DeepWind: Weakly Supervised Localization of Wind Turbines in Satellite Imagery

Sharon Zhou*, **Jeremy Irvin***, **Zhecheng Wang, Eva Zhang,**
Jabs Aljubran, Will Deadrick, Ram Rajagopal, Andrew Ng
Stanford University
{sharonz, jirvin16, zhecheng, evazhang, aljubrmj, wdead, ramr, ang}
@stanford.edu

Abstract

Wind energy is being adopted at an unprecedented rate. The locations of wind energy sources, however, are largely undocumented and expensive to curate manually, which significantly impedes their integration into power systems. Towards the goal of mapping global wind energy infrastructure, we develop deep learning models to automatically localize wind turbines in satellite imagery. Using only image-level supervision, we experiment with several different weakly supervised convolutional neural networks to detect the presence and locations of wind turbines. Our best model, which we call DeepWind, achieves an average precision of 0.866 on the test set. DeepWind demonstrates the potential of automated approaches for identifying wind turbine locations using satellite imagery, ultimately assisting with the management and adoption of wind energy worldwide.

1 Introduction

Decarbonization of the electricity sector and decreasing costs are accelerating the adoption of renewable energy, with 50% of the worldwide capacity to be renewable by 2035 (1). Wind energy in particular is growing at a rapid pace, with 591 GW of capacity deployed by 2018—a 60% increase in 5 years—and a 130% increase expected by 2035 (2). Knowing the location and type of installed wind turbines is critical for a variety of stakeholders, such as: (i) wind developers to identify the best new areas for deployments (3); (ii) electricity grid operators to integrate renewable energy, to perform real-time system operation, and to plan capacity expansion (4); (iii) utilities and city planners to plan the management of local demand (5); and (iv) policymakers to design and estimate the impact of incentives and other contracting policies (6). Currently, there are more than 398,000 wind turbines installed worldwide and only 42% are mapped with accurate locations (7; 2). An automated approach to localizing wind turbines would enable efficient identification of their locations and ultimately the construction of a more complete, up-to-date, global database of wind energy infrastructure.

Convolutional neural networks (CNNs) have demonstrated success in a variety of image-based localization tasks, including semantic segmentation (8) and object detection (9; 10). Much of this success, however, has been attributed to the use of fully supervised methods that require large datasets with detailed annotations like bounding boxes and segmentation masks, which can be expensive and time-consuming to collect. Satellite imagery offers an immense data source for image-based localization tasks, but lacks this detailed annotation which is necessary for fully supervised model development. To address this, we leverage advances in weakly supervised methods which use image-level classification labels to localize objects, without the use of detailed annotation. These methods have emerged as an effective alternative to their fully supervised counterparts, enabling advances in various fields, including healthcare (11; 12) and energy (13).

*Equal contribution.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

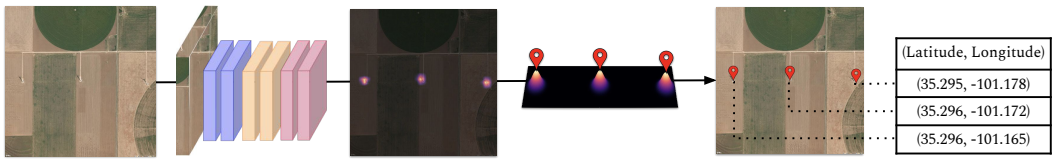

Figure 1: The weakly supervised localization models input a satellite image and output a predicted response map which localizes the turbines in the image. A peak detection module identifies local maxima in the response map and merges them based on proximity. The predicted locations in the image can be converted to the geographic locations (latitude, longitude coordinates) of the turbines.

In this work, we develop DeepWind, a deep learning model to automatically localize wind turbines from satellite imagery. We leverage satellite images annotated with image-level binary classification labels from the U.S. Wind Turbine Database, a manually curated database containing more than 50,000 wind turbine locations in the U.S. (14). Without bounding boxes for fully supervised localization models, we develop a variety of weakly supervised localization models to identify the locations of wind turbines in satellite imagery. We evaluate each model on turbine location prediction, and our best model, DeepWind, achieves an average precision of 0.866 on the test set.

2 Data

The dataset used to develop the models contains 42,255 positive images and 58,672 negative images for a total of 100,927 images. We use latitude and longitude coordinates in the U.S. Wind Turbine Database (USWTDB) for the positives, excluding those that were labeled with low location confidence or did not have corresponding imagery (14). We have two negative sets, one sampled randomly across the U.S. and the other containing visually similar objects to turbines (difficult negatives) with an open source GeoVisual search tool (15). For all inputs, we obtain high resolution satellite imagery ($\approx$ 1.2m) of size 1500×1500. We split the dataset into training (97,825 images, 41,182 positives, 3,392 difficult negatives), validation (1,427 images, 432 positives, 500 difficult negatives) and test (1,283 images, 249 positives, 500 difficult negatives). We do not match the true prevalence of wind turbines in the real world as this would require very large dataset sizes to obtain a sufficient number of positive examples. To evaluate the localization performance of the models, we manually annotate all of the turbines in the validation and test sets with bounding boxes to be used as ground truth.

3 Models

Task The wind turbine detection task is a weakly supervised localization task, where the input is a satellite image and the output is a set of coordinates indicating the locations of all turbines (if any) in the image. Only image-level binary labels are available during training.

Localization Methods We experiment with several state-of-the-art weakly supervised localization methods, including Class Activation Mappings (CAM (16)), Improved Gradient-weighted Class Activation Mappings (GradCAM++ (17)), Soft Proposal Networks (SPN (18)), WILDCAT (19), and Peak Response Maps (PRM (20)). Each method produces a single *response map* for an input image. Hyperparameters specific to each method are set to the best values described in the original work.

Training Procedure We use an ImageNet-pretrained DenseNet-121 (21) backbone architecture for all localization models. We randomly crop[2] each 1500×1500 image to 500×500, ensuring that the turbine still remains in the image, and normalize images based on ImageNet mean and standard deviation. We use an unweighted binary cross entropy loss to train all models for 5 epochs with a batch size of 4, using an Adam optimizer with standard parameters and a fixed learning rate of 0.001.

Inference Procedure During inference, the 1500×1500 image is tiled into nine 500×500 images. The network outputs a probability for each input image using a sigmoid nonlinearity, and the

[2]This cropping is performed for data augmentation and for GPU memory constraints.

Model	Localization		Counting	
	AP	F1	MAE	RMSE
CAM	0.738	0.799	0.242	0.744
Grad-CAM++	0.830	0.848	0.211	0.672
SPN	0.875	0.911	0.156	0.573
WILDCAT	**0.919**	0.941	**0.095**	**0.396**
PRM	0.916	**0.945**	0.101	0.487

Table 1: Localization and counting performances of all models on the validation set.

classification of the input image is determined by the threshold at which the model achieved the highest classification F1-score on the validation set. The models additionally produce a 15×15 response map for each of the 500×500 tiles.

To obtain the predicted wind turbine locations for the 1500×1500 image, we first upscale the response map of every input tile to 500×500 using bilinear interpolation, and then normalize the upscaled response map so that all values lie between 0 and 1. Each of the response maps are then concatenated to produce a full 1500×1500 response map over the whole image. The response map is used to produce the turbine locations through *peak detection*, consisting of the following steps. We first use a local max finding algorithm (22) to identify the predicted turbine locations and their confidence values, determined using the values of the response map at each location. Then, we merge any location predictions that occur within the same region defined by the upsampling ratio between the 15×15 and 500×500 the response maps under bilinear interpolation (34 pixels). The merging process consists of identifying the connected components in the response map and then computing the center of mass within each connected component. Finally, the predicted locations for a 500×500 tile are dropped the tile is classified as negative. These predicted locations can be converted to latitude and longitude coordinates (see Figure 1 for a simplified visual explanation).

4 Results

We evaluate each of the weakly supervised model variants on the validation set. We measure localization performance using an extension of the pointing with prediction metric (18) to the multiple predicted object locations setting. We also compute a precision-recall curve, but instead of checking whether the maximum response lies within any of the ground truth bounding boxes, we match each predicted response to the bounding boxes as done in standard object detection. We use average precision (AP) to summarize the localization performance of each model, and report the F1-score at the threshold which led to the highest F1-score on the validation set.

Counting wind turbines is also an important task for many downstream analyses, e.g. computing aggregate wind power out from a region. To measure this, we use standard metrics from crowd counting literature (23; 24), namely mean absolute error (MAE) and root mean squared error (RMSE) between the number of ground truth bounding boxes and number of predicted locations at the optimal threshold.

WILDCAT and PRM demonstrate the best localization performance on the validation set (Table 1). CAM demonstrated the lowest localization performance followed by Grad-CAM++ and SPN. We identify WILDCAT as the best performing weakly supervised model, outperforming the others on localization and counting, and evaluate this model—DeepWind—on the test set. On localization, DeepWind achieves an AP of 0.866 and an F1-score of 0.919 (precision=0.920, recall=0.917). On counting, DeepWind achieves an MAE of 0.076 and an RMSE of 0.306.

5 Conclusion

We develop deep learning models to localize wind turbines in satellite imagery using a large dataset of over 100,000 images. We explore a variety of weakly supervised localization techniques and find that the best model, which we call DeepWind, achieves an average precision of 0.866 on the test set with image-level supervision alone. In future work, we plan to construct a global database of wind turbine locations and make it freely accessible to the public. We hope that our work can encourage the development and adoption of wind energy and demonstrate the potential for machine learning to tackle challenging problems relevant to climate change (25).

References

[1] H. Kheradmand and J. A. Blanco, *Climate Change: Socioeconomic Effects*. BoD–Books on Demand, 2011.

[2] K. Ohlenforst, S. Sawyer, A. Dutton, B. Backwell, R. Fiestas, J. Lee, L. Qiao, F. Zhao, and N. Balachandran, "Global wind report - annual market update 2018," *Global Wind Energy Council (GWEC), Brussels, Belgium, Report*, 2018.

[3] P. Jain, *Wind energy engineering*. New York: McGraw-Hill,, 2011.

[4] L. Xie, P. M. Carvalho, L. A. Ferreira, J. Liu, B. H. Krogh, N. Popli, and M. D. Ilic, "Wind integration in power systems: Operational challenges and possible solutions," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 214–232, 2010.

[5] J. C. Smith, M. R. Milligan, E. A. DeMeo, and B. Parsons, "Utility wind integration and operating impact state of the art," *IEEE transactions on power systems*, vol. 22, no. 3, pp. 900–908, 2007.

[6] R. H. Wiser, A. Mills, J. Seel, T. Levin, and A. Botterud, "Impacts of variable renewable energy on bulk power system assets, pricing, and costs," tech. rep., Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2017.

[7] L. Byers, J. Friedrich, R. Hennig, A. Kressig, X. Li, C. McCormick, and L. M. Valeri, "A global database of power plants," *World Resour. Inst*, vol. 18, 2018.

[8] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431–3440, 2015.

[9] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in neural information processing systems*, pp. 2553–2561, 2013.

[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91–99, 2015.

[11] Y. Xu, J.-Y. Zhu, I. Eric, C. Chang, M. Lai, and Z. Tu, "Weakly supervised histopathology cancer image segmentation and classification," *Medical image analysis*, vol. 18, no. 3, pp. 591–604, 2014.

[12] R. Zhao, W. Liao, B. Zou, Z. Chen, and S. Li, "Weakly-supervised simultaneous evidence identification and segmentation for automated glaucoma diagnosis," in *Thirty-Third AAAI Conference on Artificial Intelligence*, 2019.

[13] J. Yu, Z. Wang, A. Majumdar, and R. Rajagopal, "Deepsolar: A machine learning framework to efficiently construct a solar deployment database in the united states," *Joule*, vol. 2, no. 12, pp. 2605–2617, 2018.

[14] J. E. Diffendorfer, L. A. Kramer, Z. H. Ancona, and C. P. Garrity, "Onshore industrial wind turbine locations for the united states up to march 2014," *Scientific data*, vol. 2, p. 150060, 2015.

[15] R. Keisler, S. W. Skillman, S. Gonnabathula, J. Poehnelt, X. Rudelis, and M. S. Warren, "Visual search over billions of aerial and satellite images," *Computer Vision and Image Understanding*, 2019.

[16] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2921–2929, 2016.

[17] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 839–847, IEEE, 2018.

[18] Y. Zhu, Y. Zhou, Q. Ye, Q. Qiu, and J. Jiao, "Soft proposal networks for weakly supervised object localization," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1841–1850, 2017.

[19] T. Durand, T. Mordan, N. Thome, and M. Cord, "Wildcat: Weakly supervised learning of deep convnets for image classification, pointwise localization and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 642–651, 2017.

[20] Y. Zhou, Y. Zhu, Q. Ye, Q. Qiu, and J. Jiao, "Weakly supervised instance segmentation using class peak response," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3791–3800, 2018.

[21] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[22] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," *PeerJ*, vol. 2, p. e453, 2014.

[23] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column convolutional neural network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 589–597, 2016.

[24] D. B. Sam, S. Surya, and R. V. Babu, "Switching convolutional neural network for crowd counting," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4031–4039, IEEE, 2017.

[25] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown, *et al.*, "Tackling climate change with machine learning," *arXiv preprint arXiv:1906.05433*, 2019.