Hyperspectral image remote sensing classification using RotBoost

I G A A D Indradewi, N L W S R Ginantra and M S Ariantini*

STMIK STIKOM Indonesia, Informatic Engineering, Denpasar, Indonesia

*suci.ariantini@stiki-indonesia.ac.id

Abstract. Hyperspectral images are able to provide more information because this kind of images have hundreds of spectral bands. In machine learning, the classification of hyperspectral data has several challenges, including high number of dimensions, number of output classes, and limited data references. The solution given to overcome the challenge is to use Ensemble Learning. The benefit of using Ensemble Learning is that we can improve the classification performance of hyperspectral data. One of the Ensemble Learning methods is RotBoost, which is a combination of the Rotation Forest and Adaboost methods. To find out the performance of the RotBoost method, this research used hyperspectral data of vegetation area in Indian pines, Indiana, USA that is provided by NASA's airborne visible infra-red imaging spectrometer (AVIRIS). The RotBoost method then compared with Rotation Forest method to find a better performance. Confusion matrix used to evaluate the accuracy of each method in the classification. The performance measured is overall accuracy obtained by doing 5-fold cross validation. This experiment also conducted to find the most optimal S (number of base classifiers) and T (number of iteration in Adaboost) values. Experimental results showed that RotBoost produces better accuracy than the Rotation Forest. S and T parameter values are also not very influential on RotBoost accuracy.

1. Introduction

Optical remote sensing has developed from grayscale imagery to multispectral and hyperspectral. The role of the development of hardware technology also supports the availability of high spatial, spectral and temporal images that are very useful for various applications. Compared to RGB color images that have 3 bands, hyperspectral images have hundreds of spectral bands. Hyperspectral images are able to provide more information than other types of images. However, the hyperspectral image itself has challenges including; high number of dimensions, number of output classes, and limited data reference. Therefore, a good method is needed to classify hyperspectral data. Ensemble learning is one promising solution considering its potential to significantly improve accuracy. One of the existing ensemble learning methods is random forest [1]. The random forest method can handle high dimensional data such as hyperspectral data properly. Improving the random forest method, namely Rotation Forest, can improve the accuracy and diversity of individuals in ensemble classifiers [2].

Chan classifies hyperspectral data using the random forest and adaboost tree based methods [1]. In their research, the random forest and adaboost tree based produce almost the same accuracy. In terms of training time, random forest produces better performance. Xia et al. performed classification using the Rotation Forest method by using cart as the base classifier [3]. The training data used has been projected...
into the new feature space using the principal component analysis (PCA) and several other methods. Experiments carried out resulted in better accuracy than bagging, random forest, and adaboost.

The RotBoost method proposed by Zhang and Zhang, is a method formed by combining the Rotation Forest and adaboost methods [4]. The RotBoost method significantly increases predictive accuracy compared to the Rotation Forest and adaboost methods when tested using 36 UCI datasets. RotBoost also produces better performance than bagging and multiboost. The purpose of this study was to try to classify hyperspectral data especially Indiana pines (AVIRIS) data using the RotBoost algorithm by using cart as the base classifier and using PCA to project data into the new feature space [5]. In this study also will look for the number of Rotation Forest (K) base classifier and the optimal number of adaboost (S) iterations.

2. Method

2.1. Proposed method
In this research we will make comparisons between RotBoost and Rotation Forest method. Our first step is to do data collection for our dataset from AVIRIS, then we split dataset to training and validation using cross validation. After splitting the data, the RotBoost and Rotation Forest method applied to the classification to get the prediction result. The next step we apply RotBoost and Rotation Forest method which we tested for predictions. Our end result computes the results of the performance of each algorithm. Our proposed method explains on figure 1.

![Figure 1. Our proposed method.](image.png)

2.2. Algorithm comparison

2.2.1. Rotation forest. Rotation Forest is a method for building an ensemble classifier based on the feature extraction [2]. To obtain the training data used by the base classifier, we perform a principal component analysis (PCA) on the K subset of the feature set that is divided randomly. All principal components are maintained to maintain variability information in the data and K axis rotations will be
a new feature for the base classifier. The idea of the Rotation Forest is to improve the accuracy and diversity of individuals in ensembles. This diversity will be pushed through feature extraction for each base classifier.

Table 1. Rotation forest algorithms.

| Input |
|-------|
| \(X\): training set (\(N \times n\) matrix) |
| \(Y\): Label of training set (\(N \times 1\) matrix) |
| \(L\): number of classifiers in the ensemble |
| \(K\): number of subsets |
| \(\{w_1, \ldots, w_L\}\): set of class label |

**Training Phase**

for \(i = 1 \ldots L\)

- Look for rotation matrix \(R_i^a\):
  - Share the set \(F\) feature to \(K\) subset: \(F_{i,j}\) (for \(j = 1 \ldots K\))
  - for \(j = 1 \ldots K\)
    - take \(X\) set data from features \(F_{i,j}\), declare it as \(X_{i,j}\)
    - Delete a random subset consisting of several classes on \(X_{i,j}\)
    - Take a bootstrap sample from \(X_{i,j}\) for \(75\%\) of the number of objects on \(X_{i,j}\). State the new set with \(X_{i,j}'\)
    - Perform PCA on \(X_{i,j}'\) to obtain coefficients \(C_{i,j}\)
  - for \(j = 1 \ldots K\), Arrange a rotation matrix \(R_i\) from \(C_{i,j}\) according to the following equation:
    \[
    R_i = \begin{bmatrix}
    C_{i,1}^{(1)} & C_{i,1}^{(2)} & \cdots & C_{i,1}^{(M_i)} & [0] \\
    [0] & C_{i,2}^{(1)} & \cdots & C_{i,2}^{(M_i)} & [0] \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    [0] & [0] & \cdots & C_{i,K}^{(1)} & C_{i,K}^{(2)} & \cdots & C_{i,K}^{(M_i)} \\
    \end{bmatrix}
    \]
    where \(M = n/K\)
  - Reorder columns from \(R_i\) according to the order of features \(F\), declare it as \(R_i^a\)

Build classifier \(D_i\) use \((X R_i, Y)\) as training set

**Phase Classification**

For an \(x\), share \(d_{i,j}(xR_i^a)\) as an opportunity from the hypothesis that \(x\) is class of \(w_i\). Calculate confidence for each class \(w_i\), with the following average combinations:

\[
\mu(x) = \frac{1}{L} \sum_{i=1}^{L} d_{i,j}(xR_i^a), \quad j = 1, \ldots, C
\]

- declare \(x\) into the class with the greatest confidence

2.2.2. RotBoost. Adaboost is a machine learning approach based on the idea of making accurate prediction rules by combining several rules that are relatively weak and inaccurate [6]. This method will manage the set of a weight \(D \alpha(i) = (i = 1, 2, \ldots, N)\) from the original dataset \(L\) and the value will be set the same at first. In the next iteration, the weight will be adjusted so that the weight of the sample that is misclassified in the previous classifier will be increased and the sample classified correctly will be lowered.

Based on the explanation of the Rotation Forest and adaboost above, the RotBoost classification technique that combines the two methods is explained by the pseudocode in table 2.

Table 2. RotBoost algorithm.

| Input |
|-------|
| \(X\): training set (\(\times p\) matrix) |
| \(Y\): Label of training set (\(\times 1\) matrix) |
| \(L\): \(\text{Training set, } L = \{x_i, y_i\}\) \(\times [X, Y]\) |
| \(K\): number of subset of attributes (or \(M\): number of input attributes for each subset) |
| \(W\): base classifier |
| \(S\): the number of base classifier |
| \(T\): the number of AdaBoost iterations |
| \(x\): data to be classified |
Table 2. Cont.

Training Phase

for \( s = 1, 2, ..., S \)
1. Calculate Rotation matrix \( R^s_T \) as done in the previous method and use \( L^s = [X R^s_T Y] \) as sebagai training set classifier \( C_s \)
2. Weight distribution initialization on \( L^a \) sebagai \( D_1(i) = \frac{1}{N}(i = 1, 2, ..., N) \)
3. For \( t = 1, 2, ..., T \)
   a. Based on distribution \( D_t \), extract random amounts of \( N \) from \( L^a \) with replacement to obtain a new set \( L^a_t \)
   b. \( \varepsilon_t = Pr_{l \sim D_t}(C^u_t(x_l) \neq y_l) = \sum_{i=1}^N I(C^u_t(x_l) \neq y_l) D_t(i) \) use \( W \) on \( L^a_t \) for practice classifier \( C_t^a \), and calculate the error from \( C^a_t \) as
   c. if \( \varepsilon_t > 0.5 \), then set \( D_1(i) = \frac{1}{N} (i = 1, 2, ..., N) \) and go to phase (a); if \( \varepsilon_t = 0.5 \) then set \( \varepsilon_t = 10^{-10} \) to continue the iteration
   d. select \( \alpha_t = \frac{1}{2} \ln(\frac{1-\varepsilon_t}{\varepsilon_t}) \)
   e. Update distribution \( D_t \) on \( L^a \) with
      \[
      D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t}, & \text{if } C^a_t(x_l) = y_l \\ e^{\alpha_t}, & \text{if } C^a_t(x_l) \neq y_l \end{cases}
      \]
      where \( Z_t \) is the normalization factor chosen and \( D_{t+1} \) aim distribution opportunities \( L^a \)
4. End for

Output

Class X labels are predicted by the final ensemble classifier \( C^* \) as \( C^*(x) = \arg\max_{y \in \Phi} \sum_{s=1}^S I(C_s(x) = y) \)

2.3. AVIRIS data collection

To find out the performance of the RotBoost method, we will experiment using a data hyperspectral vegetation area in Indian pines, Indiana, USA that is provided by Nasa's airborne visible infra-red imaging spectrometer (AVIRIS) [5]. The data consists of 200 spectral bands after removing 20 water absorptions bands. AVIRIS data measures \( 145 \times 145 \) pixels with a resolution of 20m / pixel. A detailed description of the number of classes that exist and the frequency of the number of classes in the data can be seen in table 3.

Table 3. The frequency of each class in AVIRIS data.

| # | Class            | Number of Samples | # | Class        | Number of Samples |
|---|------------------|-------------------|---|--------------|-------------------|
| 1 | Alfalfa          | 46                | 9 | Oats         | 20                |
| 2 | Corn-notill      | 1428              | 10| Soybean-notill | 972               |
| 3 | Corn-mintill     | 830               | 11| Soybean-mintill | 2455              |
| 4 | Corn             | 237               | 12| Soybean-clean | 593               |
| 5 | Grass-pasture    | 483               | 13| Wheat        | 205                |
| 6 | Grass-trees      | 730               | 14| Woods        | 1265              |
| 7 | Grass-pasture-mowed | 28           | 15| Building-Grass-Trees-Drives | 386 |
| 8 | Hay-windrowed    | 478               | 16| Stone-steel-towers | 93                |

2.4. Tools

In this research we use computer with minimum specification of processor is i5 and minimum RAM 8 Gb. We use MATLAB application to test the algorithm.

3. Results and discussion

To evaluate the accuracy of each algorithm, in this study using confusion matrix. Accuracy value is used to determine closeness of measurement to the true value.
The performance that will be measured is overall accuracy obtained by doing 5-fold cross validation. To find out the comparative accuracy between RotBoost and Rotation Forest, we will experiment using the following parameters:

- The number of input features in each subset: \( M = 10 \)
- 2. Number of base classifiers: \( S = 10 \)
- The number of iterations in Adaboost: \( T = 10 \)
- Feature extraction method: PCA
- Base Classifier: CART

The accuracy value of the RotBoost and Rotation Forest methods can be seen in Table 4. The results of comparing accuracy and standard deviation of both methods show that the RotBoost method gets better accuracy than the Rotation Forest method in classifying AVIRIS data.

| Method             | RotBoost | Rotation Forest |
|--------------------|----------|-----------------|
| Overall accuracy   | 0.87     | 0.76            |

Table 4. Accuracy of RotBoost and rotation forest.

In addition to making comparisons between RotBoost and Rotation Forest, we will also conduct experiments to find the most optimal \( S \) and \( T \) values. There are 5 types of \( s \) and \( t \) value configurations which consist of \((S = 50, T = 2), (S = 25, T = 4), (S = 10, T = 10), (S = 25, T = 25), (S = 4, T = 25), \) and \((S = 2, T = 50)\). The accuracy value and standard deviation of each configuration can be seen in Table 5. The accuracy of the RotBoost method on different \( S \) and \( T \) configurations does not show a significant difference. Except for configurations \( S = 50 \) and \( T = 2 \) which shows a considerable difference compared to other configurations.

| Overall Accuracy RotBoost | 0.84 | 0.87 | 0.87 | 0.88 | 0.88 |
|---------------------------|------|------|------|------|------|
| \( S=50, T=2 \)          | 0.84 |      |      |      |      |
| \( S=25, T=4 \)          | 0.87 |      |      |      |      |
| \( S=10, T=10 \)         | 0.87 |      |      |      |      |
| \( S=4, T=25 \)          | 0.88 |      |      |      |      |
| \( S=25, T=25 \)         | 0.88 |      |      |      |      |
| \( S=2, T=50 \)          | 0.88 |      |      |      |      |

Table 5. Accuracy of RotBoost with different \( S \) and \( T \) configurations.

From the Table 4 above shows that by using RotBoost method get the best result with accuracy 0.87. This concludes that in the case of predictions on the dataset AVIRIS, RotBoost algorithm has a better prediction rate than Rotation Forest.

![Figure 2](image-url)
4. Conclusion
A RotBoost method that forms an ensemble classifier, has been introduced to classify AVIRIS hyperspectral data. The RotBoost method combines the Rotation Forest method with adaboost. The base classifier used is CART and the data projection method used to form the rotation matrix is PCA. Compared to the experiments, the accuracy of the RotBoost and Rotation Forest methods in hyperspectral classification on AVIRIS data. Experimental results show that RotBoost produces better accuracy than the Rotation Forest. S and T parameter values are also not very influential on RotBoost accuracy.

Acknowledgment
We thank our colleagues from STMIK STIKOM Indonesia who provided insight and expertise that greatly assisted the research. We would also thank for the assistance and comments that greatly improved the manuscript.

References
[1] J Cheung-Wai Chan and D Paelinckx 2008 Rem. Sens. Of Env. 2999-3011
[2] J J Rodriguez, L I Kuncheva and C J Alonso 2006 IEEE Trans. On Patt. Anal And Mach Intell. 28 1619-1630
[3] J Xia, P Du and X He 2014 J. Chanussot, IEEE Geo. And Rem. Sens. Lett. 11 239-243
[4] C Zhang and J Zhang 2008 Patt. Rec. Lett. 29 1524-1536
[5] M F Baumgardner, L L Biehl and D A Landgrebe 2015 Pur. Univ. Res. Repo.
[6] F Yoav and R E Schapire 1997 Mach. Learn.: Proc. Of The Thirt. Inter. Conf. 1-9