International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Segmentation of Paddy Fields from A Remote Sensing Images Using Ai Based Learning

[1]Dr. Rajiv Kumar, [2]Dr. Shikha Maheshwari, [3]Dr. J. Rajendra Prasad, [4]Dr. Rajeeb Lochan Moharana, [5]Ms. Chalamalasetty Sarvani, [6]Dr. Ramesh Babu P.

Submitted: 01/10/2023 **Revised:** 20/11/2023 **Accepted:** 30/11/2023

Abstract: Hyperspectral image segmentation (HSI) is a technique that is commonly used to remove redundant and linked data from the original high-dimensional HSI spectral space while at the same time keeping the essential data in a low-dimensional subspace. The use of superpixels has been beneficial to a wide variety of applications, some of which are listed here. In this paper, for the very first time, zeroed in on how well-established super-pixel techniques can serve as a helpful first stage in hyperspectral analysis, with a concentration on classification. In addition to this, we make use of the network that is in the middle of the model, and after that, we employ the technique known as feature fusion to combine the features that originate from the various subnetworks.

Keywords: *PCA, Attention model., deep learning, paddy fields*

1. Introduction

People have been using the term superpixel to refer to a specific kind of pixel that possesses several characteristics that set it apart from other kinds of pixels. This type of pixel has received a great deal of attention since it was first used [1]. After being introduced for the first time, the concept quickly acquired traction and was implemented in a great many different contexts shortly afterward. The research literature includes descriptions of a variety of different superpixel segmentation techniques that are designed to work particularly well with natural images. These techniques are intended to work particularly well with natural images. In [2], which can be read in its entirety by clicking on the link provided above, we perform an in-depth analysis of 28 modern superpixel algorithms.

The writers of the paper [3] present a table in which they evaluate the performance of the SLIC algorithm in comparison to that of other methods that are currently considered to be the state-of-the-art. The researchers who are working in this field are spending a significant amount of their time and energy to investigating the processes that are used to divide superpixels in their research. The sector of the processing industry that is concerned with high-speed images has been paying an increasing amount of attention to superpixels over the course of the past few years.

This ability of superpixels makes the incorporation of superpixel segmentation into HSI processing intriguing. Because of this, the incorporation of superpixel segmentation into HSI processing is a desirable component. This is because superpixels can organize regular pixels into meaningful groupings based on their relative positions to one another in the picture, whereas regular pixels are unable to do this. The spatial configuration of the objects that are being depicted can also be replicated in the arrangement of the superpixels [4], which is another useful feature.

The following are just a few examples of the countless applications that have reaped tremendous rewards as a direct result of the utilization of superpixels. In many situations, standard techniques for the segmentation of pictures can be applied to superpixel images with only minimal adjustments necessary. This reduces the amount of work that needs to be done. The reason for this is that regular pictures only have a certain number of pixels in them, whereas superpixel pictures have a lot more of them all to themselves.

Even though dedicated HSI superpixel segmentation algorithms have not yet been developed, it is likely that natural image processing algorithms will be able to

[1]*Assistant Professor, School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh, India. Email ID: rajiv.kumar@bennett.edu.in*
[2]*Associate Professor, Directorate of Online Education, Manipal University Jaipur, Rajasthan, India. Email ID: shikha.maheshwari@jaipur.manipal.edu*
[3]*Professor, Department of IT, NRI Institute of Technology, Agiripalli, Andhra Pradesh, India. Email ID: rp.rajendra@rediffmail.com*
[4]*Assistant Professor, Seed Science and Technology, College of Agriculture, OUAT, Bhawanipatna, Odisha, India. Email ID: rlmoharana@gmail.com*
[5]*Assistant Professor, Department of Computer Science and Applications, Koneru Lakshmaiah Educational Foundation, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India. Email ID: sarvanich@kluniversity.in*
[6]*Associate Professor, Department of Computer Science, College of Engineering and Technology, Wollega University, Nekemte, Oromia Region, Ethiopia. Email ID: drprb2009@gmail.com*

function effectively in the new domain. This prediction is since natural image processing algorithms already exist.

In this article, we introduce a deep residual network architecture for the purpose of achieving rapid infrared denoising. This architecture makes use of attention, which is a feature that is unique to deep residual networks. One way to think about the network is as a convolution layer, which is made up of a variety of receiving fields, shortcut connections, and channel attention mechanisms. This is just one way to think about the network. Both modules are part of the multi-scale feature extraction module. In addition, a technique that is referred to as feature fusion is utilized to combine the features that are generated by the many different subnetworks.

2. Related works

Dimensionality reduction, also known as DR, is a common preprocessing tool that is used to extract pertinent information from an original high-dimensional HSI spectral space while simultaneously discarding data that is either irrelevant or correlated [5]. A small amount of data to train with, it is still possible to train the classifier in a manner that will result in accurate results being produced. According to the degree of supervision that they are subjected to, most of these methods either belong in the controlled category or the uncontrolled category.

When determining DR, these algorithms typically only consider the spectral characteristics and ignore any information that can be gleaned from the geographic features. This is because spectral characteristics are more easily measured than geographic features. In many cases, this is how events turn out to play out. It may be beneficial to provide a spatial context to spectral categorization algorithms because pixels that are nearby to one another in an image [6].

This is because it is customary for neighboring images in an image to share the same categories of interest with one another. Recent years have seen the development of superpixels, which allow DR applications to take into consideration the surrounding spatial environment more effectively. The invention of superpixels [7] paved the way for this to become a possibility rather than a pipe dream.

Analysis of the image correlation matrix is a method that has been developed and can be used to originally organize the highly correlated bands that are present in a hyperspectral image. This technique was developed and can be used because it has been proven to be effective. After being compiled into one large group in the prior phase, the bands are then put through an operation known as SuperPCA feature extraction. A spatially regularized local graph discriminant embedding (LGDE) strategy is recommended in the paper [8] to naturally incorporate spatial information into the LGDE model. This is done to achieve the desired result.

In addition, the kernel counterpart of the linear SLGDE model is extended to take into consideration the potential non-linearity. This was done to improve the accuracy of the model. This potential non-linearity could be the outcome of the complicated nature of the data acquisition process, as well as the effects of atmospheric and geometric distortions.

3. Proposed Method

In the following section, an in-depth description of the attention-based deep residual network that was recommended for use in HSI denoising is provided. This network was suggested for use because it can effectively reduce noise. Figure 1 provides a high-level overview of the infrastructure of our network. Yspatial is used to refer to an input noise band, and Yspectral is used to refer to the K bands that are immediately contiguous to it.

The multi-scale feature extraction module oversees gathering the necessary spatial context and spectral correlation data so that the data can be evaluated. This data is required for the module to perform its function. to carry out the analysis, it is necessary to have these facts. A specialized component of multilayer feature representation is utilized to build the noise.

The residual is subtracted from the topographical input, which finally ends up producing the clear signal. After that, we will move on to the subsequent step of offering a more in-depth description of the blocks and loss function that our network utilizes.

Channel Attention Block

The traditional CNN applies the same strategy to each channel of a feature map, which limits the representational power of deep neural networks. The traditional CNN is unable to learn in a discriminative fashion across channels, which is why this is the case. It has been brought to our attention that the feature maps of the spectral input have variable effects on the outcome of the denoising process.

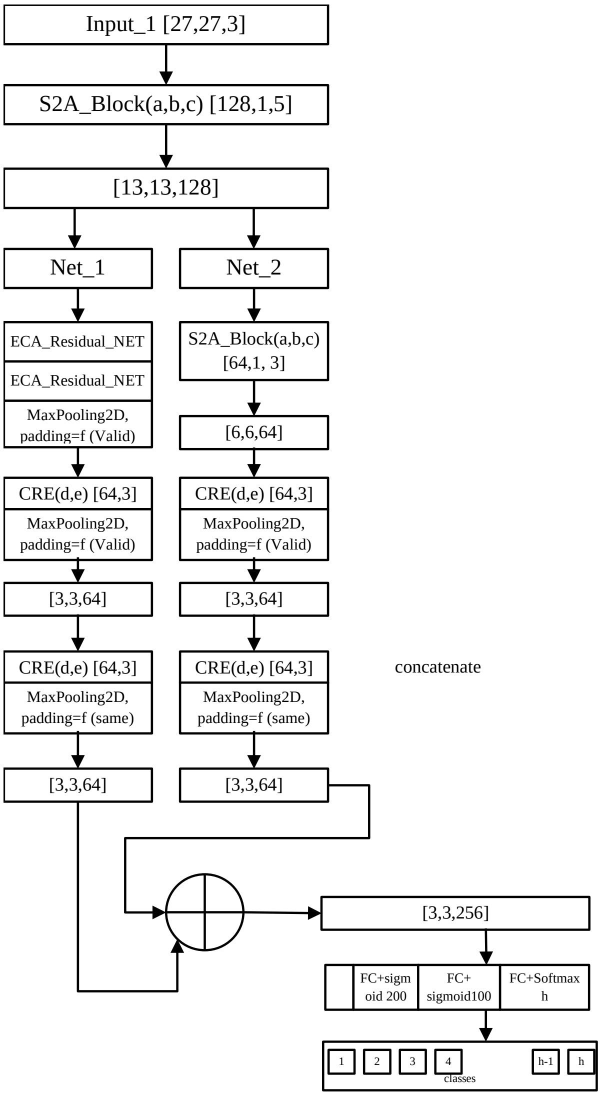

Fig 1: Attention based Deep Network

This indicates that the training that is supplied by our network ought to focus characteristics that are most notable to maximize their effectiveness. Our strategy for residual learning should also allocate greater weight to the convolution kernels so that we can improve the accuracy of our noise prediction. This will allow us to extract high-frequency information more effectively. As a result of these factors, we concluded that our system needed to be enhanced by the addition of a channel attention block, which would allow for the dynamic modification of feature representation.

Figure 1 depicts the internal structure of both our channel and our attention block.

$Fi = Fi-1 + WCA * Xi$ (1)

Where Fi and Fi1 are the input and output feature maps, respectively, and Xi is the residual component that was produced by using two stacked convolution layers with filter sizes of 3x3, respectively.

$Xi = W2 * \delta(W1 * Fi-1)$ (2)

where

Xi - residual component.

If we assume that weight sets (W1 and W2) and ReLU function, then the answer is Xi = W2(W1*Fi1). Using the Xi function known as the global average pooling will allow us to successfully complete this task.

We make use of a ReLU convolution layer that has a dimension of 1, which allows us to reduce the overall number of channels by a factor of r. Following the completion of this stage, a sigmoid convolution layer with a dimension of one is applied to bring back the initial number of channels and ensure that WCA is consistently within the range of 0 to 1.

We suggest making use of an attention-based deep residual network to acquire this mapping as quickly as possible. When utilizing the spatial-spectral information, input is received not only from the band that is being selected now but also from the K bands that are immediately adjacent to it. This ensures that the most accurate results are obtained.

The study combines the multi-scale features by using a convolution layer with different filter sizes, and we integrate the multi-level information by making use of a shortcut connection to achieve more effective noise suppression.

Both steps are carried out to improve the quality of the output. In addition, the channel attention technique is used to educate the network to pay attention to the supporting data and features that contribute the most to the denoising operation.

This is done by teaching the network to pay attention to the supporting data and features. to accomplish this, the network needs to be trained to pay attention to the data and characteristics that it is supporting. Instead of simply making a prediction, which would make the process of training more complicated, we reconstruct the outcome by using a residual mode. This makes the process much simpler.

For the model to make use of the spatial-spectral structural association, it must be simultaneously supplied with information pertaining to both the space it inhabits and the bands that surround it. Only then will the model be able to maximize its potential benefits from the association.

We can make use of the spatial-spectral structural relationship that already exists. We also make use of the centralized network that the model possesses, and after that, we use a method that is referred to as feature fusion to combine the information that we get from the different parts of the network. Feature fusion is one of the techniques that we use. In the final step, you will apply this combined feature to an entirely connected layer so that it can be used as an input for classification based on the combined feature. This will be done so that the layer can be completely connected.

4. Results and Discussions

The study maintains K as 64, r as 10 throughout the entirety of the training procedure. Both numbers, the down sample ratio and the trade-off value, correspond to their respective concepts.

The assignment of weights to nodes in the network, which are originally based on the shortened normal distribution, is the first step in the training process for the network.

The default specifications for TensorFlow are utilized in conjunction with the optimization tool Adam, which possesses a mini-batch size of 382, which is twice as many as the band number.

The training rate begins at an extremely low level of 0.0001, and it drops off very quickly at regular intervals that have been established in preparation. There are approximately 300,000 different permutations carried out in total.

Table 1: Training of Attention and non-attention models

Images	CNN	RNN	A-CNN
10	80.45	80.38	82.86
20	80.08	82.56	83.32
30	78.05	79.26	83.43
40	71.43	76.39	79.19
50	80.17	82.53	80.92
60	75.22	79.18	82.35
70	78.46	79.42	86.32
80	74.53	79.47	80.03
90	76.33	78.69	82.27
100	78.59	79.29	83.89
110	75.65	78.31	83.50
120	77.99	77.93	81.34
130	76.70	79.46	85.37
140	72.97	76.98	84.32
150	77.32	79.01	84.68
160	77.64	82.60	83.46

Table 2: Testing of Attention and non-attention models

Class	CNN	RNN	A-CNN
10	80.91	80.86	83.68
20	80.55	83.05	84.14
30	78.51	79.74	84.24
40	71.85	76.85	79.97
50	80.63	83.02	81.71
60	75.66	79.65	83.16
70	78.92	79.89	87.16
80	74.96	79.94	80.81
90	76.78	79.16	83.08
100	79.05	79.76	84.72
110	76.09	78.78	84.32
120	78.44	78.40	82.14
130	77.14	79.93	86.21
140	73.39	77.44	85.15
150	77.77	79.48	85.51
160	78.09	83.09	84.28

Table 3: Validation of Attention and non-attention models

Class	CNN	RNN	A-CNN
10	81.66	82.41	84.50
20	81.29	84.64	84.97
30	79.23	81.26	85.08
40	72.51	78.32	80.76
50	81.37	84.61	82.52
60	76.36	81.17	83.98
70	79.65	81.42	88.03
80	75.65	81.47	81.61
90	77.48	80.67	83.90
100	79.78	81.29	85.56

110	76.79	80.28	85.16
120	79.16	79.90	82.95
130	77.85	81.46	87.06
140	74.06	78.92	85.99
150	78.48	81.00	86.36
160	81.26	82.66	84.62

The results in Table 1-3 make use of the network that is in the middle of the model, and after that, we employ a technique known as feature fusion to combine the features that originate from the various subnetworks. The results show that the proposed network architecture can be used in HSI-denoising applications.

5. Conclusions

This paper proposes an attention-based deep residual network architecture for high-spatial-spectral denoising. The network is composed of a convolution layer of various reception fields, a shortcut connection, and a channel attention mechanism. In addition, the network that is in the middle of the model employs a technique known as feature fusion to combine the features that originate from the various subnetworks.

References

[1] Sun, Y., & Zheng, W. (2022). HRNet-and PSPNet-based multiband semantic segmentation of remote sensing images. *Neural Computing and Applications*, 1-9.

[2] Li, X., Li, Y., Ai, J., Shu, Z., Xia, J., & Xia, Y. (2023). Semantic segmentation of UAV remote sensing images based on edge feature fusing and multi-level upsampling integrated with Deeplabv3+. *PloS one*, *18*(1), e0279097.

[3] Su, Z., Wang, Y., Xu, Q., Gao, R., & Kong, Q. (2022). LodgeNet: Improved rice lodging recognition using semantic segmentation of UAV high-resolution remote sensing images. *Computers and Electronics in Agriculture*, *196*, 106873.

[4] Yang, L., Huang, R., Huang, J., Lin, T., Wang, L., Mijiti, R., ... & Du, X. (2021). Semantic segmentation based on temporal features: Learning of temporal–spatial information from time-series SAR images for paddy rice mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-16.

[5] Lan, Y., Huang, K., Yang, C., Lei, L., Ye, J., Zhang, J., ... & Deng, J. (2021). Real-time identification of rice weeds by uav low-altitude remote sensing based on improved semantic segmentation model. *Remote Sensing*, *13*(21), 4370.

[6] Lu, X., Jiao, L., Liu, F., Yang, S., Liu, X., Feng, Z., ... & Chen, P. (2022). Simple and Efficient: A Semisupervised Learning Framework for Remote Sensing Image Semantic Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1-16.

[7] Fan, Z., Zhan, T., Gao, Z., Li, R., Liu, Y., Zhang, L., ... & Xu, S. (2022). Land Cover Classification of Resources Survey Remote Sensing Images Based on Segmentation Model. *IEEE Access*, *10*, 56267-56281.

[8] Li, H. (2022). Multi-Scale Segmentation Method of Remote Sensing Big Data Image Using Deep Learning. *Journal of Interconnection Networks*, 2242004.

[9] Rajiv, A., Saxena, A.K., Singh, D., Awasthi, A., Dhabliya, D., Yadav, R.K., Gupta, A. IoT and machine learning on smart home-based data and a perspective on fog computing implementation (2023) Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across

[10] Chaudhury, S., Dhabliya, D., Madan, S., Chakrabarti, S. Blockchain technology: A global provider of digital technology and services (2023) Building Secure Business Models Through Blockchain Technology: Tactics, Methods, Limitations, and Performance, pp. 168-193.