Data Science Case Study

Automating Utility Pole Recognition & Inspection with Computer Vision

Mosaic designed and deployed custom computer vision models to automate asset recognition & inform inspection decisions.

Industry

Utility

Use Case

Improving and Automating Asset Inspection

Techniques

Adoption & Scale, Classification & Computer Vision

Outcome

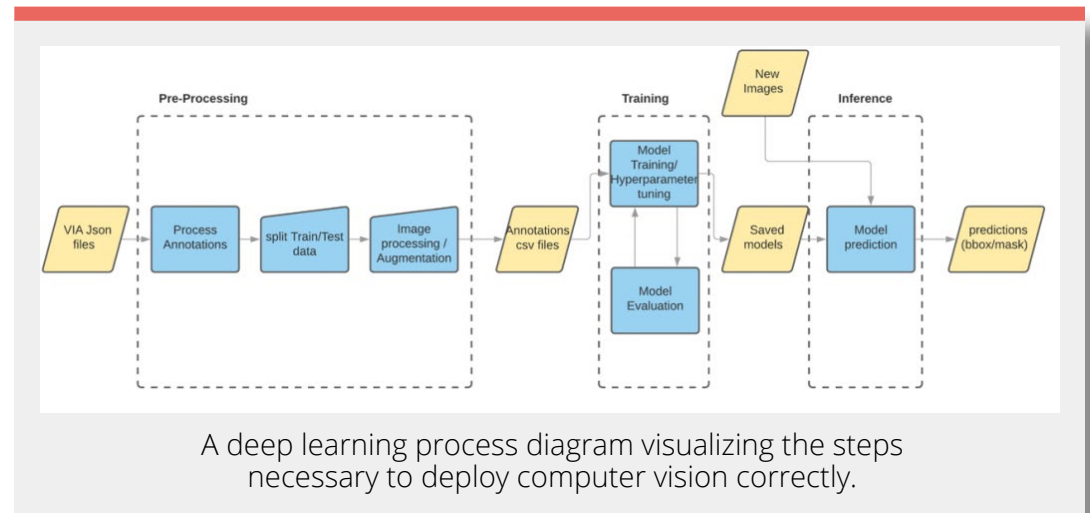Substantial Reduction in Operational Costs

COMPUTER VISION INTRO

Computer vision offers significant cost savings by way of scalability and automation. Collecting a large set of image data is only the first step. The true art of any AI application comes down to how a data scientist can tune AI to accurately label what is in the image set. Mosaic Data Science is poised to help companies make these deep learning techniques work for them[1] and, in the following case study, did just that.

NEED FOR COMPUTER VISION

One of the largest utilities in the United States needed an AI consulting partner to deploy computer vision effectively & efficiently. The company uses a fleet of drones to inspect some of their physical assets and had collected a large data set of inspection images. To maximize their return on this investment, the firm needed deep learning models to find and identify different equipment types in the images and, eventually, to help diagnose visible defects to alert their personnel to a maintenance need. What good is a set of images if a machine cannot learn and provide recommendations to operational decision makers? The company approached Mosaic to assist with this effort.

Mosaic, a leading AI consulting company, was tasked to design computer vision models that automatically identify and label various asset types in inspection images. The models would be integrated with an image inspection tool to enable analysts to quickly search for images of specific equipment types such as pole tops, crossarms, insulators, and transformers; automatically catalog the equipment installed on the pole; and potentially flag defects for closer inspection. A human-in-the-loop feedback mechanism would be implemented to allow any annotations that were approved or adjusted by analysts during the review process to be added to the training data to continuously improve model performance. The utility needed this to be completed in a span of 3 months.

A deep learning process diagram visualizing the steps necessary to deploy computer vision correctly.

DATA ANNOTATION AND IMAGE PROCESSING

Obtaining a sufficiently large sample of good quality data – accurately annotated images – is a crucial and often challenging first step in all computer vision projects. In this case, drone-captured images were being stored on Azure blob storage but did not have image annotations. Mosaic worked with the utility customer to guide the image annotation process. After a comparison of a variety of annotation platforms and an analysis of the customer's current and future annotation needs, VGG Image Annotator[2] (VIA) was chosen for the project based on ease of installation/annotation and project management functionality. Annotations could easily be generated and shared between analysts and Mosaic data scientists for collaboration and quality assurance.

Automating Utility Pole Recognition – Any machine learning model is highly dependent on the training data going in, annotating images correctly was critical to model accuracy.

COLLABORATION IS KEY

Customer SMEs worked with Mosaic to define annotation guidelines based on visibility of the asset components (e.g., is a cylinder visible in the transformer?), variability in background, model or sub-type of the asset (e.g., a transformer can look different based on the manufacturer), occlusion (e.g., asset is partially hidden by foreground objects). The training images were annotated using polygons to tightly enclose each object of interest and labeling each polygon with the asset type. The image below is a screen shot from the VIA tool showing polygon annotations of numerous asset instances on a pole.
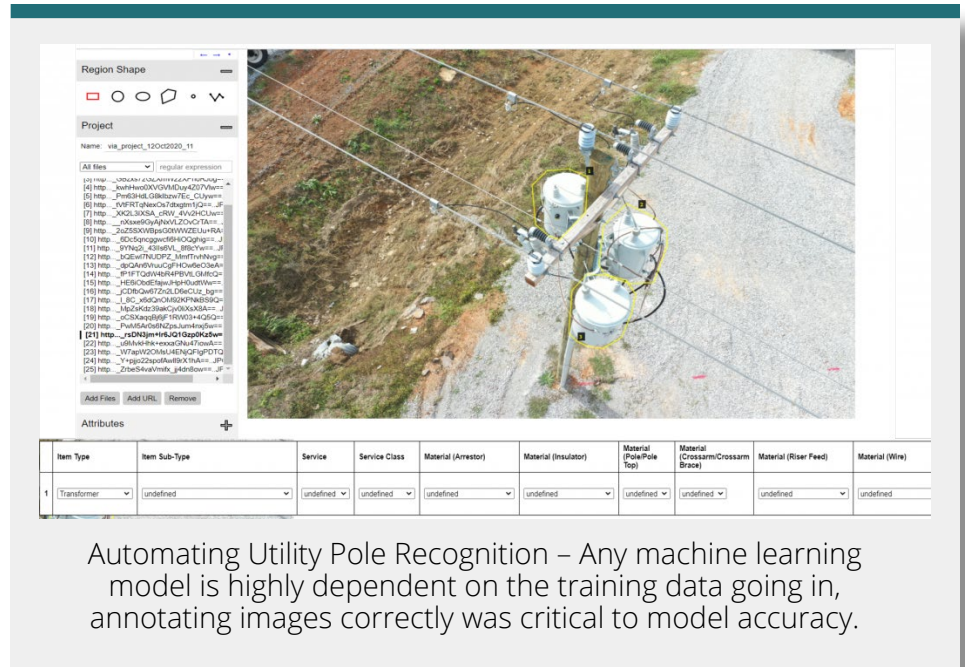
WORKING THROUGH DATA CHALLENGES

Even with pre-defined annotation guidelines, Mosaic's computer vision consultants faced problems in annotation consistency because of multiple people annotating smaller sets of images. There were differences in annotation styles with some more tightly bound polygons than others and some misclassification from variable domain knowledge between annotators. There were also other problems like determining when to exclude an image from the training set if an asset was occluded beyond a defined threshold. Such challenges, if left unaddressed, make it impossible to maximize the performance of a computer vision model. Mosaic worked with the customer in an iterative process to improve the annotations, resulting in a high-quality data set for model training.

The training data for all assets combined consisted of 1000 images, a relatively small training set for a deep neural network. Offline augmentation of the training images was used to increase the training data size by 3-fold and bring robustness to the model. Python libraries like Albumentation and ImgAug were used to create variations in background, orientation, contrast, saturation, crops, and more. Online augmentation using the Keras inbuilt function ImageDataGenerator was also used for this purpose.

DEEP LEARNING DEVELOPMENT

Mosaic developed a separate deep neural network model for each asset type based on the customer's request to make it easier to independently retrain asset models at different times based on the model performance for each asset and the availability of new training data. The models were trained on an Azure NC24 Linux environment which has 4 Tesla K80 GPUs using Keras and TensorFlow deep learning frameworks.

Object detection models, which output a rectangular bounding box and label, were developed for detecting pole tops, transformers, and insulators. An instance segmentation model, which outputs a polygonal mask, rectangular bounding box and label, was developed for crossarms. A range of model architectures including Single-Shot Detection (SSD), Faster Region-based Convolutional Neural Nets (FR-CNN), Retinanet, and Mask-R CNN were explored. All models used transfer learning techniques to initialize the model weights based on source models that were trained on the COCO dataset of 300K+ images. Different backbone network architectures were explored; however, ResNet50[3] had the best performance. Retinanet[4] showed the most promising performance in the object recognition models.

Models were tuned extensively to fit the custom training data and to achieve the best recall rates for the purpose of the application. Recall was prioritized to minimize the likelihood of any visibly defective asset escaping detection. The precision of the bounding boxes and masks was a secondary objective. All metrics were measured using thresholded intersection over Union (IoU) with PASCAL VOC standards.

AI INFORMING OPERATIONAL DECISION MAKING

Once Mosaic had tuned the computer vision models to a level of accuracy the utility was satisfied with, the models needed to be validated on the customer's image sets. Mosaic fed new inspection images into the models to make this AI actionable. In the image below, you can see Mosaic's model identifying different parts of the utility pole.

Automating Utility Pole Recognition – Mosaic specializes in taking deep learning vision techniques out of the lab and into business users' hands

The utility can now feed new images into their inspection platform and have a machine accurately tag different parts of their infrastructure, saving them valuable time and identifying components that might need to be replaced or repaired.

Endnotes

1. https://www.mosaicdatascience.com/computer-vision/

2. http://www.robots.ox.ac.uk/~vgg/software/via/

3. https://keras.io/api/applications/

4. https://keras.io/examples/vision/retinanet/

Want more information?
Click here to learn more