

Geobia approach for urban land use mapping: random forests and spatial metrics relationship for Ciudad Juarez, Chihuahua, Mexico

Ivan E. Ruiz^{1*}, Wenzhong Shi^{1, 2}

¹*The Hong Kong Polytechnic University, Department of Land Surveying and Geo-Informatics, Hung Hom, Kowloon, Hong Kong 999077, China.*

²*Professor and head of the Department of Land Surveying and Geo-Informatics. Director of the Advanced Research Centre for Spatial Information Technology.*

* 14900652r@polyu.edu.hk

Since 2007 half of the world inhabitants have been living in urban areas; and by the year 2050 is projected that the urban population will surpass 60% of the total world population (United Nations, 2014; 2015). This expansion is most likely to happen in developing countries (Graizbord, 2007). Some of these countries do not count with the resources to cope with this growing social phenomenon, thus causing segregation, slums, deficiency of infrastructure, social inequality and uncontrollable sprawl. This has become one of the main challenges for those countries all over the world that have been working on solutions to provide acceptable conditions and quality of life for the growing population in urban areas. The demographic growth has indeed raised many concerns such as, environmental, economic, hydrological, heat fluxes, micro-climatic, violence, irregular settlements, visual, sound, air and water pollution, social and psychological issues.

The main objective of this paper is to apply a novel method to extract the urban land use from remotely sensed images. This is due to the fact that there is a need to improve methodologies that can update and identify the constant complex changing patterns and spatial development of urban areas.

The central tool for good urban planning is the map of land use; however, their production is largely based in visual interpretation and census data, which in most cases, the time and money invested is very high. Currently, the processes to obtain these tools in developing countries are languid and costly, primarily because they tend

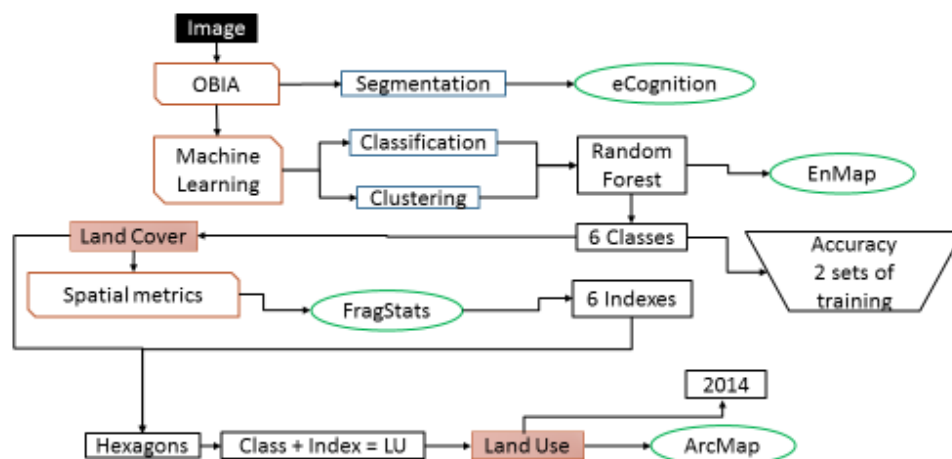
to rely on census data. Additionally, the environment of a city is moving relentlessly. The shape, size, land cover, land use and transport are changing constantly, making the periodicity of the national census information not capable for immediate response. Technologies such as remote sensing and geographic information science combined with new methodologies based on spatial metrics and deep learning can aid in the production of high quality land use map with trustworthy accuracy reliably and expeditiously (Berger, Voltersen, Hese, Walde, & Schmulius, 2012; Maeyer, Sotiaux, & Wolff, 2010; Herold, Meinel, Hecht, & Csaplovics, 2012; Hoffmann, Strobl, Blaschke, & Kux, 2008; Novack, Kux, Feitosa, & Costa, 2010).

While remote sensing has been making a constant progress with the use of better sensors and technology, the classifications are still based on concepts established over thirty years ago and most of them are based on pixels and not in objects (Blaschke, et al., 2000). Nowadays there is enough evidence that classification for urban areas must be done with an object based approach (Bakos, Lisini, Trianni, & Gamba, 2013; Chen, Hay, Carvalho, & Wulder, 2012; Dezso, Fekete, Gera, Giachetta, & Laszlo, 2012; Hu & Wang, 2013; Pu, Landry, & Yu, 2011; Tomljenovic, 2012; Zhou & Troy, 2008). GEOBIA involves three phases: segmentation, training and classification (Abbas, 2008; Li & Shao, 2013; Myint, et al., 2011).

This has generated the enthusiasm for Deep learning algorithms in remote sensing because this classification can now be done in minutes instead of days like before. Deep learning is an area of machine learning research. While the term is novel the definition dates from 1950 (Deep Learning, 2014; Bengio, Goodfellow, & Courville, 2015; Wing, 2014). It has been successfully applied in visual classification (Duin, 2012), pedestrian detection (Arnold, Rebecchi, Chevallier, & Paugam-Moisy, 2011), face recognition (Jones, 2014), transport (Hasegawa, Arimura, & Tamura, 2013) and more recently to remote sensing. Random forest being the most popular algorithm because of its simplicity, speed and accuracy. Developed by Breiman (2001) and tested by many nowadays with good results (Kamusoko & Gamba, 2015; Feng, Liu, & Gong, 2015; Rodríguez-Galiano & Chica-Rivas, 2012; Breiman, 2001). Normally the use of this algorithm involves programming languages (Evans, 2014; Breiman, Cutler, Liaw, & Wiener, 2015) but now, thanks to the software EnMAP box, random forests is

accessible to every researcher (Waske, Linden, Oldenburg, & Jakimow, 2012) A good classification it is not enough for a land use map and the spatial metrics can fill that void. These metrics are essential in urban planning and land use mapping. They can measure the structure and arrangement of the urban landscape and they have been tested greatly with acceptable results (Araújo, 2010; Badii & Landeros, 2007; Herold M. , 2004; Herold, Couclelis, & Clarke, 2005; Jaafari, Sakieh, Shabani, Danehkar, & Nazarisamani, 2015; McGarigal, Cushman, & Ene, 2012; Sapena & Ruiz, 2015). Spatial metrics can measure and model the city thus being able to analyze the relationship between the urban density and the land cover map to create a good quality land use map (Kim, 2015).

In this paper we propose a combination of techniques to achieve this, using a GEOBIA approach by means of eCognition software with a WorldView-2 image from Juarez, Mexico of the year 2014 and a machine learning algorithm (random forests using EnMAP box software) to obtain first, the land cover of the city with six classes; dark roof, bright roof, highway, bare soil, vegetation and parking lot). Then, extract the spatial metrics indexes (FRAGSTATS) to understand the relationship between these indexes and the land cover (LPI, TE, ED, CONTAG, CA and TA) to attain the land use map as explained in the next flowchart.



To test the accuracy of the classification two different sets of training are used to compare the results and create the cross validation. For the final land use map the accuracy will be tested against in site data collection.

The results will generate an accurate model of the urban land use for the city of Juarez, Mexico. This map is expected to be an essential asset for urban planners to solve the different issues to be tackled by the government. Once tested the method can be replicated in cities that share similar features with Juarez in developing countries.

The heterogeneity of factors involved in the complex dynamic and the rapid urban growth of the cities is a challenge for the city management and its government. Urban land-use it is not only crucial but, a specific need to know the characteristics of a city in an efficient, accurate and in a timely manner. These characteristics are fundamental to contribute to the improvement of the urban environment and the quality of life of its inhabitants. Currently, the processes to update the government databases are based in visual interpretation and on-screen digitizing using aerial photography, field survey and census data. This method can provide a tool to help the governments to implement strategies to guarantee that the development of the city is sustainable and the benefits communal.

[Number of characters with spaces 6838]

1 Reference list

- Abbas, S. (2008). *Study of Landscape Matriz for the Forest of Ayubia, NWFP. Master of Science dissertation*. Lahore, Pakistan: University of the Punjab.
- Araújo, C. S. (2010). *Aplicacoes de tecnicas de Sensoriamento Remoto na analise multitemporal do ecossistema manguezal na Baixada Santista, SP. (Mphil Dissertation)*. Sao Paulo, Brazil: Universidade de Sao Paulo.
- Arnold, L., Rebecchi, S., Chevallier, S., & Paugam-Moisy, H. (2011). An Introduction to Deep Learning. *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (pp. 477-488). Bruges: ESANN. Retrieved June 23, 2015, from <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2011-4.pdf>
- Badii, M. H., & Landeros, J. (2007). Measurement of the landscape fragmentation and its relation with sustainability. *Daena: International Journal of Good Conscience*, 26-38.

- Bakos, K., Lisini, G., Trianni, G., & Gamba, P. (2013). A novel framework for urban mapping from multispectral and hyperspectral data. *International Journal of Remote Sensing*, 34(3), 759-770.
- Bengio, Y., Goodfellow, I., & Courville, A. (2015, July 7). *University of Montreal*. Retrieved June 23, 2015, from Deep Learning: <http://www.iro.umontreal.ca/~bengioy/dlbook/intro.html>
- Berger, C., Voltersen, M., Hese, S., Walde, I., & Schmullius, C. (2012). USING GEOGRAPHIC OBJECT-BASED IMAGE ANALYSIS (geobia) FOR URBAN LAND COVER MAPPING AND SETTLEMENT DENSITY ASSESSMENT. *Proceedings of the 4th GEOBIA*, (pp. 503-508). Rio de Janeiro.
- Blaschke, T., Lang, S., Lorup, E., Strobl, J., Cremers, A., & Greve, K. (2000). Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. *Environmental information for planning, politics and the public*, 2, 555-570.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 5-32.
- Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2015, February 20). *Package 'randomForest'*. Retrieved June 23, 2015, from Breiman and Cutler's random forests for classification and regression: <http://stat-www.berkeley.edu/users/breiman/RandomForests>
- Chen, G., Hay, G., Carvalho, L., & Wulder, M. (2012). Object-based change detection. *International journal of Remote Sensing*, 33(14), 4434-4457.
- Deep Learning. (2014, March 5). *Deep Learning... moving beyond shallow machine learning since 2006*. Retrieved from Welcome to Deep Learning: <http://deeplearning.net/>
- Dezso, B., Fekete, I., Gera, D., Giachetta, R., & Laszlo, I. (2012). OBJECT-BASED IMAGE ANALYSIS IN REMOTE SENSING APPLICATIONS USING VARIOUS SEGMENTATION TECHNIQUES. *annales University of Science of Budapest*, 1-18.
- Duin, B. (2012, October 18). *Pattern Recognition Tools*. Retrieved June 23, 2015, from Machine learning and pattern recognition: <http://www.37steps.com/638/machine-learning-and-pattern-recognition/>

- Evans, J. (2014, January 14). *CLICKFOX*. Retrieved January 23, 2015, from Data Science Using Open Source Tools Decision Trees and Random Forest Using R: http://www.clickfox.com/ds_rcode/
- Feng, Q., Liu, J., & Gong, J. (2015). UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. *Remote Sensing*, 1074-1094.
- Graizbord, B. (2007). Megaciudades, globalización y viabilidad urbana. *Investigaciones Geográficas, Boletín del Instituto de Geografía UNAM*, 125-140.
- Hasegawa, H., Arimura, M., & Tamura, T. (2013). Hybrid Model of Random Forests and Genetic algorithms for Commute Mode Choice Analysis. *Proceedings of the Eastern Asia Society for Transportation Studies*. 9, pp. 123-136. Tokyo: EASTS.
- Hay, G., & Castilla, G. (2006). Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- Herold, H., Meinel, G., Hecht, R., & Csaplovics, E. (2012). A GEOBIA APPROACH TO MAP INTERPRETATION - MULTITEMPORAL BUILDING FOOTPRINT RETRIEVAL FOR HIGH RESOLUTION MONITORING OF SPATIAL URBAN DYNAMICS. *Proceedings of the 4th GEOBIA*, (pp. 252-254). Rio de Janeiro.
- Herold, M. (2004). *Remote Sensing and Spatial Metrics for Mapping and Modeling of Urban Structures and Growth Dynamics. PhD Dissertation*. Santa Barbara: University of California.
- Herold, M., Couclelis, H., & Clarke, K. (2005). The role of spatial metrics in the analysis and modeling of urban land use change. *Computer, Environment and Urban Systems*, 29, 369-399.
- Hoffmann, P., Strobl, J., Blaschke, T., & Kux, H. (2008). Detecting informal settlements from QuickBird data in Rio de Janeiro using an object-based approach. In T. Blaschke, S. Lang, & G. Hay, *Object-Based Image Analysis Spatial Concepts for Knowledge Driven Remote Sensing Applications* (pp. 531-554). Berlin: Springer.

- Hu, S., & Wang, L. (2013). Automated urban land-use classification with remote sensing. *International Journal of Remote Sensing*, 34(3), 790-803.
- Jaafari, S., Sakieh, Y., Shabani, A., Danehkar, A., & Nazarisamani, A.-a. (2015). Landscape change assessment of reservation areas using remote sensing and landscape metrics (case study: Jajroud reservation, Iran). *Environment, Development and Sustainability*, 1-17.
- Jones, N. (2014, January 9). *Nature*. Retrieved June 24, 2015, from The Learning Machines:
http://www.nature.com/polopoly_fs/1.14481!/menu/main/topColumns/topLeftColumn/pdf/505146a.pdf
- Kamusoko, C., & Gamba, J. (2015). Simulating Urban Growth Using a Random Forest-Cellular Automata (RF-CA) Model. *International Journal of Geo-Information*, 447-470.
- Kim, J. H. (2015). Crossing-over between land cover and land use: Exploring spatially varying relationships in two large US metropolitan areas. *Applied Geography*, 37-45.
- Li, X., & Shao, G. (2013). Object-based urban vegetation mapping with high-resolution aerial photography as a single data source. *International Journal of Remote Sensing*, 34(3), 771-789.
- Maeyer, M. D., Sotiaux, A., & Wolff, E. (2010). Comparison of standardized methods (object-oriented vs. per pixel) to extract the urban built-up area: example of Lubumbashi (DRC). *GEOBIA 2010: Geographic Object-Based Image Analysis* (pp. 1-6). Ghent, Belgium: ISPRS.
- McGarigal, K., Cushman, S., & Ene, E. (2012). *FRAGSTATS v4*. Amherst: University of Massachusetts. Retrieved October 12, 2015, from UMass Landscape Ecology Lab: <http://www.umass.edu/landeco/research/fragstats/fragstats.html>
- Myint, S., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011, May 15). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115(5), 1145-1161. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0034425711000034>

- Novack, T., Kux, H., Feitosa, R., & Costa, G. (2010). PER BLOCK URBAN LAND USE INTERPRETATION USING OPTICAL VHR DATA AND THE KNOWLEDGE-BASED SYSTEM INTERIMAGE. *GEOBIA 2010: Geographic Object-Based Image Analysis* (p. 7). Ghent: ISPRS.
- Pu, R., Landry, S., & Yu, Q. (2011). Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing*, 32(12), 3285-3308.
- Rodríguez-Galiano, V., & Chica-Rivas, M. (2012). Clasificación de imágenes de satélite mediante software libre: Nuevas tendencias en algoritmos de Inteligencia Artificial. *XV COngreso Nacional de Tecnologías de la Información Geográfica*. Madrid: AGE-CSIC.
- Sapena, M., & Ruiz, L. (2015, May 11). ANALYSIS OF URBAN DEVELOPMENT BY MEANS OF MULTI-TEMPORAL FRAGMENTATION METRICS FROM LULC DATA. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7, 1411-1418.
- Tomljenovic, I. (2012). GEOBIA methods for LiDAR obtained point clouds. *Ekscentar*, 88-92.
- United Nations. (2014). *World Urbanization Prospects*. United Nations, Department of Economic and Social Affairs. New York: United Nations.
- United Nations. (2015). *THE WORLD BANK*. Retrieved from Urban population: <http://databank.worldbank.org/data/reports.aspx?source=2&type=metadata&series=SP.URB.TOTL.IN.ZS>
- Waske, B., Linden, S. v., Oldenburg, C., & Jakimow, B. (2012). imageRF - A user-oriented implementation for remote sensing image analysis with Random Forests. *Enviromental Modelling & Software*, 192-194.
- Wing, A. (2014, December 19). *FORBES*. Retrieved from Tech 2015: Deep Learning and Machine Intelligence Will Eat the World Roster: <http://www.forbes.com/sites/anthonykosner/2014/12/29/tech-2015-deep-learning-and-machine-intelligence-will-eat-the-world/>

Zhou, W., & Troy, A. (2008). An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing*, 29(11), 3119-3135.