Creating Artificial Worlds with AI to Improve Access to Energy Data

Introduction
Worldwide, about 1.2 billion people still don’t have access to electricity in their homes. This map from the World Bank in 2017 highlights the places that are still greatly lacking in electricity. Access to electricity is becoming increasingly critical, especially for promoting economic development, social equity, and improving quality of life. For example, it has been shown that electricity access is correlated with improvements in income, education, maternal mortality, and gender equality.

One of the first steps in improving energy access is acquiring comprehensive data on the existing energy infrastructure in an area. This includes information on the type, quality, and location. This information is key for helping make decisions about where to prioritize development and more. However, this critical information for expanding energy accessibility is often unavailable or of low quality.

Methods & Data
A solution to efficiently filling this data gap is to automate the process of mapping energy infrastructure in satellite imagery. Using a deep learning object detection model, we can input satellite imagery to the model and make predictions about the characteristics and contents of the energy structures in the image.

However, properly training the object detection model requires significant amounts of already labelled images and in situations where the available training data is from a different geographical region than the region to which the model is applied, the variations in physical characteristics negatively affect the model's accuracy as the models ability to generalize across these different geographies is poor.

Our proposed solution to this is to generate synthetic images that supplements the original satellite imagery training dataset. We generate these synthetic images by cropping the existing energy infrastructure out of satellite images and placing them on top of a different real image (serves as a background) that originally features no energy infrastructure from one of the target geographic domains. By doing so, we create a larger dataset to train on. To blend the cropped energy infrastructure and the background images together, we use an image harmonization technique, specifically GP-GAN (2), to blend the images together. GP-GAN is a Gaussian-Peron Process Generative Adversarial Network that works to blend high-resolution images.

We created our augmented synthetic dataset by generating 78 images blended by GP-GAN and mixing it with 70 real images. We selected this ratio of real to synthetic images based on previous experimentation that aimed to find the ratio of images that optimized performance (4).

Then, we test our augmented synthetic dataset against the original baseline dataset and against our previous Baseline Connections Team's augmented synthetic dataset - which was created by placing cropped 3D models of wind turbines from CityEngine over background imagery. CityEngine is a 3D modeling software that contains assets of various infrastructure, including wind turbines.

References & Acknowledgments

Thank you to Wayne Wu, Dr. Kyle Bradford, Dr. Jordan Malo, and the previous Bass Connections team!