MAPPING ELECTRIC TRANSMISSION LINE INFRASTRUCTURE FROM AERIAL IMAGERY WITH DEEP LEARNING

Wei Hu[1], Ben Alexander[2,3], Wendell Cathcart[4], Atsushi Hu[5], Varun Nair[3,5], Lin Zuo[2,3], Jordan Malof[6], Leslie Collins[6], Kyle Bradbury[1]

[1]Energy Initiative, [2]Department of Statistical Science, [3]Department of Computer Science,
[4]Nicholas School of the Environment, [5]Department of Mathematics,
[6]Department of Electrical and Computer Engineering
Duke University, Durham, NC 27701, USA

IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium | 978-1-7281-6374-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/IGARSS39084.2020.9323851

ABSTRACT

Access to electricity positively correlates with many beneficial socioeconomic outcomes in the developing world including improvements in education, health, and poverty. Efficient planning for electricity access requires information on the location of existing electric transmission and distribution infrastructure; however, the data on existing infrastructure is often unavailable or expensive. We propose a deep learning based method to automatically detect electric transmission infrastructure from aerial imagery and quantify those results with traditional object detection performance metrics. In addition, we explore two challenges to applying these techniques at scale: (1) how models trained on particular geographies generalize to other locations and (2) how the spatial resolution of imagery impacts infrastructure detection accuracy. Our approach results in object detection performance with an F1 score of 0.53 (0.47 precision and 0.60 recall). Using training data that includes more diverse geographies improves performance across the 4 geographies that we examined. Image resolution significantly impacts object detection performance and decreases precipitously as the image resolution decreases.

***Index Terms*—** Electricity infrastructure, power transmission and distribution, aerial image, computer vision, object detection

1. INTRODUCTION

Over 11% of the global population lacks access to electricity. This lack of access is largely concentrated in the developing nations of Sub-Saharan Africa and Southeast Asia [1]. Studies show that electricity access positively correlates with improved economic, educational, and health outcomes [2], so identifying cost-effective pathways to electrification is important. However, this process requires data on existing electricity infrastructure, which is often unavailable or expensive.

One way to fill this information gap without requiring a huge amount of human labor is to automatically extract electricity infrastructure from aerial or satellite imagery. Past research on aerial imagery has focused on diverse topics including the classification of land use [3], building detection [4], and the segmentation of road networks [5].

Two notable studies investigated aspects of mapping transmission and distribution lines in satellite imagery. Development Seed used machine learning to augment human tracing of high voltage transmission lines [6] in Pakistan, Nigeria, and Zambia. While effective, this approach requires substantial human labor. Rohrer, et al. from Facebook Research developed an automated approach to estimate the location of medium voltage distribution lines from indirect indicators of electricity infrastructure, including lights at night data, MODIS landcover data, and OpenStreetMap highways data [7]. Rohrer, et al. use a probabilistic model they name "Pathfinder" that uses a many-to-many variant of the Dijkstra shortest distance algorithm to identify the most likely location of distribution lines. Recently, Arderne, et al. used this "Pathfinder" model to generate a composite map of the global power grid which showed a predictive accuracy of 75% for 15×15 km predictions on a validation set of 14 countries [8]. Performance increased as the size of the evaluation grid cell increased and decreased as the grid cell resolution was decreased. In contrast to past work, we propose applying object detection tools to automatically detect the locations of electricity infrastructure from aerial imagery without a human-in-the-loop and through direct observation of high resolution visible spectrum imagery. We focus on correctly identifying the precise location of the transmission towers (performing validation to ensure detections are within 10 m instead of kilometer-scale evaluations).

Object detection methods are computer vision techniques that identify discrete objects within images. Modern object detection methods are based on convolutional neural networks (CNNs), which enable automatic feature extraction, hierarchical and high-dimensional feature representation, joint optimization with several other tasks [9].

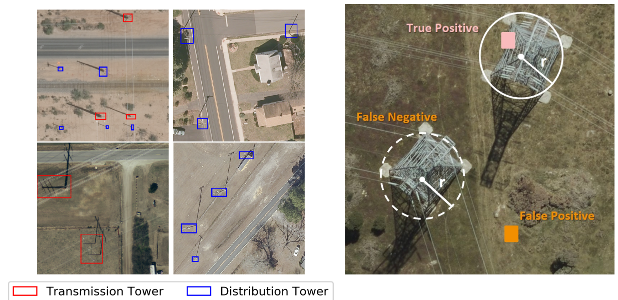
In this paper, we explore three CNN-based object detection architectures to build a model that automatically detects electricity infrastructure in satellite imagery. Furthermore, we examine two specific challenges that arise when applying these techniques at scale. (1) First, we explore the impact of applying object detection techniques across diverse locations around the world. We train transmission infrastructure detection models on four distinct geographic types (deserts, plains, forests, and suburban settings) and evaluate how well each model generalizes to the other geographic types. (2) Second, we investigate how image resolution affects model performance. The higher the resolution of satellite imagery, the greater the cost of the imagery. Intuitively, we expect that higher resolution imagery will allow for better object detection performance. We seek to identify the lowest resolution at which identifying transmission lines may still yield acceptable performance to determine the data requirements of this problem.

2. DATASET DESCRIPTION

Imagery used in this work is obtained from the Electric Transmission and Distribution Infrastructure Imagery Dataset [10]. The dataset contains RGB aerial imagery with annotated transmission and distribution infrastructure. The dataset covers 14 cities and 5 continents, and encapsulates the diversity in human settlement density and terrain type. For experiments in this paper, a subset of the dataset was selected including imagery collected from 4 US states: Arizona (AZ), Connecticut (CT), Kansas (KS), and North Carolina (NC). The 4 selected US states represent 4 diverse geographies, namely deserts, suburbs, plains, and forests, respectively. Annotations associated with the imagery indicate the type of electricity infrastructure (transmission or distribution tower) and image-wise xy coordinates of infrastructure objects. The imagery used in this work has a spatial resolution of 0.15 m per pixel and each image file (or tile) is $10,000 \times 10,000$ pixels. Examples of selected images and associated annotations are shown in Fig. 1a.

3. METHODS

Our work focuses on identifying the transmission and distribution towers/poles in satellite imagery. To identify these objects, we explore three CNN-based object detection models, Faster R-CNN [11], YOLOv2 [12], and RetinaNet [13]. All models are pre-trained on ImageNet and use augmentation techniques recommended by the authors of the corresponding methods; for YOLOv2 augmentation included random crops, rotations, and hue, saturation, and exposure shifts and for RetinaNet random horizontal flips. For Faster R-CNN, as no recommendations were provided by the author we investigated common augmentation techniques and compared their performance, selecting the best performing option: random horizontal flips. All models take images as inputs and output detected objects in the form of bounding box coordinates. Each $10,000 \times 10,000$ pixel image tile in the dataset is cropped into smaller image patches that are 512×512 pixels and then used as the input during training and testing. We use 25% edge overlap for neighboring patches to account for potential information loss at the edges of image patches.



(a) (b)

Fig. 1. (a) Example images and associated annotations from the dataset of transmission and distribution infrastructure. (b) Examples of how true positive, false negative, and false positive predictions are scored.

To score our predictions, we declare an object detected by the object detection model to be a true positive if its location falls within a radius of 10 m (approximately equivalent to 66 pixels for 0.15 m resolution imagery) of the ground truth transmission/distribution tower, otherwise it is declared a false positive. Using this scoring method, and varying the confidence threshold for declaring a detection, we construct precision-recall (PR) curves and calculate F1 scores for evaluating performance. Precision is the ratio of true positives to all detections labeled positive while recall is the proportion of objects correctly detected. An F1 score is the harmonic mean of precision and recall. In these experiments, while there are two separate classes for transmission and distribution towers in the dataset and during training, we combine those into one class during testing to evaluate the algorithm's overall ability to identify transmission and distribution infrastructure.

4. EXPERIMENTS AND RESULTS

4.1. Electricity infrastructure detection

With each of the 3 object detection architectures mentioned in Section 3, we trained a model on 0.15 m resolution aerial imagery from Arizona (AZ), Connecticut (CT), Kansas (KS), and North Carolina (NC). We use a 75-25 image-level train-test split before any further pre-processing is performed, to ensure that the training and test datasets do not overlap. This training and testing strategy also serves as a baseline strategy for the other two experiments in this work.

PR curves of three trained object detection architectures are shown in Fig. 2.RetinaNet achieves the best overall performance with a maximum F1 score of 0.53. Therefore, we use RetinaNet, the best performing architecture on 0.15 m resolution imagery, as the model of choice in the experiments that follow.

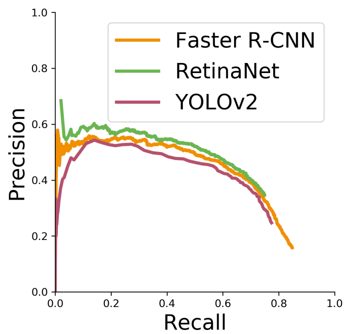


Fig. 2. Transmission/distribution object detection performance trained and tested on 0.15 m resolution imagery.

4.2. Generalizability across diverse geographies

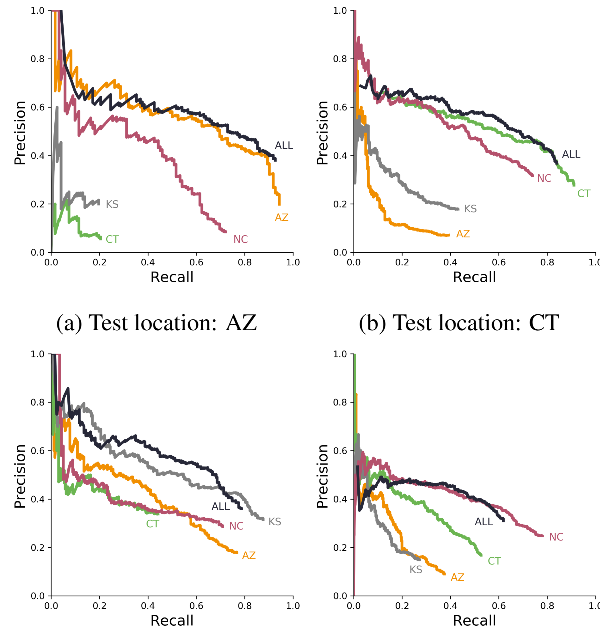
Our method would be most beneficial if it could be applied anywhere in the world, but geographies differ significantly across locations. This experiment explores the impact of training on data from different geographies. To investigate model generalizability across geographies, we first trained the RetinaNet model on only one geography (AZ, CT, KS, NC) then tested on each of those 4 geographies. Then we trained on the union of the training data from each of the four geographies, which we label the "ALL" set and tested on each of the 4 geographies.

Our results (shown in Fig. 3) showed that the model trained on the "ALL" set was generally the best performing one. The local models, whose training and testing data originated from the same region, showed similar performance compared to the union model. Other models whose training location are different from the testing location performed significantly worse than the models whose training data was from the same region as the test data and also worse than the more diverse "ALL" training dataset.

These results indicate that a model's object detection performance is heavily impacted by its training geographies. A model tends to perform well on geographies that it's seen while diverse training geographies may potentially result in even better performance. While additional efforts are always required to acquire labeled imagery for a new geography, using such imagery for fine-tuning is likely to yield better performance on the new geography.

4.3. Impact of imagery resolution

Another challenge in satellite and aerial imagery analysis is that imagery resolution varies from 0.15 m resolution (e.g.



(a) Test location: AZ (b) Test location: CT

(c) Test location: KS (d) Test location: NC

Fig. 3. Transmission/distribution object detection performance of RetinaNet trained and tested across diverse geographical locations. Lines labeled with training geography.

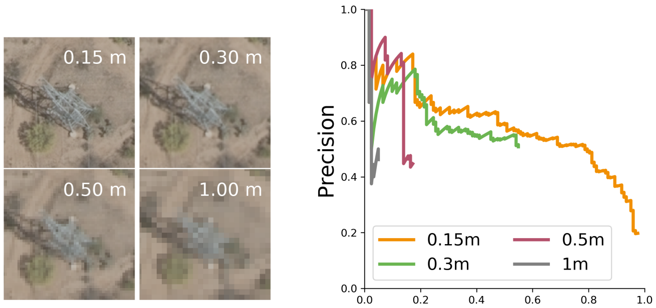
proprietary aerial imagery) to 10 m or lower resolution (e.g. publicly available Sentinel satellite imagery). Higher resolution imagery is richer in details; however, it's generally limited in availability and greater in cost. Therefore, we investigate the impact of spatial resolution on object detection performance.

We downsample the 0.15 m per pixel images in our training and test set from AZ to create 0.3 m, 0.5 m, and 1 m per pixel resolution images (examples shown in Fig. 4a).

The results showed that model performance declined significantly as the resolution decreased. When the resolution decreased to 0.3 m, the highest recall that the model could achieve was around 0.6 while achieving a high precision around 0.8, which was much worse than the model trained on the original resolution but it still showed acceptable performance. For imagery with 0.5 m or coarser resolution, the models showed much lower recall (below 0.2) which means the majority of objects of interest were not detected. These results indicate that a resolution of 0.3m would be required for identifying transmission and distribution towers/poles.

5. CONCLUSIONS

This work introduced a deep learning object detection approach to mapping electricity infrastructure in satellite and aerial imagery from 4 diverse U.S. states. This approach to



(a)

(b)

Fig. 4. (a) Example imagery with resolution varied from 0.15 m to 1.00 m per pixel; (b) Object detection performance of RetinaNet trained and tested on imagery of each resolution

transmission identification is fully-automated without a human in the loop. It relies on the direct observation of infrastructure in overhead imagery data for precise determination of infrastructure locations within 10 m of the true location rather than through lower-resolution proxies such as lights at night data. Of the models we considered, RetinaNet performed the best, achieving a maximum F1 score of 0.53. Geographic representativeness and diversity of the training data with respect to the test data were demonstrated to be vital factors for detection performance: model performance on previously unseen geographies is significantly worse than for either the case of training on data from the same geography or from a mixture of diverse geographies. Lastly, for reasonable infrastructure detection performance, the resolution of the imagery needs to be at least 0.3m to be able to detect half of the transmission and distribution lines. Of course, the connectivity of the lines and their inherent spatial resolution may be able to fill in missed detections, and that will be explored in future work.

6. ACKNOWLEDGEMENTS

Support for this work was provide by the Duke University Bass Connections program; the Duke University Data+ program; and National Science Foundation Grant no. OIA-1937137.

7. REFERENCES

[1] World Bank, "World bank open data," `https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS`, 2017.

[2] Peter Alstone, Dimitry Gershenson, and Daniel M Kammen, "Decentralized energy systems for clean electricity access," *Nature Climate Change*, vol. 5, no. 4, pp. 305–314, 2015.

[3] Grant J Scott, Matthew R England, William A Starms, Richard A Marcum, and Curt H Davis, "Training deep convolutional neural networks for land–cover classification of high-resolution imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 4, pp. 549–553, 2017.

[4] Bohao Huang, Kangkang Lu, Nicolas Audeberr, Andrew Khalel, Yuliya Tarabalka, Jordan Malof, Alexandre Boulch, Bertr Le Saux, Leslie Collins, Kyle Bradbury, et al., "Large-scale semantic classification: outcome of the first year of inria aerial image labeling benchmark," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2018, pp. 6947–6950.

[5] Teerapong Panboonyuen, Kulsawasd Jitkajornwanich, Siam Lawawirojwong, Panu Srestasathiern, and Peerapon Vateekul, "Road segmentation of remotely-sensed images using deep convolutional neural networks with landscape metrics and conditional random fields," *Remote Sensing*, vol. 9, no. 7, pp. 680, 2017.

[6] Development Seed, "Mapping the electricity grid from space," `https://developmentseed.org/projects/hv-grid/`, 2018.

[7] Brandon Rohrer, Sedar Sahin, Dimitry Gershenson, and Anna Lerner, "A new predictive model for more accurate electrical grid mapping," `https://engineering.fb.com/connectivity/electrical-grid-mapping/`, 2019.

[8] Christopher Arderne, Conrad Zorn, Claire Nicolas, and Elco E Koks, "Predictive mapping of the global power system using open data," *Scientific Data*, vol. 7, no. 1, pp. 1–12, 2020.

[9] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu, "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, 2019.

[10] Kyle Bradbury, Qiwei Han, Varun Nair, Tamasha Pathirathna, and Xiaolan You, "Electric transmission and distribution infrastructure imagery dataset," `https://figshare.com/articles/Electric_Transmission_and_Distribution_Infrastructure_Imagery_Dataset/6931088`, 2018.

[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.

[12] Joseph Redmon and Ali Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[13] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.