Classification of Remotely Sensed Images using Deep Learning and Multiresolution Analysis

Rizwan Ahmed Ansari[a]*, Rakesh Malhotra[a] and Krishna Mohan Buddhiraju[b]

[a] Environmental, Earth and Geospatial Sciences, North Carolina Central University, Durham, USA
[b] Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India
* ransari@nccu.edu

Keywords: classification, deep learning, remote sensing, multiresolution

Introduction

Remote sensing images are widely used in many applications including land cover, land use mapping and monitoring, urban-slum classification analysis, building and road extraction as well as change detection. Multiresolution analysis (MRA) has been successfully used in texture analysis for image classification. Traditionally, wavelet transforms have been very popular for MRA, but they are constrained in their ability to capture directional information. Therefore, wavelet is more appropriate for isotropic features (Mallat, 1999). A range of other directional basis function sets have been used in signal-processing domain, with properties relating to alignments, elongations, edges and curvilinear features. The developments in MRA in the context of directional decomposition such as ridgelets and contourlets are shown to overcome these limitations (Ansari et al. 2019). The use of directional MRA and its variants for image classification by texture, particularly in the context of remotely sensed image is limited. Inspired by the recent success of deep learning, this work investigates utility of directional MRA in deep learning to design a framework for classification of remotely sensed images. This paper proposes a composite architecture derived from the deep learning by processing the data obtained using directional MRA transforms. The intuition behind this is that different objects in a remotely sensed image contain features at different scales and direction. These features are important for classification therefore a multi-scale directional approach is useful for their extraction. This enhances the contextual overview of the convolutional network as the features are extracted in different directions and at multiple scales before learning.

Method

In this paper a pixel based classification method to extract features at different scales for different objects in remotely sensed images using deep learning and directional MRA is proposed. The image is decomposed using wavelet, ridgelet and contourlet transforms (Do and Vetterli, 2005) into different approximation and detailed directional sub-bands which are integrated with the layers of neural network. This supplements the network with decomposed low-frequency and high-frequency content for feature extraction which proves to be pivotal for improvement in performance for classification.

Convolutional networks are used in several applications including whole-image classification to pixel classification as thematic mapping in remote sensing (He et al. 2016; Kampffmeyer, 2016). In recent years, deep learning has become a state-of-the-art method for pixel classification for remotely sensed images (Ronneberger, 2015; Kim et al. 2018). Fully convolutional networks are adapted as effective methods for the semantic mapping of remote sensing data (Volpi and Tuia 2016; Wang et al. 2017). This paper uses the modified architecture of U-net.

The U-net is trained on MRA coefficients along with the original data set. The U-net consists of two paths: contracting and expansive. The contracting part is used to extract the features of input data, and we modified the input layer to adapt the five elements of the input data. The input layer is followed by a normalization layer and a max-pooling layer. The activation layer in the network has a rectified linear unit and a 2 × 2 max-pooling operation for the subsampling. For feature extraction, there are three stages and every stage includes several sub-blocks. The feature maps in the same block have the same size, and the feature maps in the following blocks are half that of the previous ones to implement MRA decomposition. The feature maps in different blocks have different scale features.

Using two dimensional MRA on the input image, the first and second-level decompositions are half and one-fourth the size of the original image. This allows us to concatenate the decompositions to the outputs of the max-pooling layer. There are two variants of this proposed model – first in which only the first level decomposition of the input along with the original input is being processed in the U-Net, and the second in which both first and second-level decompositions are being given to the layers of the network along with the original input. The four-channel first-level decomposition is passed through two convolutional layers and is concatenated with the output of the first max-pooling layer of the U-Net. The second-level decomposition is fed similarly using different MRA techniques.

By wavelet, ridgelet and conourlet decompositions, the features obtained at multiple scales, a concatenation with the corresponding stage from the contracting part is designed in the deep neural network. Every stage in the expansive part includes the up-sampling of the feature map, a concatenation block and a convolution block, which consists of a 3 × 3 convolution layer, a normalization layer and a rectified linear unit. A softmax layer is added at the end to calculate the classification results. This work further analyses the effects of different wavelets and contourlets for decomposition on classification performance.

Results

The proposed framework is tested using data containing an image of Kuwait City from IRS-1 (5.8mx5.8m) is used. Five classes are considered; deep water, shallow water, roads, built-up area, open area and roads for classification task.

Figures 1-3 show the original image and classified results using U-net and proposed method respectively where the classes are shown in different colors as labels. It is observed that water class is accurately segmented with fine boundary details using MRA method. The structural details of built-up areas are retained while correctly identifying shallow and deep water classes.

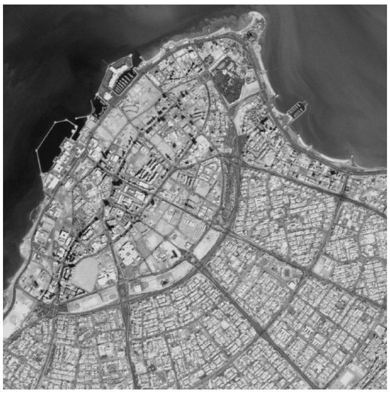

Figure 1. Original Kuwait City image (IRS data set)

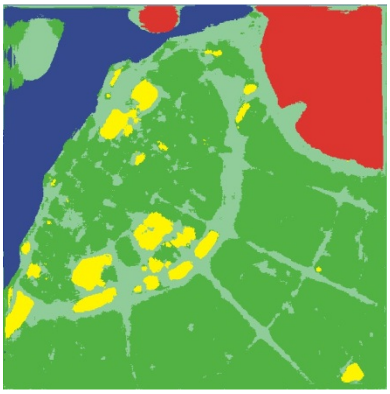

Figure 2. Classified image using traditional U-net

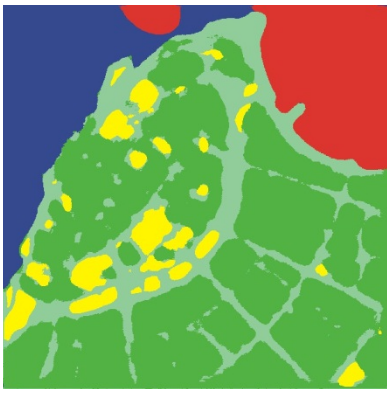

Figure 3. Classified result using proposed directional MRA assisted U-net

MRA assisted models perform better than the conventional U-Net in terms of both pixel accuracy and mean intersection over union (mIoU). The general trend observed is that the accuracies and mIOUs for the images are increasing as the order of the basis function for wavelet is increased. It is observed that contourlet model is outperforming the wavelet and the ridgelet based MRA. The pixel accuracies, intersection over union (IoU) for each class are detailed in Table 1.

Model	Pixel Accuracy	IoU				
		Shallow Water	Build Up	Open Area	Road	Deep Water
Plain U-net	82.86	79.92	77.62	56.60	46.81	88.86
U-net+Wavelet	89.22	98.44	83.08	69.66	56.85	98.69
U-net+Ridgelet	91.31	98.22	86.75	71.20	65.36	99.07
U-net+Contourlet	94.7	98.78	90.23	88.29	78.21	99.09

Table 1: Performance comparison for Kuwait image

In order to assess the effectiveness of different MRA features apart from visual interpretation, local and global consistency errors are computed as quantitative measures to evaluate the degree of matching between classified output and the reference site. The lower values of LCE and GCE demonstrate higher degree of matching for ridgelet (0.1, 0.14) and contourlet (0.094, 0.096) features when compared to wavelet-based (0.19, 0.22) and plain U-net based (0.29, 0.32).

Discussion and Conclusion

The results using proposed method shows better performance in terms of both visual interpretation and feature discrimination and are sufficiently robust against random pixels while preserving spatial arrangement and boundary continuity.

Although high resolution image is easily applied to distinguish different objects, some boundaries are not clear between objects with similar radiometric value, so it is difficult to classify these pixels. By analysis, it is clear that the performance using the directional MRA as a channel of input has improved when compared to the case without MRA. It is observed that deep water class is misclassified as shallow water and road classes at a number of places. There are no clear boundaries between built-up, road and building classes extracted, even though they are clearly visible in the input image.

Wavelet based features are able to classify water body or open areas from other classes probably due to their smooth texture compared to other land cover features. These classes normally do not have edges, directionality, spatial periodicity and characteristic scale, whereas the rest of the texture features possess complex appearances with pronounced anisotropy and spatial periodicity.

It is observed that built-up area, roads and open areas are properly classified, the corresponding producer’s and user’s accuracies are better for ridgelet and contourlets in comparison with wavelet based MRA. The overall accuracy using ridgelet and contourlet is 91.31% and 94.7% respectively as compared to 82.86% accuracy obtained from plain U-net approach. The highest user’s accuracy using wavelet method is found to be for open area class (which is in general more homogeneous as compared to other classes and does not contain structural details), however this is not as high as that obtained from ridgelet and contourlet assisted U-net.

On the basis of the above results, the directional MRA features significantly improved the classification accuracy over the plain U-net method.

Acknowledgements:

No funding sources are availed for this work.

References

Ansari, R. A., Buddhiraju, K. M., & Bhattacharya, A. (2019). Textural classification of remotely sensed images using multiresolution techniques. *Geocarto International*, 1-23.

Do, M. N., & Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on image processing*, *14*(12), 2091-2106.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Kampffmeyer, M., Salberg, A. B., & Jenssen, R. (2016). Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 1-9).

Kim, J. H., Lee, H., Hong, S. J., Kim, S., Park, J., Hwang, J. Y., & Choi, J. P. (2018). Objects Segmentation From High-Resolution Aerial Images Using U-Net With Pyramid Pooling Layers. *IEEE Geoscience and Remote Sensing Letters*, *16*(1), 115-119.

Mallat S. (1999). A wavelet tour of signal processing. Academic press, 1999.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.

Volpi, M., & Tuia, D. (2016). Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(2), 881-893.

Wang, H., Wang, Y., Zhang, Q., Xiang, S., & Pan, C. (2017). Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote Sensing*, *9*(5), 446.