Extraction of High-level and Low-level feature for classification of Image using Ridgelet and CNN based Image Classification

T Gladima Nisia and S Rajesh
Assistant Professor, Department of CSE, AAA College of Engineering and Technology, Sivakasi, India

E-mail: gladimab@gmail.com

Abstract. Remote Sensing image classification is an important research area for the recent time, because of its various application areas. Among the many available feature extraction methods, this paper uses the ridgelet based feature extraction method and those obtained features are combined with deep features obtained from the Convolutional Neural Network (CNN). Here, the Ridgelet’s are used to obtain the low-level features and CNN is used to obtain high-level feature. The system tries to construct the ridgelet filter for obtaining the low-level feature. The multi-resolution CNN is introduced using the concept of fusing high-level and low-level features via ridgelet and CNNs. Then, fused features are then classified and the output classified image is obtained. Experimental verifications are conducted on NWPU-RESISC45 dataset and the output results are provided to prove the best classification accuracies compared with the other proposed systems.

1. Introduction
Remote Sensing (RS) acts an important role in exploring the important details about the Land areas. This exploration results in obtaining useful information which can be utilized for many real time applications. Thus, RS image obtained is classified to get a clear idea about the Land cover areas present. The results of the classification depend completely on the feature extraction process. There are lot of methods present for feature extraction. There are many texture analysis methods using wavelet, ridgelet, and curvelet-based texture descriptors available.

The invention of various remote sensing sensors takes the remote sensing image classification field to the next level, thus increasing the quality of the classified images [1]-[4]. During the last decade there are lot of different classification methods are done using traditional methods such as Principal Component Analysis, independent component analysis [5], linear discriminant analysis [6], local fisher discriminant analysis [7], and local discriminant embedding [8]. The main drawback here is, the local neighborhood information is not considered and thus resulted in the noisy classification results. After this, many researches are done to introduce spatial information into land cover/ land use mapping. Then there arose a spectral-spatial based classification method include extended morphological profiles, extend multi-attribute profiles, composite kernels and multiple kernels [9], markov random field [10], and sparse representation model [11].

The wavelet-based classification is an important concept because of its time-frequency and multi-resolution framework. In addition to wavelets, numerous multi-resolution analysis tools, including
shearlet, contourlet, and curvelet, are extensively used because of its excellent capacity in representing features for classification. Qian and Ye [12] introduced 3D wavelet for texture features extraction for HSI classification. Another work proposed the 3D gabor wavelet for the same process in [13] and [14]. Another work [15], explored the multivariate models based on wavelet for supervised classification. Another work proposed, introduced a noise free classification method based on wavelet [16] and curvelet [17]. Here, the structure of filters does not depend upon the training sample count.

In recent times, deep neural network plays an important role for getting high-level features of the image. [18], [19]. Chen and Lin [20] proposed Stacked Auto-Encoders for classification of Hyperspectral Image. Then extended their research with 3D CNNs for Image classification [21]. Another work proposed [22] implemented three-dimensional convolution-recurrent networks to avoid the difficulties caused by 3-D CNNs. Zhong et al. [23] proposed new novel method namely agile CNNs for image classification. Zhao et al. [24], [25] proposed a system, that could learn multiple scale feature with the help of using different object sizes. Zhao et al. [26] classified the multispectral and panchromatic image using superpixel-level multiple local CNN. In [27], proposed a method namely multi-scale superpixel classification which could capture objects with different scales and sizes. Also [28] proposed a method for extracting hierarchical deep spatial feature for classification using hierarchical CNN. Cao et al. proposed another method for hyperspectral image classification using multi-feature fusion network [29], and then they proposed a method to obtain the relationship between high dimension sample in classification using deep learning [30].

This paper uses the ridgelet based feature extraction for low-level extraction of features. The ridgelet filter is built self-reliantly of sample available, with the requirement of setting up some parameters to attain a good representation of composite scenes. CNN figure out the high-level image features. Thus, proposed system takes advantages of both filters and it is entirely taken for extraction of feature and CNN operates by combining high-level and low-level features. Here, multiple resolution of low-level features is achieved using ridgelet filters, and multiple resolutions of high-level features are achieved using convolutional filters. Then the features obtained from both filters are stacked to decrease the necessity of Convolutional Neural Network on training samples, and thus to increase the classification ability of CNNs. The proposed method is applied on some sample VHR-RS image. The results obtained are found to exceed the classification accuracy of many existing methods.

2. Methodology

![Figure 1. Overview of the proposed system](image)

The proposed work utilizes the available deep neural network, namely CNN for classification. Here, each feature extraction layer holds both Convolutional and Ridgelet features in CNN. The convolutional
filter is applied on low-resolution features to overpower the noise, which maybe produced during classification results and ridgelet filter is applied on high-resolution features to obtain more detailed information. The overview of the system proposed is shown in figure 1.

2.1. Generating Low-Resolution Image Samples
The Very High-Resolution Remote Sensing (VHR-RS) image consists of more complete information, leading to the increase in the intraclass variation. Thus, this image resolution is condensed to reduce the intraclass variation, resulting in attaining the region uniformity within landcovers. Individual windows are applied on the image, to obtain numerous high-resolution samples. It is done by the process of nearest neighbour interpolation. Here, one high-resolution sample, may result in four low-resolution samples and the method is shown in figure 2. The equation (1) represents the interpolation.

\[ f_L(x + p, y + q) = f_H(x, y) \] (1)

![Image of interpolation process]

**Figure 2.** Process of generation of low-level feature

where, \( x = 1, 3, \ldots, M=1, y = 1, 3, \ldots, N-1, p=1,2, q=1,2 \). Here, the \( M,N \) represents the sample’s window size. The benefit of the method is, it maintains the same dimensionality for the low-resolution and high-resolution samples.

2.2. Designing of Ridgelet Filter
The designing of filter is an important step in extraction of features. The proposed system tries to design the filter with the ridgelet function. The main reason for using ridgelet is that in case of multi-variance function it has a strong ability than wavelet. Thus, the proposed method exploits the low-level feature by ridgelet filter. The specified ridgelet function is given as below in equation (2) & (3).

\[ \Psi_y(X) = a^{-1/2} \Psi(\frac{x_1 \cos \theta + x_2 \sin \theta - b}{a}) \] (2)

\[ \Psi(X) = e^{-x^2/2} - \frac{1}{2} e^{-x^2/8} \] (3)

Here, \( a \) represents scale parameter, \( b \) represents position parameter, and \( \Theta \) represents direction parameter, \( (x_1, x_2) \) is the coordinate of ridgelet filter, \( x_1, x_2 = 1, \ldots, N \), and \( N \) represents the ridgelet filter size. Ridgelet filter is designed by discretizing the parameter, \( (y,a,b) \).

2.3. A Convolutional Neural Network
The proposed system slightly modifies the general structure of the CNN. Two types of inputs are passed to the CNN. The first input is low-resolution RS image input sample and the second input is high-resolution RS image input sample. Let, \( X_{\text{High}} \) represents the input high-resolution sample and \( X_{\text{Low}} \) represents input low-resolution sample, \( l \) represents the level of the proposed CNN, and \( l=1, \ldots, L \).
stands for the depth of the networks. If \( l = 1 \), the convolutional feature \( F_{Low}^l \) and the ridgelet feature \( F_{High}^l \) are obtained by equation (4) and equation (5), respectively.

\[
F_{Low}^l = g(f(X_{Low} \ast K_l + b^l)) \tag{4}
\]
\[
F_{High}^l = g(f(X_{High} \ast R_l)) \tag{5}
\]

Here, * represents the convolutional operator, f (*) represents activation function and g(*) denotes pooling operator. The rectifier function is proposed for the activation function. \( K_l \) represents the convolutional filter and \( R_l \) represents the ridgelet filter and the value of \( l = 1 \). Ridgelet filter maintains the same size as that of the convolutional filter. Both convolutional features and ridgelet features are stacked on each other and they are produced as the output of the first layer. The output features are feed into the softmax classifier and the final classification results are obtained as output using the equation (6).

\[
\hat{C}_{j,a} = \frac{e^{W_{j}^f \cdot f_j}}{\sum_{c \in \text{class}} e^{W_{j}^f \cdot f_j}} \tag{6}
\]

Here, \( F_j \) represents the \( j^{th} \) sample’s output feature and \( W \) represents full connection weight vector and class represents a set of categories. \( \hat{C}_{j,a} \) denotes the probability of \( j^{th} \) sample belonging to class \( a \).

### 3. Experiment and Analysis

#### 3.1. Data Description

The experiments are done using NWPU-RESISC45 dataset. It is created by Northwestern Poly-technical University. The dataset consists of 31,500 remote sensing images with the total of 45 classes. Each class comprises 700 images of 256 X 256 pixels. The spatial resolution within the scene classes varies from 30.0m to 0.2m per pixel. The proposes system only considers the 10 classes for the experimental purpose namely airplane, airport, beach, bridge, cloud, desert, forest, dense_residential, commercial area and island. The sample input images are given in the figure 3.

![Sample Input images of NWPU-RESISC45](image)

**Figure 3.** Sample Input images of NWPU-RESISC45

#### 3.2. Experimental Setup

All the experiments are done using MATLAB 2018b. The window size for ridgelet filter is fixed as 4X4. The effective selection range of scale parameter, \( a \) is found within the range \([0.1, 1]\), position parameter \( b \), is found within the range \([0, N]\) and directional parameter, \( \Theta \) is found within the range \([0, \pi)\). In case, of CNN, it has 2 feature extraction layers, a max-pooling layer, fully connected layer and classification layer. The feature extraction layer - 1 contains 10 convolutional filter with of 10 X 6 size, and 6 ridgelet
filter of 10 X 6 size. The feature extraction layer - 2 contain 20 convolutional filter of 6 X 6 size, and 6 ridgelet filters of 6 X 6 size. The learning rate is set as 0.01.

3.3. Results and Discussions
Among the 31,500 samples available in the Dataset, 1:10 and 1:5 samples are considered for training and training the classifier respectively. Table 1 displays the classification results of the proposed system. The results of our system outperform the results of the existing classification system. The existing system classifies the remote sensing images with the usage of only CNN. And by doing so, the classification accuracy obtained is 89.2%. The classification accuracy displays that the proposed method excels for the given dataset and the accuracy is improved. The overall accuracy is 91.8%. The improvement in the result mainly depends on the combination of high-level feature extracted from CNN which is combined with the low-level feature extracted from the Ridgelet feature.

4. Conclusion
The proposed CNN model combines Ridgelet and CNN feature for the classification of VHR-RS images. The high-level features are obtained with the help of convolutional filters and low-level features are obtained with the help of Ridgelet filters to attain more exhaustive information. And at that time, high-level and low-level features are stacked upon one another to advance the strength and accuracy of the traditional CNN classification. By applying the proposed method, the classification accuracy is improved and the results obtained outperforms the previously available classification accuracies.

References
[1] Aderhold J, Davydov V Yu, Fedler F, Klausing H, Mistele D, Rotter T, Semchinova O, Stemmer J and Graul J 2001 J. Cryst. Growth 222 701.
[2] Ghamisi P, Maggiori E, Li S, Souza R, Tarabalka Y, Moser G, De Giorgi A, Fang L, Chen Y, Chi M, Serpico S P, and Benediktsson J A 2018, “New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning,” IEEE Geosci. Remote Sens. Mag., 6(3), 1043.
[3] Lv Z, Liu T, Benediktsson J A, Lei T, and Wan Y 2018, “Multi-scale object histogram distance for LCCD using Bi-temporal very-high-resolution remote sensing images,” Remote Sens., 10(11), 1809.
[4] Lv Z, Liu T, Benediktsson J A, and Du H 2019, “Novel land cover change detection method based on K-means clustering and adaptive majority voting using bitemporal remote sensing images,” IEEE Access, 7, 34425_37.
[5] Villa A, Benediktsson J A, Chanussot J and Jutten C 2011, “Hyperspectral image classification with independent component discriminant analysis,” IEEE Trans. Geosci. Remote Sens., 49(12), 4865_76.
[6] Xanthopoulos P and Pardalos P M 2013, “Linear Discriminant Analysis” New York, NY, USA: Springer, 237_80.
[7] Sugiyama M 2007, “Dimensionality reduction of multimodal labeled data by local Fisher Discriminant analysis,” J. Mach. Learn. Res., 8, 1027_61.
[8] Chen H T, Chang H W and Liu T L 2005, “Local discriminant embedding and its variants,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR), 5, 846_53.
[9] Wang Q, Gu Y, and Tuia D 2016, “Discriminative multiple kernel learning for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., 54(7), 3912_27.
[10] Moser G, De Giorgi A, and Serpico S B 2016, “Multiresolution supervised classification of panchromatic and multispectral images by Markov random fields and graph cuts,” IEEE Trans. Geosci. Remote Sens., 54(9), 5054_70.
[11] Liu J, Wu Z, Wei Z, Xiao L and Sun L 2013, “Spatial-spectral kernel sparse representation for hyperspectral image classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.,
[12] Qian Y, Ye M, and Zhou J 2013, “Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features,” IEEE Trans. Geosci. Remote Sens., 51(4), 2276–91.

[13] Zhu Z, Jia S, He S, Sun Y, Ji Z, and Shen L 2015, “Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework,” Inf. Sci., 298(20), 274–87.

[14] Shen L, Zhu Z, Jia S, Zhu J and Sun Y 2013, “Discriminative Gabor feature selection for hyperspectral image classification,” IEEE Geosci. Remote Sens. Lett., 10(1), 29–33.

[15] Regniers O, Bombrun L, Lafon V, and Germain C 2016, “Supervised classification of very high-resolution optical images using wavelet-based textural features,” IEEE Trans. Geosci. Remote Sens., 54(6), 3722–35.

[16] Quesadabarriuso P, Heras D B and Argüello F 2016, “Exploring the impact of wavelet-based denoising in the classification of remote sensing hyperspectral images,” Proc. SPIE, 9821, 982110.

[17] Qiao T, Ren J, Wang Z, Zabalza J, Sun M, Zhao H, Li S, Benediktsson J A, Dai Q, and Marshall S 2017, “Effective denoising and classification of hyperspectral images using Curvelet transform and singular spectrum analysis,” IEEE Trans. Geosci. Remote Sens., 55(1), 119–33.

[18] Lv F, Han M and Qiu T 2017, “Remote sensing image classification based on ensemble extreme learning machine with stacked autoencoder,” IEEE Access, 5, 9021–9031.

[19] Cheng G, Han J, Zhou P and Xu D 2019, “Learning rotation-invariant and Fisher discriminative convolutional neural networks for object detection,” IEEE Trans. Image Process., 28(1), pp. 265–78.

[20] Chen Y, Lin Z, Zhao X, Wang G and Gu Y 2014, “Deep learning-based classification of hyperspectral data,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., 7(6), 2094–107.

[21] Chen Y, Jiang H, Li C, Jia X and Ghamisi P 2016, “Deep feature extraction and classification of hyperspectral images based on convolutional neural networks,” IEEE Trans. Geosci. Remote Sens., 54(10), 6232–51.

[22] Seydgar M, Naeini A A, Zhang M, Li W and Satari M 2019, “3-D convolution-recurrent networks for spectral-spatial classification of hyperspectral images,” Remote Sens., 11(7), 883.

[23] Zhong Y, Fei F, Liu Y, Zhao B, Jiao H, and Zhang L 2017, “SatCNN: Satellite image dataset classification using agile convolutional neural networks,” Remote Sens. Lett., 8(2), 136–45.

[24] Zhao W, Guo Z, Yue J, Zhang X and Luo L 2015, “On combining multiscale deep learning features for the classification of hyperspectral remote sensing imagery,” Int. J. Remote Sens., 36(13), 3368–79.

[25] Zhao W and Du S 2016, “Learning multiscale and deep representations for classifying remotely sensed imagery,” ISPRS J. Photogramm. Remote Sens., 113, 155–65.

[26] Zhao W, Jiao L, Ma W, Zhao J, Liu H, Cao X and Yang S 2017, “Superpixel-based multiple local CNN for panchromatic and multispectral image classification,” IEEE Trans. Geosci. Remote Sens., 55(7), 4141–56.

[27] Li S, Jia X, Gao L, Zhang B and Peng L 2016, “Multi-scale superpixel spectral-spatial classification of hyperspectral images,” Int. J. Remote Sens., 37(20), 4905–22.

[28] Cheng G, Li Z, Han J, Yao X and Guo L 2018, “Exploring hierarchical convolutional features for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., 56(11), 6712–22.

[29] Cao X, Ge Y, Li R, Zhao J and Jiao L 2019, “Hyperspectral imagery classification with deep metric learning,” Neurocomputing, 356(3), 217–27.

[30] Cao X, Li R, Wen L, Feng J and Jiao L 2018, “Deep multiple feature fusion for hyperspectral image classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., 11(10), 3880–91.