A Holistic Framework for Addressing the World using Machine Learning

Ilke Demir, Forest Hughes, Aman Raj, Kaunil Dhruv,
Suryanarayana Murthy Muddala, Sanyam Garg, Barrett Doo
Facebook

Ramesh Raskar
MIT Media Lab

Abstract

Millions of people are disconnected from basic services due to lack of adequate addressing. We propose an automatic generative algorithm to create street addresses from satellite imagery. Our addressing scheme is coherent with the street topology, linear and hierarchical to follow human perception, and universal to be used as a unified geocoding system. Our algorithm starts with extracting road segments using deep learning and partitions the road network into regions. Then regions, streets, and address cells are named using proximity computations. We also extend our addressing scheme to cover inaccessible areas, to be flexible for changes, and to lead as a pioneer for a unified geodatabase.

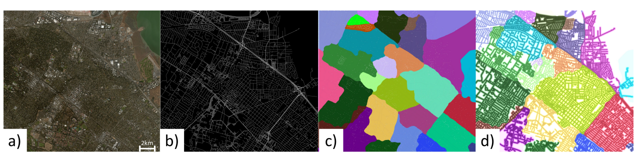

Figure 1. **Street Addresses** starts with (a) satellite imagery, predicts (b) roads, creates (c) regions, to obtain (d) addresses.

1. Introduction

Street addresses enhance precise physical presence and effectively increase the connectivity all around the world. Currently 75% of the roads in the world are not mapped, and this number is increasing in developing countries. To solve this universal problem, we propose a generative addressing system to bridge the gap between grid-based digital addressing schemes and traditional street addresses. Merging the two extents, we designed a linear, hierarchical, and intuitive addressing scheme that follows a set of properties. To automatically generate such street addresses, our system learns regions, roads, and blocks from satellite images using deep learning and graph partitioning. We also prototyped the holistic framework of the generative system supporting forward and inverse geoqueries. We compared our generated maps to existing commercial and open maps for already fully mapped areas for validating hierarchical labeling and unmapped areas for increased map coverage. We also compared travel times using old and new addresses to evaluate the intuitiveness and utility of our system.

2. Related Work

The geocoding process involves converting latitude and longitude information into a unique code. A quick investigation among popular geocoding solutions reveals that such codes are either not in human-readable form (e.g., PlusCodes, OkHi) or they tend to de-correlate from the true topological information (e.g. What3Words, Zippr, MapTags). On the other hand, automatic generation of maps is extensively studied in the urban procedural modeling world [2, 5], creating detailed and structurally realistic models.

Following the example-based generation idea, another approach is to learn from already existing data resources [4]. In our approach, to provide scalability across countries and terrains, we modified state-of-the-art image segmentation neural networks for road extraction. Processing the road topology has also been studied as an example case for clustering and graph partitioning approaches [6]. In addition to the problem being NP-hard, the underconstrained definition of regions adds another layer of complexity, thus we suggested our own partitioner in Section 4.2.

3. The Address Format

Naturally occurring addresses around the world are usually shaped by cultural dynamics, politics, and other long-term processes. We want to mimic this organic process, while still maintaining a unified representation independent of human factors. We enumerated some properties for the ideal address format. Semantic properties emphasize user-friendly features, implying linear, hierarchical, universal, and memorable addresses. Structural properties enable the format of street addresses to be computer friendly, necessitating linear, hierarchical, compressible, robust, extendible, and queryable codes. Following these design principles (Figure 2), the last field indicates version, the fourth field contains the country and state information when applicable, preceded by the city information in the third field. The second field contains the road name, which starts with the region label, followed by the road number. Lastly, the first field is composed of the meter marker along the road and the block letter from the road, animating the house number and apartment number consecutively.

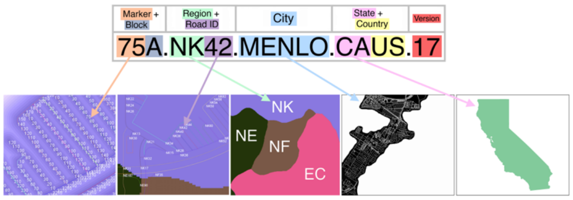

Figure 2. **Our Address Format** is composed of country, state, city, road, and house number of the cell; followed by the version.

4. The Holistic Addressing Framework

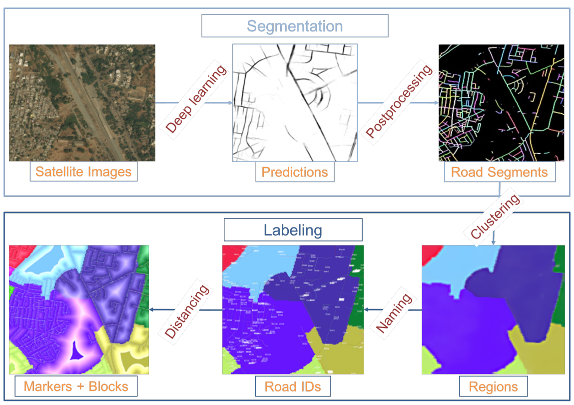

Figure 3. **System Pipeline.** Segmentation extracts roads, breaks them into road segments and clusters them into regions. Labeling names the regions and road segments, places markers, and assigns block letters to individual addressable units.

4.1. Road Extraction

The first step of our approach creates binary road prediction images from three channel satellite images. Our model is trained on satellite images of 0.5m resolution and of size $19K * 19K$, provided by DigitalGlobe. We use a modified version of SegNet [1], where the last soft-max layer is changed from the multi-class structure to have binary classes for road detection, by substituting it with a convolutional layer. We experimented with other architectures (Figure 4), like VGG , U-Net, and ResNet variations, however we achieved the best result with SegNet model resulting in 72.6% precision and 57.2% recall. We also experimented with DeepLab variations and achieved 75.4% precision and 75.9% recall (Figure 4h). After we obtain the road predictions, we threshold, binarize, and apply orientation-bucketing on the road pixels to obtain road segments.

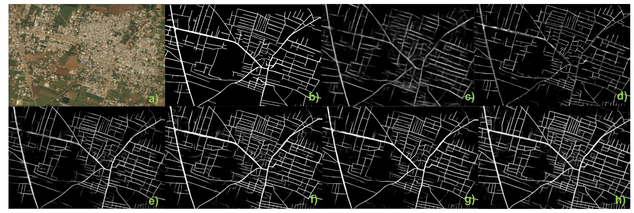

Figure 4. **Road Prediction Models.** An example (a) satellite image and (b) ground truth. Road predictions using (c) VGG, (d) U-Net, (e) ResNet50, (f) ResNet101, (g) SegNet, and (h) DeepLab.

4.2. Region Creation

We convert the road segments into a road graph where the nodes correspond to intersections (and end points) and edges correspond to streets. We weight the edges based on the segment distance and partition the road graph into communities with maximum inter- and minimum intra-connectivity. We experimented with normalized min-cut [7], Newman-Girvan, and optimal modularity based graph partitioning approaches (Figure 5), and concluded [7] efficiently produced results closest to real-world. Graph-based approaches was also better at segmenting from natural cuts (rivers, bridges) over pixel-based.

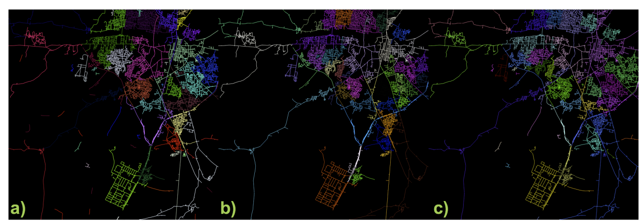

Figure 5. **Region Creation.** Experiments with (a) normalized min-cut, (b) Newman-Girvan, (c) modularity-based partitioning.

4.3. Address Labeling

We name the densest region *'CA'* for the city center. We divide all other regions based on their midpoint: *N(orth), S(outh), W(est),* and *E(ast)*. Next, we name the adjacent regions in each bucket, based on their distance from the city center. The roads in each region is ordered with two main directions of the roads, odd for north-south bound, and even for east-west bound. On each road segment, we place a virtual meter marker in every $5m$. We also compute a discretized distance field of the roads by $5m$, assign a letter to each, concluding the address generation.

5. Inaccessible Areas

Putting streets in focus creates a drawback for geocoding structures that are considerably further from any street. To cover areas that are not accessible by streets using a hierarchical hashing of spatial coordinates, we decide on the alphanumeric format with 26 letters and 10 digits. We use two levels of hashing, into 1km x 1km cell of three letters, and into 30m x 30m address cells represented by five letters – first level hash (3), second level hash (1) and the direction. Expressing it in our geocoding formulation, it becomes $f(C,(lat,lon)) = H(R(lat,2)) + H(lat - R(lat,2)) + dir(lat).H(R(lon,2)) + H(lon - R(lon,2)) + dir(lon).C$, expressed as $L_{lat}L_{lat}H_{lat}D_{lat}.L_{lon}L_{lon}H_{lon}D_{lon}.C$ where H is the hash function, and R is round.

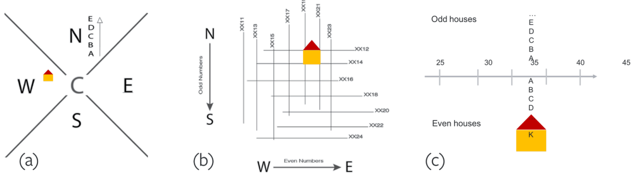

Figure 6. **Labeling System.** (a) Regions based on orientation and distance, (b) roads based on direction and order, (c) markers and blocks based on proximity. The yellow house lives at *38K WB14*.

6. Results and Applications

Our system is written in python and C++, uses Chainer, networkx, and sci-kit libraries. The results are exported in .osm format in addition to the on-the-fly computation of address cells in our prototype. We processed more than 10 cities, totaling up to more than $16K\ km^2$. The source code to convert .osm files and geotiffs to street addresses is available on our repository[1].

We evaluated our framework in a traditional US city that is already well-mapped, accomplishing to extract 95% of the roads. However, keeping the motivation of providing street addresses to the approximately 4 billion unconnected people, our results shine for unmapped developing countries as shown in Figure 7 on 3 different cities. We accomplished to automatically address more than 80% of the populated areas, which significantly improves map coverage.

We validated the usefulness of our generative maps with some user experience by comparing the travel times using the old and new addressing schemes. Overall travel times decreased by 21.7% with a 52.4 seconds improvement on the average. We used population density data to evaluate how our algorithm reflects density criteria. We evaluated our road accuracy versus manually marked roads. We designed a universal *Place Name Server* to store all addresses, including prototype web and mobile map applications that utilize PNS. For more results we refer you to our article[3].

7. Conclusions

Overall, we presented a generative system that can be applied to any given area producing a complete addressing solution. Generated street labels improve map coverage, physically connect people to services, as well as help provide immediate aid in disaster zones. More discussions and extensions about missing city boundaries, overflowing regions, 3D additions, and theoretical analysis, follow up in [3]. We would also like to acknowledge our whole team listed in the same article[3]. In future, we would like to scale up and enable large entities as cities, states, and countries to adopt our framework.

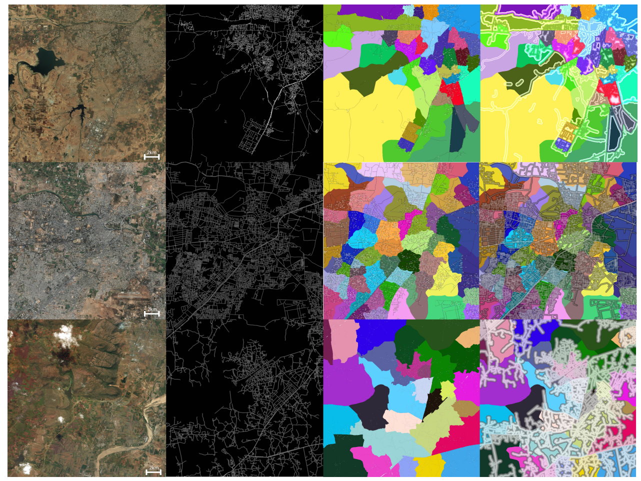

Figure 7. **Street Addresses in Developing Countries**. Satellite image, extracted roads, labeled regions and roads, and meter markers and blocks of three unmapped cities.

[1] https://github.com/facebookresearch/street-addresses

References

[1] V. Badrinarayanan, A. Kendall, and R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *CoRR*, abs/1511.00561, 2015.

[2] G. Chen, G. Esch, P. Wonka, P. Müller, and E. Zhang. Interactive procedural street modeling. *ACM Trans. Graph.*, 27(3):103:1–103:10, Aug. 2008.

[3] I. Demir, F. Hughes, A. Raj, K. Dhruv, S. M. Muddala, S. Garg, B. Doo, and R. Raskar. Generative street addresses from satellite imagery. *ISPRS International Journal of Geo-Information*, 7(3), 2018.

[4] G. Mattyus, W. Luo, and R. Urtasun. Deeproadmapper: Extracting road topology from aerial images. In *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[5] C. A. Vanegas, T. Kelly, B. Weber, J. Halatsch, D. G. Aliaga, and P. Muller. Procedural generation of parcels in urban modeling. *Comput. Graph. Forum*, 31(2pt3):681–690, May 2012.

[6] J. D. Wegner, J. A. Montoya-Zegarra, and K. Schindler. A higher-order crf model for road network extraction. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1698–1705, June 2013.

[7] S. X. Yu and J. Shi. Multiclass spectral clustering. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ICCV '03, pages 313–, Washington, DC, USA, 2003. IEEE Computer Society.