PGRID: Power Grid Reconstruction in Informal Developments Using High-Resolution Aerial Imagery

Simone Fobi Nsutezo[1]*, Amrita Gupta[1], Duncan Kebut [2], Seema Iyer [3], Luana Marotti[1], Rahul Dodhia[1], Juan M. Lavista Ferres[1], Anthony Ortiz[1]

[1]Microsoft AI for Good Research Lab, [2] HOTOSM, [3]USA for UNHCR

1. Supplementary Material

1.1. Electrical pole detection illustration

We deploy a Fully Convolutional Network (FCN8) semantic segmentation model to detect poles within the image. Figure 1 illustrated the point supervision method used to train the pole detection model. A mask with points is compared to the predicted blobs using the 4 component loss, to obtain a supervisory signal for training. The pole detection model is trained with the Adam optimizer at a small learning rate of 1e-6. Standard data augmentation techniques (rotations, vertical and horizontal flips, color jitter) are applied to regularize learning.

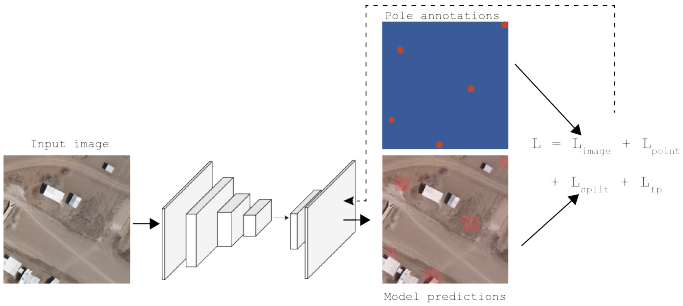

Figure 1. Illustration of the pole detection workflow, including the 4 types of losses: Image level loss is an image classification loss (L_{image}) measuring the accuracy in classifying animage as containing a pole or not, a point level loss (L_{point}) for localization of poles within the image, a split level loss (L_{split}) to ensure unique poles are obtained and a false positive loss (L_{fp}) to minimize false positive detections.

1.2. Electrical line segmentation illustration

For the line segmentation task, we deploy an asymmetric DeepLabV3 model for patch-wise segmentation. We apply a scaling factor on the ground truth mask, to create a patch-wise mask of line presence or absence. After obtaining patch-wise predictions, the predictions are resampled using bi-linear interpolation to obtain a prediction mask the size of the input image. Figure 2 illustrates the line segmentation workflow used for training. The line segmentation model is trained with the Adam optimizer at a small learning rate of 1e-5. Standard data augmentation techniques (rotations, vertical and horizontal flips, color jitter) are also applied to regularize learning.

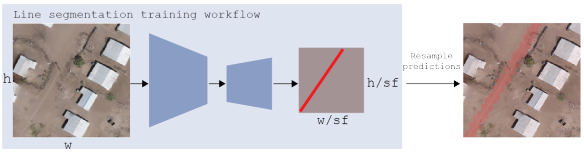

Figure 2. Illustration of the asymmetric line segmentation encoder-decoder network that outputs patch-wise predictions. The downsampling factor (*sf*) determines the patch sizes for predictions. Output masks are then resampled using bi-linear interpolation to obtain a prediction mask the size of the input image.

1.3. Sensitivity to scaling factor.

To perform patch-wise line segmentation, we experiment with three scaling factors: 1, 4 and 8. *sf* of 1 leaves the ground truth label mask size unchanged, a *sf* of 4 reduces the ground truth label raster to a quarter of its original size, while a *sf* of 8 reduces the ground truth label raster to an eighth of its original size. We observe that, an *sf* of 4 yields the best localization as measured by mIOU and detection as measured by the F1-score as shown in Table 1.

1.4. Selecting a distance threshold (*th*).

To understand the impact of the distance threshold (*th*) when evaluating the pole detection model in the Turkana Intergrated Settlement, we randomly sample 100 poles from the test set and measure the shadow lengths cast by the poles. Figure 3 shows the distribution of pole shadow

*Corresponding author: sfobinsutezo@microsoft.com

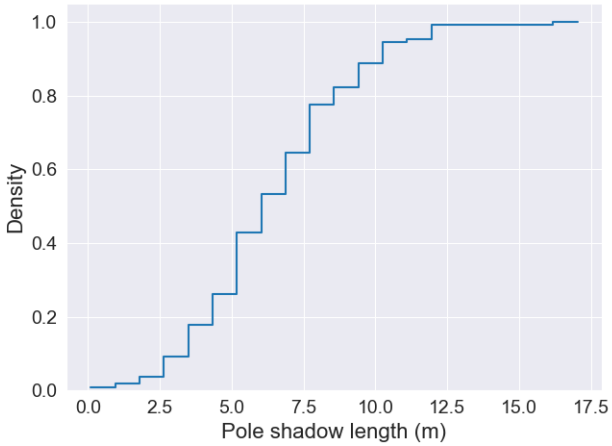

Figure 3. Distribution of pole shadow lengths, given a random samples of 100 poles from the test area.

lengths for the 100 sampled poles. We observe that less than 30 % of poles cast a shadow less than 5 meters while 90% cast a shadow of 10 meters or less. This observation is important because the model outputs blobs as predictions for pole locations. The centroids of the blob are then used as predicted pole locations. For poles with longer shadows and by consequence larger blobs, their centroids occur further away from the ground truth pole point location. Thus a correctly detected pole, might be classified as a false positive if the chosen distance threshold is too small. Understanding the distribution of pole shadows is also of importance so that unreasonably large distance thresholds are not selected. At very large thresholds, almost all poles would be reported as detected but these detections could be faulty as the predicted poles can occur very far from the ground truth point. By measuring the distribution of lengths cast by pole shadows in the dataset, a reasonable threshold can be selected to evaluate how well the model is performing and drive further model improvements. It is worth noting that electrical poles do not typically occur in dense clusters as they are almost evenly spaced out to support lines over a geographic area. This constraint in the physical placement of electrical poles reduces the likelihood of mismatches between predictions to ground truth pole, especially at a distance threshold less than 10m.

Table 1. Line segmentation model performance in Kalobeyei Camp as a function of the scaling factor (*sf*). The *sf* of 4 produces the best localization and detection results as measured by the mIOU and F1-scores. At this *sf*, the original label mask is reduced to 1/4 of its size, and the model classifies every 4x4 patch within the image as containing a line or not.

	mIOU			**F1-score**		
sf	K1	K2	K3	K1	K2	K3
1	0.68	0.61	0.68	0.82	0.76	0.81
4	**0.70**	**0.63**	0.67	0.82	**0.77**	0.81
8	0.69	0.62	0.68	0.82	0.77	0.81