Sumatra Journal of Disaster, Geography and Geography Education, June, 2024 Vol.8, No.1, pp.126-132
DISASTER, GEOGRAPHY, GEOGRAPHY EDUCATION http://sjdgge.ppj.unp.ac.id/index.php/Sjdgge
ISSN : 2580 - 4030 (Print) 2580 - 1775 (Online), Indonesia

MACHINE LEARNING TO IDENTIFY SLUM SETTLEMENTS BASED ON SATELLITE IMAGERY CASE STUDY: PADANG CITY

*Risky Ramadhan[1], Azhari Syarief[1]

[1] Departement of Geography – Universitas Negeri Padang

Email: riskyramadhan@fis.unp.ac.id

*Corresponding Author, Received: May 10,2024, Revised: May 28, 2024, Accepted: June 28, 2024

ABSTRACT: Mapping of slums in Indonesia uses a *survey-based mapping* method. This method requires a lot of time and money, especially when updating data, especially during the current pandemic, it will be risky. Based on the Minister of PUPR Number 14/PRT/M/2020. Over the last 6 years (2014-2020), the mapping of slums cost 382 billion. In addition, the main disadvantage of the survey method is the inconsistency of the results due to the interpretation of surveyors in different fields. To overcome this problem, remote sensing-based slum identification with a *machine learning approach* with the help of High-Resolution Satellite Imagery (CSRT) and *support vector machine* (SVM) features can quickly and cheaply calculate slum areas. Therefore, the objectives of this study are (1) Identifying the distribution of slums using *machine learning methods* based on Satellite Imagery in Padang City. Based on the description above, it can be concluded that the level of slums in the identified areas of slum settlements in Padang City is predominantly light category. There is only one ward, Teluk Kabung Selatan, which is at the medium category. Slums not always squatter which is proven by the ownership status of the land certificate. Pasar Ambacang Ward has the most families in slum settlements (1.112 families). Meanwhile, Tarantang Ward has the smallest slum families in the identified slum area (25 families). Based on the results of the analysis using support vector machine and field verification, slums were identified in 45 (forty-five) villages spread across 11 sub-districts with a total area of 129.16 hectares. This condition is supported by the centralization of activities in the *core* village of Padang City, which creates a development gap with the hinterland area.

Keywords: Slum Area, Machine Learning, Support Vector Machine, Padang

1. INTRODUCTION

Handling slum settlements became an international agenda campaigned on Millenium Development Goals (MDGs) [1] and Sustainable Development Goals (SDGs) [2]. Mapping slum settlements in Indonesia using survey methods (survey-based mapping) which is based on physical and social criteria [3]. Based on PUPR Ministerial Regulation Number 14/PRT/M/2020 [4], over the last 6 years (2014-2020), slum mapping has cost 382 billion. Apart from that, the main drawback of the survey method is that there are inconsistencies in area results due to different interpretations of surveyors in the field.

In 2018, the total area of slum areas in Indonesia was 87,000 ha, reported by 390 cities/regencies in Indonesia [5]. Padang City according to Regional Regulation No. 501 of 2019 has a residential area of 122, 83 ha and 2,224 houses are unfit for habitation [6]. In line with the publication of the 2020-2024 RPJMN [7] This needs to be handled quickly, both in identifying the location of slum settlements.

To overcome this problem, a remote sensing-based slum settlement identification approach is proposed by machine learning. Machine Learning is one part of artificial intelligence which focuses on the classification of an object based on image processing. Feature extraction and classification via image processing methods has become very popular, especially based on the distance between named points support vector machines (SVM) [8] [9] [10]. SVM through the testing and validation process provides higher accuracy compared to other methods such as neural network, decision tree and maximum likehood.

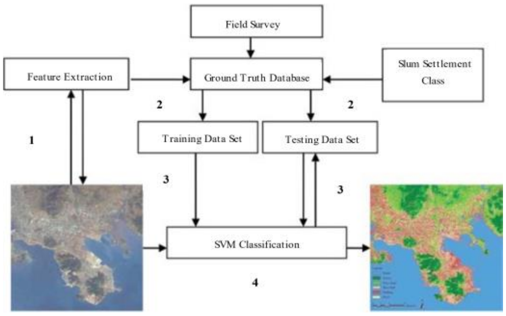

Fig.1 Steps in working on SVM Classification

Several studies have used machine learning to identify slum settlements [11] [12], but has not been integrated with slum-based indicators generic slum ontology (GSO) and local slum ontology (LSO). LSO includes object level image settlements and level of image visualization

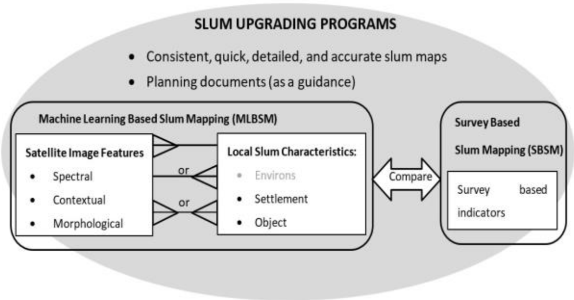

Fig.2 The relationship between LSO, GSO and Machine Learning and Slum Settlement Survey

Supported with in depth interviews with expert and field surveys to take pictures, a class/classification of slum settlements is produced. To reduce margins of errors can be assisted by morphological, textural, and spectral features. The novelty value of this research is the application of technology artificial intelligence, which using a machine learning method based on satellite imagery with GSO and LSO criteria.

Based on the description above, the importance of research with the title "**Machine Learning to Identify Slum Settlements Based on Satellite Imagery Case Study: Padang City"**. The purpose of this research is to identify the distribution of slum settlements using methods of machine learning based on satellite imagery.

2. RESEARCH METHODS

This research went through several stages and methods which are described in the points below;

1) Pleiades Image Correction (Geometric and Radiometric)

The method chosen is image to image by utilizing Ikonos satellite imagery of the 2010 research area which has been geometrically and radiometrically corrected [13]. Mark root mean square (RMS) 0 indicates there is no position error, while an RMS value of 1 means there is a distortion error of 1 pixel. Pleiades imagery has an accuracy of 0.5 meters. Geometric and radiometric accuracy is measured by the RMS value-error (Root Mean Square Error) follows the equation [14]:

$$RMS - Error = \sqrt{(x' - u°)^2 + (y' - v°)^2}$$

Information:

x : Latitude on the map y : Elongate on the map

u : Raw image v : Column in the image

2) Slum Characteristics and Image Features

Based on literature study [15], characteristics of slums and their relationship with feature images such as contextual, spectral, morphological so they can represent Local Ontology Slum Area (LSO). The following table explains the above characteristics;

Table 1. Relationship between GSO, Indicator, LSO, and Image Features

GSO Dimensions	Indicator	Local Indicator/LSO	Image Features
Environment	**Location**	Disaster area, small alley	No image features are used to describe neighborhood levels
	Neighbor Characteristics	Proximity to industry, commercial, elite housing, bus stations, and smelly and dirty areas	
Settlement	**House Form**	Irregular pattern, following the longitudinal formation of a river or railroad track	Ground Thruth and Morphological
	Density	High density (> 250 units/Ha), high roofs, coverage, less vegetation	Survey and Spectral/NDVI
Object	**Accessibility**	Roads are unpaved or poorly constructed, 2.5 m wide, closed channels or without	Ground Thruth

GSO Dimensions	Indicator	Local Indicator/LSO	Image Features
		drainage	
	Building Characteristics	Permanent and non-permanent structures, with roofs made of corrugated iron, asbestos, plastic, fiber, and clay tiles; building size 10-60 m 2; poor sanitation, using well water or buying water	Spectral/NDVI

To describe morphology, the features used are a combination of bands on satellite images with the Vegetation Index (NDVI) method, Build-up presence index (PanTex), with the Appendix method.

1) Normalization Different Vegetation Index

NDVI is used to analyze the presence and condition of vegetation in very dense slum areas (with no vegetation), as an indicator to differentiate slum and non-slum environments.

2) Build-Up Presence Index (PanTex)

PanTex is able to extract the structural characteristics of buildings, based on the fact that buildings have shadows that lead to high local contrast [16]. Brightness bands are used for input due to the absence of panchromatic bands assuming that the constructed structure is the brightest feature in the optical band [17], with the formula:

$$\text{Brightness} = \text{Max (Visible Bands)}$$

Anisotropic rotationally invariant texture measures are used by PanTex to suboptimally solve displacement problems in GLCM [17]. PanTex uses min or fuzzy to replace average, as defined in the formula:

$$tx\ (\text{built up}) = \cap_i tx_{i;} i\ \in [1 \ldots n]$$

where n is the number of displacement vectors (combination of distance and angle) [18]. Additionally, the intersection operator (min) between the texture sizes of different displacement vectors is used [16] as formulated below:

$$f\ (\text{built up}) = \wedge \{tx_{i,}\ tx_{2,}\ ilttx_n\ ; i\ u[1]$$

3) Support Vector Machine (SVM) Classification

Classification using SVM is carried out in software R for statistics. Accuracy of the Pleiades Image with an accuracy of 0.5-meter based on city planning documents. SVM classification steps [19];

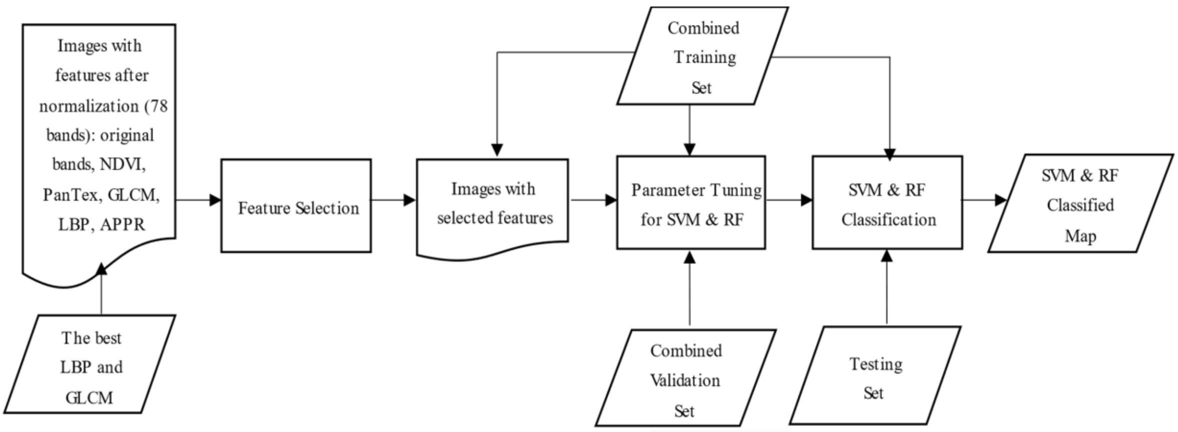

Fig.3 Feature Selection and SVM Classification Process to Identify Slum Settlements

3. RESULTS AND DISCUSSION

The characteristics of slums are identified in 2 (two) stages, namely using satellite images and field surveys. At the stage of interpretation of satellite images, NDVI analysis is carried out.

Based on the image classification that has been carried out, the level of slums in the identified areas of slum settlements in Padang City is in the light category. There is only one ward, Teluk Kabung Selatan which is in the medium category where there are 4-5 slum criteria that are met in this location. For the legality of the land, this area identified as a slum still has a land certificate, thereby confirming that it is a slum settlement (*slum*) is not always wild (*squatter*) which is proven by the ownership status of the building on its own land.

The survey results also show the number of heads of families identified as living in slum settlements. Below is a complete table of the population, area and families in the slums of Padang City in 2021(Table 2).

Table 2. Population, Area and Families (KK) in Slum Settlements in Padang City, 2021

Subdistrict	Ward	Total Population (Persons)	Area (Ha)	Number of families (KK) in slum settlements
Bungus Teluk Kabung	Bungus Barat	6408	1808	347
	Bungus Timur	6107	2581	164
	Teluk Kabung Selatan	1882	914	222
Koto Tangah	Padang Sarai	21951	1324	47
	Batipuh Panjang	15637	1432	126
	Batang Kabung Ganting	14077	332	135
	Balai Gadang	17423	10690	144
Kuranji	Pasar Ambacang	19674	503	1112
	Anduring	15025	404	288
Lubuk Begalung	Koto Baru Nan XX	9215	103	123
	Banuaran	10941	129	134
	Lubuk Begalung XX	9108	155	87
Lubuk Kilangan	Bandar Buat	16303	287	143
	Tarantang	2881	185	25
	Padang Besi	6950	491	201
	Koto Lalang	9126	332	66
Nanggalo	Tabing Bandar Gadang	8188	91	161
Padang Barat	Flamboyan Baru	4855	43	161
Pauh	Pisang	9788	399	194
	Piai Tangah	4273	497	341
	Limau Manis	8596	2486	125
	Kampung Dalam	6693	297	438
Total slum families in Padang City in 2021				**4782**

Source: Results of data analysis, 2021.

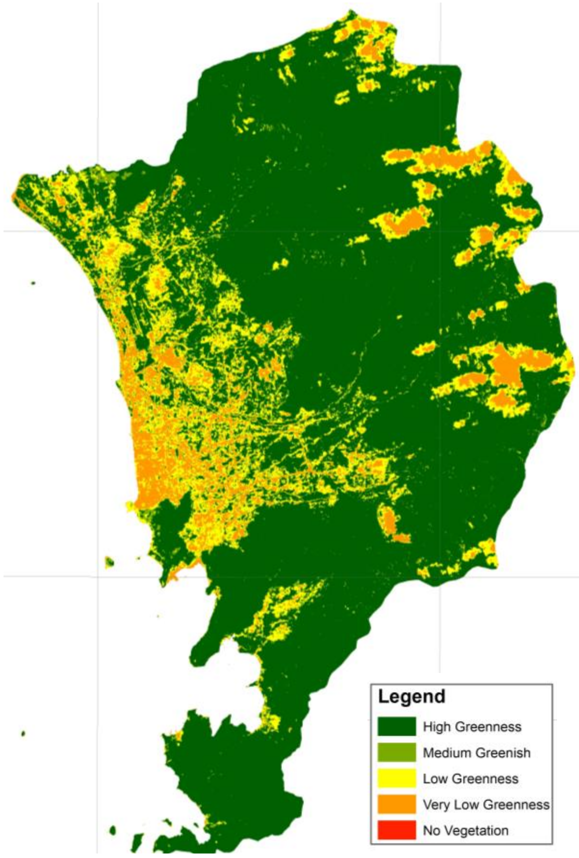

Fig.5 Normalizatin Difference Vegetation Index (NDVI)

The survey results were used as a basis for conducting training and testing in slum settlement areas in Padang City. There are 11 characteristics of settlements presented in the picture below;

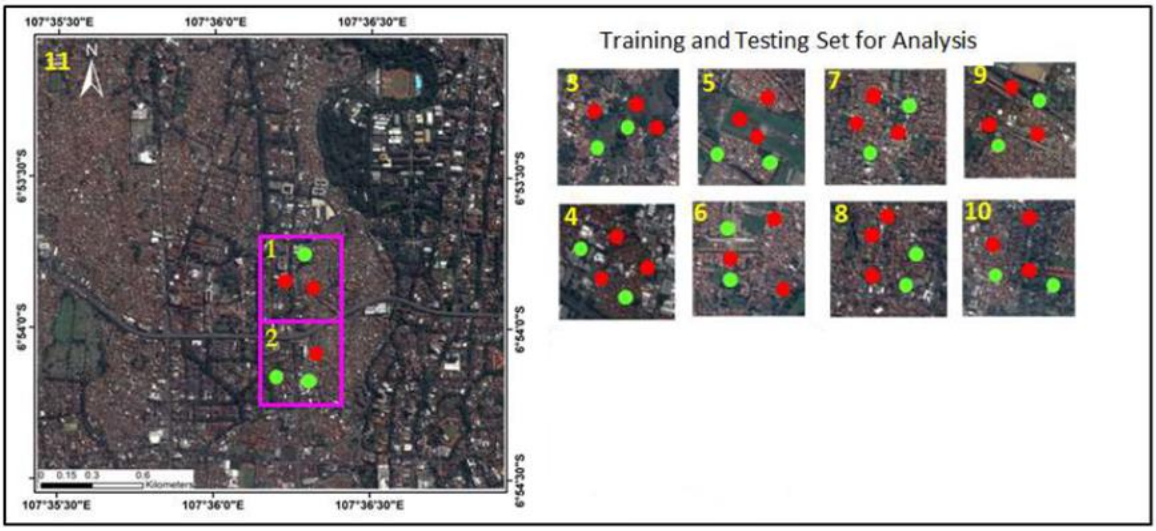

Fig.4 Training and Testing Set for Analysis

The results from the table above show that there are 4.782 families living in settlements identified as slums in Padang City. Pasar Ambacang Ward, Kuranji Subdistrict has the largest number of slum families at 1.112 families. Tarantang is the smallest contributing family in the area identified as a slum, 25 families. The following is a map of the distribution of slum settlements in 2021 in Padang City (Fig. 5);

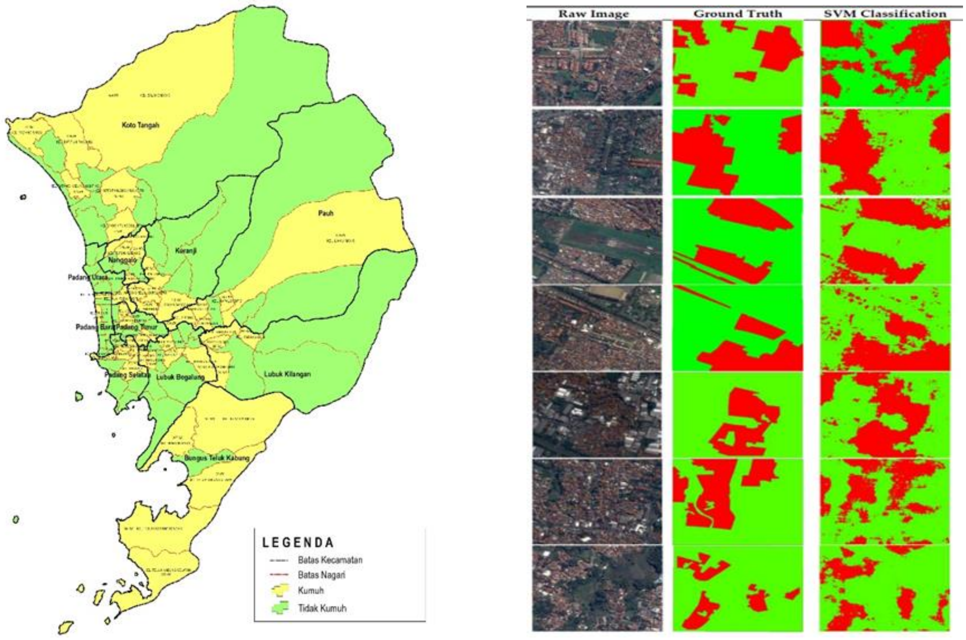

Fig.5 Map of Slums Using Machine Learning in Padang City and Sample Areas

The results of the analysis of slums using the support vector machine obtained slum delineation in 22 (twenty-two) villages in Padang City, so the total area of slums is 122.16 hectares. This result is also strengthened by the Decree of the Mayor of Padang Number 501 of 2019 concerning slums, where the area of Padang City slums in 2019 is 122.33 hectares. When compared to the Decree of the Mayor of Padang Number 163 of 2014, there was an increase in the area of 14.38 hectares with an area in 2014 of 107.78 hectares, but the location of the slum determined was different from the Decree in 2020. This means that there is a shift in slums from those that were initially not slums to slums according to the criteria for slum housing and slums according to the Regulation of the Minister of Public Works and Public Housing Number 14 of 2018.

In the slum area of Padang City, there are 8.282 families living in settlements identified as slums in Padang City. Batang Arau Ward, Padang Selatan Subdistrict has the largest number of slum families with 801 families. Tarantang Ward is the smallest contributing family in the area identified as a slum, 25 families.

4. CONCLUSION

Based on the description above, it can be concluded that the level of slums in the identified areas of slum settlements in Padang City is predominantly light category. There is only one ward, Teluk Kabung Selatan, which is at the medium category. Slums not always squatter which is proven by the ownership status of the land certificate. Pasar Ambacang Ward has the most families in slum settlements (1.112 families). Meanwhile, Tarantang Ward has the smallest slum families in the identified slum area (25 families). Based on the results of the analysis using support vector machine and field verification, slums were identified in 45 (forty-five) villages spread across 11 sub-districts with a total area of 129.16 hectares. This condition is supported by the centralization of activities in the *core* village of Padang City, which creates a development gap with the hinterland area.

5. ACKNOWLEDGEMENTS

Thanks to the Institute for Research and Community Service (LPPM) of Universitas Negeri Padang for supporting research funds in the contract with the Rector’s Decree Number 164/UN.35/LT/2021. The research has no intention or conflict of interest toward individuals or groups

6. REFERENCES

[1] Wijaya DW. 2016. Perencanaan penanganan kawasan Permukiman Kumuh Studi Penentuan Kawasan Prioritas Untuk Peningkatan Kualitas Infrastruktur pada Kawasan Permukiman Kumuh di Kota Malang. Jurnal Ilmiah Administrasi Publik Vol;2(1):1-10.

[2] SDgs Indonesia. 2020. Build Forward Better tentang Upaya Pencapaian Target SDgs 2030. http://sdgsindonesia.or.id/

[3] Peraturan Daerah Kota Padang Nomor 5 Tahun 2018 Tentang Pencegahan dan

Peningkatan Kualitas Perumahan Kumuh dan Permukiman Kumuh.
[4] Peraturan Daerah Kota padang Sumatera Barat Nomor 501 Tahun 2019 Tentang Deliniasi permukiman Kumuh 2016-2019.
[5] Bappenas. 2020. Perpres No. 18 Tahun 2020 tentang RPJM Nasional Tahun 2020-2024. http://jdih.bappenas.go.id/peraturan/detailperaturan/1037
[6] Kohli, D.; Kerle, N.; Sliuzas, R. Local ontologies for object-based slum identification and classification. In Proceedings of the 4th GEOBIA, Rio de Janeiro, Brazil, 7–9 May 2012; pp. 1–4
[7] Bachofer, F.; Quénéhervé, G.; Märker, M.; Hochschild, V. Comparison of SVM and boosted regression trees for the delineation of lacustrine sediments using multispectral ASTER data and topographic indices in the Lake Manyara Basin. Photogramm. Fernerkund. Geoinf. 2015, 2015, 81–94. [CrossRef]
[8] Gevaert, C.M.; Persello, C.; Sliuzas, R.; Vosselman, G. Informal settlement classification using point-cloud and image-based features from UAV data. ISPRS J. Photogramm. Remote Sens. 2017, 125, 225–236. [CrossRef]
[9] Bergado, J.R.; Persello, C.; Stein, A. Recurrent multiresolution convolutional networks for VHR mage Classification. IEEE Trans. Geosci. Remote Sens. 2018, 1–14. [CrossRef]
[10] Syrris, V.; Ferri, S.; Ehrlich, D.; Pesaresi, M. Image enhancement and feature extraction based on low-resolution satellite data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2015, 8, 1986–1995. [CrossRef]
[11] Heung, B.; Ho, H.C.; Zhang, J.; Knudby, A.; Bulmer, C.E.; Schmidt, M.G. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 2016, 265, 62–77.[CrossRef]
[12] Owen, K.K.; Wong, D.W. An approach to differentiate informal settlements using spectral, texture, geomorphology and road accessibility metrics. Appl. Geogr. 2013, 38, 107–118. [CrossRef]
[13] Sitorus S.R.P, Leonataris C, Panuju D.R. 2012. Analisis Pola Perubahan Penggunaan Lahan dan Perkembangan Wilayah di Kota Bekasi, Provinsi Jawa Barat. Jurnal Tanah Lingkungan. Vol;14(1):21-28.
[14] Fajeri DB, Subiyanto S, Sukmono A. 2017. Analisis Perkembangan Permukiman dan Perubahan Nilai Tanah (Studi kasus: Kec. Banyumanik Kota Semarang Jawa Tengah). Jurnal Geodesi Undip. Vol; 6(1):179-188.
[15] Pidora D, Pigawati B. 2014. Keterkaitan Perkembangan Permukiman dan Perubahan Harga Lahan di Kawasan Tembalang. Jurnal Wilayah dan Lingkungan. Vol;2(1):1-10.
[16] Panuju DR, Rustiadi E. 2013. Teknik Analisis Perencanaan Pengembangan Wilayah. Bogor (ID): Departemen Ilmu Tanah dan Sumberdaya Lahan, Institut Pertanian Bogor.
[17] Pidekso MD, Subiyanto S, Amarrohman FJ. 2017. Analisis Faktor Aksesibilitas Terhadap Zona Nilai Tanah dengan Pendekatan Penilaian Massal dan Sistem Informasi Geografis (Studi Kasus: Kecamatan Pemalang, Kabupaten Pemalang). Jurnal Geodesi Undip. Vol; 6(1):285-292.
[18] Surtiani E.E. 2006. Faktor-Faktor yang Mempengaruhi Terciptanya Kawasan Permukiman Kumuh di Kawasan Pusat Kota (Studi Kasus: Kawasan Pancuran, Salatiga). [Tesis]. Semarang (ID): Universitas Diponegoro.
[19] Sidauruk AS, Subiyanto S, Sukmono A. 2016. Analisis Pengaruh Perubahan Penggunaan Lahan Terhadap Zona Nilai Tanah (Studi Kasus: Kecamatan Kaliwungu Kabupaten Kendal Tahun 2010-2015). Jurnal Geodesi Undip. Vol; 5(2):44-54.