Land Cover Classification through Ontology Approach from Sentinel-2 Satellite Imagery

Miranda, E.,[1,2] Mutiara, A. B.,[3] Ernastuti, E.[3] and Wibowo, W. C.,[4]

[1]Information Systems Department, School of Information Systems, Bina Nusantara University, Jakarta Indonesia, E-mail: ekamiranda@binus.ac.id

[2]Doctoral Programs in Information Technology, Gunadarma University, Indonesia

[3]Faculty of Computer Science and Information Technology, Gunadarma University, Jakarta, Indonesia E-mail: amutiara@staff.gunadarma.ac.id, ernas@staff.gunadarma.ac.id

[4]Computer Science Faculty, Universitas Indonesia, Jakarta, Indonesia, E-mail: wibowo@cs.ui.ac.id

Abstract

This study proposed a land cover classification through an ontology approach from Sentinel-2 satellite imagery. Five steps were conducted as the research workflow, those were (1) Determine reference data and land cover classes; (2) Determine visual image interpretation features; (3) Feature extraction; (4) Ontology rules design; and (5) Land cover classification using ontology approach. The image feature of ontology was acquired by semantic reasoning approach from Indonesia National Standard RSNI-1 Land Cover Class in Medium Resolution Optical Imagery document. The features comprised of NDVI, Brightness, GLCM homogeneity and Rectangular fit. Image segmentation produced 2072 segments/objects by using eCognition software. Overall accuracy was used to evaluate the performance of the classification result. In addition, classification without ontology approach was carried out using CNN to compare the ontology result. Classification with ontology rule produced overall accuracy 99.8% whereas classification without ontology rules produced overall accuracy 98%.

1. Introduction

Regional land use planning and monitoring remain an issue related to the need for fast data acquisition and wide data coverage. In the past two decades, both tasks largely depend on remote sensing technology to capture and analyze remote sensing data on the region of interest (Arvor et al., 2013). Remote sensing is a technology to obtain and analyze information over the earth. This technology facilitates faster data acquisition and wider data coverage compared to conventional technology or field observations (Arvor et al., 2013). Frankly, the land cover classification is the classification of land cover objects into a class based on particular criteria (Arvor et al., 2013). Satellite imagery like a map contains worthwhile and interesting information, for example shape, texture and spectral information (Arvor et al., 2013). Hence, image analysis and classification are dynamic research fields (Forghani et al., 2018). Researchers from the past decade have proposed various approaches for these fields (Arvor et al., 2013). Image classification of satellite imagery is the process of assigning pixels (or groups of pixels) to a land cover class (Arvor et al., 2013). In the past decade, many researchers worked mainly focusing on classifying individual pixels or objects by identifying low-level image features only (Arvor et al., 2013 and Manzoor et al., 2015). Jiaoa and Liua (2012) recommended shape as a component in land use classification. This research also recommended a rule-based reasoning system integrating spectral, textural and shape information to improve the object-oriented classification of land use classes as future work. Hurni et al., (2013) suggested texture feature to classify satellite image. The classifier that employed one feature could merely classify low-level objects/classes (for example, road, and soil). The content of the satellite image contains not only low-level features but also high-level features (Upadhyaya and Dixit 2016). The content could be categorized into two types. First, visual content referring to the low-level features of the image, such as spectral information, shape, texture, color and indices. Second, semantic description referring to the high-level features, such as the theme or the expert contextual knowledge of the image. These contents derived from the document or the expert knowledge (Upadhyaya and Dixit, 2016). Semantic technologies like ontology propose an auspicious approach for image classification such as trying this technology to map the low-level image features to high-level image features (Arvor et al., 2013,

Online ISSN 2673-0014/ © Geoinformatics International

Manzoor et al., 2015 and Upadhyaya and Dixit, 2016).

There are thousands of potential features that can be selected from the image; the selection of the relevant features to recognize certain classes was still a trial-and-error process yet. The derivation of high-level information itself was not a trivial task. It mainly relied on the expert knowledge about the semantics of real-world objects. The gap between the visual appearance content of a digital image and semantic descriptions of the image itself was known as the semantic gap (Upadhyayab and Dixit, 2016 and Belgiu et al., 2013). The approach which was introduced as a semantic technology (ontology) could be employed to reduce the semantic gap (Belgiu et al., 2013). Ontology offers a potential solution to conceptualize and formalize the a priori knowledge on the evaluated domain (Hao and Guoping, 2011). The concept of domain knowledge is expressed in the form of machine-understandable rule sets and employed for semantic modeling (Hao and Guoping, 2011). Ontology was originated in Western philosophy and subsequently introduced into GIS (Geographic Information System). Recently, researchers have begun to explore the potential of ontology for land cover classification (Agarwal, 2007). Belgiu et al., (2014) presented an ontology-based classification method for extracting types of buildings. Huang et al., (2017) built coastal zone ontology to extract coastal zones using background and semantic knowledge. Lao et al., (2016) proposed ontology-based framework that was used to model the land cover extraction knowledge and interpret remote sensing images at the regional level. Kohli et al., (2012) developed ontology of slums for image-based classification. Wei et al., (2008) studied on geographic ontology based on object-oriented remote sensing analysis for forestry and prairie domains.

All these studies focus on a single thematic aspect based on expert knowledge. Previous studies did not provide comprehensive methods for object modeling in the ontology for satellite imagery classification. Therefore, this study would develop a land cover classification model based on ontology approach. This model would classify not only single thematic aspect (classes) but also seven thematic aspects (primary dry forest, secondary dry forest, planting forest, grassland, settlement, water body, bare land). In addition, this model would use not only single image feature but also several image features (color, hue, texture and shape). This study also would provide a comprehensive explanation of semantic reasoning from expert contextual knowledge to extract the image features and to design ontology rules. The image feature of ontology was acquired by semantic reasoning approach from Indonesia National Standard RSNI-1 Land Cover Class in Medium Resolution Optical Imagery document (National Standardization Agency, 2014). This research produced a contribution namely Visual image interpretation features based on semantic reasoning approach; Ontology rules and Land cover classification using ontology approach to classify seven land lover classes: primary dry forest, secondary dry forest, planting forest, grassland, settlement, water body, and bare land. This contribution was expected to answer regional land use planning and monitoring issue. The data was obtained from The Sentinel-2 satellite imagery.

2. Related Work

2.1 Ontology

Ontology is a formal representation or classification concept including a relationship among concepts within a particular domain. It usually refers to an ontology domain that can be associated with the ontology domain on the higher level. A domain itself is a set of the semantic terms that are designed only for one area of focus (Agarwal, 2007). Over the past decade, ontology has been the center of research in the computer science community such as knowledge representation, information integration, information extraction and retrieval. In the context of computer science and information, the ontology defines a set of primitive representations that can be used to model the domain of knowledge. These primitive representations are in the form of class (or set), attribute (or property), and relationship (or relationship between class members). The IS-A relationship used to describe the hierarchical structure. The IS-A hierarchy is usually property base shared by similar concepts (Agarwal, 2007). Ontology is explicit descriptions of concepts, which include classes (concepts), class properties that describe various features and attributes (slots or properties), and slot limits. Component of ontology consists of two parts, namely: (1) Class (formal representation of concepts) and (2) Slots (properties) and slot restrictions (role restrictions). Classes explain concepts in the domain, as well as properties of each concept that describe the various features and attributes of the concept. For example, the land cover class consists of forests, water bodies, open land and settlements. Meanwhile, the slot describes the properties of a class. For example, the class property of land cover of the remote sensing satellite image consists of visual image interpretation features, such as color, texture and shape (Agarwal, 2007).

2.2 Convolutional Neural Network

Convolutional neural network (CNN) is one of the deep learning methods, which consists of a number of convolutional and pooling layers and a fully connected layer (FCL) as the classifier (Weng et al., 2018). CNN is employed for image processing, natural language processing, and other kinds of cognitive tasks (Weng et al., 2018). CNN is one of the neural network models for deep learning, which is characterized by three specific characteristics, namely locally connected neurons, shared weight, and spatial or temporal sub sampling (Weng et al., 2018). Generally, CNN is composed of two major parts, namely (1) feature learning, which contains alternating convolutional and down-pooling layers. The output of each layer is the input to the next layer. The feature vectors extracted from the original input images by the hierarchical neural network are the output of the last layer, (2) FCL is the second part as the classifier, and it is a typical feed-forward neural network (Weng et al., 2018). A typical CNN architecture for land cover classification shown in Figure 1.

Each neuron in convolutional layer functions as a two-dimensional convolution with a certain filter and its input are locally connected to the output of the previous layer. The Convolutional operation can be seen in equation 1 (Weng et al., 2018).

$$\mathrm{aij} = \sigma(\mathrm{FxX}) = \sigma\left(\sum_{i'=1}^{h}\sum_{j'=1}^{w} f_{i'j'} x_{i+i'j+j'} + b\right.$$

Equation 1

F is an h x w weight matrix of the convolutional filter, X is the activations of the input neurons connected with the neuron (i,j) in the convolutional layer, and aij is the corresponding activation.

2.3 Confusion Matrix

A confusion matrix is a table with two rows and two columns that reports the number of false positives, false negatives, true positives, and true negatives. This table allows a more detailed analysis than a mere proportion of correct classifications (accuracy) (Novakovic et al., 2017). True Positive (TP) shows the prediction is positive and it's true, True Negative (TN) shows the prediction is negative and it's true, False Positive (FP) shows the prediction is positive and it's false, and False Negative (FN) shows the prediction is negative and it's false (Novakovic et al., 2017). Confusion matrix table is shown in Figure 2.

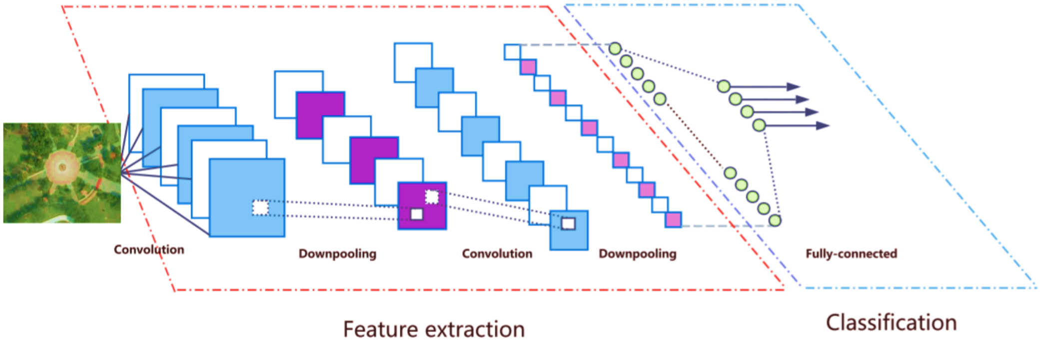


Figure 1: A typical CNN architecture for land cover classification (Weng, 2018)

Actual Values

Predicted Values		Positive (1)	Negative (0)
	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 2: Confusion matrix table (Novakovic, 2017)

Overall accuracy tells how many proportions are mapped correctly from all reference data. The overall accuracy formula is shown in equation 2.

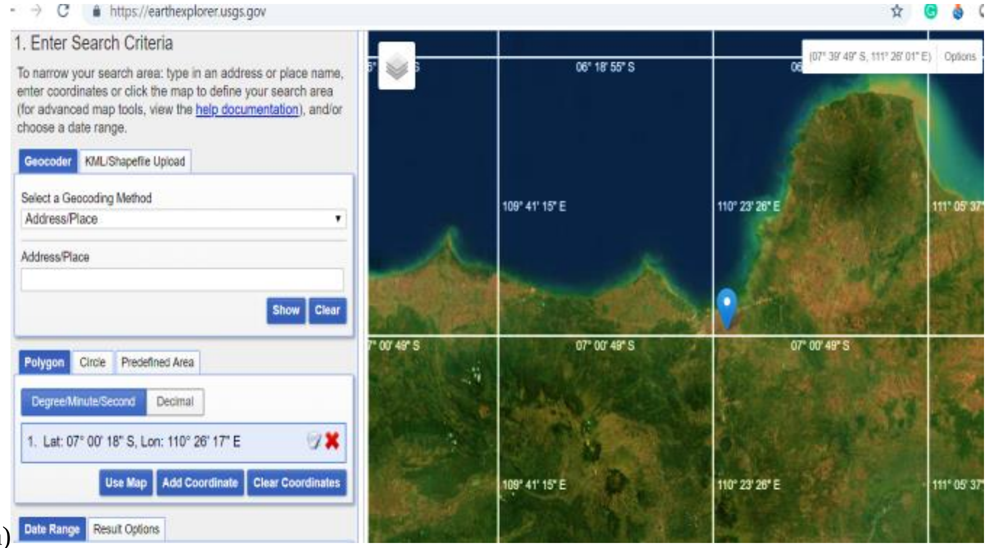
$$\text{Overall accuracy (\%)} = \frac{\text{the number of objects that are correctly classified}}{\text{number of samples for test accuracy}} x\ 100$$

Equation 2

3. Methods

3.1 Study Area

The test site for this research was located in Semarang area, Central Java, Indonesia. The part of the city selected for the study was mainly characterized by classes identified as primary dry forest, secondary dry forest, planting forest, grassland, settlement, water body, and bare land. The seven land covers were defined based on Indonesian National Standard (SNI) 7645 2010 for Land Cover Classification (National Standardization Agency, 2010). Image data was obtained from the Sentinel-2 satellite. This satellite has 13 spectral bands ranging from the Visible (VNIR) and Near Infra-Red (NIR) to the Short Wave Infra-Red (SWIR) (Esa Sentinel). Sentinel satellite imagery has taken as part of strategic planning for delivery of spatial information to support environmental and community safety (Thankappan et al., 2008). Steinhausen et al., (2018) showed the potential use of the Sentinel-2 satellite imagery land use and land cover map. The Sentinel-2 satellite imagery was used for the following reason: (1) Worldwide coverage in 10 m resolution and full color; (2) The temporal resolution of Sentinel-2 is 10 days performed by one satellite and also 5 days performed with two satellites, therefore large amounts of observational data were produced (Esa Sentinel). The satellite imagery for this research composed of five bands, namely: band 4 (red), band 3 (green), band 2 (blue), band 8 (Near-Infrared) and band 11 (SWIR, Short-Wave Infrared).



(a)

(b)

Figure 3: (a) the Sentinel-2 Satellite for Semarang, Central Java, Indonesia, (b) False color image fusion result (band 4, 3, 2)

The band 4 was worthwhile for identifying types of vegetation, soil and urban features, band 3 provided excellent contrast between clear and turbid (muddy) water, band 2 was useful for land and vegetation identification, forest type mapping, and identifying human-made features, and band 11 was useful for measuring soil moisture and vegetation, and provided good contrast between various types of vegetation (Sentinel 2 EO). Sentinel-2 satellite image was retrieved through a website with URL https://earthexplorer.usgs.gov/. Sentinel-2 captured the image on August 27, 2017, at 14:55:47 with the cloud cover in the observed area at 8.49%. The image was acquired in the sunny season, with high exposure of sun, dry temperature and very low humidity. The satellite imagery of Semarang area is shown in Figure 3.

3.2 Reference Data

Indonesia National Standard RSNI-1 Land Cover Class in Medium Resolution Optical Imagery document (National Standardization Agency, 2014) was used to determine visual image interpretation features and to extract the image feature of ontology. The image features of ontology were acquired by semantic reasoning approach from RSNI-1 document. The land covers definition of each land cover class on RSNI-1 document is shown in Table 1.

3.3 Model Development

This study consisted of five steps: (1) Determine reference data and land cover classes, (2) Determine visual image interpretation features using semantic reasoning approach, (3) Feature extraction from the image, (4) Ontology rules design and (5) Land cover classification using ontology approach. Overview of the research flow followed is shown in Figure 4.

Step 1. Determine reference data and land cover classes. This step included: (1) determining reference data as input for building ontology rules, (2) determining land classes and (3) choosing area of interest. The thematic map of Indonesia, Indonesia National Standard RSNI-1 and the spectral indices of Sentinel-2 (Sentinel 2 EO) were used for the references in this study. This study used seven land cover classes. The seven land-covers, namely primary dry forests, secondary dry forest, planting forest, grassland, settlement, water body, and bare land, were defined based on Indonesia National Standard (SNI) 7645 2014 Land Cover Classification (National Standardization Agency, 2014).

Step 2. Determine visual image interpretation features. Visual interpretation of the image consists of four basic features, namely: color, hue, texture and shape (Agarwal, 2007 and Wei et al., 2008). Feature selection is an important step for land cover classification from satellite imagery. There are thousands of potential features describing objects. Indonesia National Standards RSNI-1 document describes the definition of each land cover classes based on the visual image feature. Semantic reasoning approach from RSNI-1 document was used to determine selected feature (classifier attributes) of each land cover class. This step produced two results: (1) definition of each class based on optical interpretation and (2) image feature from the definition of each class. The image features would be used as the classifier attributes for ontology rules. The land cover definition on RSNI-1 document and image features (classifier attributes) produced using semantic reasoning approach are shown in Table 1. Table 1 that shows the keywords like green, light blue, whitish-blue, black, pink to the dark red area, brown, and white were mapped to color and hue. Color consists of the basic color (red, green and blue). Light, dark and medium are the appearances of Hue. Coarse, fine textured were mapped to a texture. Clustered, neatly arranged, certain pattern groups of a dense building were mapped to shape (regular or irregular). Basically, land cover classes were divided into two main areas: primarily vegetated area (vegetation area) and not primarily vegetated area (not vegetation area). Both classes were identified by vegetation index (NDVI). The keyword dark green was mapped to dense vegetation and green was mapped to medium vegetation.

Step 3. Feature extraction from the image. Feature extraction began with the segmentation process. Image segmentation is the process of partitioning a digital image into multiple segments. Multi-resolution segmentation technique was applied in this study. This technique merged the same pixels into an object. Pixels (or groups of pixels) would be classified into one class only. Feature extraction and segmentation of satellite imagery were carried-out using eCognition software. eCognition software could be used for image segmentation and object-based image analysis classification (Forghani et al., 2007). There were 2072 objects as a result of the image segmentation and image extraction process. Each object contained spectral information for image features. The 2072 objects were subsequently used as the training data and testing data (80%:20% respectively).

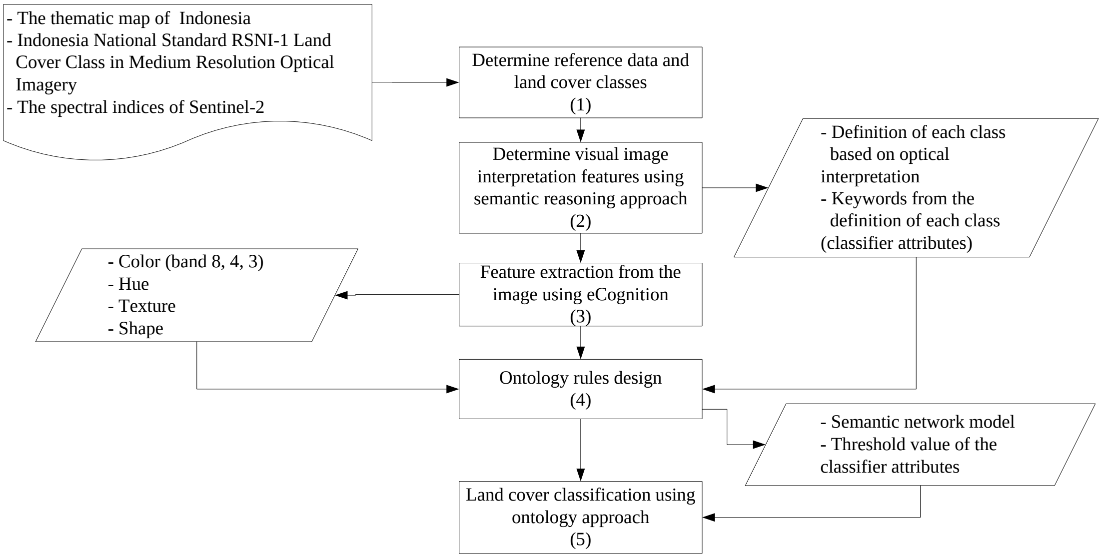


Figure 4: Overview of the research flow

Table 1: Land covers definition of each land cover class (National Standardization Agency, 2014)

Land cover class	Definition	Image features
Primary dry forest	The primary dry forest characterized by the presence of dark green objects (in bands 8, 4, 3), tend to dark, coarse texture with clustered tree canopy. There are no logged marks.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Secondary dry forest	The primary dry forest characterized by the presence of dark green objects (in bands 8, 4, 3), tend to dark, coarse texture with clustered tree canopy. There are logged marks.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Planting forest	Green (on the bands 8, 4, 3). Neatly arranged and have a certain pattern.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Grassland	Characterized by thin lines of very fine textured vegetation in moss green (on the bands 8, 4, 3).	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Settlement	Characterized by a group of dense building patterns in urban settlements and sparse building pattern in a rural settlement. The road network looks solid.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Water body	Objects are indicated by existence light blue area, whitish blue or black (on bands 8, 4, 3) covers a fairly wide area.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape
Bare land	Objects (in bands 8, 4, 3) characterized by *pink to the dark red area* sometimes *brown*, depending on the content of the soil material, and *white* when the material composed by lime.	1.Color (band 8, 4, 3) 2. Hue 3. Texture 4. Shape

The feature extraction process produced the numerical value of image feature on each object for every land cover classes. Color, Hue, Texture, and Shape were mapped into the eCognition feature to be an NDVI (Normalized Difference Vegetation Index), Brightness, GLCM (the Gray-Level Co-occurrence Matrix) homogeneity, and Rectangular fit respectively. NDVI (Normalized Difference Vegetation Index) is the most known vegetation indices. It is simple but effective for quantifying green vegetation (Sentinel 2 EO). NDVI value was calculated from the formula as shown in equation 4 (Sentinel 2 EO). This formula utilized bands 4 and bands 8. The NDVI formula is shown in equation 3.

$$NDVI = \frac{(B08 - B04)}{(B08 + B04)}$$

Equation 3

Step 4: Ontology rules design using semantic reasoning approach. The features that had already been identified in step 3 were used as classifier attributes and subsequently used to design the ontology rules. Ontology rules were designed through a semi-automatic approach. Semantic network model and the threshold value for each classifier attribute were defined to design the ontology rules. Semantic network model was formed through the construction of the land cover, image object features and classifiers using ontology (Gu et al., 2017). Semantic network model was designed manually, while the threshold value of the classifier attributes was got from the experimental process using eCognition software. The ontology rules were formed from the image features/classifier attributes that have already defined on Step 3. The entire semantic network model was formed through the construction of the land cover class, image object features and classifiers using ontology approach (Agrawal, 2007). The entire semantic network model of land cover image features is shown in Figure 5. The threshold value for each image features (classifier attributes) was obtained from the experimental process. The threshold value of the features is shown as follow:

NDVI >= 0.6 → dense vegetation
NDVI < 0.6 AND NDVI >= 0.5 → medium vegetation
NDVI < 0.5 → light vegetation
Brigthness < 1500 → dark
Brigthness <=1600 AND Brigthness > 1500 → medium
Brigthness > 1600 → light
GLCM homogeneity > 0.052 → rough
GLCM homogeneity <= 0.052 → smooth
Rectangular fit < 0.9 → irregular
Rectangular fit >= 0.9 → regular

Subsequently, the ontology rules were formed from the land cover definition (defined by RSNI-1, see Table 1), the classifier attributes/image features and the threshold value of the features. The ontology rules of seven land cover are shown in Table 2.

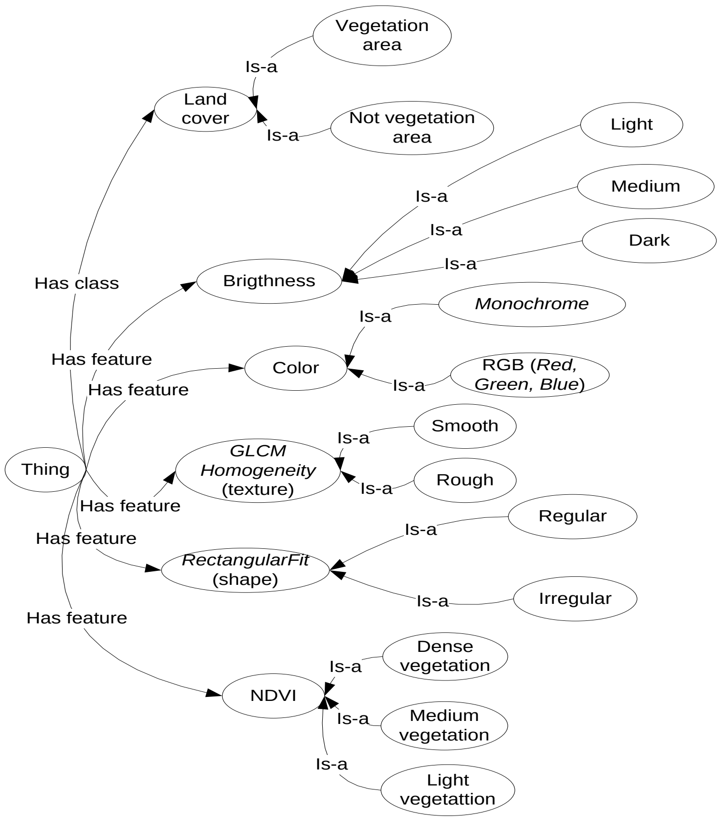


Figure 5: Semantic network model of image features

Table 2: The ontology rules of seven land-covers

Land cover class	Decision rules	Ontology rules
Primary dry forest	NDVI >= 0.6 AND *Brightness* < 1500 AND *Homogeneity* > 0.052 AND *RectangularFit* < 0.9	Primary dry forest ≡ dense vegetation ∩ dark ∩ rough ∩ irregular
Secondary dry forest	(NDVI < 0.6 AND NDVI >= 0.05) AND *Brightness* < 1500 AND *Homogeneity* > 0.052 AND *RectangularFit* < 0.9	Secondary dry forest ≡ medium vegetation ∩ dark ∩ rough ∩ irregular
Planting forest	(NDVI < 0.6 AND NDVI >= 0.05) AND (*Brigthness <=1600 AND Brigthness > 1500)* AND *Homogeneity* > 0.052 AND *RectangularFit* < 0.9	Planting forest ≡ medium vegetation ∩ medium ∩ rough ∩ irregular
Grassland	(NDVI < 0.6 AND NDVI >= 0.5) AND (*Brigthness <=1600 AND Brigthness > 1500)* AND *Homogeneity* <= 0.052 AND *RectangularFit* < 0.9	Grassland ≡ medum vegetation ∩ medium ∩ smooth ∩ irregular
Settlement	NDVI < 0.5 AND (*Brightness* >= 1500 AND *Brightness* <= 1600) AND *Homogeneity* <= 0.052 AND *RectangularFit* < 0.9	Settlement ≡ light vegetation ∩ medium ∩ smooth ∩ irregular
Water body	NDVI < 0.5 AND *Brightness* > 1600 DAN *Homogeineity* <= 0.052 AND *RectangularFit* < 0.9	Water body ≡ light vegetation ∩ light ∩ smooth ∩ irregular
Bare land	NDVI < 0.5 AND *Brightness* < 1500 AND *Homogeineity* <= 0.052 AND *RectangularFit* < 0.9	Bare land ≡ light vegetation ∩ dark ∩ smooth ∩ irregular

Step 5. Land cover classification using ontology approach. Ontology rules were used to classify land cover. The classification results were matched to the Thematic Map WebGIS Indonesia Ministry of Forest and Environment to produce confusion matrix and calculate the accuracy value. As a comparison to the classification based on ontology approach, on the other hand, land cover was classified without ontology rules. CNN was used to process and classify land cover without ontology approach. CNN is a supervised approach. Dataset (group of objects) was classified using an object-oriented segmentation method, into landscape theme (Kazemi et al., 2009). Data were separated into two subsets: learning process data (training data) and validation/evaluation data (testing data). Models were trained by a subset of learning data and validated by a validation subset. The class label of training data was given by opening image objects on ArchMap software subsequently matching image objects location on ArchMap to image object location on the Thematic Map WebGIS Indonesia Ministry of Forest and Environment. Evaluation based on the confusion matrix was employed to assess the classification result both for the classification using ontology rules and without ontology rules.

4. Result

Ontology rules used to classify land cover from 2072 image objects of satellite imagery. In addition, CNN was used to process and classify land cover without ontology approach.

The 2072 image objects were separated into training data for the learning process and testing data for the validation process. An accuracy assessment based on the confusion matrix was carried out both for classification using ontology rules and without ontology rules. A sample-based error matrix was created and used for classification accuracy assessment. The confusion matrixes of the two methods (with ontology and without ontology) are shown in Figure 6.

Overall accuracy with ontology = 99.8%, overall accuracy without ontology (CNN training data) = 98.4% and overall accuracy without ontology (CNN testing data) = 98%. The classification results yielded small improvements of the accuracies when involving ontology. However, the ontology approach helped out in understanding the complex structure of the land cover classification. The ontology approach helped out in understanding how to interpret the image feature from expert contextual knowledge (high-level information) and how to correspond the interpretation result to low-level information which was automatically extracted from images itself. The ontology approach had already helped out reduce the gap between both image content's (low-level feature and high-level feature). The classification based on an ontology proposed in this study had already eliminated the trial-and-error process on a selection of relevant features to classify image object through semantic reasoning of the expert contextual knowledge.

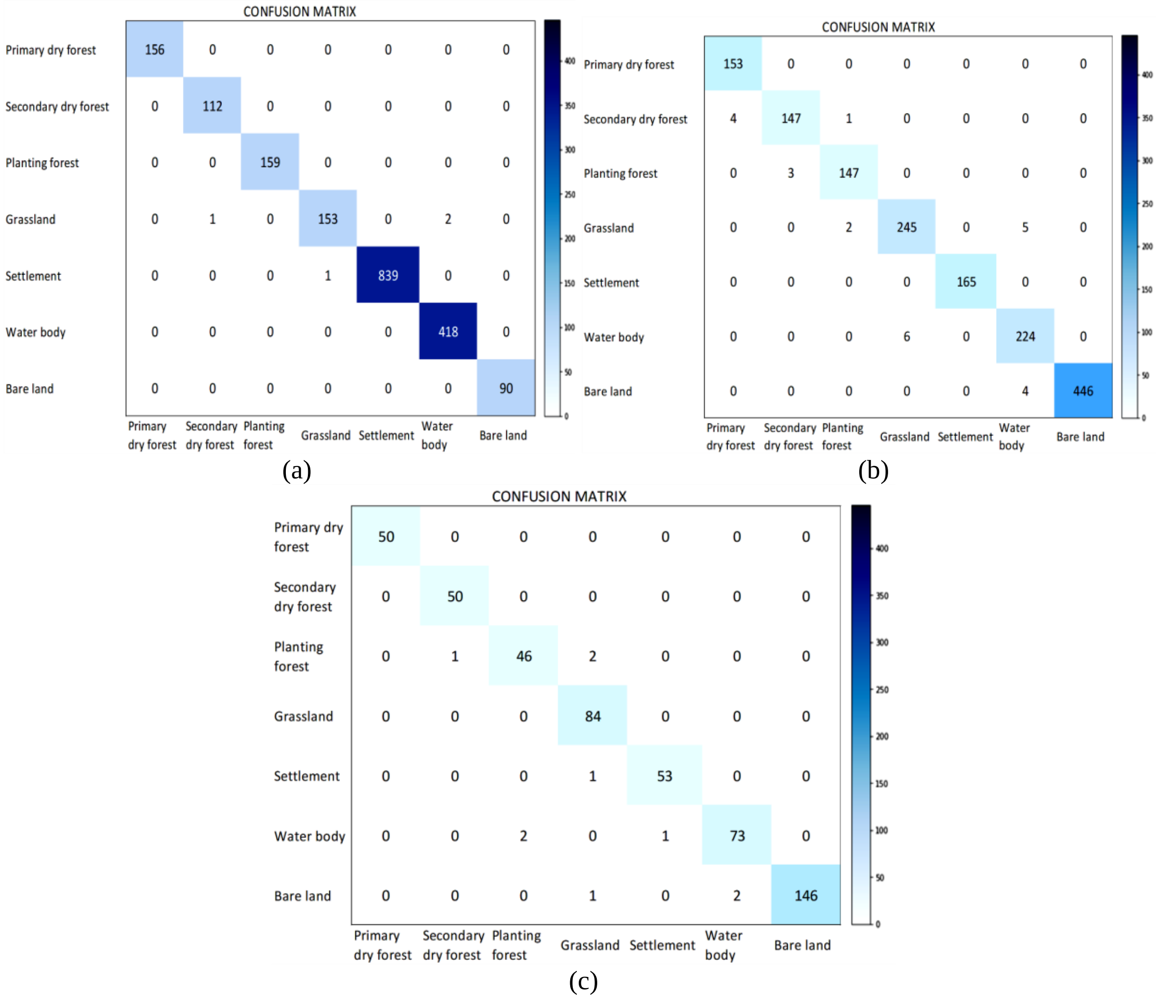


Figure 6: (a) Confusion matrix: classification with ontology, (b) Confusion matrix: classification without ontology (CNN for training data), (c) Confusion matrix: classification without ontology (CNN for testing data data)

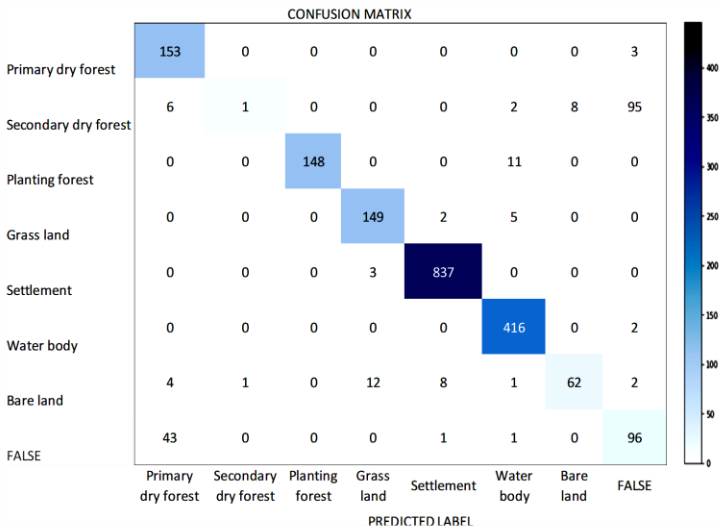


Figure 7: Confusion matrix: classification with ontology including unclassified/false class label

Table 3: The current research against Steinhausen's research

Steinhausen's research	Current research
- Satellite imagery; Sentinel 1 and 2 - Land cover class: Mango, Paddy, Palm, Sugarcane plantation, Wood, Forest, Grass & waste, Marsh land, Rock, Shrub land, Urban, Water - Classification method: Random forest - Parameter: Multi-sensory classification, mono-sensoral RF-classifications - Accuracy: Combination Sentinel 1 and 2 = 91.53% This is an improvement of 5.68 pp over a classification with Sentinel-2 data only	- Satellite imagery; Sentinel-2 - Land cover class: Primary dry forest, Secondary dry forest, Planting forest, Grassland, Settlement, Water body, and Bare land - Classification method: Ontology rules, CNN - Parameter: NDVI, brightness, GLCM homogeneity, Rectangular-fit - Accuracy: Ontology rules = 99.8

This research also provided a comprehensive explanation of semantic reasoning from expert contextual knowledge to design ontology rules. This research also showed Sentinel-2 could classify land cover with high accuracy value. A study by Steinhausen used Sentinel 1 and 2 to classify land cover whereas this research used Sentinel-2 to classify land cover. The result of this research against Steinhausen's research is shown in Table 3.

One issue was found in this research. The accuracy value that decreased was reaching 89.9 % when adding unclassified/false class for class observed. There were 96 objects belonged to unclassified class. Unclassified class is a class for classifying objects that do not have the same spectral value as one of the seven classes observed. The potential cause of this issue is the determination of the threshold value for each image features (classifier attributes). These values were obtained from the experimental process. An advanced experiment to determine the threshold value for each image features (classifier attributes) needs to be explored to answer this issue. Confusion matrix from classification with ontology including unclassified/false class label is shown in Figure 7. Another issue was found in this research that was the ontology rules have not supported fuzzy classification yet.

5. Conclusion

This study proposed land cover classification method based on ontology and semantic reasoning approach. It aimed to exploit the advantages of ontology. A detailed workflow has been introduced that includes five phases: Determine reference data and land cover classes, Determine visual image interpretation features using semantic reasoning approach, Feature extraction from the image, Ontology rules design and Land cover classification using ontology approach. This study demonstrated that the classification based on the ontology approach has already enhanced remote sensing image classification, particularly to reduce the semantic gap. Semantic technologies like ontology proposed an auspicious approach for image classification as trying this technology to map the low-level image features to high-level image features. As explained in the introduction section, several previous studies have already demonstrated the potential of ontologies for a particular regional or specific thematic aspect. Our study went one step ahead: the study developed classification method based on ontology and semantic reasoning approach not only for specific thematic aspect but also for seven thematic aspects (classes), namely primary dry forest, secondary dry forest, planting forest, grassland, settlement, water body, and bare land. This study showed small improvements of the accuracies when involving ontology for classifying the land cover from the satellite imagery. A future improvement to design ontology rules is still an open challenge, particularly for fuzzy classification. Moreover, this research merely produced ontology rules for seven main land cover classes; in fact, there are 22 land cover classes on RSNI-1 document. Thus, ontology rules design for other land cover classes is still an open challenge. Our method could be reusable to design ontology rules for other thematic aspects. In principle, this method can be adapted to train land cover classification model at any location as long as the thematic map standard published by local government is available to determine the land cover class for each segment.

Acknowledgments
We thankfully acknowledge the supporting from Research and Technology Transfer office (RTTO), Bina Nusantara University which facilitated the funding for this research. Our appreciation is also addressed to Christina Eli Indriyani, M.Hum. who helped to proofread the article.

References

Agarwal, P., 2007, Ontological Considerations in GIScience. *International Journal of Geographical Information Science*, Vol. 19, 501-536.

Arvor, D., Durieux, L., Andres, S. and Laporte, M. A., 2013, Advances in Geographic Object-Based Image Analysis with Ontologies: A Review of Main Contributions and Limitations from a Remote Sensing Perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 82, 125–137.

Belgiu, M., Lampoltshammer, T. J. and Hofer, B., 2013, An Extension of an Ontology-Based Land Cover Designation Approach for Fuzzy Rules, *Proceedings of GI_Forum Creating the GISociety*, 59-70.

Belgiu, M., Tomljenovic, I., Lampoltshammer, T., Blaschke, T. and Höfle, B., 2014, Ontology-based Classification of Building Types Detected from Airborne Laser Scanning Data. *Remote Sensing*, Vol. 6, 1347–1366.

Esa Sentinel, Resolution and Swath, https://sentinel.esa.int/web/sentinel/missions/sentinel-2/instrument-payload/resolution-and-swath.

Forghani, A., K., Cechet, B. and Nadimpalli, K., 2007, Object-based classification of multi-sensor optical imagery to generate terrain surface roughness information for input to wind risk simulation, *Proceedings IEEE International Geoscience and Remote Sensing Symposium*, 3090-3095.

Forghani, A. K., Nadimpalli, R. P. and Cechet, 2018, Extracting Terrain Categories from Multi-Source Satellite Imagery. *International Journal of Geoinformatics*, Vol. 14(2), 1-10.

Gu, H., Li, H., Yan, L., Liu, Z., Blaschke, T. and Soergel, U., 2017, An Object-Based Semantic Classification Method for High-Resolution Remote Sensing Imagery Using Ontology. *Remote Sensing*, Vol. 9, 1-21.

Hao, F. and Guoping, G., 2011, Integrating Low-Level and Semantic Features for Object Consistent Segmentation, *Proceedings of Sixth International Conference on Image and Graphics*, 39-44.

Huang, H., Chen, J., Li, Z., Gong, F. and Chen, N., 2017, Ontology-Guided Image Interpretation for GEOBIA of High Spatial Resolution Remote Sense Imagery: A Coastal Area Case Study. *International Journal of Geo-Information*, Vol. 6, 1-17.

Hurni, K., Hett, C., Epprecht, M., Messerli, P. and Heinimann, A., 2013, A Texture-Based Land Cover Classification for the Delineation of a Shifting Cultivation Landscape in the Lao PDR Using Landscape Metrics. *Remote Sensing*, Vol. 5, 3377-3396.

Indonesia Ministry of Forest and Environment, 2018, The Thematic Map, WebGIS Indonesia Ministry of Forest and Environment, http://webgis.dephut.go.id:8080/kemenhut/index.php/id/peta/peta-interaktif (2018).

Jiaoa, L. and Liua, Y., 2012, Analyzing the Shape Characteristics of Land Use Classes in Remote Sensing Imagery, *Proceedings of ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 135-140.

Kazemi, S., Lim, S. and Rizos, C., 2009, Interactive and Automated Segmentation and Generalisation of Raster Data. *International Journal of Geoinformatics*, Vol. 5(3), 65-73.

Kohli, D., Sliuzas, R. and Kerle, N., 2012, An Ontology of Slums for Image-based Classification. *Computers Environment and Urban Systems*, Vol. 3, 154-163.

Lao, H., Li, L. and Zhu, H., 2016, Land Cover Extraction from High Resolution ZY-3 Satellite Imagery Using Ontology-Based Method. *International Journal of Geo-Information*, Vol. 5, 1-16.

Manzoor. U., Balubaid, M. A., Zafar, B., Umar, H. and Khan, M. S., 2015, Semantic Image Retrieval: An Ontology Based Approach. *International Journal of Advanced Research in Artificial Intelligence*, Vol. 4, 1-8.

National Standardization Agency, 2010, Indonesian National Standard (SNI) 7645 Land Cover Classification Part 1: Small and Medium Scale, http://www.big.go.id/assets/download/sni/SNI/15.%20SNI%207645-2010%20Klasifikasi%20-penutup%20lahan.pdf

National Standardization Agency, 2014, Indonesian National Standard RSNI-1 Land Cover Class in Medium Resolution Optical Imagery and the Its Interpretation, http://appgis.dephut.go.id/appgis-/download/Peraturan%20Terkait/RSNI_1b_KELAS%20PENUTUPAN%20LAHAN%20DALAM%20PENAFSIRAN%20CITRA%20OPTIS%20RESOLUSI%20SEDANG.pdf

Novakovic, D. J., Veljovic, A., Ilic, S. S., Papic, Z. and Tomovic, M., 2017, Evaluation of Classification Models in Machine Learning. *Theory and Applications of Mathematics and Computer Science*, Vol. 7, 39-46.

Sentinel 2 EO, Sentinel Hub, https://www.sentinel-hub.com/develop/documentation/eo_products/Sentinel2EOproducts

Steinhausen, M. J., Wagnerc, P. D., Narasimhand, B. and Waske, B, 2018, Combining Sentinel-1 and Sentinel-2 Data for Improved Land Use and Land Cover Mapping of Monsoon Regions. *International Journal of Applied Earth Observation and Geoinformation*. Vol. 73, 595-604.

Thankappan, M., Reddy, S., Ackerly, S. and Forghani, A., 2008, Sentinel Floods Prototype: A System for Rapid Dissemination of Satellite Derived Information on Flooding. *Proceedings of the 14th Australasian Remote Sensing and Photogrammetry Conference*, 1-10.

Upadhyaya, N. and Dixit, M., 2016, A Review: Relating Low Level Features to High Level Semantics in CBIR. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol. 9, 433-444.

Wei, C., Liping, G., Wang, L. and Li, D., 2008, Study on Geographic Ontology Based on Object-Oriented Remote Sensing *Analysis, Proceedings of International Conference on Earth Observation Data Processing and Analysis*, 1-12.

Weng, Q., Mao, Z., Lin, J. and Liao, X., 2018, Land-Use Scene Classification Based on A CNN Using A Constrained Extreme Learning Machine. *International Journal of Remote Sensing*, Vol. 39, 1-18.