PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Road network mapping from aerial images

Jain, Suchit, Mittal, Rohan, Mishra, Prakamya, Sinha, Aakash

Suchit Jain, Rohan Mittal, Prakamya Mishra, Aakash Sinha, "Road network mapping from aerial images," Proc. SPIE 11139, Applications of Machine Learning, 1113917 (6 September 2019); doi: 10.1117/12.2529005

SPIE.

Event: SPIE Optical Engineering + Applications, 2019, San Diego, California, United States

Road network mapping from aerial images

Suchit Jain, Rohan Mittal, Prakamya Mishra, and Aakash Sinha

Shiv Nadar University, Greater Noida, India

ABSTRACT

Building and expansion of an efficient transportation network are essential for urban city advancement. However, tracking road development in an area is not an easy task as city planners do not always have access to credible information. A road network mapping framework is proposed which uses a random forest model for pixel-wise road segmentation. Road detection is followed by computer vision post-processing steps including Connected Component Analysis (CCA) and Hough Lines method for network extraction from high-resolution aerial images. The custom dataset used consists of images collected from an urban settlement in India.

Keywords: Road network mapping, Random forest, Aerial images

1. INTRODUCTION

Aerial image analysis and road network mapping have applications in fields such as urban development, forest planning, disaster relief and climate monitoring. Emergence of high resolution imagery from drones has led to an explosion in the availability of aerial images. Despite such ubiquity of data, most of the road mapping is still done manually by humans. Road mapping is important for urban planning and automotive navigation. As changes in global road maps occur on a daily bases, they have to be kept up to date, which when done manually can be expensive, time consuming and error prone. In this paper we propose a framework that can tackle this problem automatically and reliably.

Any road network extraction method can be broadly classified into manual, semi-automatic or fully automatic. Most of the work in road network mapping comprises of a multi-stage approaches. One of the first paper in this domain[1] found roads (tracks) in gray-scaled noisy images by growing the road tree iteratively. Another paper[2] uses an automatic road seeding method based on rectangular approximations of road footprints and a toe-finding algorithm to classify footprints for growing a road tree. Here[3] it is proposed that road detection step is based on shape classification of a local homogeneous region around a pixel. This[4] and many other recent papers use CNNs and FCNs to detect roads in high resolution aerial orthomosaics using manually labeled data. Deep learning has proven to be a powerful tool for remote-sensing.

Many ground-air hybrid approaches use expensive sensors mounted on top of cars. These approaches do produce results with high accuracy but this type of solution is often very expensive and has a narrow coverage. Here[5], an approach is proposed that directly estimates road topology from aerial images and also solves the problem of expensive sensors and narrow coverage. CNNs are again used here to segment images into the categories of interest.

The main challenge in using CNN for image segmentation tasks is the requirement of a huge amount of training data. Annotating such a massive dataset is very time consuming and expensive task. In this paper we propose a framework for road network mapping which uses random forest model for road segmentation. The model requires very little training data and is computationally efficient to train.

1.1 Image acquisition and Dataset

Our custom dataset consists of aerial images of the city of Lucknow in Uttar Pradesh, India (26.8467° N, 80.9462° E). We sampled a total of 999 images, each having a resolution of 3648×5472. As part of the pre-processing step the images were resized to a resolution of 320×480 to allow for efficient computation. All the images were acquired using a DJI Phantom 4 Pro (FC6310) drone. Fig. 1 shows some examples of the acquired images used for training.

Further author information: (Send correspondence to Suchit Jain)
Suchit Jain: E-mail: sj535@snu.edu.in, Telephone: +91 9810606539
Rohan Mittal: E-mail: rm677@snu.edu.in, Telephone: +91 9873091520

Applications of Machine Learning, edited by Michael E. Zelinski, Tarek M. Taha, Jonathan Howe, Abdul A. S. Awwal, Khan M. Iftekharuddin, Proc. of SPIE Vol. 11139, 1113917
 · doi: 10.1117/12.2529005

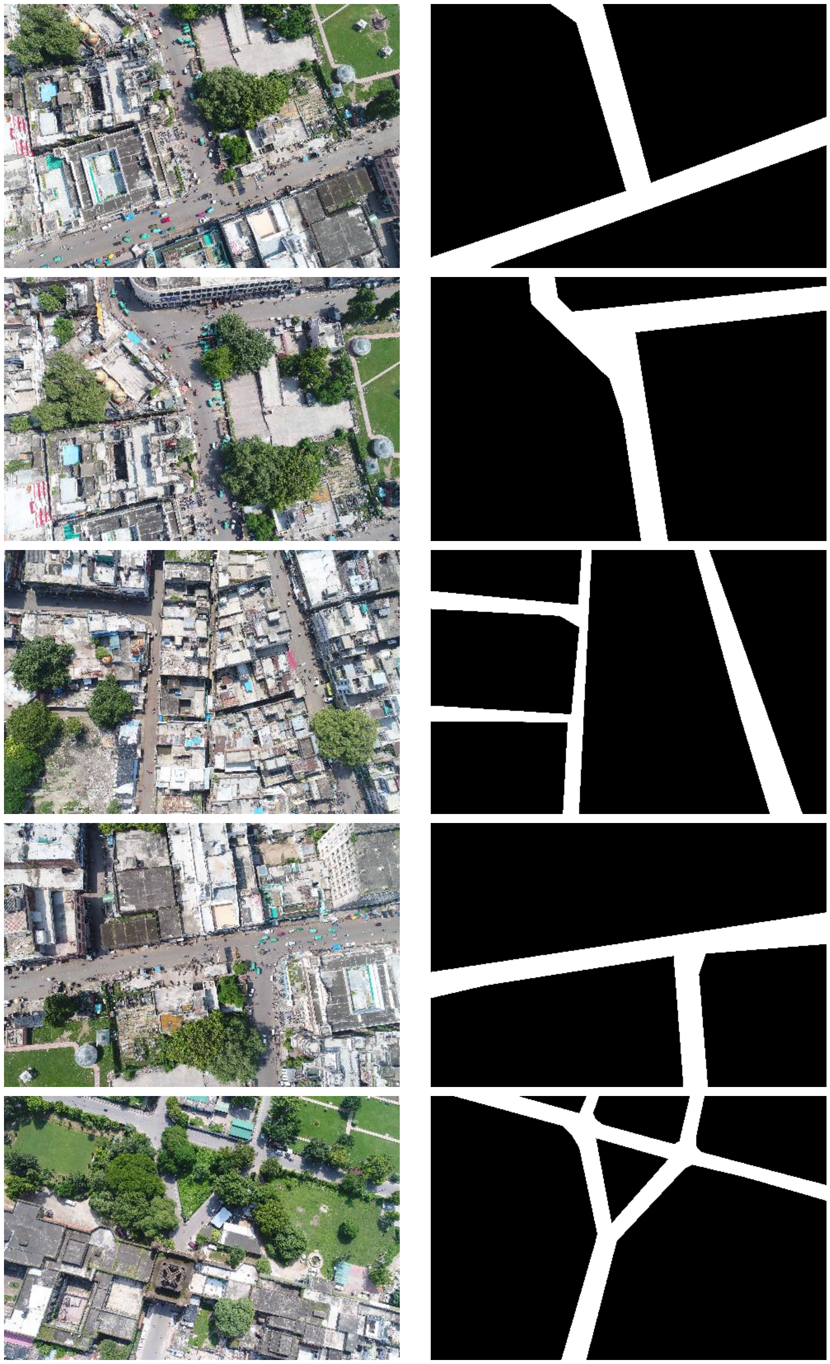

Figure 1. Training images and their ground-truths (best view in color) **Original train image (left)**, **Train image ground-truth masks (right)**

1.2 Random Forest for pixel classification

Random Forest[6] is an ensemble learning method, used for classification as well as regression tasks, which consists of individual decision trees. The final prediction of a random forest is the mode of the predictions of the individual trees. The high accuracy and low error rate of the random forest model can be attributed to the low correlation between constituent trees and the high classification strength of each individual tree in the forest. Low correlation and variation between constituent trees primarily results from Bagging, which is the process of training each tree on a random sample of the dataset (chosen with replacement) and Feature Randomness. which allows for having a random subset (of a fixed length) of all features to chose from while splitting a node. In this study, we have used the random forest model for classifying each pixel in a given image as road or not road. We used (r,g,b) pixel values of the 8 adjacent pixels (up, left, down, right, upper left, upper right, bottom left and bottom right) to a given pixel as attributes for classifying that particular pixel. Thus we used a total of 24 attributes for this classification task. In our model we have used Gini Impurity as the criteria for choosing the best split at each node. Gini Impurity is the probability that a randomly chosen element of a dataset would be classified incorrectly if it were randomly classified according to the class distribution of the dataset. Thus the gain (Impurity before spilt - Impurity after split) quantifies the quality of a split and the attribute which maximises this gain is chosen as the splitting attribute. Fig. 2 shows the raw inference images (column 1).

1.3 Connected Component Analysis (CCA)

Connected component labelling (or analysis) is the process by which image pixels are grouped into components based on pixel intensity and connectivity. Hence after the application of this algorithm we obtain different groups of pixels with pixels in a particular group/component sharing the same intensity and connected to each other in some way. The algorithm[7] works in two passes. In the first pass the algorithm scans each pixel in the image row - wise from left to right and assigns a label to each pixel based on the labels of the pixels above and to the left of it. In the second pass the algorithm identifies connected regions with multiple labels (which may have formed after pass 1) and assigns the parent label to all the pixels in the region. In this study we have used Connected Component Analysis to remove small patches of false positives that result after the application of random forest algorithm for road classification. All the patches/components that are smaller in pixel number than some pre-defined threshold are discarded and hence we obtain a significantly better result. Refer to Fig. 2 for the CCA post-processed inference images (column 2).

1.4 Hough line transform

After performing the post processing step (connected component analysis) on the road segmentation output, we use the *HoughLines* method provided in the *OpenCV library*[8] for marking the road network. It is desirable to perform an edge detection pre-processing step before applying the hough transform. We used Canny edge detector[9] for the edge detection. The hough lines method uses polar coordinates (r, θ) for representing a line. For any given point in the image $(x0, y0)$ we can obtain a graph $(r\ vs\ \theta)$ of all the lines passing through the point. If we obtain similar graphs for all points in the image, the intersection points between these graphs indicate lines passing through the respective points. Thus we set a pre-defined threshold and state that if the number of graphs intersecting at a point (r', θ') exceeds the threshold then there exists a line with parameters r' and θ' in the original image. We applied the hough lines method on individual road patches obtained after road segmentation. Using this method we are able to obtain reasonable road network mapping from the road segmentation output. Refer to Fig. 3 for the hough lines post-processed inference (column 2).

2. EXPERIMENTS

2.1 Experimental setup

We randomly sampled 25 images from our original dataset of 1000 images for training the random forest model. We then selected 75 images from the remaining images for testing. Ground-truths for all images were generated by pixel wise annotation using Photoshop (refer to the ground-truth images in Fig. 1). We adjusted the values of two hyper-parameters for tuning the random forest model. The number of trees (estimators) in the forest was set to 50 and the maximum number of features to be considered at each node for splitting was set to 9. The pixel area threshold for patch rejection in CCA was set to 90. The threshold for minimum number of points on a line for hough lines method was set to 10. We used *accuracy*, *precision* and *recall* in order to evaluate the model performance.

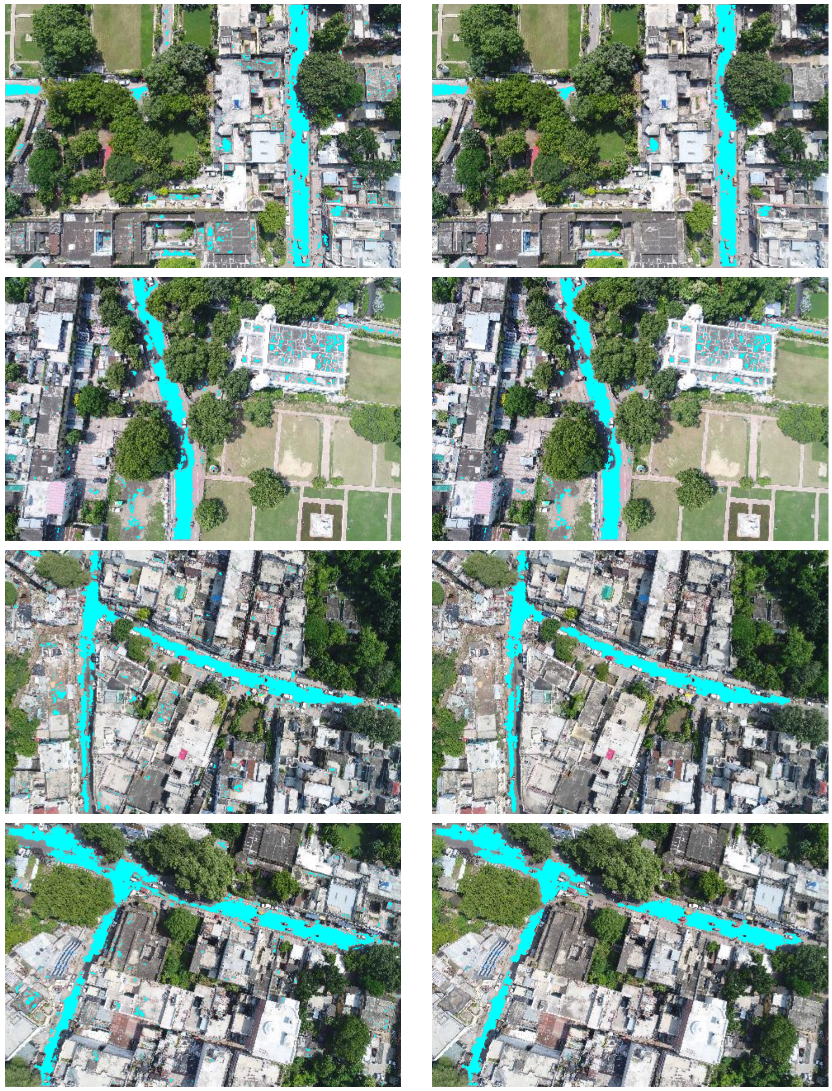

Figure 2. Raw output images and the outputs after performing connected component analysis post processing step (best view in color) **Raw output (left)**, **CCA post-processed output (right)**

2.2 Results and discussion

We present the model evaluation results in Tab. 1. Even though we obtained a good accuracy, we observed that the precision of our model suffers due to the presence of a high number of false positives. In many cases the model incorrectly predicts parts of rooftops as road. This most certainly is due to the model only working with a local context, with features being values of adjacent pixels, as opposed to a global one. This is also the reason behind a relatively high number of false negatives. The model is not able to identify the presence of road beneath vehicles and other structures.

Table 1. Evaluation Metrics

Accuracy	0.92
Precision	0.71
Recall	0.84

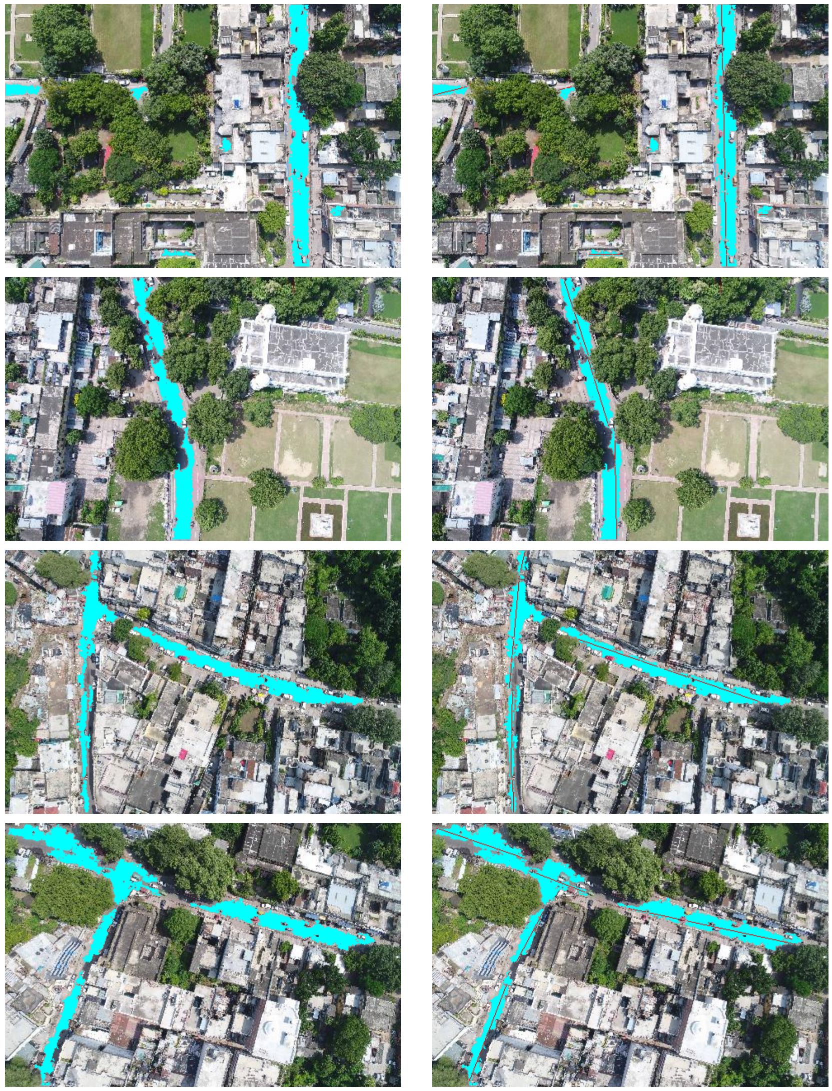

Figure 3. CCA post-processed images and the output after hough line transform post-processing step (best view in color) **CCA post-processed image (left)**, **Hough line post-processed image (right)**

3. CONCLUSION

In this study we implemented a framework for road network mapping from aerial images. We implemented a random forest model for road detection followed by post processing steps involving connected component analysis and hough line transform for extracting the road network. Though we were able to obtain some good results, our work has much room for improvement. An interesting direction for future work would be to experiment with a wider feature set for our model and use operations such as morphological thinning for obtaining a more accurate road network.

REFERENCES

[1] Bajcsy, R. and Tavakoli, M., "Computer recognition of roads from satellite pictures," *IEEE Transactions on Systems, Man, and Cybernetics* **6**, 623–637 (1976).

[2] Hu, J., Razdan, A., Femiani, J., Cui, M., and Wonka, P., "Road network extraction and intersection detection from aerial images by tracking road footprints," *IEEE Transactions on Geoscience and Remote Sensing* **45**, 4144–4157 (12 2007).

[3] Xie, S. and Tu, Z., "Holistically-nested edge detection," *CoRR* **abs/1504.06375** (2015).

[4] Cira, C.-I., Alcarria, R., Manso-Callejo, M.-., and Serradilla, F., "A deep convolutional neural network to detect the existence of geospatial elements in high-resolution aerial imagery," *Proceedings* **19**(1) (2019).

[5] Mattyus, G., Wang, S., Fidler, S., and Urtasun, R., "Enhancing road maps by parsing aerial images around the world," in [*Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*], *ICCV '15*, 1689–1697, IEEE Computer Society, Washington, DC, USA (2015).

[6] Breiman, L., "Random forests," *Mach. Learn.* **45**, 5–32 (Oct. 2001).

[7] Rosenfeld, A. and Pfaltz, J. L., "Sequential operations in digital picture processing," *J. ACM* **13**, 471–494 (Oct. 1966).

[8] Bradski, G., "The OpenCV Library," *Dr. Dobb's Journal of Software Tools* (2000).

[9] Canny, J., "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**, 679–698 (Nov 1986).