Using Image Classification with Space Syntax Model to Predict Pedestrian Volumes and Vehicular Trip Lengths

Vidit Kundu
Faculty of Planning
Centre for Environmental Planning and Technology
Ahmedabad, India
vidit.kundu.mplan17@cept.ac.in

Sabyasachi Purkayastha
Faculty of Technology
Centre for Environmental Planning and Technology
Ahmedabad, India
sabyasachi.purkayastha.mgeo17@cept.ac.in

Bindi Dave
Faculty of Technology
Centre for Environmental Planning and Technology
Ahmedabad, India
bindi.dave@cept.ac.in

***Abstract* – The current study focusses on utilizing satellite imagery and classification techniques for extracting features using machine learning models and further integrating with space syntax accessibility model for predicting pedestrian and vehicular volumes. The workflow involves satellite imagery processing such as segmentation and image classification using Support Vector Machines and extracting built-up area within the urban areas of Gurugram, Haryana, India. Population density in the census wards is redistributed within the extracted built-up area and effort is made to augment the space syntax model to better predict pedestrian and vehicular volume.**

Index Terms* – *Space Syntax, Image Classification, Support Vector Machine, Image Segmentation, Population density, Geographic Information System, Network analysis.

I. Introduction

Satellite imagery can be used to delineate built and non-built/ sparsely built regions within a city. The boundary of urban and rural can also be arrived at using the same. However, extracting exact boundaries of the built up area footprint within a dense urban fabric has not been a concern of most projects. Here, we deal with a use case where this delineation is going to help in augmenting other urban models. Our effort here is to explore and utilize techniques for image classification of satellite imagery to compensate for the lack of data within Indian cities while also looking for ways to advance existing urban accessibility models which have proven to be highly relevant for movement predictions within a city. We, here, focus particularly at Space Syntax model and how we can make predictions about pedestrian and vehicular volumes with the benefit of better distribution of population density. The context chosen is the city of Gurugram, Haryana.

II. Space Syntax and Its Application

Space Syntax is a suite of modeling tools and simulation techniques used to analyze pedestrian movement and to predict pedestrian and vehicular trip volumes [1](Noah Raford 2003).Space Syntax model analyzes layout and connectivity of urban street grids to generate “movement potentials”. A key assumption of this approach is that it assumes even population distribution across a network, and all-point-to-all-point travel throughout the pedestrian grid. Space Syntax defines various parameters. These include integration, choice, total depth, connectivity etc., which are calculated at various radiuses corresponding to movement at that scale. For instance, parameters calculated at radius ‘n’ would correspond to movement happening all the way across the city, and similarly parameters calculated at radius of ‘800 metres’ would correspond to movement at that scale. The latter is generally associated with the pedestrian movement.

Out of these parameters Integration and Choice (at different scales) have been empirically proven to have direct correlation with pedestrian and vehicular movement density [2] (Lerman 2014). Integration is an estimate of how accessible a point is and has been further defined to be an intuitive factor. The formula for calculating the integration of a network node is found in Equation 1,

$$I = 2 (MD - 1) k -2$$

Where MD equals the mean depth of the entire system, and k equals the number of nodes within the system. This formula compares an ideally connected graph (one where each point connected to every other point) with the properties of the graph in question to determine a measure of accessibility for each node or intersection. Integration is derived from this value for each node in the system.

On the other hand, Choice is an estimate of how frequented a route is. It is considered to be an emergent factor which takes form over a period of time and is found to correlate with the commercial distribution pattern in a setting. It basically calculates the potentials for each segment element to be selected by pedestrians as the shortest path (when considering a small radius) or selected by drivers (when considering a large radius) or both. So choice signifies the through-movement potential of a segment in a spatial system. The calculation is defined according to Turner as the following, “for all pairs of possible origin and destination locations, shortest path routes from one to other are constructed. Whenever a node is passed through on a path from origin to destination, its choice value is incremented.”[3] (Turner 2001)

However, even though accessibility dictates a large part of movement within the city, the assumption of an even population distribution across the network may prove to be skewed at times. Population within Indian cities is categorized by highly varying densities based on the kind of income group residing within a particular region. In such a scenario, it may be useful to incorporate population density to augment predictions made by space syntax model.

III. IMAGE CLASSIFICATION FOR EXTRACTING BUILT UP AREA

Learning efficient image representation is at the core of scene classification tasks in remote sensing. However, due to a lack of data availability [4] (Martin Längkvist 2016), machine learning models have not been a concern for most projects. With the release of large amounts of both satellite (Sentinel and Landsat) and air borne images, it is possible to make use of machine learning techniques in classification tasks. Machine learning, is being intensively used for extracting features such as buildings and roads. The methodology is termed as feature extraction In the following study we make use of Support Vector Machine as our classification method which is a particular class of machine learning and has been very successful with feature extraction from visual imagery [5](Foody & Mathur, 2004). The aim is to arrive at a built area footprint to integrate with Space Syntax predictions of “movement potentials”.

SVM is characterised by an efficient hyperplane searching technique that uses minimum training area and therefore is efficient with respect to processing time. SVMs can avoid over fitting problem and requires no assumption on data type. Although non-parametric, the method is capable of generating efficient decision boundaries and therefore can minimize misclassification. This is done through mapping of optimal separating hyperplanes between classes by feeding on the training samples (support vectors) that lie at the edge of the class distributions, with the other training cases being excluded. This study aims to carry out SVM classification with and without segmentation on land covers over Gurugram, Haryana, India for built area extraction. Here; we make use of a combination of techniques which include Super-Pixel Segmentation and Support Vector Machine Multiclass classifier to extract built area footprints from Google static satellite imagery. Our use case employs Support Vector Machine Algorithm on both segmented and unsegmented images and compared for best outcomes. The super pixel segmentation creates clusters for similar pixels, and saves on processing time during image classification. The results are validated using the Kappa index, overall Accuracy, Producer’s and User’s accuracy.

The above methodology puts forward a few constraints. There can be errors associated with smaller building footprints due to the spatial resolution of the imagery used [6](Körner 2016). Commercially available high resolution data sets like World View 3 from Digital Globe can solve this problem and accuracy can increase. Secondly, the algorithm works well on multispectral data which has a lot of information within the different bands but due to lack of high resolution open source multispectral data, we have resorted to Google static satellite images and hence our accuracy will be compromised up to a certain extent.

IV. METHODOLOGY – IMAGE CLASSIFICATION

First of all the image is collected from Google server through the Google Static Image API. The satellite image generated is further georeferenced and segmented using Simple Linear Iterative Clustering. The segmentation algorithm generates super-pixel clusters based on spatial and spectral characteristics of the pixels. We used a minimum of 100 pixels for generating each segment. The minimum number of pixels is usually dependent on the resolution of image available and the feature to be extracted. Since we are extracting built area from a high resolution satellite image, it is more logical to go for small pixel grouping sizes so as to minimize mixing of adjacent features.

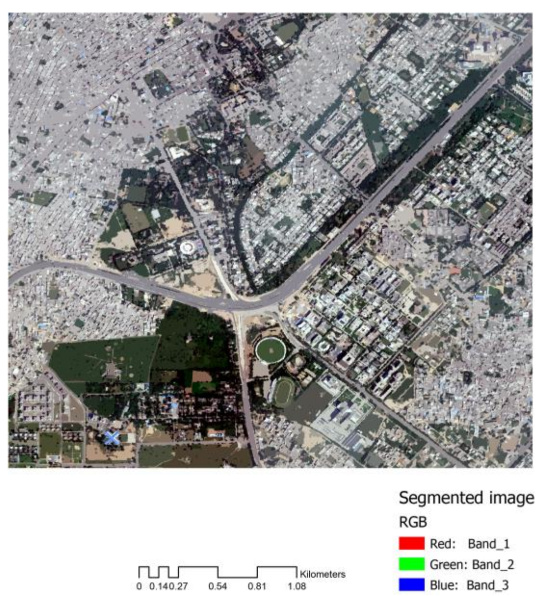

Fig. 1 Segmented image

Thereafter, we prepared two images for classification- segmented and original image.

It was done to analyse the impact of segmentation on built area extraction. The next steps involve training sample generation, training the classifier and finally classifying the image using Support Vector Machines. The above procedures were carried out in ArcGIS Pro proprietary software with an education license.

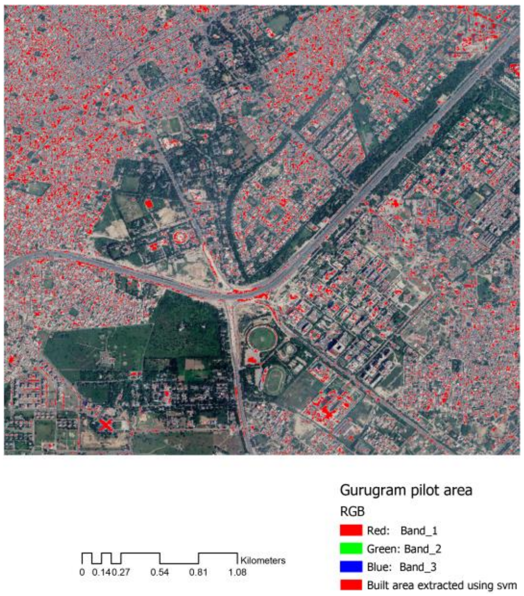

Fig. 2 Built area extracted by SVM

The accuracy of classification was interpreted using pixel samples which were collected for each class maintaining uniform distribution and reducing bias. The number of samples collected for each class was given by-

Pi_{tot} = (No. of classes) X 10

Pi_{tot} = 5 X 10 = 50

Where Pi_{tot} is the total number of sample collected for each class. Since, there were five distinct classes identified using SVM, the total number of samples collected for the confusion matrix was 250.

The confusion matrix was generated for the following cases cases-

1. Classification with segmentation
2. Classification without segmentation

	Without Segmentation	With Segmentation	Selection
Overall Accuracy	76.37	67.578	Without Segmentation
Kappa Coefficient	0.7	0.59	Without Segmentation
Built-Area error commission	27.78	47.37	Without Segmentation
Built-Area error omission	22	20	With Segmentation
Built-Area Producers acc.	78	80	With Segmentation
Built-Area Users acc.	74.07	52.632	Without Segmentation
Time(in seconds)	463	66	With Segmentation

Fig. 3 Table comparing the performance of classification

V. METHODOLOGY – INCORPORATING SYNTAX VARIABLES

The street network was digitized by drawing street – centre lines in GIS and was subsequently brought to UCL Depthmap. Depthmap reads the network grid as an axial map. However, evidence favours conversion of this network into segments and then running a segment analysis to obtain accessibility measures [7]. Choice and Integration values were calculated for different radius. These were weighted by the length of the line segments. Along with this, the network was also weighted with the location of Transit nodes (Metro stations) to better explain pedestrian volumes.

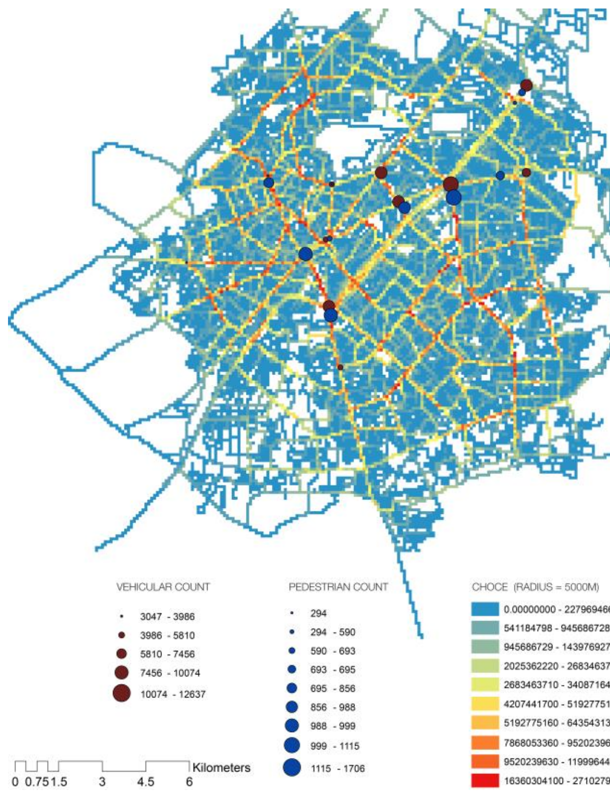

Fig. 4 Street Choice (Radius 5000m) on100m Cells

Ward-wise population density was obtained from 2011 Census data. For the sample of extracted built-up area, the population density was recalculated considering the redistribution of population within the extracted built pockets instead of the entire ward.

Pedestrain and vehicular count at 10 prominent locations was obtained from the “Integrated Mobility Plan for the Gurgaon Manesar Urban Complex” prepared by the Department of Town and Country Planning (DTCP), Government of Haryana.

A cell size of 100m by 100m grid was considered appropriate for further analysis. All layers were sampled on this grid for finding out the correlation between volumes observed and accessibility measures and population density.

VI. RESULTS

Correlation with the pedestrian count was found to be extremely high for accessibility measures including Integration ($R^2 = 0.8$, $p < 0.001$) for radius 4000m and 5000m, and Choice ($R^2 = 0.76$, $p < 0.001$) for radius 5000m and 6000m. The correlation was much higher for the network weighted by the location of transit nodes (Fig.5). This could be due to the fact that metro is used by a large number of people commuting to and from work (which is the case because of large number of people travelling from New Delhi to work in Gurugram). The correlation with population density was found to be relatively low. However, the results were slightly better with the redistributed population density in the sample region.

Correlation of accessibility measures with the vehicular count was also found to be high. Here also, the correlation was much higher for the network weighted by the location of transit nodes. This time the vehicular count coincided with the maximum radius n and 20,000m ($R^2 = 0.86$, $p < 0.001$). This highlights the fact that majority of traffic in Gurugram is due to long distance trips reaffirming the fact that it is a satellite city where people travel from neighbouring city of New Delhi to and from work . Here also the population density did not correlate well.

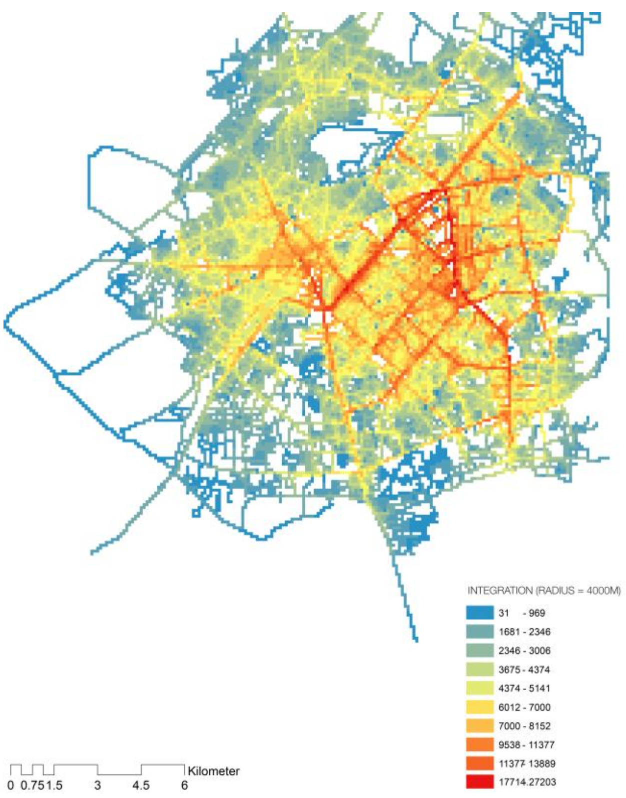

Fig. 5 Integration (Weighted with Transit Nodes) at 4,000m radius

VII. CONCLUSION

We find that in the current context, Space Syntax parameters have been significantly successful in predicting both pedestrian and vehicular volumes. The initial hypothesis was that population density parameter would prove to be an equally important role in predicting volumes as was established by previous studies [1] (Noah Raford 2003). In our context, however, population density, both ward-wise and redistributed, had nominal correlation with pedestrian and vehicular volumes. This could be due to the fact that pedestrian and vehicular movement within Gurugram is more dependent on global trips and most thorough-fare can be explained by long distance work – home trips than locally generated trips. It would be interesting to see in future research the impact of redistributed population density (through image classification) in a context where population density already plays a significant role in predicting movement volumes.

ACKNOWLEDGEMENT

We would like to thank the Faculty of Technology at CEPT University for providing us with adequate resources, and technical mentorship for writing the paper.

REFERENCES

[1] Noah Raford, David R. Ragland. 2003. "Space Syntax: An Innovative Pedestrian Volume Modeling Tool for Pedestrian safety." *U.C. Berkeley Traffic Safety Center, Institute of transport studies.*

[2] Lerman, Yoav & Rofé, Yodan & Omer, Itzhak. 2014. "Using Space Syntax to Model Pedestrian Movement in Urban Transportation Planning." *ResearchGate.*

[3] Turner, Alasdair. 2001. "Space syntax based agent models." London: Researchgate.

[4] Martin Längkvist, Andrey Kiselev, Marjan Alirezaie, Amy Loutfi. 2016. "Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks." *mdpi (Multidisciplinary Digital Publishing Institute)* 1-2.

[5] Foody, G.M. and Mathur, A. (2004) A Relative Evaluation of Multiclass Image Classification by Support Vector Machines. IEEE Transactions on Geoscience and Remote Sensing, 42, 1335-1343.

[6] Körner, L. Liebel and M. 2016. "single-image super resolution for multispectral remote sensing data using convolutional neural networks." *International Society of Photogrammetry and Remote Sensing* 2-3.

[7] Turner, Alasdair. from axial to road center lines: A new representation for space syntax and a new model of route choice for transport network analysis