, Xiuping Jia, Senior Member, IEEE, and Pedram Ghamisi, Member, IEEE, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks", October 2016.