Abstract— Remote sensing technology plays an important role in agriculture applications, especially for paddy growth stages classification, which is a critical process in predicting crop production. The analysis of multi bands data covering very large swath areas using iterative methods such as neural network or SVM will certainly cost much computation time. This paper addresses this problem by taking advantage from a non-iterative tuning capability of Extreme Learning Machine (ELM) for paddy growth stages classification using MODIS (Moderate Resolution Imaging Spectroradiometer) remote sensing images. The accuracy of classification is measured by Cohen’s kappa. Seven classes are used in classification, with consist of six classes for paddy growth stages and one class for dominated cloud. The contribution of this study is a new ensemble incremental approaches based on random bootstrap resampling for basic ELM (B-ELM) and Error Minimized ELM (EM-ELM) are applied to build multi-class classifier using two types of hidden nodes function, i.e. additive and radial basis function (RBF) hidden nodes. The classification results were compared each other with these two types of hidden nodes. Our ensemble incremental approach successfully classify seven paddy growth stages and significantly improve the overall kappa coefficient to 10.2% higher with only in average 7 nodes addition overhead.

Index terms— Remote sensing, MODIS images, classification, extreme learning machine (ELM), an ensemble.

I. INTRODUCTION

Direct human observations based statistical calculation and estimation of rice production currently used in Indonesia, often causes irregularities since the results tend to be excessive or over-estimated. More reliable harvest area estimation for paddy fields is a critical issue to support the National Food Security Program that have been promoted and coordinated by Indonesian government. In addition, an accurate and timely rice conditions monitoring and rice harvest area estimation are certainly needed.

In our previous work, we proposed a rice yield prediction method based on airborne hyperspectral images with spectral domain using Genetic Algorithms based New Sequence Principal Component Regression (GA-NSPCR) [1]. Hyperspectral airborne campaign is very expensive, whereas MODIS images are largely available for free. In the absence of hyperspectral imagery, we extended our previous method to work with 1,000 meter spatial resolution of MODIS images to classify nine paddy growth stages classification with One Against All (OAA) strategy and balanced branches strategy (BB-OAA) of Support Vector Machines (SVM) [2].

In current work, we further extend our previous work on paddy growth stages classification with a new ensemble incremental approach of ELM. We used the MODIS surface reflectance product computed from six bands of MODIS with 500 meter spatial resolution to build some classifiers using ELM which is directly applied for multiclassification.

ELM, originally proposed by Huang et al. [4], have been successfully applied in remote sensing such as [7] for classification of five soybean varieties using Hyperion hyperspectral data and [9] for land cover classification using ETM+ multispectral data set in comparison to a backpropagation neural network. ELM has also been used for paddy growth stages classification from hyperspectral data without dimensionality reduction [3], where the performance was measured by Cohen’s kappa for five activation functions, i.e. sigmoidal, triangular, sine, hardlim, and radial.

Extending basic ELM, Feng et al. [5] proposed an incremental approach referred to as Error Minimized Extreme Learning Machine (EM-ELM) to determine the number of hidden nodes automatically in generalized Single-hidden Layer Feedforward Networks (SLFNs) by updating their output weights during hidden nodes addition into the network, which in theory will bring the ELM significantly faster to reach convergence and reduces its computational cost. We view such a fast and efficient learning would be beneficial in dealing with high dimensional and high volume of remote sensing data.

This paper addresses the implementation of ensemble incremental approaches for both B-ELM and EM-ELM, that will increase the accuracy of paddy growth stages classification. In addition, this is also very important and urgent issue on how to predict the harvest area effectively through growth stages classification using MODIS remote sensing images.
II. DATA SAMPLES

A series of MODIS images used in this study are taken during March to July 2012. The acquisition is done in parallel with field campaign of paddy growth stages observation around Karawang and Indramayu District of West Java.

We processed MODIS surface reflectance product computed from the MODIS bands 1, 2, 3, 4, 5, 6, and 7 (centered at 648 nm, 858 nm, 470 nm, 555 nm, 1240 nm, 1640 nm, and 2130 nm, respectively) with 500 meter spatial resolution, and all images were also preprocessed in two stages for geometric and atmospheric corrections. In this study, we neglected band 5 of MODIS due to an error caused by its sensor, and used the rest of six bands for analysis.

A standard paddy field map 2010 issued by Indonesian Ministry of Agricultural was used for separating paddy field area images from non-paddy field areas. The terminology of nine paddy growth stages that refers to International Rice Research Institute (IRRI) is used in previous work, where all samples are obtained by reconstructing hyperspectral data into MODIS data with 1,000 meter spatial resolution. Since MODIS image has lower spectral resolution than hyperspectral, it is very difficult to distinguish nine paddy growth stages in actual. Therefore, this study used the terminology of six paddy growth stages for analysis purpose, that commonly used by Indonesian Ministry of Agriculture, i.e. vegetative-1, vegetative-2, generative, harvesting, bareland (harvested), and ploughing.

Like hyperspectral remote sensing data[4], in spectral domain, each pixel of MODIS image has specific spectral signature depends on the responds of reflectance of sunlight toward to all object covered by an area 500m x 500m on the ground which representing any dominated paddy growth stages according to ground reference data. By this way, we collected 6 classes for paddy growth stages from MODIS images along paddy field area around Karawang and Indramayu District of West Java. Figure 1 shows the spectral signatures of a sample pixel from each class in the normalized intensity range [0-1] after calibration.

The total number of available data samples was 259 which consists of six classes for paddy growth stages and one class for cloud. A summary of the data samples can be seen in Table I.

<table>
<thead>
<tr>
<th>No</th>
<th>Class</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vegetative-1</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>Vegetative-2</td>
<td>69</td>
</tr>
<tr>
<td>3</td>
<td>Generative</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>Harvesting</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>Bareland</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>Ploughing</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>Cloud</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Total samples</td>
<td>259</td>
</tr>
</tbody>
</table>

III. METHODOLOGY

A. An Ensemble

The main idea of ensemble methods is to build a classifier by integrating multiple models of machine learning as much as possible, and in such a way, it can be used for improving its performance. The original ensemble method is Bayesian averaging [8], but recently many methods have been developed. [10] proposed an ensemble method for ELM by implementing feature segmentation and nonnegative matrix factorization (NMF) to the original data to improve the classification efficiency. Meanwhile, an ensemble method used in this paper is “Manipulating the training samples”. The learning algorithm is run several times, each time with a different subset of training samples based on random bootstrap resampling (RBR). In order to maintain the ratio of data samples for each class to be balanced during learning, the data samples were divided into two parts (80% for training and the rest 20% for testing) for each class data samples and randomly recurrence at least 20 times using RBR.

B. Extreme Learning Machine (ELM)

In general, a major disadvantages of traditional feedforward neural networks is due to considerably slow gradient-based learning algorithms that are extensively used to train the neural networks. In such learning algorithms, all the parameters of the networks are tuned iteratively and leads further to slower computation and larger memory usage to maintain previous iterations information. In contrast, ELM originally developed by Huang et al. [4], is designed as a generalization of the single-hidden-layer
feedforward neural networks (SLFNs) where all the parameter of networks do not necessarily to be tuned iteratively. The ELM randomly chooses instead a number of hidden nodes and analytically determines the output weights of SLFNs, so that it can learn extremely faster than traditional SLFNs. The decision function of ELM with \( n \) hidden nodes can be expressed as

\[
f(x) = \sum_{i=1}^{n} \beta_i g_i(x), \quad x \in \mathbb{R}^d, \beta_i \in \mathbb{R}
\]  

where \( g_i(x) \) is the output hidden layer corresponding to the input samples \( x \), and \( \beta_i \) is the output weight vector between output hidden layer and output. According to the network structures, there are two types of hidden nodes used in SLFNs. The first one is additive hidden nodes, which is defined as

\[
g_i(x) = g(a_i \cdot x + b_i), \quad a_i \in \mathbb{R}^d, b_i \in \mathbb{R}
\]  

where \( a_i \) is the weight vector connecting the input layer to \( i \)th hidden nodes, and \( b_i \) is the bias of \( i \)th hidden nodes. And the second one is Radial Basis Function (RBF), which defined as

\[
g_i(x) = g(b_i \|x - a_i\|), \quad a_i \in \mathbb{R}^d, b_i \in \mathbb{R}^+
\]  

where \( a_i \) and \( b_i \) are the center and impact factor of the \( i \)th hidden nodes respectively. Equation (1) can be written in matrix form as

\[
F = G\beta
\]

If the number of hidden layer is equal to the number of training samples, equation (4) can be easily accomplished by using conventional least square method. However, in most cases the number of hidden nodes much less than the number of training samples, then this problem can be solved by

\[
\beta = G^+F, \quad G^+ = (G^T G)^{-1}G^T
\]

where \( G^+ \) is the moore-penrose generalized inverse of matrix \( G \).

C. Proposed Incremental approach algorithms

The main objective of this paper is to provide optimal number of hidden nodes of ELM for multiclass classification with iteratively generated nodes. This paper presents 2 kinds of ensemble increment approaches. The first one is by using B-ELM as a learning, which iteratively train along with increasing number of hidden nodes, which explained above by equation (1) – (5).

The second one is by using error minimum extreme learning machine (EM-ELM) proposed by [5]. In contrast, EM-ELM is a simple and effective approach to automatically determine the number of hidden nodes in generalized SLFNs [5]. This approach can add random hidden nodes to SLFNs one by one or group by group, during the growth of the networks, the output weight \( \beta \) are updated incrementally by

\[
\beta_{k+1} = G_{k+1}^T F = [U_k] F
\]

where,

\[
D_k = (I - G_k G_k^+ \delta G_k)^+ \quad U_k = G_k^T (I - \delta G_k^T D_k)
\]

and \( k \) indicates a number of hidden nodes in iteration progress. The algorithm can be summarized by figure 2.
IV. RESULTS AND DISCUSSION

Incremental approaches of ELM are applied to MODIS remote sensing image with 500 meter spatial resolution. As shown in Figure 3a and 3b, some hidden nodes added to the network may instead increasing the residual error (red dots) and do not affect the accuracy improvement in case of B-ELM. To tackle this problem and to improve the accuracy, an ensemble approach of random bootstrap is used to provide some alternative models of learning in order to select the optimal number of hidden nodes in ELM algorithm.

In general, figure 3 shows that both in case of B-ELM and EM-ELM, by combining random bootstrap into algorithm, the incremental nodes lead to smoother convergence and increased the accuracy (blue dots) in comparison to conventional method. The accuracy improvement for each method shows in table II, where the best accuracy is reached by B-ELM RBF. The improvement ratio for B-ELM additive, B-ELM RBF, and EM-ELM additive are 12.23%, 14.79%, and 19.13% respectively, meanwhile EM-ELM RBF could not improved the accuracy significantly (only 0.051%).

Table II. Improvement of accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Conventional method</th>
<th>An ensemble incremental approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kappa</td>
<td>Number of node</td>
</tr>
<tr>
<td>B-ELM Additive</td>
<td>0.775</td>
<td>13</td>
</tr>
<tr>
<td>B-ELM RBF</td>
<td>0.818</td>
<td>19</td>
</tr>
<tr>
<td>EM-ELM Additive</td>
<td>0.716</td>
<td>42</td>
</tr>
<tr>
<td>EM-ELM RBF</td>
<td>0.825</td>
<td>45</td>
</tr>
</tbody>
</table>

Fig. 3. An ensemble incremental approach effects
Table III. Experimental results

<table>
<thead>
<tr>
<th></th>
<th>Optimal Number of hidden nodes</th>
<th>Accuracy (Kappa)</th>
<th>Computational cost (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-ELM Additive</td>
<td>27</td>
<td>0.883</td>
<td>32.65</td>
</tr>
<tr>
<td>B-ELM RBF</td>
<td>40</td>
<td>0.939</td>
<td>48.78</td>
</tr>
<tr>
<td>EM-ELM Additive</td>
<td>46</td>
<td>0.853</td>
<td>31.67</td>
</tr>
<tr>
<td>EM-ELM RBF</td>
<td>34</td>
<td>0.867</td>
<td>45.89</td>
</tr>
</tbody>
</table>

According to Table III, the computational cost between B-ELM and EM-ELM is not so much different. However, in terms of optimal number of hidden nodes and accuracy, B-ELM is much better than EM-ELM. Meanwhile, the individual accuracy for generative class ($\omega_3$) in Table IV looks lower than the other along all methods, due to less number of data set for $\omega_3$ seen in Table I. This phenomena proved that ELM is significantly sensitive and affected by the structure of data set used in learning. To improve the accuracy, it needs to provide a proportional number of data set among classes.

Table IV. Individual Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Class</th>
<th>$\omega_1$</th>
<th>$\omega_2$</th>
<th>$\omega_3$</th>
<th>$\omega_4$</th>
<th>$\omega_5$</th>
<th>$\omega_6$</th>
<th>$\omega_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-ELM Additive</td>
<td>0.923</td>
<td>0.986</td>
<td>0.588</td>
<td>0.913</td>
<td>1.000</td>
<td>0.857</td>
<td>0.977</td>
<td></td>
</tr>
<tr>
<td>B-ELM RBF</td>
<td>0.942</td>
<td>1.000</td>
<td>0.588</td>
<td>1.000</td>
<td>0.885</td>
<td>0.964</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>EM-ELM Additive</td>
<td>0.865</td>
<td>1.000</td>
<td>0.118</td>
<td>0.870</td>
<td>1.000</td>
<td>0.857</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>EM-ELM RBF</td>
<td>0.923</td>
<td>1.000</td>
<td>0.118</td>
<td>0.870</td>
<td>0.846</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

In terms of incremental number of hidden nodes (from 1 up to 50), all methods have deal with competitive accuracy (Fig. 4) and tend to have similar performance profile. However, as shown in Figure 5, by considering incremental bands of MODIS data from 2 bands up to 6 bands, B-ELM with RBF hidden nodes can reach much better accuracy than the rest. Figure 6 shows the distribution map of paddy growth stages for each method using MODIS image with 500 meter spatial resolution.

V. CONCLUSION

In this study, an ensemble incremental approach based on random bootstrap resampling of B-ELM and EM-ELM are applied for multi-class classification using MODIS images data. Experiments show that this approach which is applied to both additive and RBF hidden nodes can reach smoother convergence and better accuracy. The main disadvantage of ELM is that it significantly sensitive to unbalanced training data set used in learning where the ratio for each class is different one from another. By providing more balanced ratio number of each class, ELM may be more robust for multi-class classification. However considering the optimal number of hidden nodes, an ensemble of B-ELM method with RBF hidden nodes is more suitable and more recommended for real applications in MODIS images for multi-class classification.
VI. REFERENCES


