Hyperspectral Image Classification using Random Forests and Neural Networks

B. T. Abe, O. O. Olugbara and T. Marwala

Abstract—Spectral unmixing of hyperspectral images are based on the knowledge of a set of unknown endmembers. Unique characteristics of hyperspectral dataset enable different processing problems to be resolved using robust mathematical logic such as image classification. Consequently, pixel purity index is used to find endmembers from Washington DC mall hyperspectral image dataset. The generalized reduced gradient algorithm is used to estimate fractional abundances in the hyperspectral image dataset. The WEKA data mining tool is selected to construct random forests and neural networks classifiers from the set of fractional abundances. The performances of these classifiers are experimentally compared for hyperspectral data land cover classification. Results show that random forests give better classification accuracy when compared to neural networks. The study proffers solution to the problem associated with land cover classification by exploring generalized reduced gradient approach with learning classifiers to improve overall classification accuracy. The classification accuracy comparison of classifiers is important for decision maker to consider tradeoffs in accuracy and complexity of methods.

Index Terms—Generalized reduced gradient, hyperspectral image, land cover classification, classifiers.

I. INTRODUCTION

The advancement in remote sensing technology has brought about new perspectives into image processing. The image objects are generally entrenched in a single pixel and hence cannot be detected spatially [1]. The traditional spatial based image processing techniques therefore, cannot be used. The importance of image processing that affords abundant data to be interpreted into useful information is crucial in image classification [2]. The preprocessing of remote sensing images before features extraction is important to remove noise and increase the ability to interpret image data more accurately. All generated images after preprocessing procedure must be as if they were obtained from the same sensor [3].

Hyperspectral instruments developed through remote sensing technology are capable of collecting hundreds of images corresponding to wavelength channels, for the same area on earth surface [4]. Hyperspectral images provide abundant spectral information to identify and differentiate between spectrally similar materials. The advantage of using hyperspectral sensors is the ability to provide a high-resolution reflectance spectrum for each pixel in the image [5]. Hyperspectral image data has found applications in areas such as mineral exploration, urban processes, agriculture, risk prevention, land cover mapping, surveillance, resource management, tracking wildfires, detecting biological threats and chemical contamination [6]-[7]. As a result, researchers have developed interest in the mathematical analysis of hyperspectral images [8]-[9].

The ability to obtain useful information from hyperspectral image data has stimulated researchers to use data mining methods to identify valid, novel, potentially useful and ultimately understandable patterns in data [3]. In remote sensing technology, there are varieties of earth objects present in the direct view of sensors because of the complexity of target objects and the limited spatial resolution of remote sensors. The information available in a certain pixels of a remote sensing image is a mixture of information on various ground objects, resulting into mixed pixels [4], [8]-[10]. The presence of mixed pixels has significant effects on some practical applications of remote sensing images such as information extraction, image classification and object detection. It is therefore, an important task in remote sensing study to discover objects and corresponding quantity present in the mixed pixel. This has led to the invention of hyperspectral remote sensing techniques to proffer solutions to the mixed pixel problem in remotely sensed imagery.

This work considers the problem of land cover classification of hyperspectral images by using a linear spectral mixture analysis technique, which is a commonly accepted approach to mixed-pixel classification. Land cover refers to the physical surface of the earth, including various combinations of vegetation types, soils, exposed rocks, water bodies and anthropogenic elements, such as agriculture and built environments [8], [11]. Our objectives are to (i) identify a collection of spectrally pure constituent spectral, which are referred to as the endmembers [12]-[13]. Thereafter, we express the measured spectrum of each mixed pixel as a linear combination of endmembers weighted by fractional abundances that indicate the proportion of each endmember present in the pixel [8], [13], (ii) explore Generalized Reduced Gradient (GRG) optimization algorithm to estimate the fractional abundance in the dataset thereby obtaining the numeric values for land cover classification [14]-[15] and (iii) to experimentally compare...
the performance of random forests and neural networks classifiers to examine the suitability of GRG algorithm for solving land cover classification problem. The application of machine learning techniques such as random forests and neural networks is usually to predict land cover of unsampled map units that help to retrieve important information from the scene. This work establishes that in comparison, though the classifiers’ performances are remarkable, random forests give better classification accuracy when compared to neural networks.

II. PROBLEM STATEMENT

The task of land cover classification can be formulated as a linear spectral unmixing problem. The linear spectral unmixing is a sub-pixel classification process that decomposes mixed pixels and determines the combination of fractional abundances. Based upon the linear unmixing model assumptions, each pixel at spatial coordinate \((l,p)\) for a particular band in a remotely sensed hyperspectral image \((I)\) having \(M\) number of bands can be formally expressed as [8], [10]:

\[
x(l, p) = \sum_{k=1}^{z} a_k(l, p) e_k + n(l, p)
\]  

(1)

The component \(x(l, p)\) is the measured reflectance value at the spectral coordinate \((l, p)\), \(e_k\) is the spectral response of the \(k^{th}\) endmember, \(a_k(l, p)\) is fractional abundance of the \(k^{th}\) endmember, \(n(l, p)\) denotes the spectral band error and \(z\) is the total number of endmembers. For linear spectral mixture analysis, each image pixel is a mixture of various endmembers and the spectrum recorded by the sensor is a linear combination of endmembers spectral [16].

Eq. (1) operates under two physical constraints on fractional abundances to account for the full composition of a mixed pixel [8], [12]. These are (i) Nonnegativity constraint, all abundances should be no negative, that is

\[
a_k(l, p) \geq 0, \forall \quad k = 1,2, ..., k
\]  

(2)

(ii) The abundances sum to one constraint,

\[
\sum_{k=1}^{z} a_k(l, p) = 1
\]  

(3)

The endmembers \(e_k, k = 1,2, ..., z\) can be extracted from the image \((I)\) by using a certain algorithm such as Pixel Purity Index (PPI) and Automated Morphological Endmember Extraction (AMEE) [8], [12]-[13] before the equation can be solved for a set of fractional abundances. Extant works on linear spectral unmixing problem [8], [9]-[10] have explored the least square method to estimate a set of fractional abundances as follows.

\[
a_{LS}(l, p) = (e^T e)^{-1} e^T x(l, p)
\]  

(4)

Eq. 4 can only satisfy the sum to unity of abundances, but the non-negativity of fractional abundances cannot always be guaranteed. The results obtained by the least square method therefore, are generally not optimal in terms of material quantification [8], [12].

In order to find a set of fractional abundances that simultaneously satisfy these two constraints, the following fully constraint linear spectral unmixing GRG optimization problem formulation has to be solved to minimize the spectral band error \(ml(p)\).

Minimize:

\[
a_{LS}(l, p) \in \Delta\{(x(l, p) - a(l, p)e)^T (x(l, p) - a(l, p)e)\}
\]  

Subject to

\[
\Delta_1 = \{a(l, p)| \sum_{k=1}^{z} a_k(l, p) - 1 = 0\}
\]  

(5)

\[
\Delta_2 = \{a(l, p)| 0 \leq a_k(l, p) \leq 1\}
\]  

(6)

(7)

In Equation (5), the expression \((X)^T\) represents the transpose vector of the vector \((X)\). To solve Equations (5-7) by applying GRG algorithm [14]-[15], to obtain an estimate of a set of fractional abundances, the PPI algorithm is first applied to extract endmembers from the hyperspectral image. The PPI method efficiently handles hyperspectral imagery as it provides a convenient and physically motivated decomposition of an image in terms of a relatively few components [17]. After a set of endmembers \(e = \{e_k\}\) is determined, the corresponding fractional abundances \(a(l, p) = \{a_k(l, p)\}_{k=1}^{z}\) in a certain pixel vector \(x(l, p)\) on the image \(I\) is estimated by using the GRG algorithm.

III. DESIGN METHODOLOGY

The design methodology of the study entails arrangement of steps that the input hyperspectral image undergoes for its land covers to be classified into one of the desired multiple classes. The input data has to be taken through four steps to obtain the desired classification result. Figure 1 shows the block diagram of the land cover classification process implemented in this study.

![Fig 1: Hyperspectral image classification procedure](image)

Before discussing the vital steps used in our design methodology, the input dataset is first introduced.

A. Dataset

Figure 2 shows the input airborne hyperspectral image of Washington D.C. dataset [18]. The sensor used measured pixel response in 210 bands in the 0.4 to 2.4µm region of the visible and infrared spectrum. It has 1208 scan lines with 307 pixels in each scan line, which is approximately 150 Megabytes. Bands in the 0.9 and 1.4µm region where the atmosphere is opaque have been omitted. The remaining 191 spectral bands are used for this study. The dataset shown in figure 2 contains seven ground cover types namely: Roofs, Street, Path, Grass, Trees, Water and Shadow.
Endmember Extraction (AMEE) [20]

Purity Index (PPI), N-FINDR and Automatic Morphological
locating appropriate endmembers. These include Pixel
algorithms have been developed to execute the task of
fractional abundances. For real hyperspectral data, various
construct their special distribution, associations and
element is a crucial task in hyperspectral data
identification of image
to map the data into lower dimension from higher
dimension at the same time preserve the main features of the
original data. The process is carried out to reduce the time
isolated during the processing of the hyperspectral data.
The algorithm does not generate an image different from the
original image. Instead, it is designed to reduce error by
finding minimum representation of the original image that
adequately keeps the original information for successful
unmixing in the lower dimension [19]. Among various
algorithms normally used for dimension reduction are;
Principal Component Analysis (PCA) and Maximum Noise
Fraction (MNF) transform. The aim is to ease computational
complexity and for compact information in transformed
components. MNF is used in this study because it is more
effective than PCA [20].

C. Endmember Determination
An endmember is known as a spectrally pure pixel that
portrays various mixed pixel in an image [21]. The method
of feature selection involves identifying the most
discriminative measurements out of a set of D potentially
useful measurements, where d ≤ D. Endmember extraction
has been widely used in hyperspectral image analysis due to
significantly improved spatial and spectral resolution
provided by hyperspectral imaging sensor also known as
imaging spectrometry [20]. Identification of image
endmember is a crucial task in hyperspectral data
exploitation, especially classification [13]. After
endmembers selection, various methods can be used to
construct their special distribution, associations and
fractional abundances. For real hyperspectral data, various
algorithms have been developed to execute the task of
locating appropriate endmembers. These include Pixel
Purity Index (PPI), N-FINDR and Automatic Morphological
Endmember Extraction (AMEE) [20].

This study applies the PPI algorithm [7], [22]-[23], which is
available in the Environment for Visualizing Images
(ENVI) to determine endmembers from the hyperspectral
image. The choice of the algorithm is motivated by its
publicity in ITTVIS (http://www.ittvvis.com/) ENVI software
that was originally developed by Analytical Imaging and
Geophysics (AIG) [24]. PPI generates a large number of n-
Dimensional vectors called “skewer” [7] [22], through the
dataset. N-FINDR fully automated method locates the set of
pixels with the largest possible volume by “inflating” a
simplex within the image data [21], [25]. In order to
generate the endmembers from the data, “noise whitening”
and dimensionality reduction are performed using MNF
transform [21], [22]. Then pixel purity score is obtained in
the image cube by producing lines in the n-dimension space
containing the MNF-transformed data. The spectral points
are projected on the lines and the points at the extremes of
each line are counted. Bright pixels in the PPI image
generally are image endmembers. The highest-valued of
these pixels are input into the n-dimensional visualizer for
the clustering process that develops individual endmember
spectrum.

D. Fractional Abundance Estimation
After determining the endmembers using PPI procedure, per
pixel fractional abundances of various materials is estimated
using GRG optimization method. This study presents six
endmember models to characterize the land cover structure
which are; Roofs, Street, Path, Grass, Trees, Water and
Shadow. Normalized numerical values of the fractional
abundant generated were calculated from the spectral
signatures of the land cover label signatures. The values
obtained were used to train the random forests and neural
networks classifiers for land cover classification.

E. Land Cover Classification
Random Forests (RF) and Neural Networks (NN) classifiers
are experimentally compared to examine their performances
in the field of land cover classification. The WEKA [26]
data mining software is selected as a tool to build the
classifiers from a training dataset of 3355 instances and 191
band features.

The RF ensemble classifier builds several decision trees
randomly as proposed [27] for classification of multisource
remote sensing, geographic data and hyperspectral imaging.
Various ensemble classification methods have been
proposed in recent times and they have been proven to
considerably improve classification accuracy. The most
famous and widely used ensemble methods are boosting
and bagging [27]. The boosting method is based on sample
re-weighting technique, but a bagging method uses
bootstrapping. RF classifier uses bagging or bootstrap
aggregating to yield an ensemble of classification and
regression trees. The method works by searching only a
random subset of the features for a split at each node to
minimize the correlation between the classifiers in the
ensemble. The method selects a set of features randomly and
creates an algorithm with a bootstrapped sample of the
training data [27]. This method provides a potential benefit
that it is insensitive to noise or overtraining because
resampling is independent of the weighting scheme
employed. For our experiment, 10 trees were constructed.
Out of bag error was 0.0605 while considering 192 random
features.

NN methods are general classifiers that can handle
problems with lots of parameters and can classify objects
even when the distribution of object in n-dimensional
parameters space is very complex. Research activities have
established that NN are promising alternatives to numerous
conventional algorithms [28]-[29]. They are data driven self-
adaptive technique that adjust themselves to data under
investigation without any explicit specification of functional
or distributional form. They are also universal functional
approximators that can approximate any function with

Fig 2: Hyperspectral image of Washington D. C. mall [18].
arbitrary accuracy [28]. Furthermore, they are capable of estimating the posterior capability that provides the basis for creating classification rule and carrying out statistical analysis [28]. Various NN models are available for classification purposes [28], but this paper focuses on MultiLayer Perceptron (MLP) that uses back propagation scheme to classify instances. The nodes in the networks are all sigmoid.

IV. RESULTS AND DISCUSSION
This section presents the results and discussion of our experiment for endmember determination and classification accuracy of the classifiers investigated.

A. Result of Endmember Determination
The first experiment performed aimed to obtain endmembers from image dataset using the ENVI software application. The MNF transformation of input hyperspectral image was performed for dimension reduction. The next stage of the endmember determination is to select a set of endmembers by applying the PPI algorithm on the extracted Region of Interest (ROI) pixels. Figure 3 shows this result, wherein the extreme pixels corresponding to the endmembers in each projection are recorded and total number of times each pixel is marked as extreme is noted. A threshold value of 1 is used to define how many pixels are marked as extreme at the ends of the projected vector.

![Endmember 1](image1.png)
![Endmember 2](image2.png)
![Endmember 3](image3.png)
![Endmember 4](image4.png)
![Endmember 5](image5.png)
![Endmember 6](image6.png)

Fig 3: Purest pixels occur at edges of the projected vector

Table I displays the land cover classes and the number of pixels extracted from the original image based on the ROI. The values of these pixels are input into the ENVI visualizer for the clustering process that develops individual endmember spectral. The pixels extraction mechanism enables the image spectral to accurately account for any errors in atmospheric correction.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Number of Pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roof</td>
<td>724</td>
</tr>
<tr>
<td>Paths</td>
<td>211</td>
</tr>
<tr>
<td>Water</td>
<td>703</td>
</tr>
<tr>
<td>Street</td>
<td>404</td>
</tr>
<tr>
<td>Trees</td>
<td>398</td>
</tr>
<tr>
<td>Shadow</td>
<td>97</td>
</tr>
<tr>
<td>Grass</td>
<td>818</td>
</tr>
</tbody>
</table>

![Endmember 1](image1.png)
![Endmember 2](image2.png)
![Endmember 3](image3.png)
![Endmember 4](image4.png)
![Endmember 5](image5.png)
![Endmember 6](image6.png)

Fig 4: Fraction images for each endmember

At the completion of specified iterations, a PPI image is created in which the value of each pixel corresponds to the number of times that a pixel was recorded as extreme. The bright pixels in the PPI image are generally the image endmembers to characterize the land cover structure.

B. Result of Land Cover Classification
RF and NN classifiers are evaluated using the error confusion matrix method, which is a representation of the entire classification result. According to [29], the error confusion matrix can be used to compute the overall accuracy and the individual class label accuracy. The error confusion matrix is a widely accepted method to report error of raster data and to assess the classification accuracy of a classifier. The matrix expresses the number of sample units allocated to each land cover type as compared to the reference data. The diagonal of the matrix designates agreement between the reference data and the interpreted land cover types [30].

Table II shows the result of the error confusion matrix for the performance of RF classifier. This result shows that roofs, paths, water, streets, trees and grass have 100% classification accuracy because none of their pixel’s member
is misclassified while shadow has one of the pixels’ members misclassified.

Table IV shows the result of error, Kappa statistics and Relative Absolute Error, RRSE with 0 implying no agreement between the classified land reference or ground truth data [33]. It ranges from 0 to 1, where: C = Confusion matrix are zero, then it means no area in the map is misclassified while shadow has one of the pixels’ members misclassified.

Table III records the result of the error confusion matrix for the performance of NN. From the table, it can be observed that roofs and grass are 100% classified while other land cover classes have some of their pixels misclassified.

Generally, the two classifiers performed excellently well. Considering individual class label, RF produces a higher level of classification accuracy per class label as compared to the NN. The entire accuracy assessment procedure is that otherwise there are certain percentages of misclassified instances, incorrectly classified instances, unclassified instances and the Kappa statistic are identified from the error confusion matrices [30]-[33].

If all the non-major diagonal elements of the error confusion matrix are zero, then it means no area in the map has been misclassified and the map accuracy is 100 percent. Otherwise there are certain percentages of misclassified instances [33]. In our experiment, RF as compared to NN has only 1 instance misclassified, while NN has 25 instances misclassified.

The Kappa coefficient of agreement is a measure of how well the accuracy of the classifier compares with the reference or ground truth data [33]. It ranges from 0 to 1, with 0 implying no agreement between the classified land cover and ground truth and 1 indicates complete agreement. Table IV shows the result of error, Kappa statistics and overall accuracy classification.

According to this result, there are no unclassified instances during the RF and NN classification procedures and the overall classification accuracies of the classifiers are seen to be comparable. It can be deduced from the predictions that RF outperformed NN. In addition, RF is more computational effective as compared to NN.

V. CONCLUSION

This study aimed to establish performance comparison between RF and NN classifiers for land cover classification. The performance assessment was done, giving overall accuracy and error confusion matrix. Experimental results demonstrate that the generation of RF and NN based land cover classification systems significantly improve overall accuracy. As a result, the classifiers can significantly contribute to land cover classification system as a source of analysis and increase its accuracy. The comparability and high accuracy performance of RF and NN indicates that the GRG method introduced in this study is effective for solving a linear spectral unmixing problem of land cover classification.

REFERENCES


TABLE II: RANDOM FORESTS ERROR CONFUSION MATRIX

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>724</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>a = Roofs</td>
</tr>
<tr>
<td>0</td>
<td>211</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>b = Paths</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>703</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>c = Water</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>404</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>d = Streets</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>398</td>
<td>0</td>
<td>0</td>
<td>e = Trees</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>f = Shadow</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>818</td>
<td>g = Grass</td>
<td></td>
</tr>
</tbody>
</table>

TABLE III: NEURAL NETWORKS ERROR CONFUSION MATRIX

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>724</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>a = Roofs</td>
</tr>
<tr>
<td>1</td>
<td>210</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>b = Paths</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>699</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>c = Water</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>401</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>d = Streets</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>398</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>e = Trees</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>f = Shadow</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>818</td>
<td>g = Grass</td>
<td></td>
</tr>
</tbody>
</table>

TABLE IV: CLASSIFICATION ACCURACY

<table>
<thead>
<tr>
<th>C</th>
<th>CCI</th>
<th>CIU</th>
<th>KS</th>
<th>MAE</th>
<th>RMSE</th>
<th>RAE (%)</th>
<th>RRSE (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>3354</td>
<td>1</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>99.9992</td>
</tr>
<tr>
<td>NN</td>
<td>3331</td>
<td>25</td>
<td>0</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>99.9990</td>
</tr>
</tbody>
</table>

Where: C= Classifier, CCI = Correctly Classified Instances, CIU = Incorrectly Classified Instances, KS = Kappa Statistic, MAE = Mean Absolute Error, RSE = Root Mean Squared Error, RAE = Relative Absolute Error, RRSE = Relative Squared Error
Bolante Tolulope Abe received her Master of Engineering at the Federal University of Technology, Akure, Ondo-State Nigeria in 2003. She is currently working towards her PhD degree in Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa. She is a lecturer at the Department of Electrical Engineering, Tshwane University of Technology, Pretoria, South Africa.

Oluayo, O. Oluhbara received B.Sc (Hons) in Mathematics with Cum laude, M.Sc in Mathematics with specialization in Computer Science and PhD in Computer Science. He joined the Department of Information Technology, Durban University of Technology in South Africa, where he is currently a Full Professor of Computer Science and Information Technology.

He is a member of the Association for Computing Machinery (ACM), New Zealand and Computer Society of South Africa (CSSA). He is a University Scholar and a holder of several academic awards and scholarships, including the International Federation of Information Processing (IFIP) TC2 sponsored by Microsoft Research Cambridge. He is a member of Marquís Who’s Who in the World, United States of America. He was awarded honorary refugee of Maejo International Journal of Science and Technology, Thailand in 2007-2010 and 2011.

Thilidziki Marwala holds a Bachelor of Science in Mechanical Engineering (Magna Cum Laude) from Case Western Reserve University, a Master of Engineering from the University of Pretoria, a Ph.D. in Engineering from the University of Cambridge and completed a Program for Leadership Development at Harvard Business School.

He was previously a Full Professor of Electrical Engineering, the Carl and Emily Fuchs Chair of Systems and Control Engineering, as well as the DST/NRF South Africa Research Chair of Systems Engineering at the University of Johannesburg, South Africa.