COMBINATION OF SPECKLE DIVERGENCE AND NEIGHBORHOOD ANALYSIS TO CLASSIFY SETTLEMENT FROM TERRASAR-X DATA

M. Rokhis Khomarudin[1]* and Agung Indrajit[2]
[1]National Institute of Aeronautics and Space (LAPAN), Jakarta
[2]National Coordinating Body for Survey and Mapping (Bakosurtanal). Cibinong. Bogor
*e-mail: Rokhis.Khomarudin@lapan.go.id

Abstract. The objectives of this research were to develop and improve methods for determination of settlements area with focus on synthetic aperture radar (SAR) data. Remote sensing settlement classification has made great progress, both for optical and radar data as well for their fusion. Yet, in radar imagery, settlement classification still contains some problems. Several studies on application of radar imagery have been conducted using techniques such as textural analysis, multi-temporal analysis, statistical model, spatial indexes, and object-based classification. Most of the development methods have several problems in the specific area especially in the tropical country. Several studies also showed that settlement classification accuracies were just below 60%. This was not sufficient enough to classify settlement areas using SAR imagery. Therefore, in this research, we proposed a new method i.e., the combination of the speckle divergence and the neighborhood analysis. The proposed method was applied to classify settlement area in Cilacap and Padang Districts of Indonesia. The results showed that the proposed method produced a good accuracy i.e., 85.5% for Cilacap Districts and 78.1% for Padang Districts.

Keywords: *Settlements areas, Speckle divergence, Neighborhood, SAR*

1 INTRODUCTION

Settlement area is a key parameter on the urban growth analyses in which it always contain some problems especially in large and fast developing country like Indonesia. The availability data of settlement area is important on the management of urban area and also on the disaster management. Settlement refers to human gathering location on earth. Hence, detecting a settlement area is also detecting the number of people (Tobler, 1969). To detect a settlement area can be done by using remote sensing technology specifically with optical dan radar sensors.

Remote sensing settlement classification has made great progress, both for optical and radar data and their fusion. Yet, settlement classification with radar imagery is the most challenging. Several studies on application of radar imagery have been conducted using techniques such as textural analysis, multi-temporal analysis, statistical model, spatial indexes, and object-based classification. With the availability of high-resolution synthetic aperture radar (SAR) sensor systems, e.g. of the German TerraSAR-X earth observation satellite, new opportunities of settlement mapping are available. SAR systems are capable of acquiring data during the day and the night time, independently on weather or environmental conditions. Hence, sophisticated and efficient methodologies for settlement classification using TerraSAR-X are needed for the mapping and updating of land use and settlement areas.

The most common method for radar imagery classification for earth surface detection is textural analysis. Dekker (2003) was one example of using textural analysis using ERS SAR imagery. Modern texture measures such as histogram measures, wavelet energy, fractal dimension, lacunarity and semivariogram were applied and developed. The result improved the accuracy of the classification compared to widely used texture measures such as mean intensity, variance or entropy. Despite this improvement, the classifications were still not in satisfactory because of the low separability of some classes, especially of urban area and forest, urban area and industry/greenhouse. Acqua and Gamba (2003) proposed to use the histogram density index (HDI) to identify urban characteristics of high-resolution imagery. However, similar to the finding of Dekker (2003), the approach failed to classify main parts of urban areas like city centers, residential, and suburban areas. Therefore, the texture analysis method seemed to be insufficiently accurate for the urban areas detection. The low separability of some features is the main barrier of using textural analysis for settlement classification.

Being dissatisfied with the previous result, Acqua *et al.* (2003) modified the methodology with multi-temporal analysis to improve the urban characterization. The main problems of urban area classification by radar imagery are multiple bouncing, layovers, and shadowing. Multi-temporal imagery produces a multi-angle look on urban areas and thus allows identifying different details of the same object. Hence, the discrimination of different classes in urban areas is facilitated. Acqua *et al.* (2003) applied the HDI method using multi-temporal imagery and produced better classification results than in the previous study. Additionally, the extraction of streets was also improved by this approach. Another comprehensive research on settlement classification by using multi-temporal analysis was done by Pellizeri *et al.* (2003). They compared statistical analysis and the neural kernel-based approach to classify urban area. The result from the analysis showed that the neural kernel-based approach slightly outperforms the statistical approach when classifying mono-temporal imagery. By extending the data to multi-temporal and multi-frequency, both approaches achieved similar accuracy of classification. This showed that multi-temporal data can slightly improve the accuracy of single data classification method.

Another research using SAR imagery was conducted by Tison *et al.* (2004). This research revealed that classes can be discriminated based on their statistical properties, which require an accurate statistical model. Hence, they investigated various statistic distribution models such as Weibull, Log-Normal, Nagatani-Rice and Fisher, and analyzed how these distributions fit the urban area characteristics in SAR imagery. The Fisher distribution proved the most appropriate distribution for urban characteristics in their study area Toulouse, France. This was an important result related to settlement classification using SAR imagery. It might be used as a reference for future applications. However, one has to be aware that the best fitting statistical distribution for one region is not necessarily transferable to other regions.

In the other development, Stasolla and Gamba (2008) proposed spatial indices to extract settlements from high-resolution SAR imagery. After reviewing several indices, they defined three spatial indices for their purposes, Moran's Index, Geary's Index, and the Getis-ord Index. The basic principal of these indices was a local indicator of spatial association (neighborhood). Inhomogeneous objects in settlement were the reason to apply these indices. Hence pixel-by-pixel classification was inappropriate for the classification. This method resulted in good overall accuracy, but omission and commission errors were still high. Considerable some improvement of the method that was especially needed on indexing thresholding, the problem still exist when transferring to other areas.

The latest research on settlement classification was an object-based approach. Thiel *et al.* (2008) and Esch (2010) describe the new methodology to detect settlement area using TerraSAR-X. They proposed a pre-processing method called "speckle divergence analysis", with a comparably straight forward approach of analyzing local statistics. They achieved more appropriate noise suppression compared to establish filtering routines, which preserve true structures with texture and contour information (Thiel *et al.*, 2008). This filter was based on a sigma probability of the Gaussian distribution of speckle noise. It removed speckle and suppresses noise, but preserved the texture, structure and contour information. This was indispensable since many regions were characterized by small-scale structures. The speckle divergence image was used for object-based settlement classification. The result of this classification was satisfying and had good accuracy. It is promising to implement the method for analyzing the study area of this research to produce improvement result.

The accuracy improvement can also be done by neighborhood analysis. However, in the neighborhood analysis, the settlement pixel which cannot be detected by the speckle divergence analysis, it can be classified as settlement area. The basic principle of neighborhood analysis is classifying pixel as settlement. Therefore, the objective of this research was to develop and improve methods for determination of settlements area with focus on SAR data using the combination methods of the speckle divergence and the neighborhood analysis.

2 MATERIAL AND METHODS

2.1 Data Requirement

TerraSAR-X imageries were used in the development of new improved settlement classification method. Table 1 describes the information of those data including the acquisition date, type of data, data source, and the spatial resolution. There were three scene imageries covered Cilacap district and one scene of TerraSAR-X covered Padang District. Three times of acquisition data for Cilacap were on 4, 15, and 24 January 2009 and on 3 April 2008 for Padang. Beside the TerraSAR-X data, this research used the settlement map data from the high resolution imagery data as reference for the accuracy assessment.

StripMap (SM) was the basic SAR imaging mode as known in ERS-1 and other radar satellites. The ground swath was illuminated with continuous sequence of pulses while the antenna beam was fixed in elevation and azimuth. These conditions resulted in an image strip with a continuous image quality (in flight direction). StripMap dual polarisation data have a slightly lower spatial resolution and smaller swath than the single polarisation data. In StripMap mode, a spatial resolution of up to 3 m can be achieved. The standard scene size was 30 km x 50 km (width x length) in order to obtain manageable image files; however, acquisition length was extendable up to 1,650 km (Infoterra, 2009).

2.2 Data Analysis

The settlement classification method was an improvement over the approach of Esch *et al.* (2010). This research proposed the neighborhood analysis from the “real settlement” image using the speckle divergence and intensity image. The process of settlement classification was performed in two steps: pre-processing and classification (neighborhood). To ensure efficiency and quality of the model, an accuracy assessment and transferability analysis were carried out. The steps of the settlement classification are shown in Figure 1.

In compact steps, the classification procedure was stated as follows:

- pre-processing (speckle divergence analysis);
- threshold analysis (threshold determination for real settlement and potential settlement); and
- neighborhood analysis (classification of settlement areas).

2.2.1 Pre-processing

The pre-processing of the data included speckle suppression to enhance images degraded by speckle noise. This speckle phenomenon in TerraSAR-X data significantly hampered the analysis of radar data. The approach of speckle suppression consists of determining the difference between the local image heterogeneity and the theoretical, scene-specific heterogeneity of developed speckle.

The calculation of speckle divergence was based on kernel size of 9 x 9 pixels (the kernel size have been determined by the trial and error analysis until the best fits classification and resulting the best accuracy of settlement classification), and the formulas.

Table 1. TERRASAR-X IMAGERIES

No.	Satellite name	Location	Date of acquition	Type of data	Data source	Spatial resolution
1.	TerraSAR -X	Cilacap	04.01.2009	StripMap	DLR	2.7 m
2.	TerraSAR-X	Cilacap	15.01.2009	StripMap	DLR	2.7 m
3.	TerraSAR-X	Cilacap	24.01.2009	StripMap	DLR	2.7 m
4.	TerraSAR-X	Padang	03.04.2008	StripMap	DLR	2.7 m

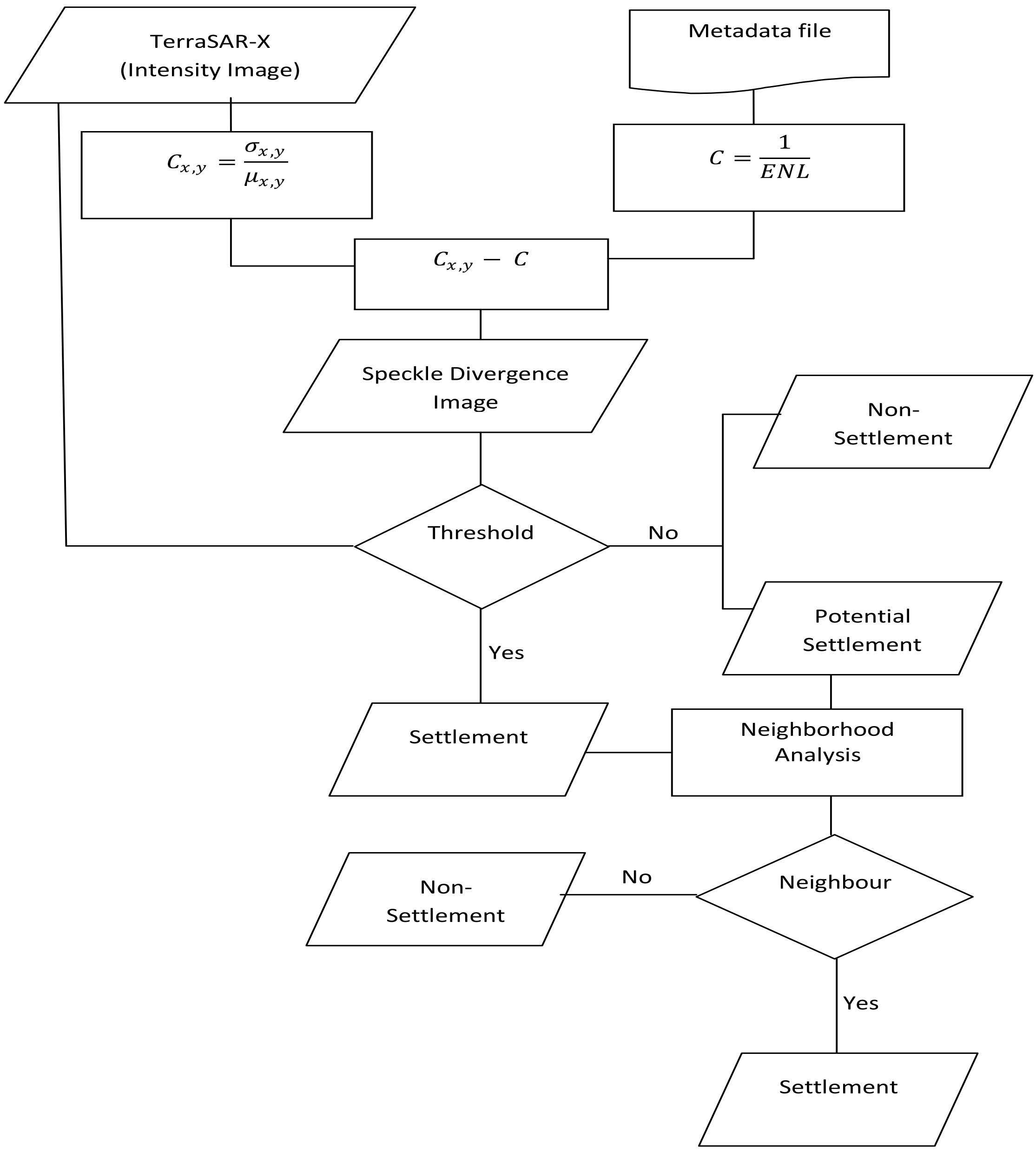


Figure 1. The steps of settlement classification by using TerraSAR-X (the chart has been re-formated)

were stated as follows (Esch *et al.*, 2010):

$$D_{x,y} = C_{x,y} - C \; with \; C_{x,y} = \frac{\sigma_{x,y}}{\mu_{x,y}} \quad (1)$$

where:

$D_{x,y}$: the speckle divergence

C : the theoretical heterogeneity due to developed speckle; it was calculated from the inverse of equivalent number of look (ENL), $ENL = L_a + L_r = 1/C$, L_a and L_r defining the effective number of looks in the azimuth and range (stated in TerraSAR X metadata)

$C_{x,y}$: the local coefficient of variation defined by the local, it is calculated in kernel size of 9 x 9 pixels

$\sigma_{x,y}$: the local standard deviation, it was calculated in kernel size of 9 x 9 pixels

$\mu_{x,y}$: the local mean, it was calculated in kernel size of 9 x 9 pixels.

2.2.2 Threshold analysis

The threshold of real and potential settlement from the intensity and speckle divergence image was defined by analyzing samples of pixels in settlement and non-settlement areas. The threshold can be decided by analyzing the histogram of samples (Figure 2). The histogram resulted from the sample data that acquired from the intensity/speckle divergence image for settlement and non-settlement.

By using the threshold, the settlement and non-settlement area can be classed as follows:

- If DN Speckel Divergence < Threshold for potential Settlement, then Non-Settlement
- If DN Speckel Divergence > Threshold for real Settlement, then Settlement
- If Threshold for potential Settlement < DN Speckel Divergence < Threshold for real Settlement, then the neighborhood analysis was needed.

2.2.3 Neighborhood analysis

The settlement classification can be performed by generating a simple threshold of the speckle divergence, but this method cannot eliminate false settlement detections in hilly areas that have similar speckle divergence values. To avoid this error, it was proposed to use neighborhood analysis (Figure 3) for the detection of settlement areas from TerraSAR-X imagery with steps as follows:

- identification of "real settlement" pixels from speckle divergence and intensity images by using the threshold;
- identification of "potential settlement" pixels from speckle divergence and intensity images by using the threshold;
- a combination of "real settlement" and "potential settlement" pixels to identify settlement areas by conditional parameters (settlement mask).

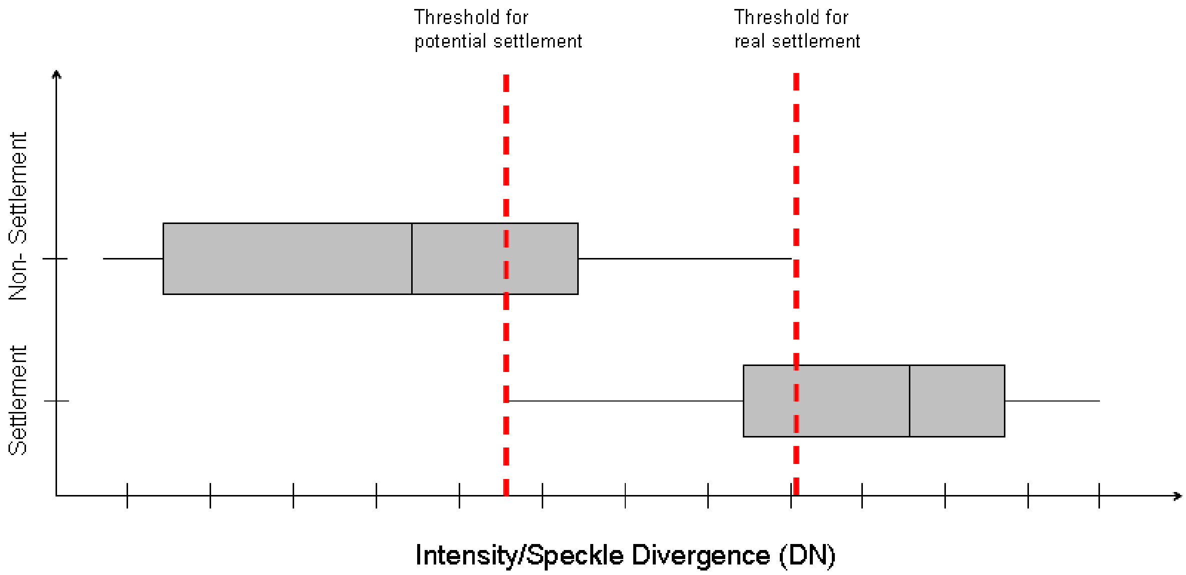


Figure 2. The threshold decision for potential settlement and real settlement.

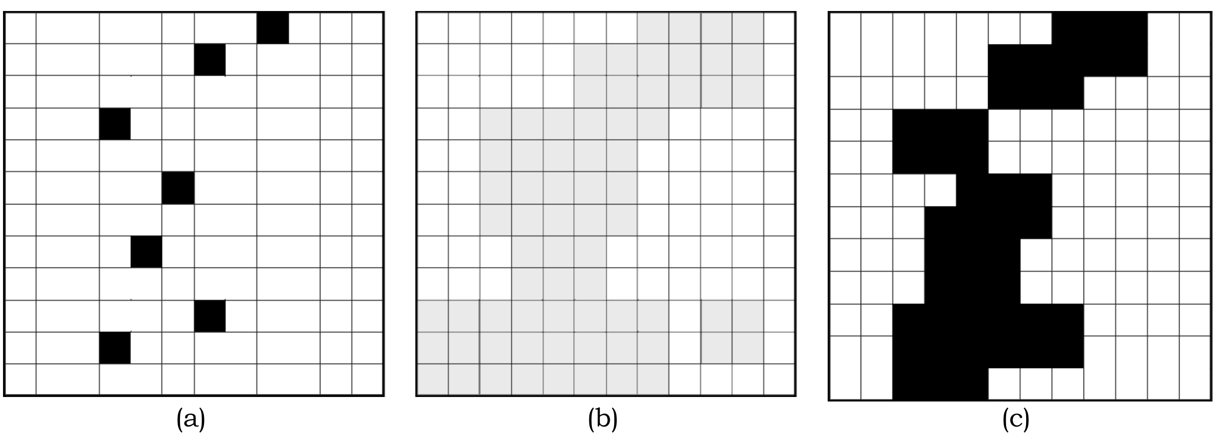


Figure 3. Illustration the neighborhood analysis for settlement classification (a) real settlement detection (b) potential settlement and (c) the settlement mask by neighbourhoods a kernel size of 3 x 3 pixels

2.2.4 Accuracy assessment

An accuracy assessment was performed to check the quality of the classification. To conduct a proper accuracy assessment, references from the field were needed. The confusion matrix accuracy assessment by Story and Congalton (1986) was used in this research. This methodology allowed identifying the omission and commission error of settlement based on the real condition. The concept of calculating the overall accuracy, omission and commission errors with a numerical example of the confusion matrix accuracy assessment is illustrated in Figure 4 (Story and Congalton, 1986; Naesset, 1996; Liu *et al.*, 2007).

		Reference Data			
		F	U	W	Row Total
Classified Data	F	40	9	8	57
	U	1	15	5	21
	W	1	1	20	22
	Colum Total	42	25	33	100

Land Cover Categories

F = Forest
U = Urban
W =Water

Omission Error	Commision Error	Overall Accuracy
F = (42-40)/42 = 5 %	F = (57-40)/57 = 30%	= (40 + 15 + 20)/100
U = (25-15)/25 = 40%	U = (21-15)/21 = 29%	= 75/100 = 75%
W = (33-20)/33= 39%	W = (22-20)/22= 9%	

Figure 4. Confusion matrix accuracy assessment calculation (modified from Congalton and Green, 1999).

2.2.5 Transfer of the approach to other areas

The same classification methodology was applied to the Padang area. The accuracy assessment was performed accordingly to check the quality and transferability of the methodology, and an error map was created.

3 RESULT AND DISCUSSION

3.1 Speckle Divergence

The result of speckle divergence analysis in Cilacap is shown in Figure 5. Using this analysis, settlement areas and their characteristics can be distinguished better than using the intensity image. The settlement areas look brighter and clearer after speckle divergence analysis. Therefore the result was used for further settlement classification with TerraSAR-X imagery.

3.2 Neighbourhood Analysis

The result of the neighborhood classification to detect the settlement area is shown in Figure 6 in which the settlement area more obvious than only using speckle divergence analysis. The classification of settlements was performed automatically and the result was visualized. After this process, the accuracy assessment was carried out to check the quality of the remote sensing classification compared to the settlement areas.

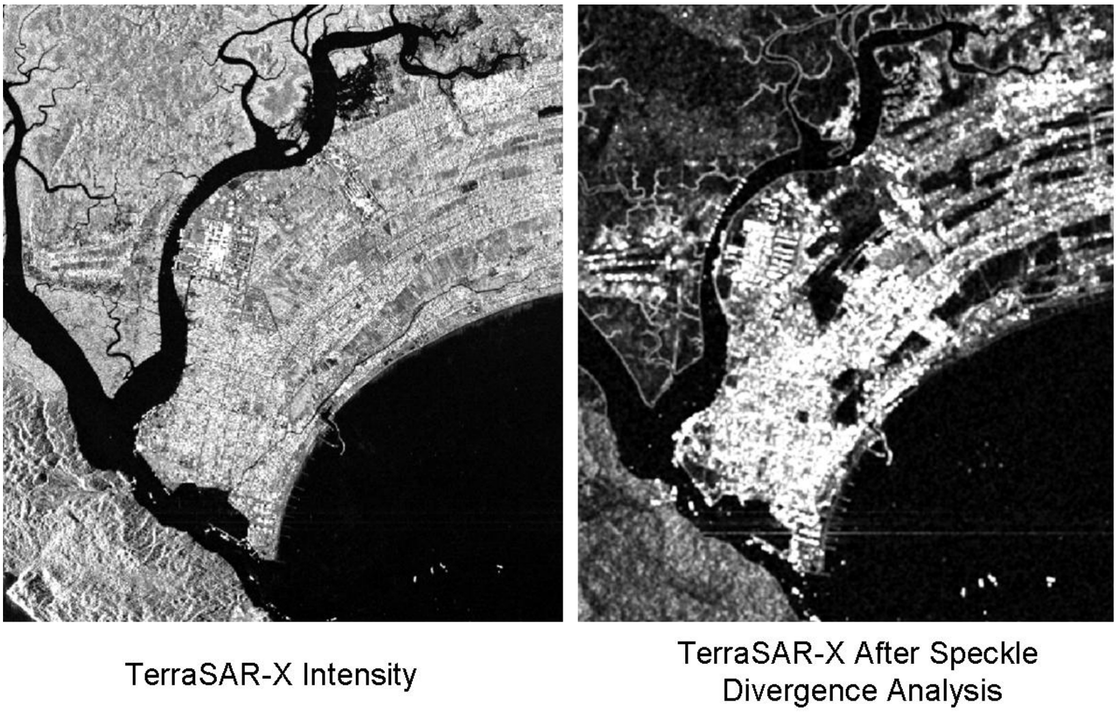
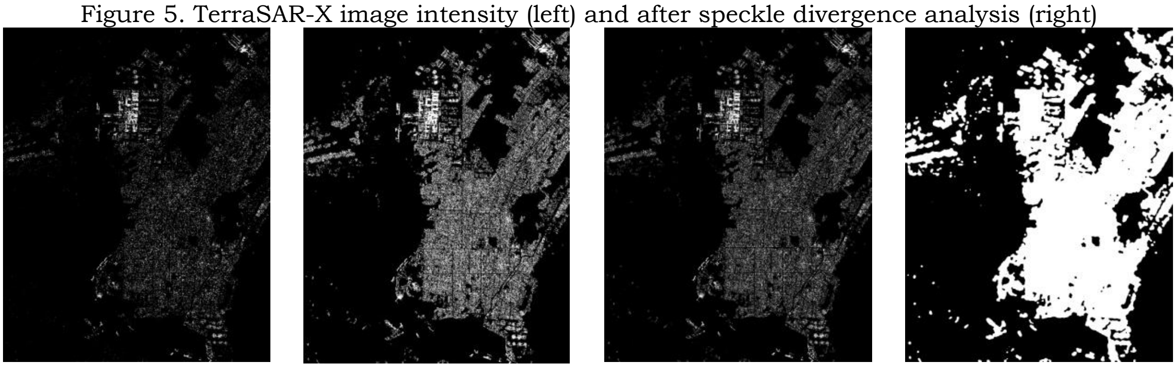


Figure 5. TerraSAR-X image intensity (left) and after speckle divergence analysis (right)



Real settlement | Settlement mask | Urban scatter | Settlement class

Figure 6. Settlement classification results and real settlement by using TerraSAR-X

Table 3. ACCURACY ASSESSMENT OF SPECKLE DIVERGENCE AND NEIGHBORHOOD ANALYSIS FOR CILACAP DISTRICT

Reference / Classification	Settlement (km2)	Non-settlement (km2)	Total (km2)	Commission error (%)
Settlement (km2)	101.5	38.4	139.9	27.5
Non-settlement (km2)	67.5	521.8	589.3	11.5
Total (km2)	168.9	560.2	729.2	
Omission Error (%)	39.9	6.9		
			Overall accuracy (%)	85.5

3.3 Classification results

Settlement areas within Cilacap District and Padang District have been classified by using the speckle divergence and neighborhood analysis from TerraSAR-X data. The overall accuracy was 85.5% for Cilacap and 78.14% for Padang. The detailed results of settlement classification in these areas by using TerraSAR-X were described in the following sub-chapters.

3.3.1 Cilacap district

The accuracy assessment of the speckle divergence and neighborhood analysis for Cilacap was shown as confusion matrix in Table 3. The commission error of built-up detection in Cilacap area was 27.5% and the omission error was 39.9%. The overall accuracy of settlement classification by using TerraSAR-X in Cilacap area was 85.5%. This accuracy can be categorized in high accuracy because more than 75%. Comparing to other research, this result has shown a significant improvement because other methods on settlement classification using SAR imagery have accuracy below 60%.

The result of speckle divergence and neighborhood analysis for settlement detection in Cilacap District is visualized in Figure 7.

Settlements in rural areas of Cilacap were difficult to detect. The misdetection (blue color) appears more often than the good detection (red color). The occurrence of this error is due to the effect that the settlements in the rural areas are often mixed with vegetation. The settlement with surrounding vegetation is common and typical for almost all rural areas in Indonesia. The vegetation sometimes, covers the roof of a house and is therefore classified as vegetation, not as settlement. In urban areas, the classification of settlements delivers a good detection (see red colors in Figure 7).

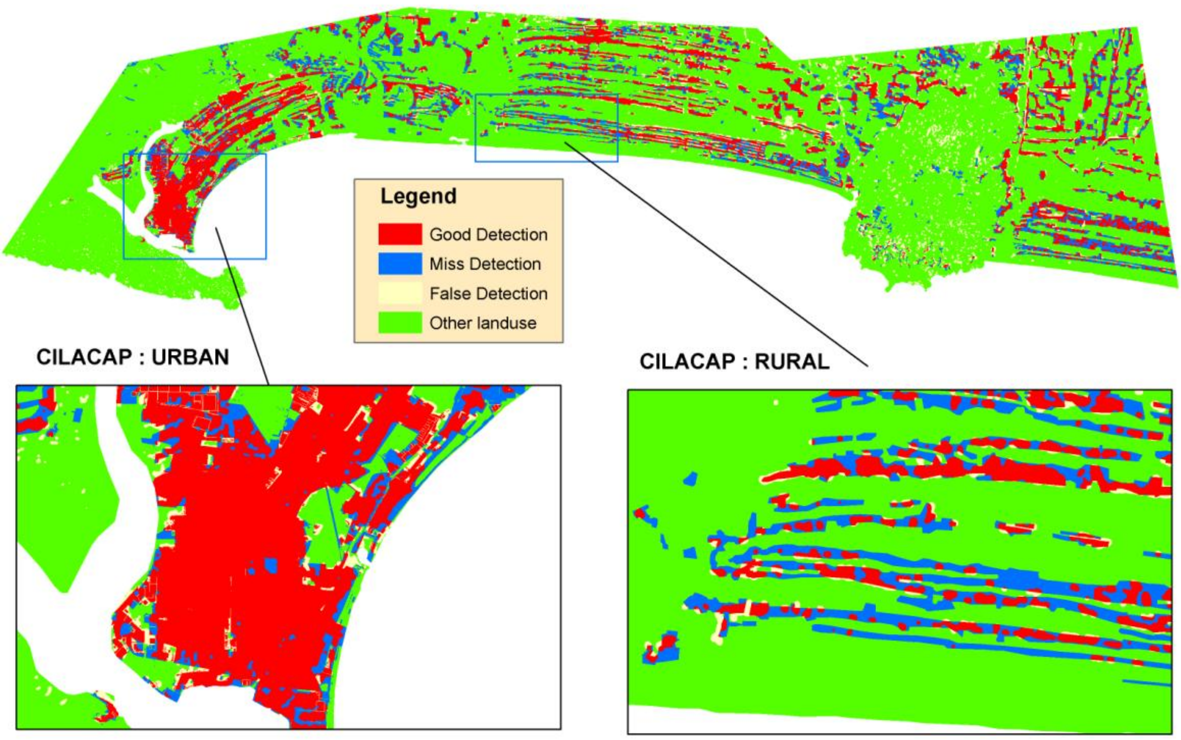


Figure 7. Result of speckle divergence and neighborhood analysis for settlement detection by using TerraSAR-X on Cilacap District

Table 4. ACCURACY ASSESSMENT OF SPECKLE DIVERGENCE AND NEIGHBORHOOD ANALYSIS FOR PADANG DISTRICT

Reference / classification	Settlement (km2)	Non-settlement (km2)	Total (km2)	Commission error (%)

Settlement (km2)	36.8	9.9	46.7	21.1
Non-settlement km2)	18.0	62.9	81.0	22.3
Total (km2)	54.8	72.8	127.7	
Omission error (%)	32.9	13.5	Overall accuracy (%)	78.1

3.3.2 Padang District

The same model as that used for Cilacap District to classify settlements using Terra-SAR X data was also implemented to Padang area. The results of classification and accuracy assessment are shown in Figure 8.

The results of the accuracy assessment of the neighborhood analysis provided by a confusion matrix are shown in Table 4. The commission error of built-up detection in Padang area is 21.1% and the omission error is 32.9%. The overall accuracy of settlement classification using TerraSAR-X imagery was 78.1% for the Padang area. Based on this result, with more than 75% accuracy, it showed that the model can be implemented to the other area.

The result of the speckle divergence and neighborhood analysis for settlement detection in the Padang District is shown in Figure 8.

Figure 8 shows the spatial distribution of false built-up detection (yellow), settlement misdetection (blue), good built-up detection (red), and other land use classes (green). Similar to the Cilacap region, mis- and false detection in the rural areas of Padang is higher than in the urban areas.

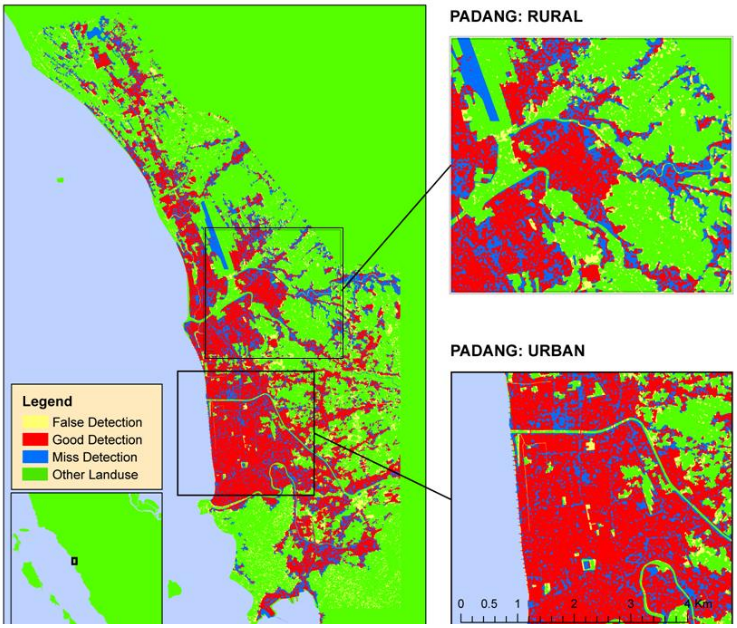


Figure 8. Result of speckle divergence and neighborhood analysis for settlement detection using TerraSAR-X of the Padang District

Figure 9. Examples of settlements in the rural area of Cilacap District (settlements surrounded by vegetation)

The methodology of settlement classification of TerraSAR-X imagery by using a combination of speckle divergence and the neighborhood analysis was applied in order to improve the settlement detection, especially for hilly areas. In Cilacap districts, the result of the accuracy assessment for the settlement classification in the study area was over 85%. The high omission and commission errors appeared in the result of settlement classification, especially in the rural areas. This was because of the settlement conditions in the rural area, which were usually surrounded by vegetation. The texture of this area in the TerraSAR-X imagery was more likely vegetation than settlement area. Figure 9 shows pictures of settlement areas with vegetation surrounding, taken in the rural area of Cilacap.

The transferability of this method has been also successfully investigated in the Padang area, which resulted in an accuracy of over 75%, which can be categorized in the high accuracy. The omission and commission errors for the rural area were also high, such as for the Cilacap District. The characteristics of settlement areas in the rural parts of Padang were mostly similar to the conditions of the rural areas in Cilacap District. Both result of accuracy can be categorized in the high accuracy. The different of Padang in Cilacap was mostly in the hilly area. There were some settlement areas in Padang located in the hilly area. The method on settlement classification using SAR Imagery in the hilly area was still a challenge to be improved.

4 CONCLUSIONS

A method of settlement classification using radar data has been further developed using TerraSAR-X imagery. The improvements included the concept of neighborhood analysis. The settlement areas in the study area were mapped with high accuracy, more than 75%. Limitations occur in rural areas where settlement structures were interspersed with vegetation. This results in relatively high omission and commission errors.

The method of settlement classification has been successfully transferred to the Padang District. Here, the same problem regarding rural settlements occurred, also resulting in high omission and commission errors in rural areas. Thus, future research on settlement classification using radar imagery must focus on improving the methodology for the classification of rural settlements and hilly areas.

REFERENCES

Acqua, F.D., P. Gamba, and G. Lisini, 2003, Improvements to urban area characterization using multitemporal and multiangle SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1996-2004.

Acqua, F.D. and P. Gamba, 2003, Texture-based characterization of urban environments on satellite SAR images, *IEEE Transactions on Geoscience and Remote Sensing*, 41(1):153-159.

Congalton, R. and K. Green, 1999, Assessing the accuracy of remotely sensed data: principles and practices, CRC/Lewis Press, Boca Raton, Fl. 137p.

Corbane, C., J.F. Faure, N. Baghdadi, N. Villeneuve, and M. Petit, 2008, Rapid urban mapping using SAR/Optical imagery synergy, *Sensors*, 8:7125-7143.

Dekker, R.J., 2003, Texture analysis and classification of ERS SAR images for map updating of urban areas in the Netherlands. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (9):1950-1958.

Esch, T., M. Thiel, A. Schenk, A. Roth, A. Müller, and S. Dech, 2010, Delineation of urban footprints from TerraSAR-X data by analyzing speckle characte-ristics and intensity Information, *IEEE Transactions on Geoscience and Remote Sensing,* 48(2):905-916.

Gamba, P. and F.D. Acqua, 2007, Fusion of radar and optical data for identification of human settlements. *In*: Weng, Q. (Editor): Remote sensing of impervious surface. Taylor and Francis Groups, LLC.

Li, J. and M. Zhao, 2003, Detecting urban land-use and land-cover changes in Mississauga using Landsat TM images. *Journal of Environmental Informatics*, 2(1):38-47.

Mesev, V., 1998, The use of census data in urban image classification. *Photo-grammetric Engineering & Remote Sensing*, 64(5):431-438.

Mubareka, S., D. Ehrlich, F. Bonn, and F. Kayitakire, 2008, Settlement location and population density estimation in Rugged Terrain using information derived from Landsat ETM and SRTM data. *International Journal of Remote Sensing*, 29(8):2339-2357.

Pellizzeri, T.M., P. Gamba, O. Lombardo, and F.D. Acqua, 2003, Multitemporal/ multiband SAR classification of urban areas using spatial analysis: Statistical versus neural kernel-based approach. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10):2338-2353.

Pozzi, F. and C. Small, 2001, Exploratory analysis of suburban land cover and population density in the U.S.A., IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas. Paper 35.Rome, Italy 8-9 November, 2001.

Stasolla, M. and P. Gamba, 2008, Spatial indexes for the extraction of formal and informal human settlements from high-resolution SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1 (2):98-106.

Story, M. and R.G. Congalton, 1986, Accuracy assessment: a user's perspective, *Photogrammetric Engineering and Remote Sensing*, 52 (3):397-399.

Tobler, W.R., 1969, Satellite confirmation of settlement size coefficients, *Area*, 1: 30-34.

Thiel, M., T. Esch, and A. Schenk, 2008, Object-oriented detection of urban areas from TerraSAR-X data. *In:* Proceeding of ISPRS 2008 Congress (37), Part B8, Commission VIII, Beijing, 23-27pp.

Tison, C., J.M. Nicolas, F. Tupin, and H. Maître, 2004, A new statistical model for Markovian classification of urban areas in high-resolution SAR images, *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2046-2057.

Xu, H., 2007, A new index for delineating built-up land features in satellite imagery, *International Journal of Remote Sensing*, 29(14):4269-4276.

Zha, Y., J. Gao, and S. Ni, 2003, Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3):583-594.

Zhang, Q., J. Wang, X. Peng, P. Gong, and P. Shi, 2002, Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. *International Journal of Remote Sensing*, 23(15):3057-3078.