AN UNDERSTANDING RICE HYPERSPECTRAL REMOTE SENSING IMAGERY CLASSIFICATION FRAMEWORK

PURWADI, NANNA SURYANA, NOR AZMAN ABU
1Department of Informatics Engineering, STMIK AMIKOM Purwokerto, Purwokerto, Indonesia
2Pusat Teknologi Pengkomputeran Termaju (C-ACT), Fakulti Teknologi Maklumat dan Komunikasi (FTMK), Universiti Teknikal Malaysia Melaka (UTeM)
E-mail: 1purwadi@amikompurwokerto.ac.id, 2{nsuryana, nura}@utem.edu.my

ABSTRACT

The staple food for the Indonesian people is rice because in rice it contains a large number of calories for the intake of more than 200 million people. Hyperspectral is the sensors that can be used in a variety of applications, one of which is for rice monitoring. Hyperspectral is a sensor that is very well used to support precision agriculture because the information obtained is more detailed. One method of monitoring rice is to use a classification method. Many classification methods were carried out in previous hyperspectral studies such as unsupervised, supervised, statistical-based and so forth. Some methods have their own advantages and disadvantages. However, hyperspectral imagery has a large number of bands, requires sophisticated analytical methods to analyze it and requires a long process to extract priority information so as not to burden computing. In this paper discusses the state-of-the-art framework and step by step regarding the classification methods commonly used to rice monitoring. From the results of the review, it was found that the RBFN classification technique has the best accuracy compared to other classification techniques.

Keywords: Rice Monitoring, Classification, Hyperspectral, Remote Sensing, Framework

1. INTRODUCTION
1.1. Main Enterprise and Security Rice Production Issues

Indonesia is an agricultural country that has a very fertile agricultural land and is suitable to grow rice. Rice is a staple food for people in Indonesia because it contains a large number of calories for the intake of more than 200 million people [1]. The harvest period in this country is short enough for approximately 3 months. The problem is that rice production can be affected by climate, where in the dry season the farmland becomes empty and abandoned [2]. Rice is also a potential production in Indonesia which is a leading export commodity in its glory era. In 1997-1998, Indonesia and many countries in the world experienced a monetary crisis. In 1997-1998, Indonesia and many countries in the world experienced an economic crisis. The economic crisis in Indonesia itself has an impact on all sectors including the agricultural sector, the government has stopped the national program in the field of agriculture, which is known as the five farms, which is a program in order to increase rice production in a sustainable manner [3].

Rice is also a significant contributor to income generation for the country and absorbs a lot of manpower [4]. The decline in the number of rice fields causes rice stocks in Indonesia to decrease, thus importing rice from other countries [5]. The stability of rice production definitely increase food security in this country. Not only food security, but rice prices can also play a major role in shaping economic stability [6]. Indonesia needs an increase in agricultural productivity, especially rice. One way to increase rice production is to use new technologies that can sense (sensing) and monitor developments in paddy fields [7].

1.2. Hyperspectral Remote Sensing

The land use is classified into four types land cover namely housing, agriculture, commercial, and industrial. The type of agriculture is the most widespread of the four types. Vegetation
changes in agricultural land are essential for human survival. Because of changes in the area affects the weather, the composition of air, atmosphere, and disaster [8].

In the early 1980s hyperspectral remote sensing was the most superior technology in remote sensing. Promising technology in studying the material of the earth's surface spectrally and spatially. To obtain an inaccessible geochemical information from the earth's surface, the remote sensing technology developed worked by dividing visible and infrared broadband into hundreds of spectral parts [9]. Hyperspectral land and air sensors are now widely available, but for hyperspectral space sensors are available very little. In table 1, table 2 and table 3 some space agencies develop various types of hyperspectral sensors for air and space.

Table 1: Airborne Hyperspectral Sensors

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Spectral coverage (nm)</th>
<th>No. of Bands</th>
<th>Band width (nm)</th>
<th>Spatial Resolution (m)</th>
<th>Image tech</th>
<th>Country</th>
<th>Launched/Developer</th>
</tr>
</thead>
<tbody>
<tr>
<td>GERIS (Geophysical Environment Research Imaging Spectrometer II)</td>
<td>400 - 1000</td>
<td>24</td>
<td>25.4</td>
<td>1-10</td>
<td>Whisk broom</td>
<td>USA</td>
<td>1987/GRE corp.</td>
</tr>
<tr>
<td>AVIRIS (Airborne visible infrared imaging spectrometer)</td>
<td>380-2500</td>
<td>220</td>
<td>10</td>
<td>5-20</td>
<td>Whisk broom</td>
<td>USA</td>
<td>1987/JPL</td>
</tr>
<tr>
<td>CASI (Compact Airborne Imaging Spectrometer)</td>
<td>400-800</td>
<td>288</td>
<td>1.8</td>
<td>30</td>
<td>Pushbroom</td>
<td>Canada</td>
<td>1988/ITRES research Ltd</td>
</tr>
<tr>
<td>DAIS (Digital Airborne Imaging Spectrometer)</td>
<td>400-1200</td>
<td>72</td>
<td>15-30</td>
<td>1-10</td>
<td>Pushbroom</td>
<td>Europe</td>
<td>1995/GRE corp.</td>
</tr>
<tr>
<td>HYDICE (Hyperspectral Data Image Collection Experiment)</td>
<td>400 - 2500</td>
<td>10.2</td>
<td>210</td>
<td>3</td>
<td>Whisk broom</td>
<td>USA</td>
<td>1996/Naval research lab</td>
</tr>
<tr>
<td>HyMAP</td>
<td>400 - 2500</td>
<td>16</td>
<td>125</td>
<td>3-5</td>
<td>Whisk broom</td>
<td>Australia</td>
<td>HyVista Corp</td>
</tr>
<tr>
<td>AisaEAGLE</td>
<td>400 - 970</td>
<td>5</td>
<td>200</td>
<td>&lt;1</td>
<td>Whisk broom</td>
<td>Spectir Corp</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Spaceborne Hyperspectral Sensors

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Spectral coverage (nm)</th>
<th>No. of Bands</th>
<th>Band width (nm)</th>
<th>Spatial Resolution (m)</th>
<th>Swath (km)</th>
<th>Launch Year</th>
<th>Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate Resolution Imaging Spectrometer (MODIS)</td>
<td>400 – 800</td>
<td>32</td>
<td>250-1000</td>
<td>1500</td>
<td>May 2002</td>
<td>NASA</td>
<td></td>
</tr>
<tr>
<td>– AQUA</td>
<td>800 - 1455</td>
<td>36</td>
<td>250-1000</td>
<td>2300</td>
<td>Dec 1999</td>
<td>NASA</td>
<td></td>
</tr>
<tr>
<td>MERIS (Medium Resolution Imaging Spectrometer)</td>
<td>410 to 1050</td>
<td>15</td>
<td>10</td>
<td>Ocean: 1040x1200, Land &amp; Coast: 260x300</td>
<td>1150</td>
<td></td>
<td>ESA</td>
</tr>
<tr>
<td>Hyperion on EO-1</td>
<td>400 - 2500</td>
<td>220</td>
<td>10</td>
<td>3</td>
<td>7.5</td>
<td>Nov 2000</td>
<td>NASA</td>
</tr>
<tr>
<td>CHRIS (Compact High Resolution Imaging Spectrometer on PROBA-1)</td>
<td>438 - 1035</td>
<td>18-64</td>
<td>1.25-11</td>
<td>18-36</td>
<td>14-18</td>
<td>Oct 2001</td>
<td>ESA</td>
</tr>
<tr>
<td>HypSI (Hyperspectral Imager) on IMS-1</td>
<td>400 - 950</td>
<td>64</td>
<td>&lt;15</td>
<td>550</td>
<td>128</td>
<td>Apr 2008</td>
<td>ISRO</td>
</tr>
</tbody>
</table>
In Table 1, Tables 2 and 3 describe various information related to sensors, spectral coverage, number of bands, bandwidth, spatial resolution, image technology, country, plot, year of launch, launch/developer and agency. Hyperspectral remote sensing technology provides very detailed spectrum information for individual pixel images that mostly refer to remote (remote) sensing. [10].

### 1.3. Hyperspectral and Spatial Remote Sensing Spectral Signature Characteristics

Sensor Hyperspectral works by utilizing the photoelectric effect to work by collecting free electron-hole pairs in the detector element. The function of the incident photon is the number of electrons collected. In the radiometric calibration process, each detector element determines the illumination function. [11].

Hyperspectral data based on how to get it is categorized into two, namely taking from the air (for example HySpex) and taking from the ground (for example ASD Spectrometer). Measurements using the point spectrometer are strongly influenced by the angle of the position taken and the intensity of the electromagnetic radiation emitted. Techniques Taking the hyperspectral data from the air can be done based on the push broom line principle so that it can be a full spatial domain image of the 3-dimensional array. The easy principle is also called a scanner. Other retrieval techniques can be done by shooting electromagnetic radiation at certain angles in one area 2.

Remote sensing is one way to represent the surface of the earth through numeric numbers in the form of an array. Each pixel in the array/matrix represents the intensity of the electromagnetic wave radiation emitted by each band. To create a classification map, the image taken by the sensor is processed by forming a certain sign that reflects the difference between the pixel classes in one image. Certain reference pixels are trained to be able to guide other pixels so that they get pixels that match the class as in the thematic map in Figure 1 [12].

![Figure 1: Spectral Signatures of example 10 classes [12]](image)

### 1.4. Hyperspectral Application Into Paddy Field

Hyperspectral sensors can be used in a variety of applications, one of which is to monitor rice growth. Studies have been conducted to monitor rice growth by near-ground hyperspectral using a spectrometer. The reason used near-ground hyperspectral is because it has an excellent spatial resolution [13]. Rice canopy that has been studied previously that consists of two parts, namely sunlit and shaded. The detailed components studied are sunlit panicles, shaded panicles, leaves that are shaded and sunlit leaves. Shadows due to lighting are often a nuisance when the rice assessment process is done using hyperspectral [14]. Shadows due to lighting are often a nuisance when the rice assessment process is done using hyperspectral.
At present, the quality of rice is a major concern related to nutritional content. Protein is one thing that affects the good quality of bad rice. The estimated quality of rice based on protein content has been studied using near-ground hyperspectral technology. This study aims to estimate the protein content to assess the quality of rice [15].

Another hyperspectral application is controlling the growth and health of rice based on nitrogen content. Monitoring of nitrogen content can assist farmers in knowing the status of rice growth as well as monitoring the level of fertility [16].

Rice health is very influential on rice productivity. One application of the near-ground based hyperspectral application is to detect the leaves of rice plants. Pest attacks such as caterpillars that often attack rice leaves can be monitored with ASD Spectrometer. Pest-infected rice can be identified by the type of disease based on spectral characteristics that occur based on hyperspectral insect index of rice leaf folder (HIIRLF) [17].

1.5. Hyperspectral Limitation

Hyperspectral imagery has a large number of bands, requires sophisticated analytical methods to analyze it and requires a long process to extract priority information so as not to burden computing. The process is complex and requires a very accurate classification system in the classification of remote sensing images. The hyperspectral image is segmented into several homogeneous regions to derive class characteristics needed for mixed pixel decomposition which allows extracting the spectral and spatial features associated with each homogeneous region. One pixel from a remote sensing image possibly covers more than one object on the ground. Furthermore, there is the problem on recognizing and classify a particular geographical object due to overlapping of two or more associated spectral properties. Mixed problem may include uncertainty indicated by low classification accuracy [14].

Multispectral has much fewer bands than hyperspectral [18]. The amount of spectra generated by multispectral images is between 4 and 11, whereas in the hyperspectral image it can have tens to hundreds of spectral.

Some multispectral remote sensing images have a different spatial resolution in their spectral range, for example, Sentinel-2 and MODIS. The underlying issues are design considerations, hardware limitations, and further effects such as atmospheric absorption, requiring the use of different resolutions from multiple multispectral channels, so as to achieve satisfactory SNR (Signal to Noise Ratio) [19].

The use of multispectral can be said to be a cheap method for analyzing a land that is not reachable from a distance. But the multispectral image has its drawbacks, ie we can not order images at any instantaneous time associated with satellite orbit time. In addition, multispectral images often have an insufficient resolution in the analysis process. There are several cases that require immediate treatment, for example in cases of disaster and analysis that require a fast time can be done with a hyperspectral image [20].

2. RELATED WORKS

Research related to rice monitoring in hyperspectral imagery has been done using several approaches, namely supervised, unsupervised, statistical, semi-supervised, knowledge-based and feature selection as shown in Table 4. The most common method is supervised based method because this method has a higher level of accuracy compared to other methods. In detail, the classification methods that have been studied in various sector fields are presented in Table 5.

Table 4: Classification Methods Of Previous Studies In Hyperspectral Remote Sensing

<table>
<thead>
<tr>
<th>Approach</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supervised</td>
<td>[21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44]</td>
</tr>
<tr>
<td>Unsupervised</td>
<td>[45], [34], [46], [47]</td>
</tr>
<tr>
<td>Statistical</td>
<td>[48], [49], [50], [39], [40], [41], [44], [16]</td>
</tr>
<tr>
<td>Feature Selection</td>
<td>[22], [23], [24], [25], [28], [30], [31], [51], [35], [36], [52], [15], [49], [50], [53], [41], [42]</td>
</tr>
<tr>
<td>Semisupervised</td>
<td>[54], [55]</td>
</tr>
<tr>
<td>Knowledge Based</td>
<td>[39]</td>
</tr>
</tbody>
</table>

Table 5: Classification Methods of Previous Studies in Hyperspectral Remote Sensing
<table>
<thead>
<tr>
<th>Cit.</th>
<th>Techniques</th>
<th>Output</th>
</tr>
</thead>
</table>
| [45] | Nonlocal Total Variation  
Primal-Dual Hybrid Gradient Algorithm  
Stable Simplex Clustering  
K-Means  
Nonnegative Matrix Factorization  
Hierarchical Rank-2 NMF  
Merriman–Bence–Osher  
NLTV2  
NLTV1 | The Total Non-Local Allocation Algorithm is proposed. This algorithm can be consistent with high accuracy in a collection of synthetic data and urban data sets (SDA, Urban, and Salinas-A), both data produce finer results by identifying easier visual segmentation, and differentiating material classes that fail to distinguish algorithms others and can work well in anomalous detection scenarios with proper initialisation. |
| [21] | Nonlinear Support Vector Machines Classifier  
Random Forest Classifier  
Gaussian Maximum Likelihood Classifier | Results obtained:  
1) Working effectively on the classification of boreal tree species with HySpex VNIR 1600 sensor obtained kappa accuracy of more than 0.8 (obtained by manufacturer accuracy is higher than 95% in Pine and Spruce).  
2) The HySpex-SWIR 320i sensor has a limited role, but the band can precisely separate the Spruce and Pine species.  
3) The strong influence possessed by spatial resolution affects the level of classification accuracy which results in a 20% decrease in accuracy at a spatial resolution between 0.4 m and 1.5 m.  
4) In the SVM or RF classification method there is no significant difference. |
| [22] | Spectral Unmixing  
Support Vector Machines (SMVs)  
Unmixing-Based Feature Extraction | - |
| [23] | Support Vector Machine  
Random Forest  
LiDAR Feature Extraction, Feature Selection | High Kappa accuracy is effectively obtained in hyperspectral data for general macro class 93.2%, forest type 82.1% and single species 76.5%. Accuracy Classification decreases sharply in single tree species and is greatly reduced in forest type when applied to multispectral data while in general, macro class multispectral data is still very accurate reaching 85.8%. Experiments on LiDAR data that have high data densities provide more information for classification of tree types that have lower data density when combined with multispectral data or hyperspectral data. |
| [24] | Classification And Regression Tree  
Bagging  
AdaBoost  
Random Forest  
Principal Component Analysis  
Independent Component Analysis  
Maximum Noise Fraction  
Local Fisher Discriminant Analysis  
Support Vector Machines  
Logistic Regression Via Variable Splitting And Augmented Lagrangian | Experiments were carried out by combining the Rotation Forest and PCA, make the results of the classification obtained more accurate than the other methods (Random Forest and Bagging). Rotation Forests are a promising approach to producing classification of ensemble in hyperspectral remote sensing. |
| [48] | Gaussian Blur Radius  
Rolling Ball Radius  
Threshold Value  
Particle Size | The algorithm parameters are chosen to reduce false negative results. Potential quarantine treatment for evaluation of the feasibility of important methods is carried out, resulting in 1.0% for false error rates, with a false positive error rate of 11.1% and for the full sample obtained an overall error rate of 6%. The lowest overall error rate in the same sample was 2.0% with a false negative rate of 3.0% and a false positive rate of 1.0%. The error rate is much higher in the sample which has a very low infestation rate, obtaining the lowest overall error rate of 12.3%. |
| [25] | Parallelepiped Classification  
Principal Component Analysis | The results of hyperspectral imaging experiments in detecting hidden bruises on the fruit obtained classification results of 14.5% for the total error rate. For the acquisition of classification rates on positive errors where normal fruit is considered as a fruit that is bruised by 16.2% while for the classification of false errors where the bruised fruit is considered normal fruit at 12.6%. |
| [26] | Support Vector Machine (SVM)  
Bidirectional Reflectance Distribution (BRDF) | From the results of the experiment using the SVM classification method, it was obtained a high level of predictive accuracy for 15 species of species, namely 76%. Flight artifacts with a heterogeneous savanna landscape |
spectrum on-air spectral data can be reduced by applying a two-way reflectance distribution model. Crown level data and pixel integration, anisotropy reflectance, and depiction of crown over a very large area allow ecosystem scales to conduct species mapping and important starting points in monitoring biodiversity and ecosystem functions are the three challenges found.

[27] Random Forest Classification
 Classification Tree

The Random Forest Classification Algorithm is applied to classify land cover obtained by accurate results, 92% for overall accuracy and 0.92 for Kappa index. In training the addition of noise value and data reduction greater than 20% and 50% due to significant differences in kappa values, the Random Forest can be used. From the McNemar test a significant level of 0.00001 was obtained, which showed that overall performance was better than a random forest model in a single decision tree.

[49] Principal Component Analysis
 Wavelet Transform
 Feature Band Set
 Dimension Reduction
 Feature Selection
 Stepwise-Multilinear Regression
 Minimum Noise Fraction Transform
 Band Ratio, Asymmetric Second Difference Method
 Partial Least Squares Discriminant Analysis
 Partial Least Squares
 Multiple Linear Regression
 Band Difference
 Partial Least Squares Regression
 Integrated Principal Component Analysis With Neural Network
 Linear Discriminant Analysis
 Integrated Principal Component Analysis-Fisher’s Linear Discriminant
 Principal Components Regression
 Quadratic Discriminant Analysis
 Artificial Neural Network
 Support Vector Machines
 Wavelet Ahalanobis

At wavelengths of 606 and 636nm, different packaging in the atmosphere with a variety of spectra can be observed. The classification of packaging fillets used was successfully achieved with >80% use of hyperspectral imaging but was highly dependent on wavelength 606 and 636nm on spectral characteristics, which may occur because in muscle there is a difference in haem protein oxidation.

[28] Spectral-Spatial Classification
 Dimension Reduction
 SVM Classifier
 PCA Based Spectral

Applications for detecting head and neck cancer using the HSI sensor approach are described and create a spectral-spatial classification model framework. The proposed classification algorithm obtains average results in differentiating tumours and normal tissue in the animal head and neck cancer models, by 93.7% for sensitivity and 91.3% for specificity. Non-invasive accurate and quantitative cancer detection can be done by combining spectral-spatial classification methods with Hyperspectral Imaging.

[29] Feature Selection
 Hyperspectral Image Classification
 Kernel-Based Feature Selection Method
 Radial Basis Function
 Support Vector Machines
 BAHSCIp Method
 FSFS
 RFE

The subset feature in the KFS Method, BAHSCIp, RFE, FSFS and FS is proposed compared to classification performance, for the SVM method to use a multi-class strategy for all (OAA). The FSFS, KFS, BAHSCIp and RFE methods are tested for classification accuracy using specific numbers from the PAVIA dataset. Test results The KFS method is better than other classification methods in all cases. The KFS method has 95 higher classification accuracy features than the BAHSCIp method with 102 features, the FSFS method with 100 features, and the RFE method with 101 features. In the IPS dataset, the KFS classification method using half the features of the 0.937 classification accuracy level approaches the highest
<table>
<thead>
<tr>
<th>Reference</th>
<th>Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[32]</td>
<td>Extended Multi Attribute Morphological Profiles Generalized Composite Kernels Hyperspectral Imaging Multinomial Logistic Regression</td>
<td>Presented the SMLR and LORSAL Classification Method, SVM is adopted to produce final classification results.</td>
</tr>
<tr>
<td>[33]</td>
<td>Spectral Angle Mapper Support Vector Machine</td>
<td>Lighting change conditions such as the complex characteristics of the topographic environment and variations in incident lighting need to be considered in the performance appraisal of the classification method because it can influence the classification results so that there is the fact that techniques Spectral Angle Mapper not sensitive to Albeno. SVMs method it was found to be insignificant though, the ongoing classification on spectral training was selected from two data of the same population taken using the same sensor and in the same condition worked well but vice versa when both population data were taken with sensors and the conditions of the classification process different doesn't work properly.</td>
</tr>
<tr>
<td>[34]</td>
<td>BPNN Model Principal Component Analysis</td>
<td>Experimental results using seven optimal wavelengths in spectral in various discriminations of BPNN model the accuracy level of classification was 89.19% working better than the PCA model, the accuracy level was only 89.18%. IBRA model classification accuracy rate on data fusion is 94.45%, spectral data is 89.91% while image data is 88.09%. Data fusion is the best result from the IBRA model.</td>
</tr>
</tbody>
</table>
| [35] | Wavelet Transform K-Nearest Neighbors Support Vector Machine CNN Models | Using two different spectral ranges of hyperspectral images in the identification of rice seed varieties. CNN, KKN, and SVM which are three methods of machine learning are tested. CNN and SVM work more than KNN models whose accuracy is lower than 60% in spectral accuracy test 1 and spectral accuracy test 2. 3000 training samples from spectral range 2 are used to test CNN performance and obtained training classification accuracy results of 89.6% and set test...
<table>
<thead>
<tr>
<th>References</th>
<th>Methods and Models</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>[37]</td>
<td>Support Vector Machine, Random Forest</td>
<td>Rice seeds are classified using two methods, namely RF and SVM methods. Rice seed classification by combining spectral-based features can improve 84% precision from multi-label classification compared to only using visual features with a precision rate of only 74%.</td>
</tr>
<tr>
<td>[55]</td>
<td>The Least Square SVM, K-Means Clustering Algorithm, LS-SVM Algorithm Based On Clustering Combination Kernel Function, RBF Kernel Function</td>
<td>Identification of rice seeds with high precision utilises hyperspectral remote sensing technology and combines two classification methods: the Kernel Function method and the RBF Kernel method using the addition operation to achieve the highest accuracy of 91.95%.</td>
</tr>
<tr>
<td>[15]</td>
<td>Principal Component Analysis, Linear Discriminant Analysis Models</td>
<td>Leaves were categorised into two infected and healthy datasets. Validation sets were classified using the LDA / Linear Discriminant Analysis Model obtained an overall accuracy of 92%. The classification model by selecting five wavelengths: 1188.1339, 1377, 1432 and 1614nm obtained comparable results with the full spectrum image database.</td>
</tr>
<tr>
<td>[16]</td>
<td>Successive Projections Algorithm</td>
<td>Vegetation indices are recommended as a potential indicator in calculating the estimated amount of arsenic in the soil using three bands (R716 - R568) / (R552 - R568) which have just been developed using Photochemical Reflectance index (PRI) and Red Edge Position (REP).</td>
</tr>
<tr>
<td>[46]</td>
<td>SVM Linear Naïve Bayes, Principal Component Analysis, Kernel PCA</td>
<td>The accuracy of the classification of rice growth stages using PCA / KPCA to reduce the high-dimensional curse in hyperspectral images obtained the highest results of 93.33%, using linear data kernels and Naïve Bayes as classifiers.</td>
</tr>
<tr>
<td>[38]</td>
<td>ELM (Extreme Learning Machine)</td>
<td>Accuracy 92% using triangular basis function can produce good testing. Accuracy using 1000 node. In this research use 100 node, the result are:</td>
</tr>
<tr>
<td></td>
<td>Accuracy : 0.84</td>
<td>• Accuracy : 0.84</td>
</tr>
<tr>
<td></td>
<td>P_o : 0.8424</td>
<td>• P_o : 0.8424</td>
</tr>
<tr>
<td></td>
<td>P_e : 0.1176</td>
<td>• P_e : 0.1176</td>
</tr>
<tr>
<td></td>
<td>P_o - P_e : 0.7248</td>
<td>• P_o - P_e : 0.7248</td>
</tr>
<tr>
<td></td>
<td>1 – P_e : 0.8824</td>
<td>• 1 – P_e : 0.8824</td>
</tr>
<tr>
<td></td>
<td>Kappa: 0.8214</td>
<td>Kappa: 0.8214</td>
</tr>
<tr>
<td>[39]</td>
<td>Feature Band Set Object-Oriented Classification, PCA, CPS</td>
<td>The method proposed in the first image to differentiate vegetation was 97.84% for overall accuracy and 0.96 for kappa coefficient while for distinguishing rice varieties in the second figure the highest accuracy was 98.65% and 0.98 for the highest kappa coefficient.</td>
</tr>
<tr>
<td>[40]</td>
<td>Object-Based Spectral Features: Object Spectra Mean, Standard Deviation, Object Texture, After Haralick LCM Homogeneity, LCM Dissimilarity, LCM Entropy, Image Segmentation Based On Spectral Reflectance And The Scale Factor, Nearest Neighbour Classifier Integrated In eCognition Developer, GLCM: Contrast, Homogeneity, Dissimilarity, Energy and Entropy</td>
<td>Overall accuracy 3ac scenario is 91.3%</td>
</tr>
<tr>
<td>[41]</td>
<td>Seven Models Based On Spectral, Combined Spectral And Texture, Morphological, Texture, Combined Spectral And Morphological, Combined Morphological And Texture And</td>
<td>91.67% the highest level of accuracy obtained from a combination of spectral features, morphology and texture in this study, feature combinations have great potential to</td>
</tr>
</tbody>
</table>
### 3. THE HYPERSONTICAL CONCEPTUAL FRAMEWORK ASSOCIATED WITH CLASSIFICATION ALGORITHM

#### 3.1. Hyperspectral Ancillary Data Derived from Spectrometer/Field Radiometry

In post-classification processing, ancillary data is needed which is an important element that is determined by expert rules in the process of modifying the classification image. Many factors contribute to successful image classification. The most important example is the presence of high-quality ancillary data and the availability of imager tools, experience and skills of an analyst, and the design of appropriate classification procedures. In addition the classification performance is one of the factors that enable to enhance image classification with the combination of ancillary data, such as water, topography, road, soil, building, and census data with remotely sensed data [56].

Usually ancillary data is available in Geographical Information System format and able to be incorporated before, during, and after classification. The conceptual basis for the use of ancillary data is that the additional information is collected independently from the remote sensed data, increase the information available for separating the classes and for performing other kinds of analysis. For example in some regions vegetation patterns are closely related to topographic elevation, slope and aspect. The combination of elevation data with remote sensing data forms a powerful analytical tool because the two kinds of data provide separate, mutually supporting contributions to the subject being interpreted.

#### 3.2. Pixel-Based Vs Object-Based Agricultural Classification

The complexity of the high differences of spectral reflectance is factors that involve in pixel-based classification. It may produce the “noise” in classification result. This problem can be overcome thru object-based classification. The object-based classifier analyzes image based on image segments and extracts real word objects from those segments. Hence it makes more sense to analyze precise area or targets on the ground.

There are two main steps included in object-oriented classification, namely image segmentation and image classification. Steps in segmentation image, there are some strategies to generate objects; first is the integration of vector and raster data which vector data as the thematic layer. It is able to split image into segments and classification process is carried out based on this segmentation [56]. In addition, pixels can be merged into objects depending on the homogeneity of pixel values within an area, if there is no vector data available the classification based on objects will follow.

Image analysis through object-based classification involves the extraction of real-world objects based on their properties for example shape.
and size. It cannot be performed through the pixel-based classifier. Therefore, to analyze objects and to produce better result, some features related to objects might be used and clustered as physical features, topological feature, and context features. The disadvantage of this approach is evidenced in dealing with the meaningful image object. It is because there are no standard rules for image segmentation.

3.3. Preprocesing of Hyperspectral Data

Hyperspectral sensors have advantages over multispectral sensors in their ability to identify objects on the ground with more detailed features. The higher spectral resolution of hyperspectral data is suitable for the detection, identification and quantification of surface materials, as well as identifying natural, biological and chemical events. The resolution of hyperspectral image can also affect accuracy of classification. the complexity of an area can also be the cause of the problem. However, HIS data has a large narrow bands, so based on this specification it can be a challenge in the classification process [57].

3.4. Unsupervised and Supervised Classification

3.4.1. Unsupervised Classification

In hyperspectral image classification, one of the most widely used and common methods is the unsupervised classification methods. W. Zhu et al. [45] researched an unsupervised method with Primal-Dual Hybrid Gradient Algorithm and Nonlocal Total Variation. This study aims to solve optimization problems in quadratic and linear models. The NLTV algorithm applied in urban data sets and synthetic data sets consistently work with high accuracy, in both data produces smoother results by identifying easier visual segmentation, and differentiating material classes that fail to distinguish other algorithms and work well in the anomaly detection scenario with the correct initialization. This method can maintain the edge of the image well when minimizing and in the iteration, no matrix invasion is involved, but NLTV and other unsupervised classification methods in datasets with a large number of clusters do not satisfy the results achieved.

Classification of hyperspectral data with low resolution can be performed. Villa et al. using K-means clustering to overcome problems that highlight structural detection and classify unsupervised hyperspectral data that have a low level of spatial resolution. From the experiments conducted on the proposed method obtained results that are superior to the classical unsupervised classification method both from the quantitative side and from the perspective of the visual point of view when the area with mixed materials is located at the scene. It takes a higher number of iterations, higher resolution improvement factors and a larger pixel count in this study. The advantage of this method is that no other source is needed in addition to HSI [58].

To classify HIS, unsupervised classification can be combined with feature extraction methods. AVIRIS is used to take hyperspectral images. Experiments on real hyperspectral images as many as two images obtained the BCFE method is superior in simplifying dimensions compared to conventional feature extraction techniques such as LDA and PCA [59].

3.4.2. Supervised Classification

In the study of Dalponte et al. classification has been carried out on hyperspectral images using two sensor data. In the evaluation process uses an accuracy method to classify boreal forest. The spectral range used ranges from 400 nm to 1700 nm. The results of this study discuss him spatial level namely pixel map versus the tree-level map. By using HySpex VNIR images targeted at boreal tree species, kappa accuracy is above 0.8, while for Pine and Spruce targets the accuracy obtained is very good, which is 95%. Classification using SWIR bands can be used to separate Pine and Spruce plant species even if only using this band itself. The use of spatial resolution turns out to have a strong effect on the level of accuracy in classification. In this study, more than 20% decreased accuracy at spatial resolution between 0.4 m and 1.5 m. While using the SVM or RF method, there is no significant difference [21].

Comparisons of the three satellite data are multispectral, hyperspectral, and LiDAR have been analyzed by applying different densities points. The results show that effective hyperspectral data is used for general macro-classes, single plant species, and forest types with kappa accuracy values of 93.2%, 76.5% and 82.1%, respectively. Based on observations it can be concluded that hyperspectral data is good for general macro-classes. Whereas for multispectral and LiDAR data has an accuracy level under hyperspectral data [23].
In other ways, Support Vector Machines (SVM) is a powerful method that often use in HIS classification. Many researchers [26][29][32][33] have used the SVM method for the classification process on hyperspectral imagery. The results showed that by using SVM on hyperspectral imagery, good accuracy was obtained for each minor classes. SVM is free from distribution algorithms that can overcome poor statistical estimates. SVM can work very well if the training spectra used comes from the same population data during the classification process (both data are taken using the same sensor and in the same conditions). Many problem of hyperspectral imaging can be solved by SVM such as, a complex area [26], Huges Phenomenon Problem [29][28], homogenous area [22], and low density data problem [23].

However, SVM also has many disadvantages, SVM performance is very sensitive to the training samples used. SVM does not have important features of the model solution in nonlinear cases and only functions as a black box. Therefore, various methods to improve accuracy have been developed by combining various support methods such as, feature selection for enhance the accuracy or dimension reduction, and statistical improvement method. The combination SVM with feature selection method, many researcher [22][28][31][53] have proposed to avoid the common problem in hyperspectral classification such as a large narrow bands.

Computational complexity in image classification can be reduced based on Rotation Forest. This method can be used for supervised classification for hyperspectral images that have low spatial resolution and can be used to detect the structure of hyperspectral images. Hyperspectral remote sensing can be classified using the Rotation Forest method then the results are compared with other method approaches such as Bagging, Random Forest, SVM and AdaBoost. The Rotation Forest method with PCA transformation in the experiments carried out obtained more accurate results than the Random Forest, AdaBoost and Bagging methods. This shows that in hyperspectral remote sensing methods that are very promising to produce ensemble classifiers are the Rotation Forests method [24].

Land cover classification is a complex area that have to explored. Rodriguez et.al. proposed Random Forest (RF) method to resolve a complex area in HIS. The experimental results on land cover classification using RF algorithms obtained an overall accuracy of 95% and Kappa index 0.92, this shows the RF method is very accurate in the land cover classification. In training data reduction and noise greater than 50% and 20% due to significant differences in kappa values, the Random Forest can be used. From the McNemar test, a significant level of 0.00001 was obtained, which showed that overall performance was better than a random forest model in a decision tree. The ability to determine important variables is non-parametric high classification accuracy properties. In producing the final classification, RF has difficulty understanding the rules because the same resampling dataset produces several classification trees [27].

In hyperspectral classification, hughes phenomenon and dimension problem are common problems that to be faced. One of methods is by using Principal Component Analysis (PCA) [28][53][49][50]. To find the main components in the hyperspectral imaging system and the classification and quality assessment features used by PCA. The large data or number of narrow bands and high computational time of HIS can be handled by using PCA. PCA will extract the main features from the spectral images.

The same research was carried out [52], feature mining is the important rule and commonly used as a recommendation method. This research proposes a method of reducing advanced and conventional features, with details of several commonly used techniques for hyperspectral data analysis. Developed feature mining techniques that include linear and nonlinear, parametric and nonparametric methods, supervised and unattended, all methods attempt to identify information space.

Feature mining can be simplification measurement including important data and can be used to ensure classifiers avoid Hughes problems, works well and reliably this process is easy to do if there is enough expert knowledge or laboratory research. Difficulties in feature mining include excessive calculation, overfitting in learning, the meaning of curves full of physical interpretation, and noise sensitivity has been investigated, and spatial information must be incorporated into 3D-DWT texture features [51].

Variations in the spectrum contained in the class cause excessive classification of spectral homogeneous areas which can result in classification of these areas there is noise of salt and pepper, this is one of the challenges faced in the classification of hyperspectral images [53], but this problem is better solved using segmentation.
method. Segmentation is usually used as a pre-classification process.

Extreme learning machine (ELM), differential evolution (DE), feature extraction of hyperspectral images have been done [30]. DE-ELM aims to minimize training errors and output weight norms. Existing learning methods are used for comparison, binary, regression and multiclass problems are characteristic of integrated formulations in ELM. DE-ELM uses a multi output node configuration, where the number of classes must equal the number of nodes. In terms of classification and computational accuracy time, the DE-ELM method is very effective. ELM provides better classification accuracy relative to SVM which is up-to-date and faster because the solution is very simple, only requires kernel matrix inversion obtained from the calculation of training samples.

Statistical method in hyperspectral classification can be done by using partial least squares regression (PLS), Discriminant PLS, Stepwise MLR, Gaussian blur radius, and rolling ball radius [49][50][48]. The advantages by using statistical method are easy computational, but limited availability of the number of bands, and difficult to be used for on-line inspection.

3.5. Satellite Acquisition in Rice Field

Based on the way of data collection, Hyperspectral can be acquired in two ways, namely near-ground and airborne-based. One way of acquisition of near-ground hyperspectral acquisition is by using a tool called ASD field spectrometer or HySpex, while hyperspectral data can also be taken from space by aircraft (Airborne HyMAP). Table 6 explains that the spectral range taken for the purposes of rice field analysis is varied, for example the range between 380-1030 nm in the category of Visible / Near-Infrared has been studied to identify rice varieties. In the spectral range 874 to 1734 with the NIR Region category can also be used to explore the feasibility of identifying rice seed varieties and multivariate data analysis. Specific spectra can also be used for rice and rice canopy analysis, such as RGB and NIR.

Data retrieval in hyperspectral can be done in 2 ways, namely near-ground and airborne-based. At near-ground, the ASD device can be used. Field spectrometers or HySpex are used to monitor the stage of the rice growth stage, while the Airborne using Aircraft (Airborne HyMAP) is used for fertility, detail components, rice canopy, rice assessment, rice health.

Hyperspectral imagery has a large number of bands, requires sophisticated analytical methods to analyze it and requires a long process to extract priority information so as not to overload computing. The method commonly used in the pre-classification hyperspectral data pre-processing process includes Wavelet Transform (WT), Artificial Neural Network (ANN), Partial Least Squares Discriminant Analysis, Principal Component Analysis (PCA), Feature Band Set (FBS), Dimension Reduction, Feature Selection, Stepwise-Multilinear Regression, PLSR, Minimum Noise Fraction Transform, Band Ratio, Asymmetric Second Difference Method, Partial Least Squares (PLS), Multiple linear regression (MLR), Band Difference, Integrated PCA With Neural Network, Integrated Principal Component Analysis (PCA) - Fisher’s linear discriminant (FLD), Principal Components Regression (PCR), Wavelet, Quadratic Discriminant Analysis, Ahalanobis, Linear Discriminant Analysis, Support Vector Machines (SVM).

Weaknesses and strengths in hyperspectral: complex processes. One pixel of a remote sensing image may include more than one object on the ground. Resolution and complexity of the image of a hyperspectral area can also affect classification accuracy, problems in recognizing and classifying certain geographical objects because of the overlap of two or more spectrum related properties.

4. DISCUSSION

4.1. Hyperspectral Band Selection for Rice Field

The spectral commonly used for hyperspectral classification on agricultural land is 380-1030nm and 874 to 1734. 380-1030nm spectral is included in the Visible/Near-Infrared category, used for identification of rice varieties while for spectral 874 to 1734nm is included in the NIR Region category used for exploring the feasibility of identifying rice seed varieties and multivariate data analysis. Specific spectra can also be used for rice and rice canopy analysis, such as RGB and NIR.
spectral information and images to improve rice varieties and quality discrimination at HSI. Spectral range 380 - 1030nm and 874 - 1734nm is used for rice identification to identify rice seed varieties using a combination of HSI and convolutional neural networks.

Spectral range 874 - 1734nm and 1039 to 1612nm are used to examine the feasibility of rice seed varieties using HSI and multivariate data analysis.

Spectral range is 900 to 1700 to detect explosive rice in the hatchery stage at the initial infection stage. Spectral range of 400 to 850 is used for classification on rice plants.

The combination of spectral ranges 440-510nm, 520-590nm, 63-685nm, 690-730nm, 760-850nm is used to classify rice plants into certain classes starting from the growth stage to the harvest period and to determine the level of productivity of cultivars. The spectral range of 390 to 1050nm is used to provide a fast, precise and non-destructive information base identifying the origin of rice in real-time system development.

Spectral range 360 - 1025nm is used to check the spectral leaves and panicles of rice in the sunlit canopy section and not to evaluate the relationship of the spectral index of leaves and chlorophyll content in the shadow effect.

4.3. Potential Techniques on Rice Classification Using Hyperspectral Imagery

The most widely used classification method is SVM. This method has several advantages, namely: SVM has a regularization parameter, which makes users think about avoiding over-fitting; SVM can use kernel tricks, so you can build expert knowledge about any problem using the kernel; SVM is defined by a convex optimization problem (no local minima) that has an efficient method (e.g. SMO); SVM is an estimate for limits on the level of test error, and there is a substantial body of theory behind it that shows it must be a good and better idea; SVM shows the advantages of handling small, non-linear and high-dimensional sample data; SVM is based on minimum structural risk (SRM), and SVM has a high generalization capacity and can provide flexible and easily calculated solutions. SVM also has many disadvantages, SVM performance is sensitive to training samples. SVM functions as a black box and does not have important features of the model solution in nonlinear cases. High computational computing influences hyperspectral large-scale data classification.

In research [29], SVM is a powerful method because it can be applied to kernels such as RBF in research with 0.947 highest accuracies. In another study, the method that began to be developed based on artificial neural networks is the Neural RBF network. In the study [62] the results of accuracy obtained reached 100% using RBFN. The 400-900nm acquired spectrum range applied to cross-linked weed, wheat and broad bean plants taken under field conditions for four years can be classified with varying success using MLP and STEPDISC analysis and RBF networks. This model selects twelve wavelengths (480, 485, 490, 520, 565, 585, 590, 595, 690, 720, 725, 730 nm), three wavebands (B, G, NIR), and five spectral vegetation indices (B/G), G, R/B, R/G, NIR/B, RVI) So, this is a potential method that can be used in hyperspectral classification.

The advantage of ELM being able to build classifier model classifiers that can produce good testing accuracy as we expected, the model can be built using data without handling the curse dimensions, does not mean that pre-processing is not important if we can preprocess the data into optimal band choices and use models this, maybe we can get improvised results and faster. The lack of accuracy in testing the sigmoidal activation function, sin function, and radial is still below 70%.
Figure 2: The Framework of Rice Classification on Hyperspectral Imagery

The framework in Figure 2 is the basis of the process of classifying rice in hyperspectral images. Five steps are done, step 1 is very important for the continuation of the process in the future, for example, the selection of AVIS/GVIS sensors, and the band is selected based on the objectives to be achieved. Supporting data such as ground truth and rice outlook are needed as a comparison or can also be processed into the material.

Step 2 discusses pre-processing data before entering the core classification process, for example, area study selection and mosaicking process if needed. Hyperspectral data is data that has many bands or high resolution, therefore it is necessary to optimize the resolution and feature extraction/selection. However, every data in the form of images has the opportunity to have noise and requires a filtering process to improve image quality.

Step 3 is the core of the rice classification process using this hyperspectral data. In this process, the information extraction process or classification of rice and non-rice can be done by means of unattended / supervised classification, signature analysis, changing the index threshold or combining several data sources, for example between hyperspectral data and Economic Research Services (ERS) data series. This aims to increase the validity of the information produced.

Step 4 is a part that can be done to improve the performance of step 3, for example by filtering, extracting spatial feature or information, object-based voting that can be done by experts, and re-learning.

Step 5 is GIS analysis/presentation, including the classification mapping process, expert knowledge, rice information related to information about the quantity or quality of rice. In this step, a performance test is also carried out as an ingredient to analyse the durability and accuracy of the proposed method.

5. CONCLUSION & RECOMMENDATION

Hyperspectral remote sensing is a tool that can be used for the classification process in all circumstances. Especially in agriculture,
hyperspectral is used to assess all rice areas. Rice fields or even non-rice fields can be classified using a number of recommended methods, namely, RBFN, SVM, ELM. However, of the many RBFN methods the most widely used method because they have a high level of accuracy that is 100% using STEPDISC and can be used in hyperspectral data that has mixed pixel problems.

The bottleneck of hyperspectral imagery is high resolution and very large number of bands. This condition can cause the computing process to run slowly. Therefore, to overcome this, dimension reduction, resolution optimization can be done at the pre-processing stage. The filtering process is an important process and is usually done at the stage of pre-processing and post-processing which is used to improve the quality of hyperspectral images.

There are five steps commonly used in processing the classification in the rice area, namely the GIS combination/selection of data, hyperspectral RS Pre-processing, hyperspectral RS Processing, hyperspectral post-processing, and GIS Analysis/Presentation. In these five categories, what distinguishes rice processing from non-rice is an additional analysis of rice observation. Rice observation can be done to calculate rice estimates or improve system accuracy and can also function as validation. Validation can be done with expert knowledge and compared to field conditions.

Hyperspectral data can be obtained through two types based on how they are taken, namely AVIS and GVIS. AVIS imagery is usually used for the purpose of a very wide land cover classification, while GVIS is used as a close observation with a limited area. GVIS is usually used for more detailed observation of a small sphere of rice.

ACKNOWLEDGEMENT

The authors would like to thank the financial support from STMIK AMIKOM Purwokerto and Universiti Teknikal Malaysia Melaka for their assistance in this research.

REFERENCES:


SPECTRAL VARIATION BETWEEN RICE CANOPY COMPONENTS USING SPECTRAL FEATURE ANALYSIS OF NEAR-GROUND HYPERSPECTRAL IMAGING DATA National Engineering and Technology Center for Information Agriculture, Jiangsu Key Laboratory for Information Agric,” in 2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2016.


