International Journal of Engineering & Technology
CENTRE OF PROFESSIONAL RESEARCH PUBLICATIONS
ISSN 2049-3444

International Journal of Engineering and Technology Volume 3 No. 5, May, 2013

Early Detection for Forest Illegal Logging at River Basin Area

Syafruddin Syarif[1], N. Harun[2], M. Tola[3], M.W. Tjaronge[4]
[1]Department of Civil and Environmental Engineering, Hasanuddin University.
[2,3]Department of Electrical Engineering, Universitas Hasanuddin, 90245 Tamalanrea-Makassar, Indonesia
[4]Department of Civil Engineering, Universitas Hasanuddin, Indonesia.

ABSTRACT

The purpose of this study was to classify the use of multiple methods of expert systems, to determine the optimal model of early detection of illegal logging and changes in a forest area. Illegal logging is one of the ecological concerns related to the forest protection in Indonesia. In some cases, the illegal logging is almost impossible to be early identified by only physical inspection from the forestry staffs due to the field difficulty to reach the location. For this reason, the image processing by means Expert System using Support Vector Machine (SVM) method might be helpful to early identify the illegal logging activity. The sample image used in this research is taken from the Google Earth for one location of river basin area in Indonesia with the size area of 560m x 375m. The procedure of identification system is started by cropping the digital image and then classified for retrieving two-halves of the image in the form of forest and non-forest. These forms will be the input for the support vector machine method. Meanwhile, the testing data consists of five pairs of initial and final data. The simulation result shows that SVM method is able to early identify of illegal logging activity in protected forest areas with a level of accuracy of 80%. The testing process is basically similar to the image processing; only process of saving data sets it apart. The main point is the illegal logging indicator is obtained through the segmentation stage. Further analysis may determine the output whether the area is illegal logging or not.

Keywords: *Illegal logging, SVM, digital image processing, forest area, identification systems.*

1. INTRODUCTION

Indonesian forest is one of the diverse biodiversity where in the global scale is the third of the seven countries called mega forest diversity country. In terms of ecological system, our forest is the home to thousands of species of flora and fauna including some rarely found species. However, it is recently realized that the human activities related to the forest utilization become uncontrollably, therefore threatening to the forest extinction.

One of the dangerous human activities in the forest is the illegal logging. The situation of illegal logging has reached the heart of the protected forest conservation and forest production. In Indonesian, the illegal logging is the main cause of the deforestation in with the speeds almost 1.6 – 2.0 million hectare per year. According to the government data in 2006, the extensive forest damaging and cannot function optimally has reached 59.6 million hectares out of 120.35 million hectares of forest areas in Indonesia. The rate of deforestation in the last five years reaches 2.83 million hectares per year.

Prosecute without heed for the forest management rules to ensure the sustainability of forest resources has led to a variety of negative effects in various aspects. In fact, the forest resources are destroyed as a result of illegal logging in great abundance. The loss of tree area due to illegal logging affecting many dimensions, not only with respect to the economic issues but also social, political and cultural environment (Wijaya, 2005). From the perspective of economic activities of the illegal logging has reduced the country's foreign exchange revenues and incomes of the country. Various sources starting that the loss due to illegal logging has reached 30 trillion Rupiah per year. The loss in terms of the environment the most important is the loss of a certain number of trees so it is not to guarantee the micro climate change, declining land productivity, erosion and flooding as well as loss of diverse biodiversity. The impact of what's worse is the damage to forest resources due to illegal logging without heed for the forest management rules can reach a point where efforts to restore to its original state to be no longer possible *(irreversible).*

In terms of law enforcement, the illegal logging is an extraordinary act of crime. The violation may occur in all stages of the production of wood, namely at the stage of timber felling, transporting, processing and marketing of logs. In addition, the manipulation of some data in order to covering the ways without receiving any complaints and obtaining access to the National Department of Forestry and financial transgressions may happen, for instance, not paying taxes. The law violations occurs due to unclear administrative boundaries of the national forest area and the number of units of production forests are certified nationwide that operate inside the forest area and demarcated line along on the field by involving the local community is not well-implemented.

The flow of this detection system (illegal logging) is divided into pre-processing stage, neural network stage, and system testing phase. Pre-processing stage is the stage of data collection and training. At this stage of systems manufacturing, the neural networks that will be used as a detector of illegal logging was formed and trained with data obtained from the previous stage. System testing phase to test the extent to which the capabilities and accuracy of the neural network in detecting illegal logging [Syarif, et.al., 2012; Yang, et. al.,2000; Klobucar,et.al., 2010].

Adaptive Neuro Fuzzy Inference System (ANFIS) method can be used to identify a region with 70% accuracy rate by using the difference between the value of Green and Blue, the end and the beginning (Syarif, et. al., 2011).

The study found that the synthetic images derived from stable greenness, delta greenness and delta brightness of multi temporal principal component analysis (MPCA) summarized the information of landslides effectively producing accuracy of 88% for Teradomari and 91% for Tochio and Sidata Mura. The vegetation index Differencing (VIDN) provides relatively lower accuracies than those from MPCA, i.e., only 62.5% for Teradomari and 64% for Tochio and Sidata Mura (Surati Jaya, 2005).

The advantages of SVM compared to other methods, namely: (1) A generalization, a variety of empirical study shows that SVM and SRM approach gives the generalization error is less than the ERM strategy on Neural Network as well as other methods. (2) The Curse of dimensionality, one of the right method is used to solve the problem of high-dimensional, in a limited sample of data. (3) The business feasibility, SVM implemented relatively easy, because the process of determining support vector can be formulated in the QP problem.

2. CONFIGURATION OF PROPOSED SYSTEM

Image processing is a method used to process or to manipulate image for the purpose of repairing, analyzing, or modifying the image. The basic concept of image processing is taken from the ability of the human sense of sight that is associated with the ability of the human brain. In general, the objective of image processing is transforming or analyze an image so that new information about the image is made more clear [Danneberg, et.al, 2007; Diallo, et.al., 2009].

In this research, the image segmentation process is the image of the forest. The purpose of segmentation is to separate the forest image from the overall image. Then, the conversion of the separated image to the gray scale binary image is needed for the isolated pixel. The isolated pixel is a pixel of black color found in the dominant area of pixels are white, and vice versa. Pixel removal methods on isolated one is done by filling the holes on a binary image. The hole itself is a collection of pixels on a *background* that is not reachable by filling in the *background* from the image (Gonzales et.al. 2002).

The image processing results will be the input of support vector machine (SVM) method. The SVM method is one of the expert system methods that receives much attention as a state of the art in pattern recognition systems (Byun et.al., 2003). This method is firstly introduced by Boser, Guyon and Vapnik in the Annual Workshop on Computational Learning Theory in 1992. The basic principle of SVM method is trying to find the best hyperplane in the input space. In the initial development, the SVM is focused on linier classifier problems, however, it is recently found that several non-linier problems have been successfully solved by inserting the concept of the Kernel trick on high-dimensional workspace (Tsuda., 2000). The main concept is to obtain the best hyperplane that separates these two classes, then trying to find the most maximal margin. The margin is the distance between the hyperplane with the nearest pattern of each class that is called the support vector.

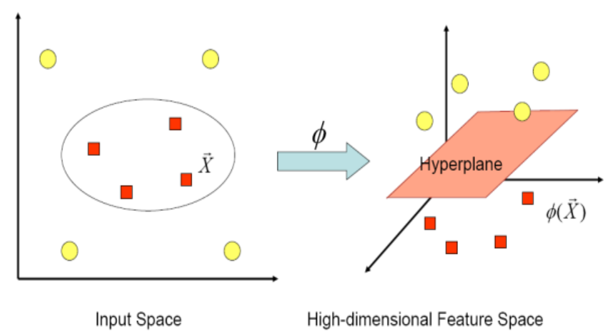

Fig.1 Both classes can be separated by a finite hyperplane

The concept of SVM can be described simply as attempt to find the best hyperplane that serves as a dividing two classes in the input *space*. Figure 1a shows some of the pattern that is a member of two classes: +1 and –1. Pattern incorporated on class –1 is symbolized by the blue color (box), while the pattern in class +1 is symbolized by the red color (triangle). The problem of classification can be translated with the effort to find a hyperplane separating lines between the two groups. Various alternative boundaries discrimination is shown in Fig. 1a.

The best hyperplane separation between these two classes can be found by measuring the margin hyperplane and looking for the maximum points. The margin is the distance between the hyperplane with the nearest pattern of each class that is called the support vector. The solid line in Fig. 1b shows the best hyperplane, which is located right in the middle of the second class, while the red and yellow dots in the dark circles are called support vector. Attempts to find the location of this hyperplane is the core of the learning process in the SVM. The available data are denoted asc$_i$ and labeled with each class denoted asy$_i$ ∈ {−1,+1} for I = 1,2,…, l, where is the number of data.

The advantages of support vector machine method, among others are the generalization capability, curse of dimensionality and the feasibility [Vapnik V. N., 1999]. Generalization is defined as the ability of neural network method of SVM to classify a pattern which does not include the data used in the learning phase. Vapnik explained that the generalization error is influenced by two factors, i.e., error on the training set and the factor influenced by the VC (Vapnik-Chervokinensis) dimension. Learning strategies in neural network and learning machine methods are generally focused on the effort to minimize the error on the training set. This strategy is called *empirical risk minimization* (ERM). While in the SVM method, the minimization of the second factor is added. The strategy is called *structural risk minimization* (SRM) and the SVM is realized with the biggest margin hyperplane with the picks. A variety of empirical study shows that SVM in approach to SRM provides a generalization error is smaller than that of obtained from the ERM strategy on neural network as well as other methods.

Meanwhile, the curse of dimensionality is defined as a problem that is faced by a method of pattern recognition related to parameter estimation, for example, the number of the hidden neurons in the neural network, stopping criteria in the process of learning and so on because the number of sample data is relatively small compared to the dimensional vector space of the data. The higher the dimension of the vector space of information processed, bringing the consequences to the high amount of data in the learning process. In fact, it often happens that the data processed is limited and the collection for more data is not possible because the cost and technical difficulties constraints. If the method implemented on the data amounted to very little in relative dimensions, it continues to very difficult process parameter estimation method. Curse of dimensionality is often experienced in the application in the field of biomedical engineering, because usually the biological data that are available are very limited and the arrangement takes high cost. Vapnik proved that the generalization level obtained by the SVM is not affected by the dimensions of the input vector. This is the reason why the SVM is one the proper methods used to solve the problem of high-dimensional within the limitations of the sample data.

In associated with the feasibility aspect, the SVM method can be implemented relatively easily because the process of determining support vector machine can be formulated in the Quadratic Programming (QP) problem. Thus, if the solution for QP problem is searched then, the SVM can be implemented easily. Moreover, it can be solved by sequential method for the high accuracy solutions.

Nevertheless, the SVM method has several weaknesses. For instance, the method has difficulty to use in large-scale problems with too many number of samples process. In addition, the SVM method was theoretically developed for classification problem with two classes, although the SVM method is currently modified in order to resolve the problem with the number of classes' more than two classes, namely one versus rest strategy and strategy tree structure. However, each of these strategies has its disadvantages, so that research and development on SVM multiclass is still widely opened (Vapnik V. N., 1999).

3. TESTING OF PROPOSED SYSTEM

The sample of forest imagery which is obtained from *Google Earth* is an area of one of protected forest areas in Indonesia. Protected forest is a national forest area that functions as the principal buffer protection system to regulate the life of the water, prevent floods, control erosion, prevent sea water and nourish the soil. Therefore, the Government set regulalations that prohibited conduct the poll results in the form of forest wood on protected forest areas. All sample images of the forest area has a size of 560 m × 375 m. This measurement is used as an input for testing system because using this measure of forest areas can still be clearly visible and each image has a resolution 150 × 150 pixel. One pixel represents only 9.33 m^2 of the image, thus the changes may result in better identification. Testing data consists of 10 images, where 5 images for initial data and the rest of 5 images for final data. These images were taken in different forest areas location.

This imagery sample will be processed through image preprocessing and image segmentation then use the SVM method for the detection of illegal logging. The phase of system testing is mainly intended to extent capability and accuracy of support vector machine in early detection of illegal logging. Therefore, the method may be able to identify whether or not illegal logging activity in the area. The system architecture and flowchart are shown in Fig.2 and Fig.3. After the desired image is obtained, then segmentation process of grayscale image and filter process are conducted in order to eliminate the noise in the image. After the binary image obtained with the black color inside the white color, then binary data is saved after the process of segmentation (Vapnik V. N., 1999).

In associated with the annual production ratio of logging plan for one year, the ratio is about 10 trees per 100 m^2 or 10 m^2/tree. The size of the tree itself has an average size of 6.25 m^2/tree. After the cropping image process, the sample images taken from Google Earth has now an average size of 100m x 170m. Therefore, how large a minimum area of area that can be used to cut down per year is calculated. Based on this assumption, the extensive forested areas on the image is 210,000 m^2, the number of trees to be cut down is 21,000 trees, the extensive forest that may be suspected cutting illegally is 131,250 m^2, with the ratio of extensive forest to be cut down is 62.5%. This calculation means that if the ratio more than 62.5%, it indicates the illegal logging activity in the area.

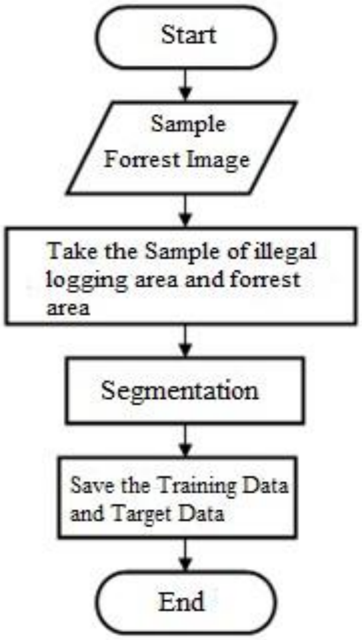

Fig.2 Flowchart pre-processing

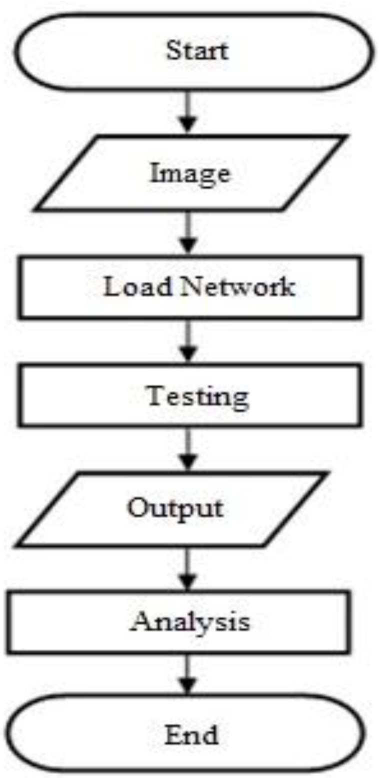

Fig.3 Flowchart Testing System

The testing process is basically similar to the image processing; only process of saving data sets it apart. The main point is the illegal logging indicator is obtained through the segmentation stage. Further analysis may determine the output whether the area is illegal logging or not.

In order to know the capability design SVM system, the testing phase is conducted. The detection results using SVM can be confirmed by administering the circle on the image output. The area is characterized by the circle area is illegal logging activity. In the determination of whether those areas illegal logging or not, done for the initial image detection and image of late. The difference between the area of illegal data on end with the area of illegal generates value as follows:

[1] If the reduction is positive (+) then the illegal logging is identified by means the area of trees on the initial data is reduced in the final data.
[2] If the reduction is zero (0) then the illegal logging is not identified, by means the area of trees on the preliminary data and the final data are still the same without undergoing any changes.
[3] If the reduction that is worth negative (–) then there is no activities of illegal logging that has been identified after seeing the area of trees is growing.

In addition, to show more detailed results of detecting system for the granting of illegal logging images, the circle or *mark* also makes it easy to be calculated the percentage of illegal on the image. The calculation results will be shown on GUI systems based Matlab command window in Fig.4.

To find out which area of the *illegal logging* in the testing stage, firstly the method performs extensive calculations of forest areas on the image. So the input image with a size of 560 m × 375 m is processed, which is about 210,000 m^2.

Then, the calculation of coverage area per pixel in the picture is done using eq. (1) as follows:

$$ECP = \frac{A}{NP} \quad (1)$$

where ECP is the extensive coverage per pixel, A is the area in the image (m^2) and NP is the number of pixel of the image (pixel). If the input image resolution is 150 x 150 pixel, the retrieved extensive coverage per pixel is 9.33 m^2/pixel. To obtain the area of illegal logging, the eq. (2) is utilized as follows:

$$AIL = DNFxECP \quad (2)$$

where AIL is the area of illegal logging and DNF is the difference in number of pixel suspected as illegal logging.

In the beginning of testing, the SVM system is carried out to detect *illegal logging* on a satellite image for off-line simulation. It due to the image through the process before testing the program, for instance the look for resolute and decisive areas experiencing *illegal logging* then the change in the appropriate image pixel used. There are 5 pairs of image is considered for this off-line test by taking a sample value of the initial data and final data. To continue in real-time simulation, the GUI system as in Fig.4 is developed. The simulation results by running GUI.m shows that the level of illegal logging can be determined in the group of tested images. As results, location of logging activity can be seen from the output mark marking system shown in Fig. 5 and Fig. 6 consecutively.

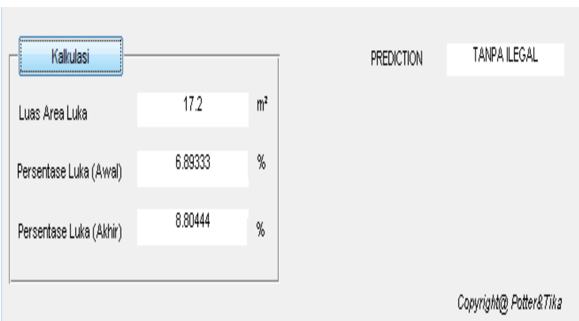

Fig.4 GUI System Prediction

Fig.5 Input System

Using the SVM method, the early detection of illegal logging can be obtained about 62.5% following the accurate capability of system to detect the color pattern changes. The results indicate that the pattern recognition is still very good based on the image quality. In associated with the size of the image, one *pixel* in the image with true scale is 9.33m^2 wide, so there is a relationship between the minimum area that can be detected with use of spatial resolution imagery.

The accuracy level is calculated following the eq. (3). In this calculation, there is a single test data that does not comply with the validation, retrieved the value of the accuracy of the system. The accuracy level of this method reaches 80%.

Fig.6 Output System

$$Accuracy = \frac{NCD}{ND} x100\% \quad (3)$$

Table 1 shows comparison of performance among two methods applies in this research. It can be seen that SVM outperform the other methods and reach 80% accuracy.

Table 1. The Comparison of Performance and Accuracy of the Expert System Method

Method	Performance	Accuracy (%)
ANFIS	Rule establishment using anfis, with partition method which has general structure called curse of dimensionality because amount of rule that shaped will be increased exponentially equivalent to the increasing of input	70.0
SVM	Image processing used is edge detection technique, which is a proccess that produce edges of object that supposed to increase the appearance or sharpen the line of an object or an area in an image.	80.0

4. CONCLUSION

The image processing by means expert system using support vector machine (SVM) method might be helpful to early identify the illegal logging activity. The sample image used in this research is taken from the Google Earth for one location of river basin area in Indonesia with the size area of 560m x 375m. The procedure of identification system is started by cropping the digital image and then classified for retrieving two-halves of the image in the form of forest and non-forest. These forms will be the input for the support vector machine method. Meanwhile, the testing data consists of five pairs of initial and final data. The simulation result shows that SVM method is able to early identify of illegal logging activity in protected forest areas with a level of accuracy of 80%. Nevertheless, the SVM method still has some drawbacks such as limitations in distinguishing between areas that are experiencing damage to forests and the trees at high brightness levels, as well as the lacking classification for the type of forests, areas such as shrubs, grasses, aquatic area, as well as other areas due to limited image used. The use better image quality may reduce such limitations.

5. FUTURE STUDY

Further research by combining other methods such as methods of Huffman and watermarking in wavelet domain may improve the accuracy level as well as the identification capability.

REFERENCES

[1] Byun, H., Lee S. W., 2003. A Survey on Pattern Recognition Applications of Support Vector Machines, *International Journal of Pattern Recognition and Artificial Intelligence, Vol.17, No.3,* 2003, 459-486.

[2] Danneberg, Roger B., et.al. 2007. A comparative evaluation of search techniques for query-by-humming using the MUSART test bed, *Journal of the American Society for Information Science and Technology, Vol.58, no.2*, 2007, 687–701.

[3] Diallo, Yacouba, Guangdao Hu, Xingping Wen, 2009. Aplications of Remote Sensing in land use/land cover Change Detection in Puer and Simao Counties, Yunnan Province, *Journal of American Science, 5(4),* 2009, 157-166.

[4] Gonzales, R. C., Richard W. E., 2002. *Digital Image Processing*, Second Edition (New Jersey: Prentice Hall,).

[5] Klobucar, Damir, R. Pernar, S. Loncaric, M.Subasic, A. Seletkovic, Mario A., 2010. Detecting Forest Damage in Cir Aerial Photographs Using a Neural Network, *Original scientific paper – Izvorniznanstveni, Croat. j. for. eng. 31* (2010) 2: 157-163.

[6] Syarif,,S., N. Harun, M. Tola, Indrabayu A, R. Parung, Dionasius K.., 2012 Illegal Logging Detection With Neural Network Backpropagation Model, *Proceedings of The 3rd Makassar International Conference on Electrical Engineering and Informatics*, MICEEI, Indonesia, 2012, 225-230.

[7] Syarif, S., Pingkan W., Trimurti F. R., Ady W. Paundu, 2011. Utilization of Adaptive Neuro Fuzzy Inference System to Identify Illegal Logging. Proceedings of the National Conference of Electrical Engineering Forum Indonesia *(FORTEI),* 2011, 245-252

[8] Surati Jaya, I N.,2005. Land slide Detection Technique using Multi date SPOT Imageries: A Case Study in Teradomari, Tochio and Shitada Mura, Niigata, Japan, *Jurnal Manajemen Hutan Tropika Vol. X No.1:* 2005, 31-48.

[9] Tsuda, K., 2000. Overview of Support Vector Machine, *Journal of IEICE, Vol.83, No.6,* 2000, 460-466.

[10] Vapnik, V. N., 1999. *The Nature of Statistical Learning Theory,* 2nd edition (New York Berlin Heidelberg: Springer-Verlag,).

[11] Wijaya, Arief., 2005. *Application of Multi-Stage Classification to Detect Illegal Logging with the Use of Multi-Source Data; A Case Study in Labanan Forest Management Unit, East Kalimantan, Indonesia*, Thesis Planning and Coordination, International Institute for Geo-information Science and Earth Observation Enschede, The Nederlands

[12] Yang, C.C., S.O. Prasher, J. A. Landry, H.S. Ramaswamy, and A.Ditommaso,2000. Application of Artificial Neural Networks in image recognition and classification of crop and weeds, *Canadian Agricultural Engineering, Vol. 42, No.3,* 2000, 147-152.