

A Deep Learning Approach for Detection, Semantic Segmentation and Density Classification of Smoke in Satellite Imagery

Rey Koki ^{1, 2, 3}

Michael McCabe ¹

Jebb Stewart ²

Christina Kumler ^{2,3}

Jed Brown ¹

Wilfrid Schroeder ⁴

¹Computer Science Department, CU Boulder, Boulder, CO

²NOAA's Global Systems Laboratory, Boulder, CO

³Cooperative Institute for Research in Environmental Sciences, CU Boulder, Boulder, CO

⁴NOAA NESDIS, College Park, MD



University of Colorado Boulder

Contributions

1. Created database of labeled satellite imagery using smoke annotations
2. Built deep learning arch to perform detection, classification and segmentation
3. Introduced thermometer encoding for ordinal ranking of smoke density
4. Calculated metrics and analyzed performance

Motivation

Because of NOAA's large quantity of historical, analyst-labeled data, we can deploy deep learning methods that leverage big data in order to create an automated algorithm for localizing smoke in satellite imagery. Our goal is to create a human in the loop algorithm that will lower the required labor.

Smoke Labels

The Hazard Mapping Systems (HMS) team is continuously providing fire and smoke products for North America. Smoke plumes are identified, outlined and classified by their particulate density as either heavy, medium or low (figure 1). The HMS smoke labels are created using a timeseries of imagery that can span up to 5 hours, complicating the data if there are shifts in smoke position.

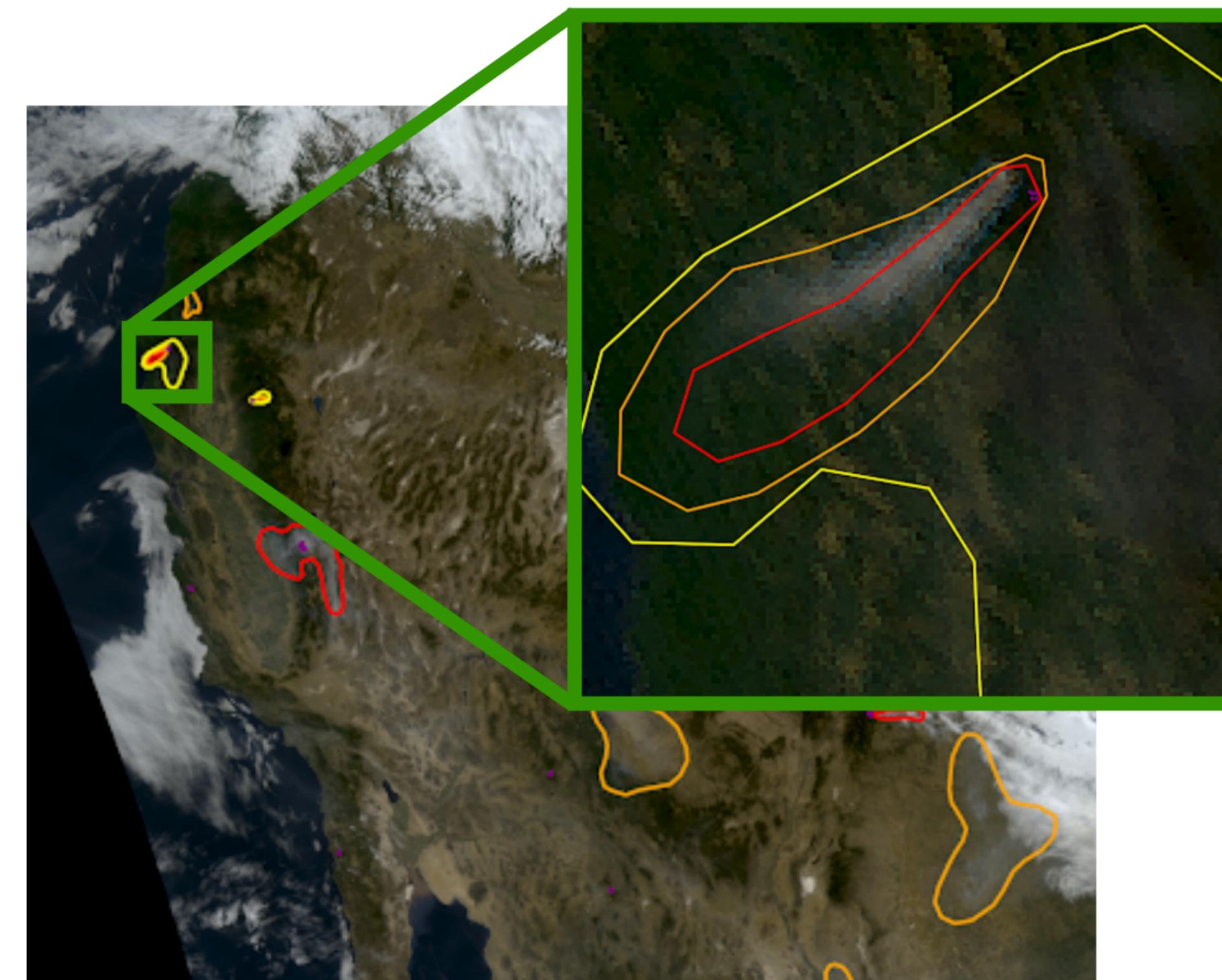


Figure 1. GOES-EAST imagery with corresponding HMS smoke annotations. The color of the outlines correspond to the density.

Satellite Imagery

NOAA's GOES satellites provide uninterrupted imagery that we use as input to our model. Leveraging the physics of the problem, we apply feature engineering in order for our model to better interpret our data.

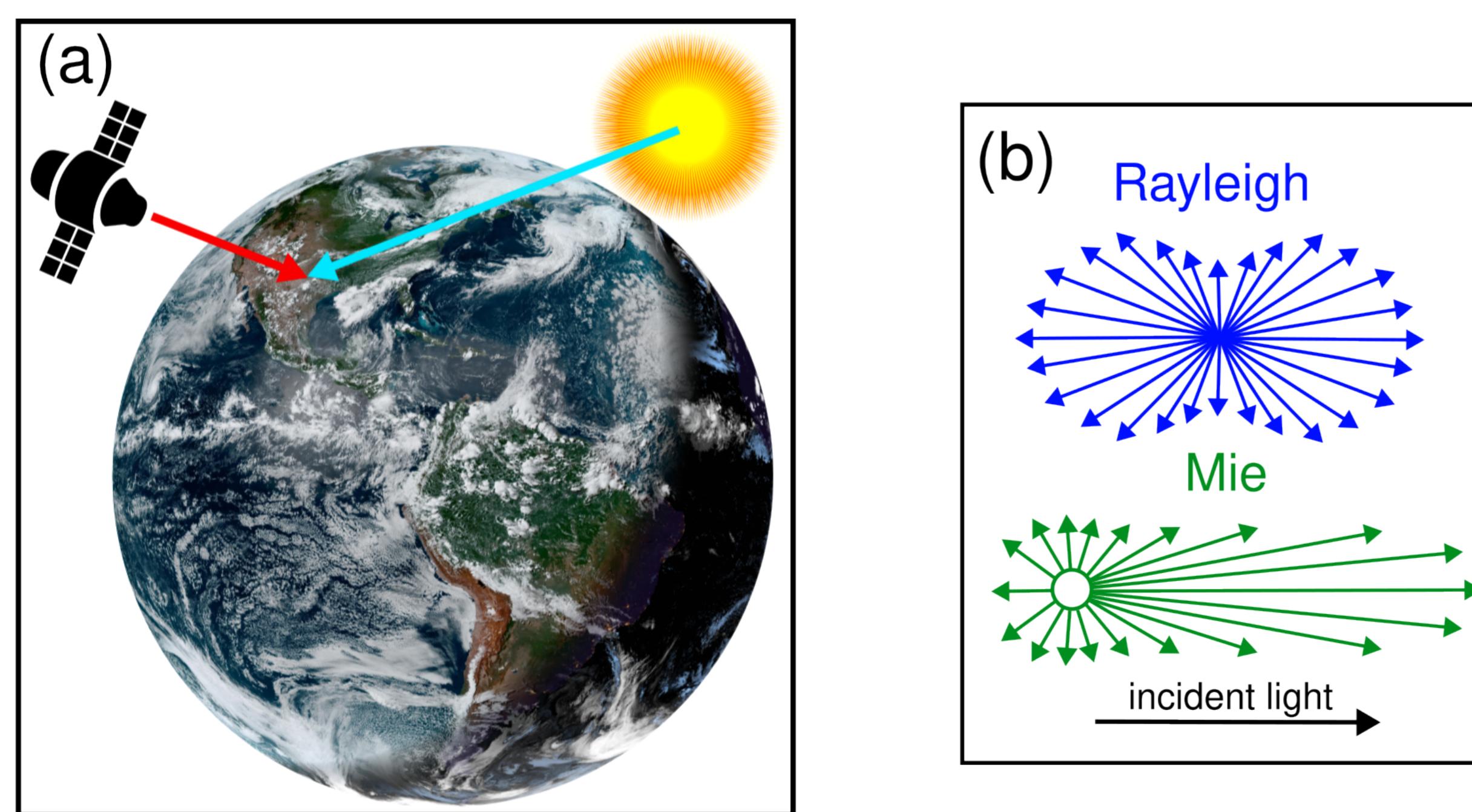


Figure 2. (a) GOES imagery with line of sight illustrated for the sun and satellite. (b) Illustration of how Mie scattering favors forward rather than backward scattering.

Semantic Segmentation

We use semantic segmentation (figure 3) to classify the smoke on a pixel by pixel basis.

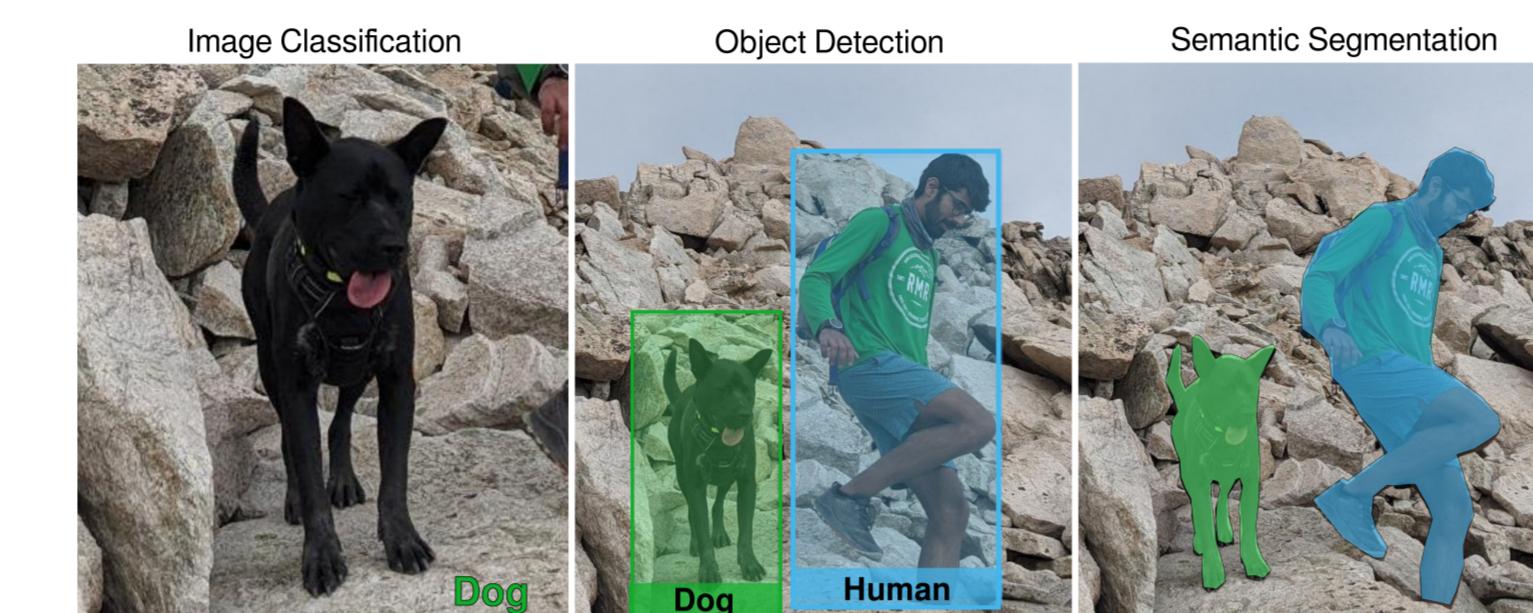


Figure 3. Comparison to traditional image or object classification.

Dataset and Workflow

Training set included 63,076 samples from 2019, 2000 and 2022 data. Validation set contain 8,227 samples from odd weeks in 2021, test set contains 7,729 samples from even weeks in 2021. All data is within the continental United States.

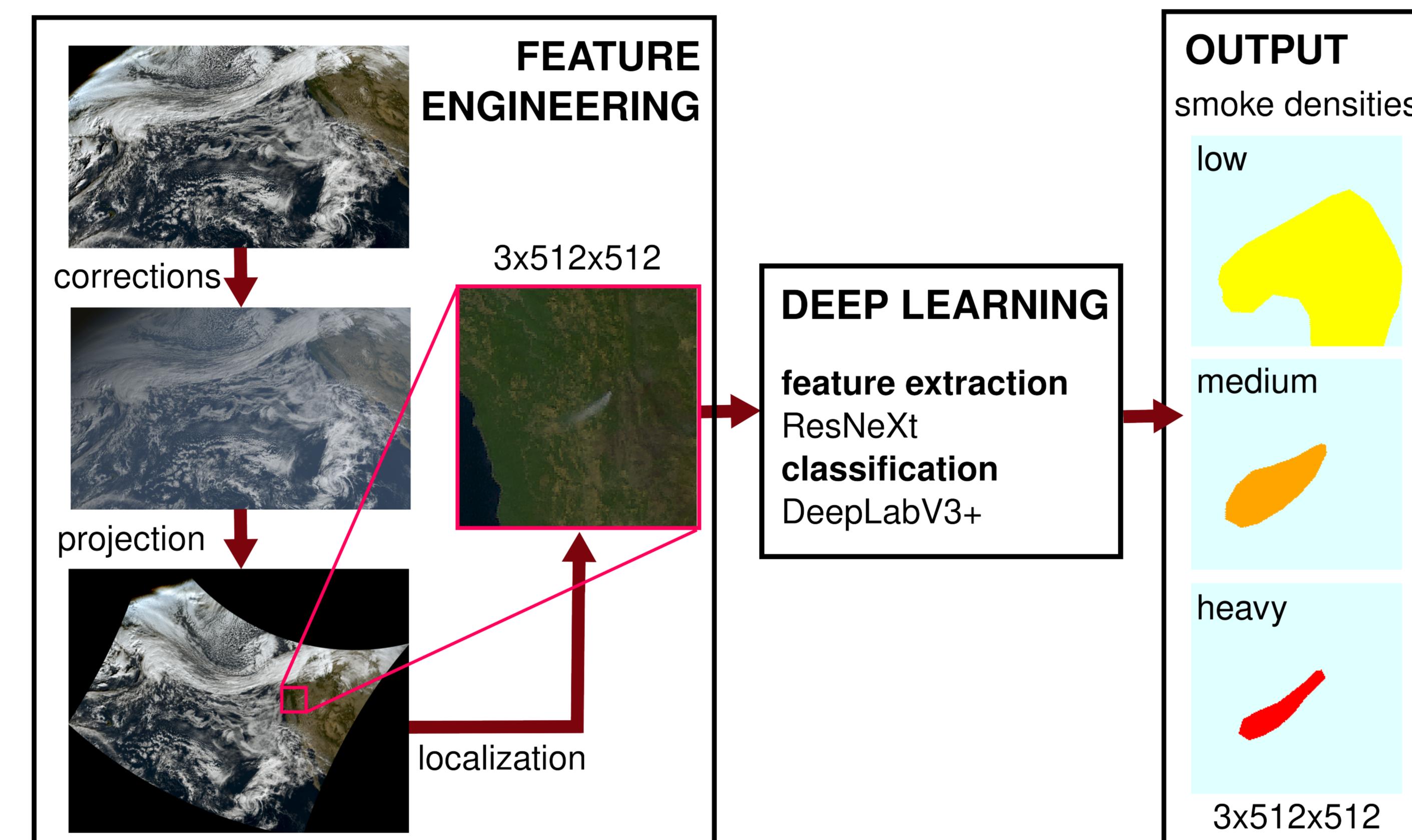


Figure 4. Workflow from feature engineering to the output of the model. The deep learning architecture consists of ResNeXt [?] that extracts features from the data and DeepLabV3+ [?] that performs semantic segmentation.

Thermometer Encoding

no smoke	[0, 0, 0]
low smoke	[0, 0, 1]
medium smoke	[0, 1, 1]
heavy smoke	[1, 1, 1]

Table 1. We apply ordinal rankings so that the penalization is representative of density proximity.

Discussion and Future Work

By taking advantage of physics based feature engineering, domain knowledge based architecture design and thermometer encodings, we were able to show state-of-the-art performance. Future work includes feature attribution to identify additional bands and pretraining the neural network weights.

Intersection over Union

$$\text{IoU} = \frac{\text{Overlap of TRUTH and PREDICTION}}{\text{Combined TRUTH and PREDICTION Regions}} = \frac{\text{INTERSECTION}}{\text{UNION}}$$

Table 2. IoU and Class Averaged IoU results by density and over entire dataset.

Figure 5. To measure performance, we use the IoU metric.

Example Results

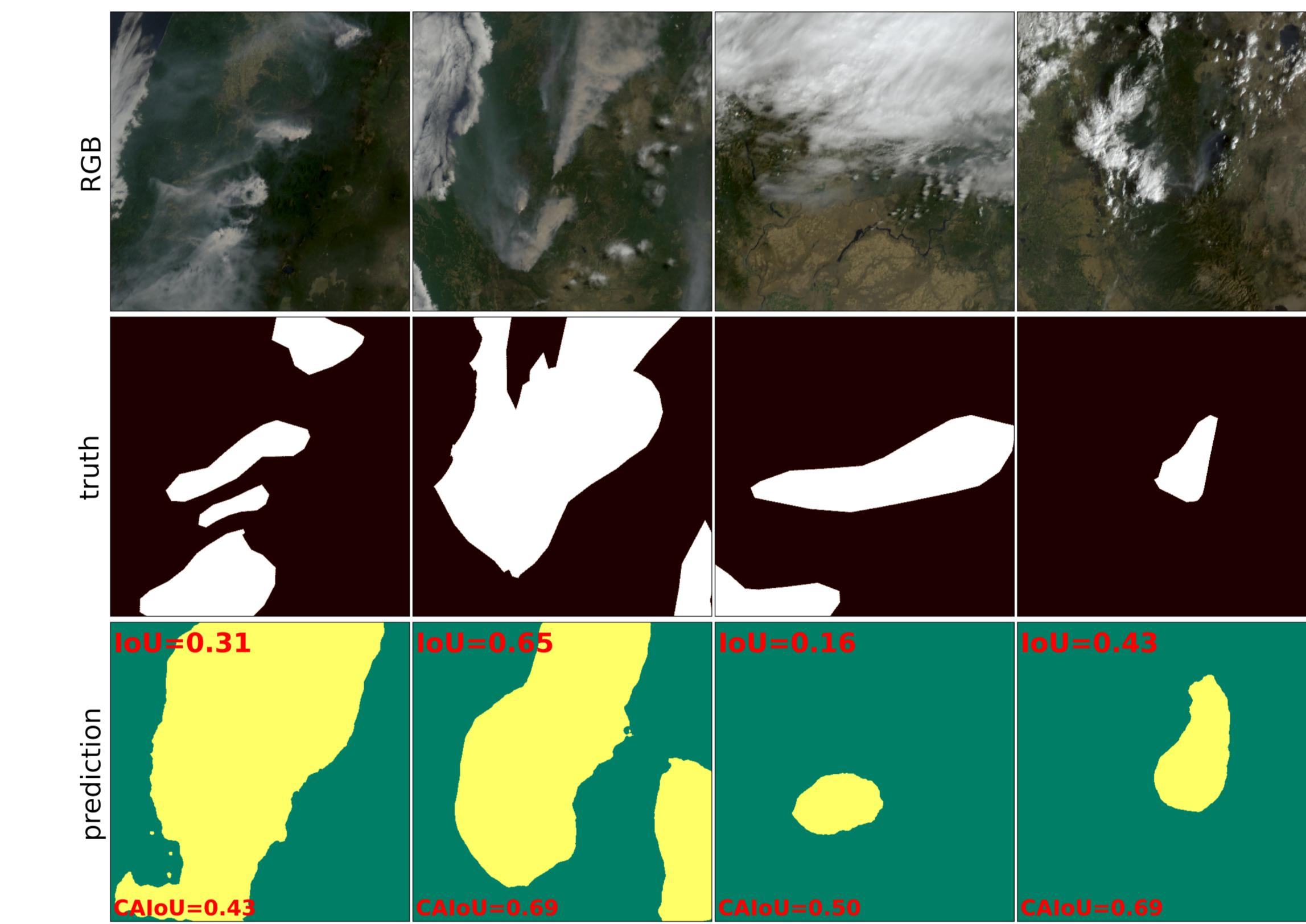


Figure 6. Visualization of the results for 4 test samples, IoU is calculated over all densities.

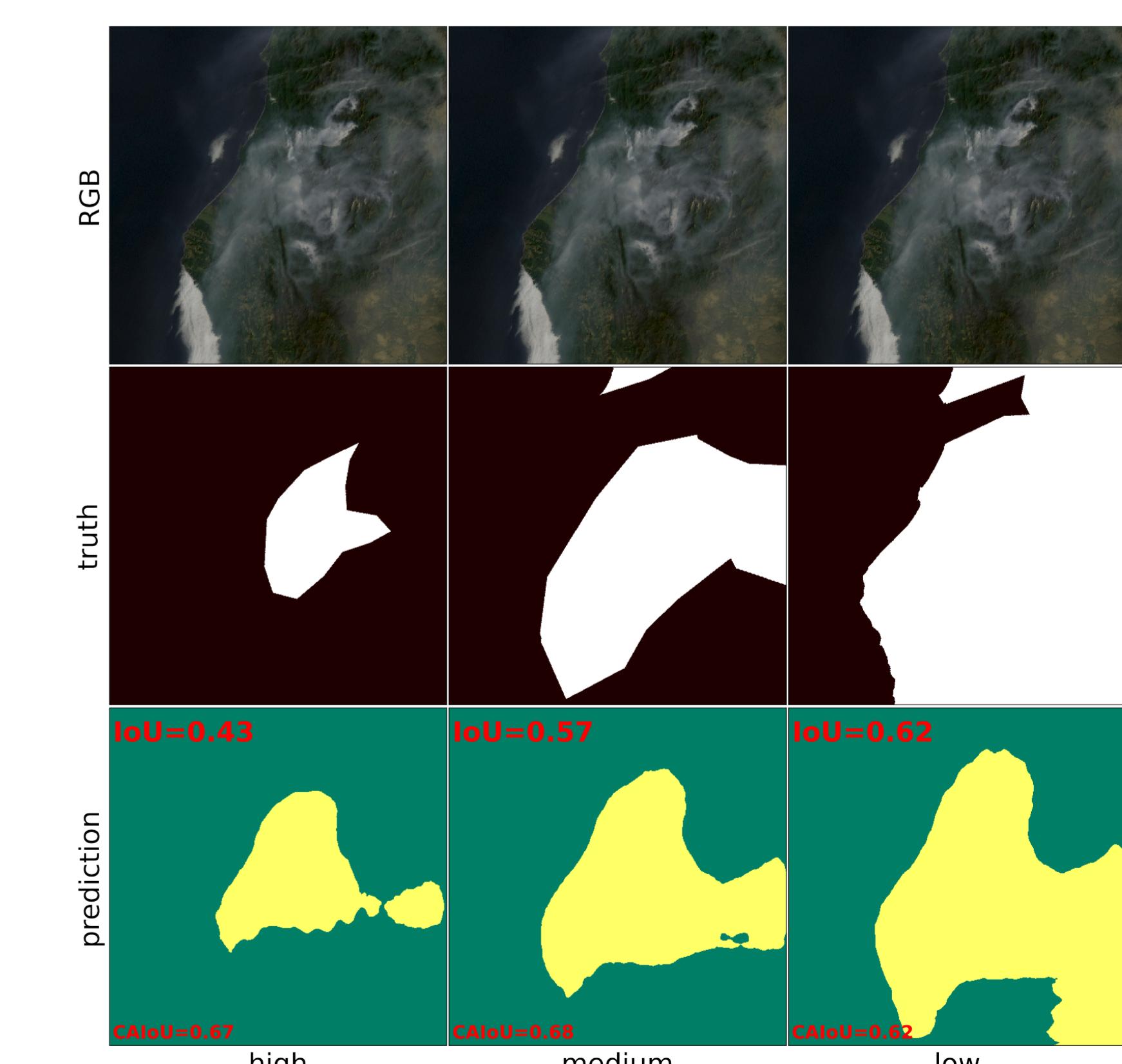


Figure 7. Visualization of the results for 1 test sample, IoU is calculated per density.

References