Identifying Electrical Structures Through Image-Detection Using Street View Imagery

Mateo Ignacio mjignacio@ucsd.edu
Kelly Park khpark@ucsd.edu
Brandon Luong baluong@ucsd.edu
Derek Wen d3wen@ucsd.edu

Phi Nguyen pnguyen@sdge.com

Abstract

As climate change escalates the prevalence of wildfires, utility risk management has become an integral aspect of community safety in recent years. Much of the existing utility management systems overlook the data available through public imagery and the corresponding impact of computer vision, and this has led to missed opportunities in timely and accurate hazard identification. Our project aims to address this gap by utilizing Google Street View imagery to detect and identify different utilities out in the open through image detection technologies. In doing so, we have trained an object detection model on image data sourced from Google's API to accurately classify underground and overhead utility structures. By optimizing a pre-trained machine learning model with our dataset, our project sets itself up for further development in detecting more specific hazards and assets in future endeavors. This advancement offers utility companies a novel tool to enhance decision-making processes and mitigate risks more proactively.

Code: https://github.com/Derek-Wen/DSC180A-Q1-FinalProject

1 Introduction

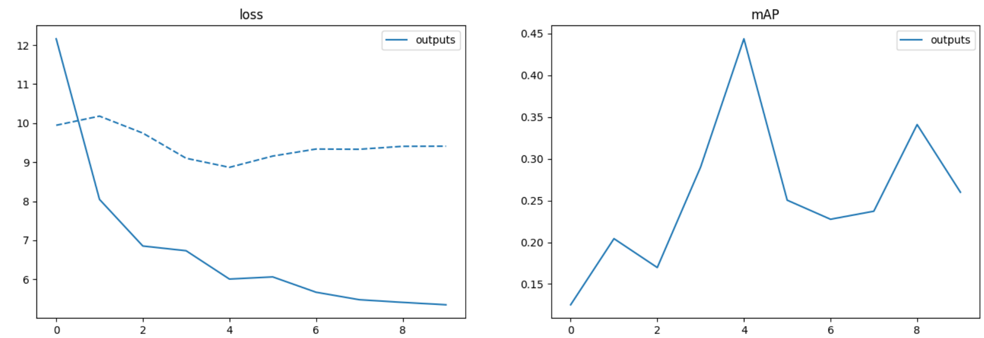
The increasing frequency and intensity of wildfires pose a significant risk to communities and the utility infrastructure that serves them. As climate change exacerbates, the importance of effective utility risk management becomes important for ensuring public safety and preventing catastrophic damage. This paper introduces an approach to utility detection, leveraging the potential of publicly available Google Street View imagery to improve the identification. Historically, utility risk management has relied on traditional methods such as manual inspections and historical data analysis. These methods are resource-intensive and often are too slow to provide real-time risk assessments in light of the cost of inspections and asset population. Recent advances in machine learning and computer vision have opened up new possibilities, allowing for more proactive risk management strategies. Despite these advancements, risk assessment models lack integration with publicly accessible data sources, such as Google Street View. Previous attempts to incorporate such data have been limited by the quality of annotations and the generalizability of the models. In this project, we utilize Google Street View images to detect and classify utility assets, such as poles and transformers, to identify potential hazards. Google Street View API offers visual information about the physical state of utility infrastructure and its surroundings. With this imagery, a dataset was built for our models by taking 15 images of 120 different underground and overhead structures as seen on Google Street View. By analyzing this data through various object detection algorithms, we can replicate and improve the hazard detection capabilities of utility companies.

2 Methods

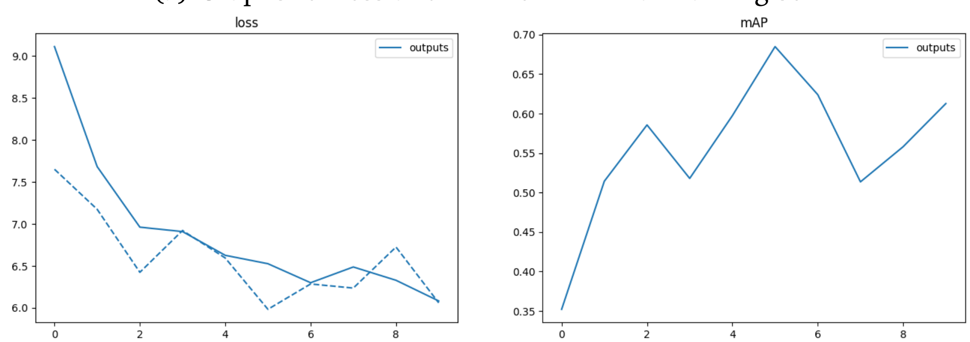
For our data collection, our method leveraged Google Street View imagery as the primary data source. Each team member was responsible for acquiring a set of images, focusing on capturing underground and overhead utility structures. In total, we gathered 150 images from each team member totaling to 1800 images. Once the data set was collected, we proceeded with the annotation and labeling process. We each explored various image labeling tools to find the most suitable program for our needs, such as CVAT. The program allowed us to accurately label the utility structures in our images and was an important step for training our object detection model (Nguyen 2023).
For our model selection, we utilized DETR, a pre-trained object detection model and Facebook's Detection Transformer. We chose this model as it is a direct set prediction model that uses a transformer encoder-decoder architecture to predict all objects at once. This makes it more simple and efficient compared to traditional object detectors such as YOLO, while producing accurate results. In specific, the DETR architecture takes in and processes the input data in the form of a COCO dataset. The three main components of this architecture are a convolutional neural network backbone, a transformer encoder-decoder, and a feed-forward neural network. We then fine-tuned the model using our labeled COCO data set by training the model on our custom data on 10 epochs, and adjusting its parameters and configurations accordingly (Rustamy 2023).

Our final step involved assessing our model's performance in accurately detecting and classifying structures within our validation dataset. For the purposes of observing the effectiveness of our model, we additionally trained the model on sets of 75 images and the combined total. By observing their accuracies against the same validation set and additional performance metrics, this allowed us to measure the effectiveness of utilizing varying sizes of custom data on DETR.

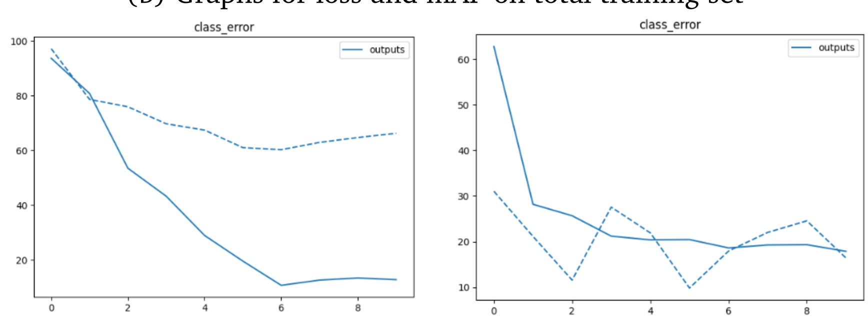
3 Results



(a) Graphs for loss and mAP on minimal training set



(b) Graphs for loss and mAP on total training set



(c) Graphs for class_error for minimal and total training set

(a) Label prediction on minimal training set

(b) Label prediction on total training set

4 Discussion

As shown in first loss graph for DETR performance when trained on a minimal image set, we see the largest decrease to loss in the second epoch. Interestingly, the same metric the more extensively trained model in Figure (b) in largely similar. Loss measures the model's prediction accuracy. This means that if the loss is high then the model is more frequently predicting the wrong class, and we thus we aim to minimize it. Loss can further be likened to class_error, shown in Figure (c) wherein the two models show comparatively different trends. For the minimally trained model, we see a continued drastic decrease and minimum at 6 epochs. For the more extensively trained model, class error takes a sharp drop at epoch 1, and has a more stable decrease, yet even at a minimum, its error is higher than the other model. This may be due to overfitting.

Mean average precision, or mAP, is an advanced metric that leverages multiple other metrics to identify the average performance on the correct true predictions for each class. For the first maP graph in Figure (a), two classes of overhead (OH) and underground (UG) structures and a mAP of around .45 at 4 epochs indicates that DETR is most balanced at predicting the correct class and averse to overfitting around the 4th epoch. For the second mAP visualization in Pigure (b), this peak appears to occur at the 5th epoch at a mAP of approximately 0.68. The metrics display generally well working models given the size of the image sets and considering the numerous other factors that present risks to accurate prediction.

Lastly, the output images for both DETR models display the limitations and capabilities of differently sized image sets with training done on 10 epochs. In the first image of both Figures (a) and (b), while there are no overhead poles within the photo, it can be observed that both models attempted to identify one with moderate confidence levels of 0.79 and 0.72 accordingly. While this displays faulty classification, it can be seen that the more comprehensively trained model placed a bounding box for a pole over a more concrete object that the other. In addition, for the second image, the minimally trained DETR is correct in identifying no overhead structures within the image. This is in contrast to the second model identifying a tree as a pole. As many poles in the training data were wooden poles, the model that was trained on a greater amount of data may have fallen prey to overfitting, where any object resembling a wooden pole will be identified as one.

5 Conclusion

Our project marks a step in the field of computer vision and utility management. By using machine learning models in identifying and classifying utility structures from Google Street View images, we have opened a new usage in how computer vision can be employed for public safety and infrastructure monitoring. The versatility of computer vision demonstrated in our project highlights its potential as a useful tool despite the observed danger of overfitting. This methodology provides a scalable, efficient, and cost-effective approach to utility risk management, particularly in the context of increasing environmental challenges such as wildfires. Looking ahead, there are several promising directions for advancing our current project. Enhancing the model and improving its abilities to handle diverse and challenging imaging conditions, such as lighting, weather, and obstructed views. Expanding the model's capabilities to detect specific types of hazards related to utility structures, such as signs of wear, damage, or how the environment affects the structures. Partnering with other utility companies for a pilot program in order to provide valuable real-world data and feedback.

References

Nguyen. 2023. “StreetWatch.” In *Streetwatch: Utility and Wildfire Risks Detected From Street View Imagery*. GitHub. [Link]

Rustamy. 2023. “DEtection TRansformer (DETR) vs. YOLO for object detection..” [Link]