Moving Object Detection with Deep Convolutional Networks and LSTM

Gregoire Robinson
Advisors: Dong Zhang and Dr. Mubarak Shah

Background

- Objective is to detect moving vehicles on roads and highways in Wide Area Surveillance Imagery

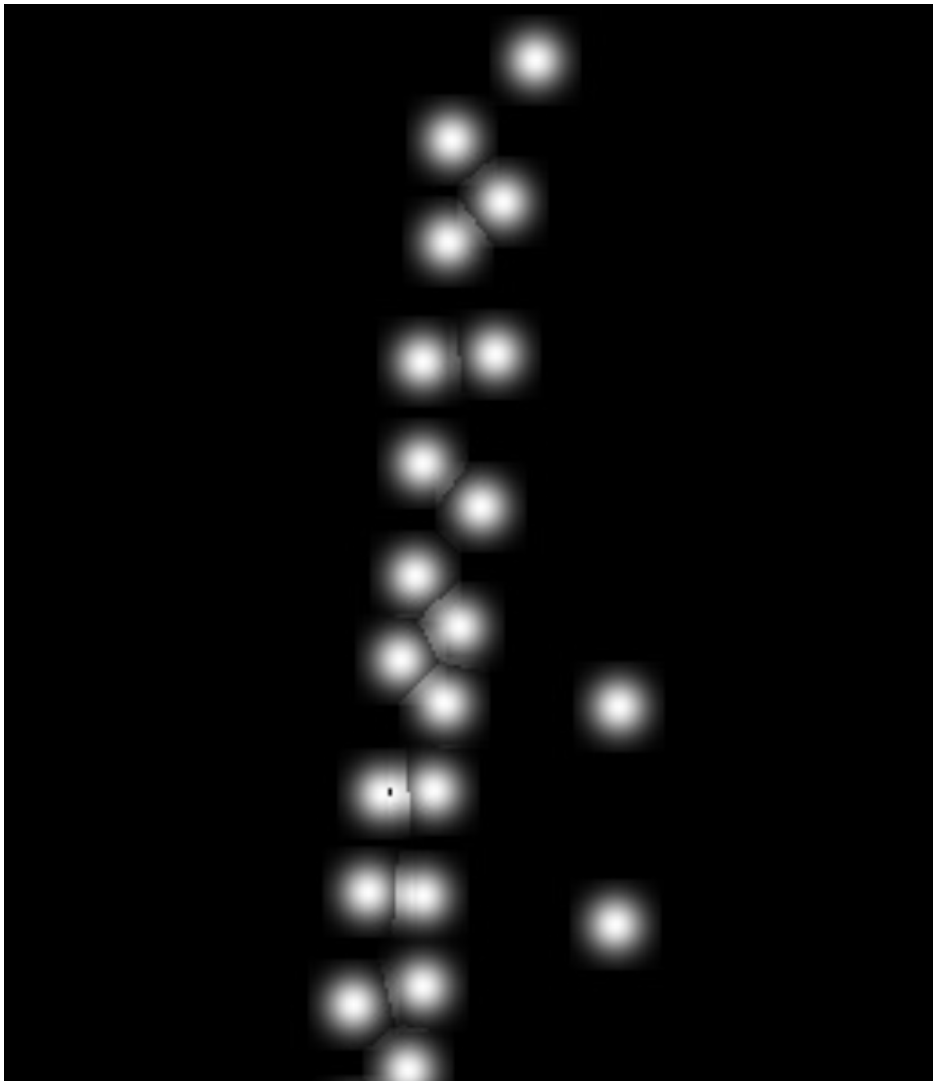


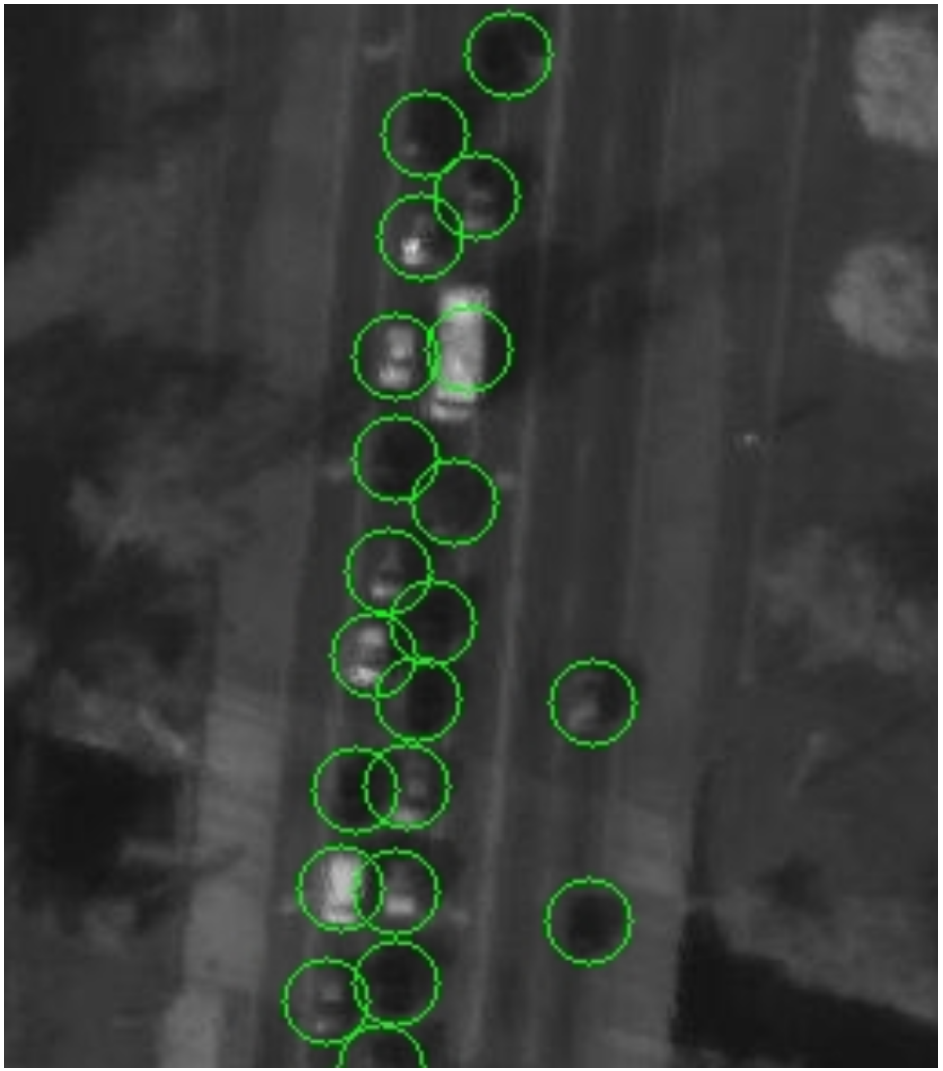
Overview of Previous Work

- Using WPAFB 2009 dataset. Images are fairly well annotated, but missing some vehicles on roads as well as many parked vehicles

- From these annotations, a heatmap can be generated

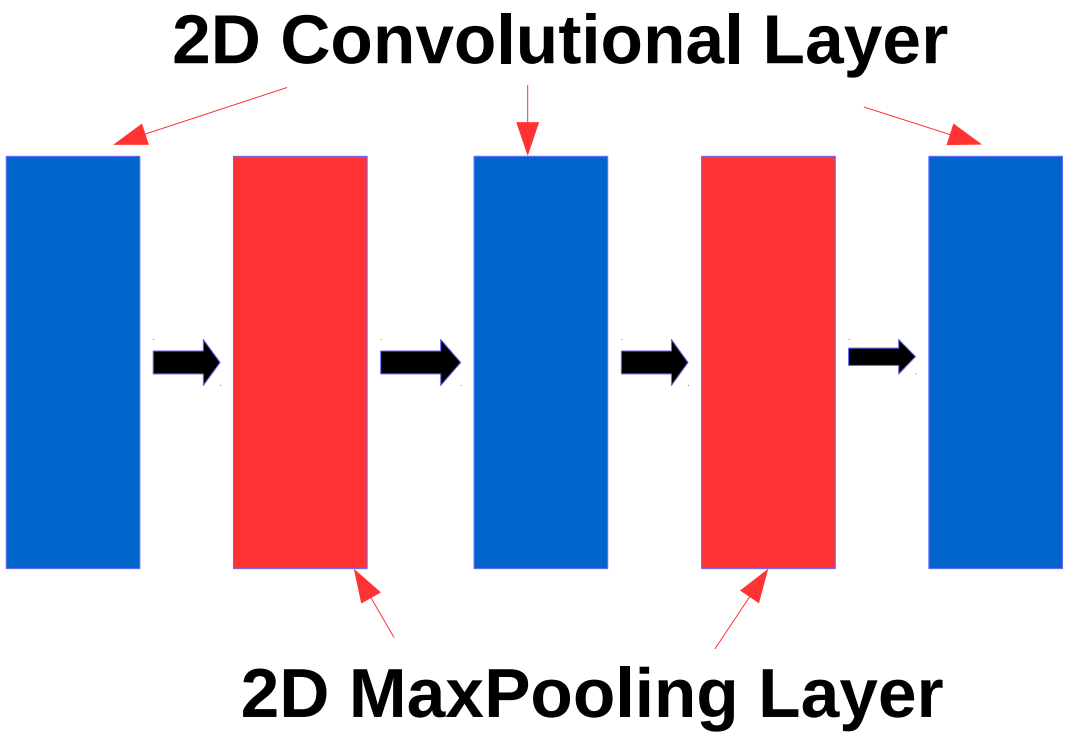
Overview of Previous Work





Overview of Previous Work

- Use the heatmap as a training label for a convolutional neural network



Overview of Previous Work

- Desired goal is for the network output to closely resemble heatmaps, with each detection point corresponding to a vehicle
- Results from previous weeks have not been meaningful
- This was likely due to insufficient activation layers in the network, problems with how image borders were handled during training, lack of data preprocessing, and too many images with parked vehicles

Convolutional Kernel Size 5

First 10 images in dataset, 10 iterations

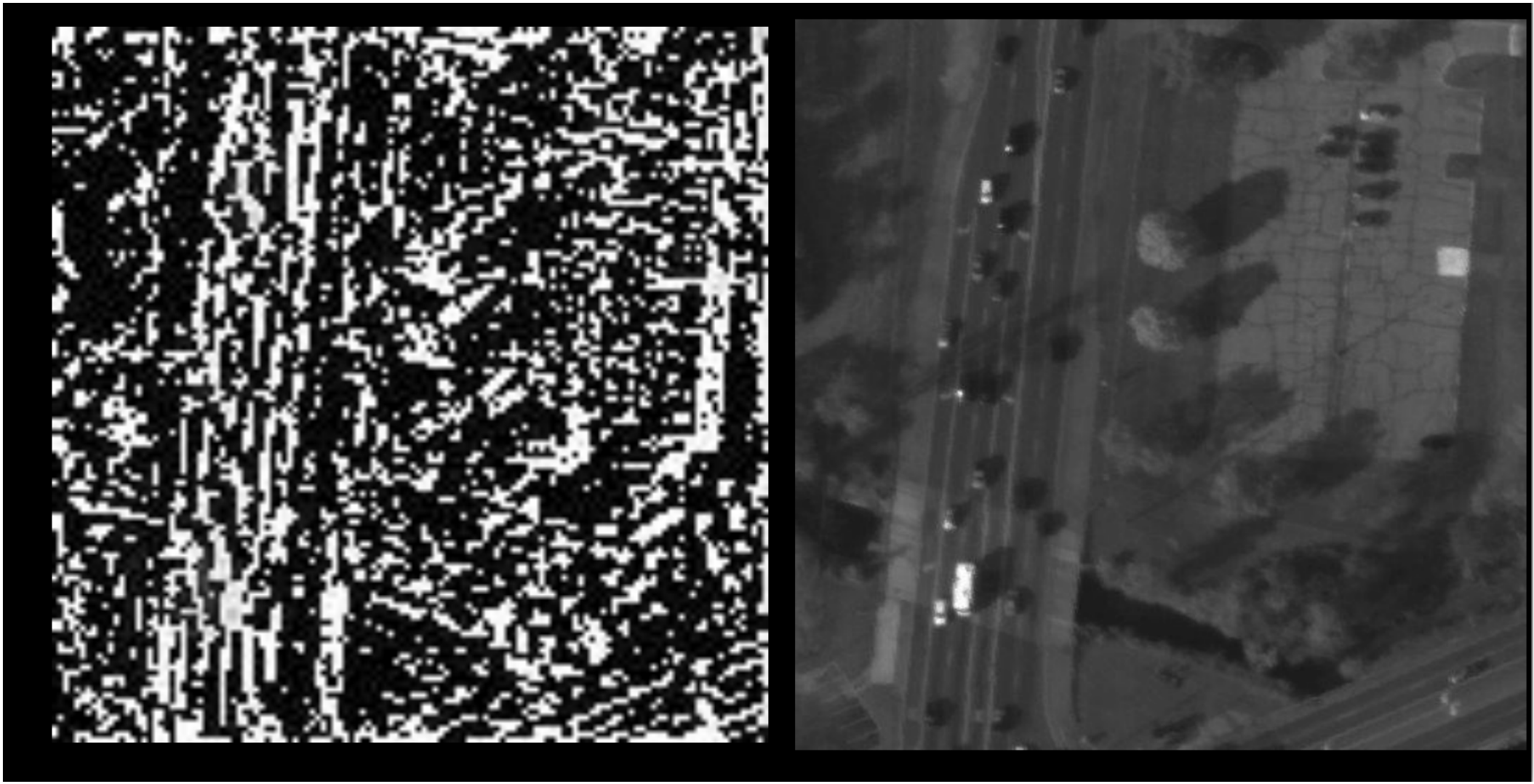




Convolutional Kernel Size 5

Training Time:
~24 hours

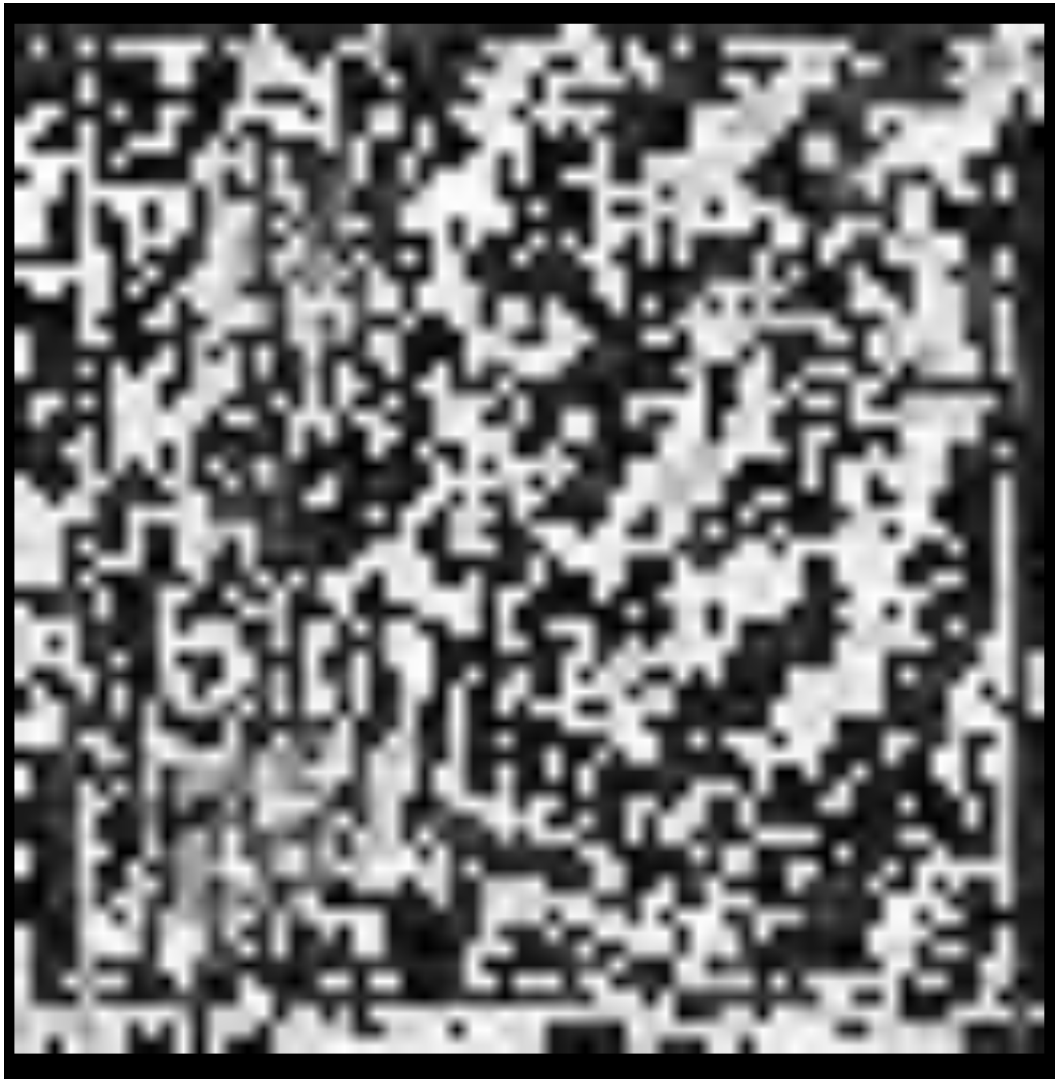
First 10 images in dataset, 200 iterations



Added MaxPool Layer, Increased Convolutional Kernel size from 5 to 15

Training Time: ~18.75 hours

Batch Size = 3000, 1 iteration per batch, 20 iterations



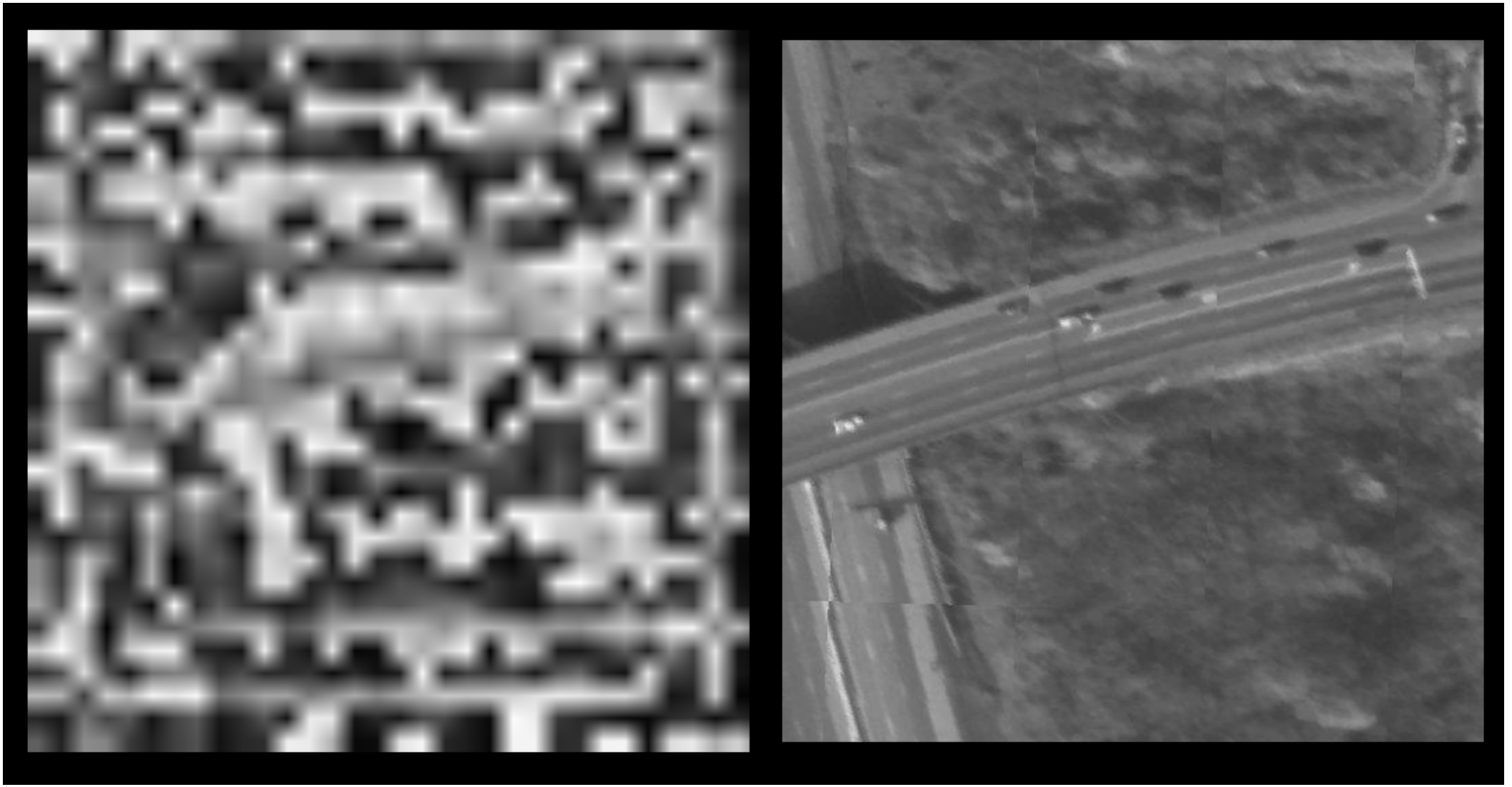


Convolutional Kernel Size 15, Train on 100 images





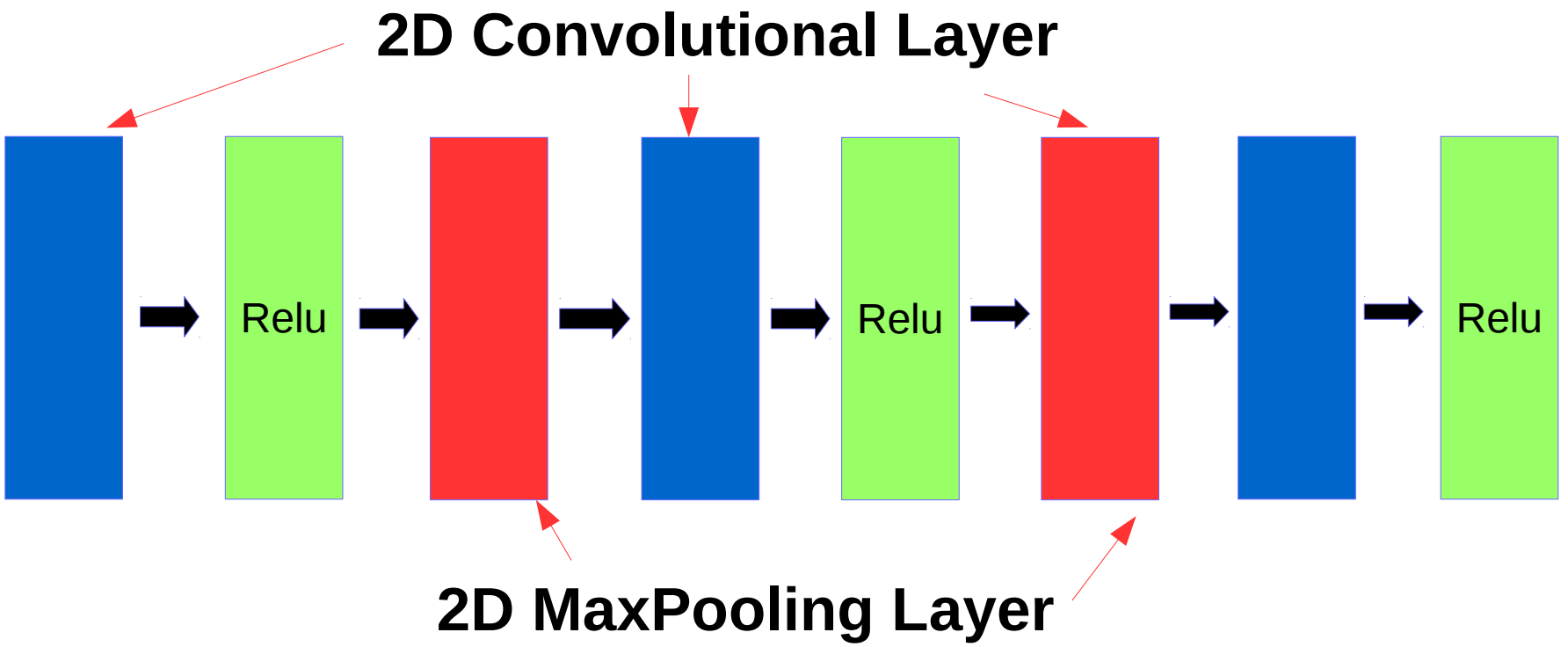
Convolutional Kernel Size 15, 100 images, Input Size 128 x 128



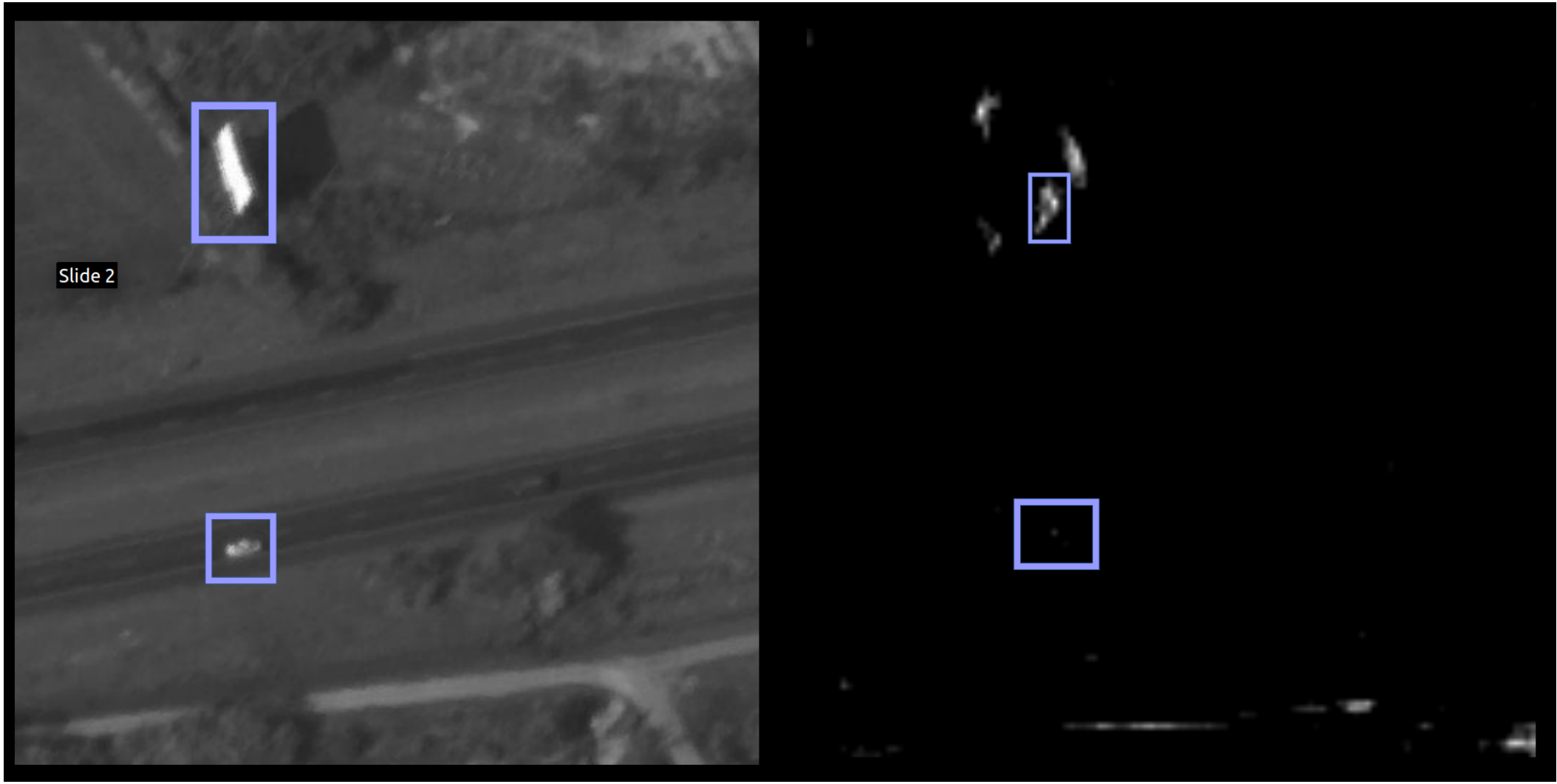
This Week...

- Added multiple relu activation layers
- Used a smaller, better-annotated dataset
- Changed the way image borders were handled in the network
- Eliminated images with many parked vehicles
- Preprocessed the data

Added Relu Activation Layers

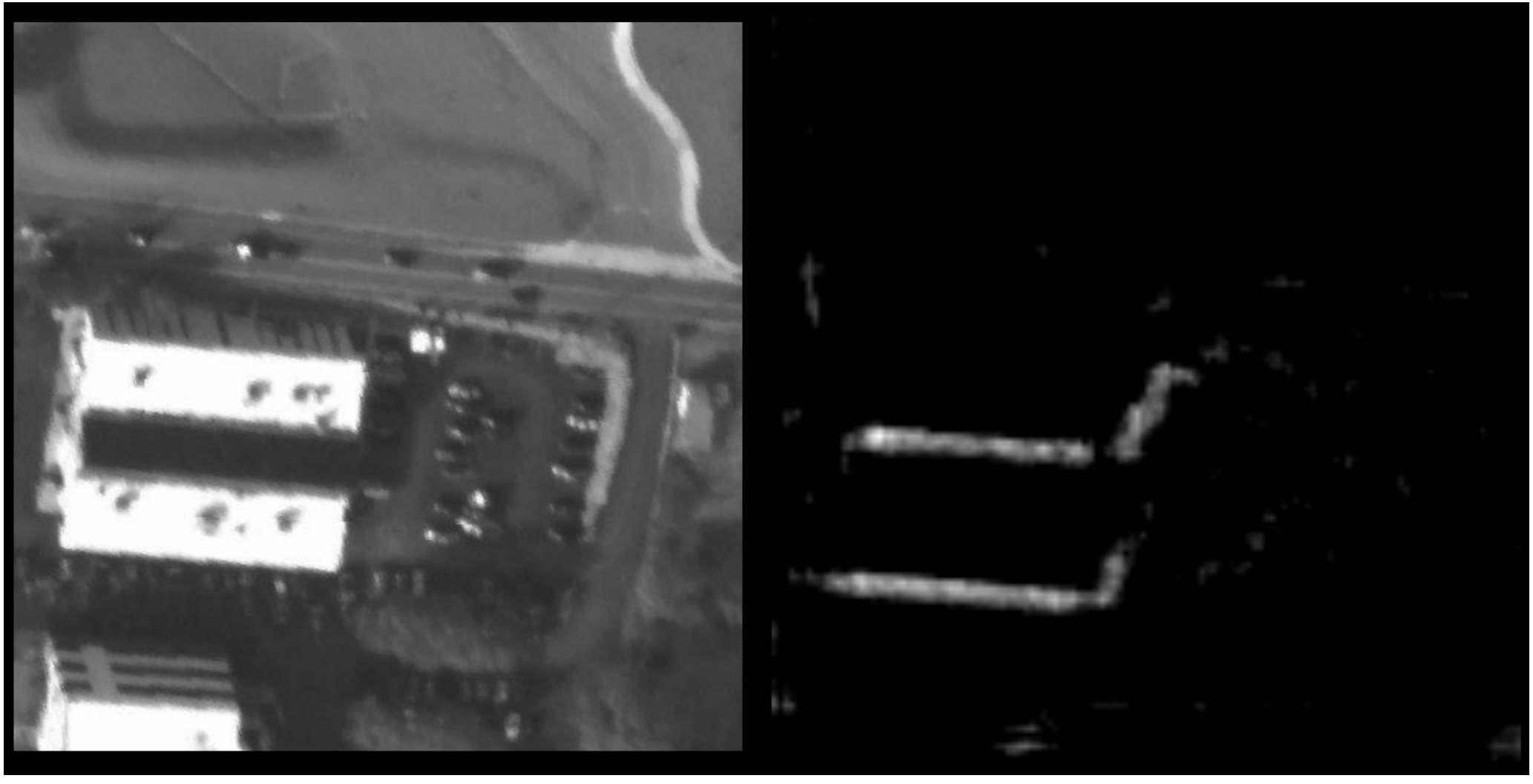


Experiment 1: 3000 images, <20 epochs



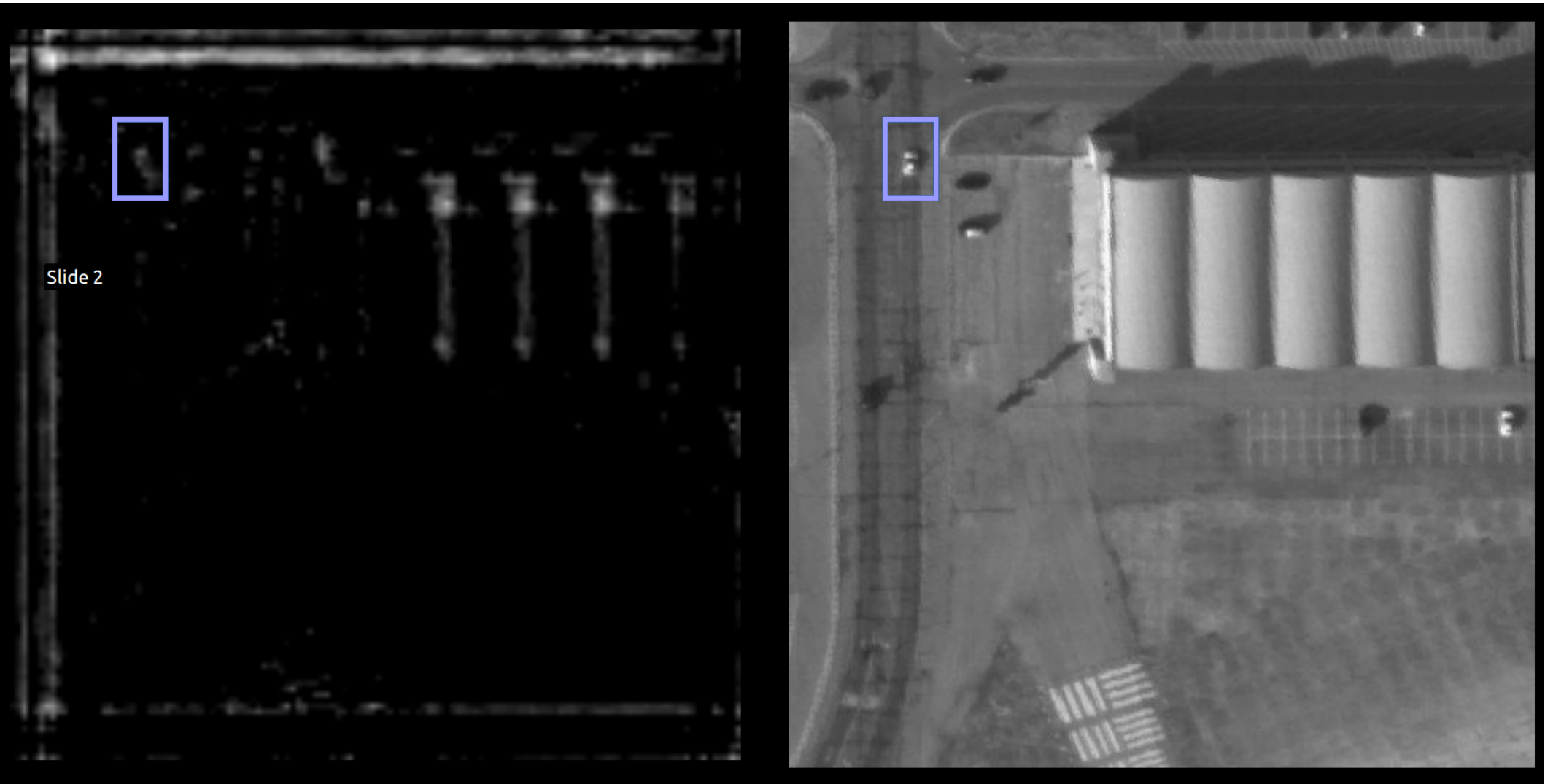
Experiment 1: 3000 images, <20 epochs

Detects corners and edges strongly if they are pronounced



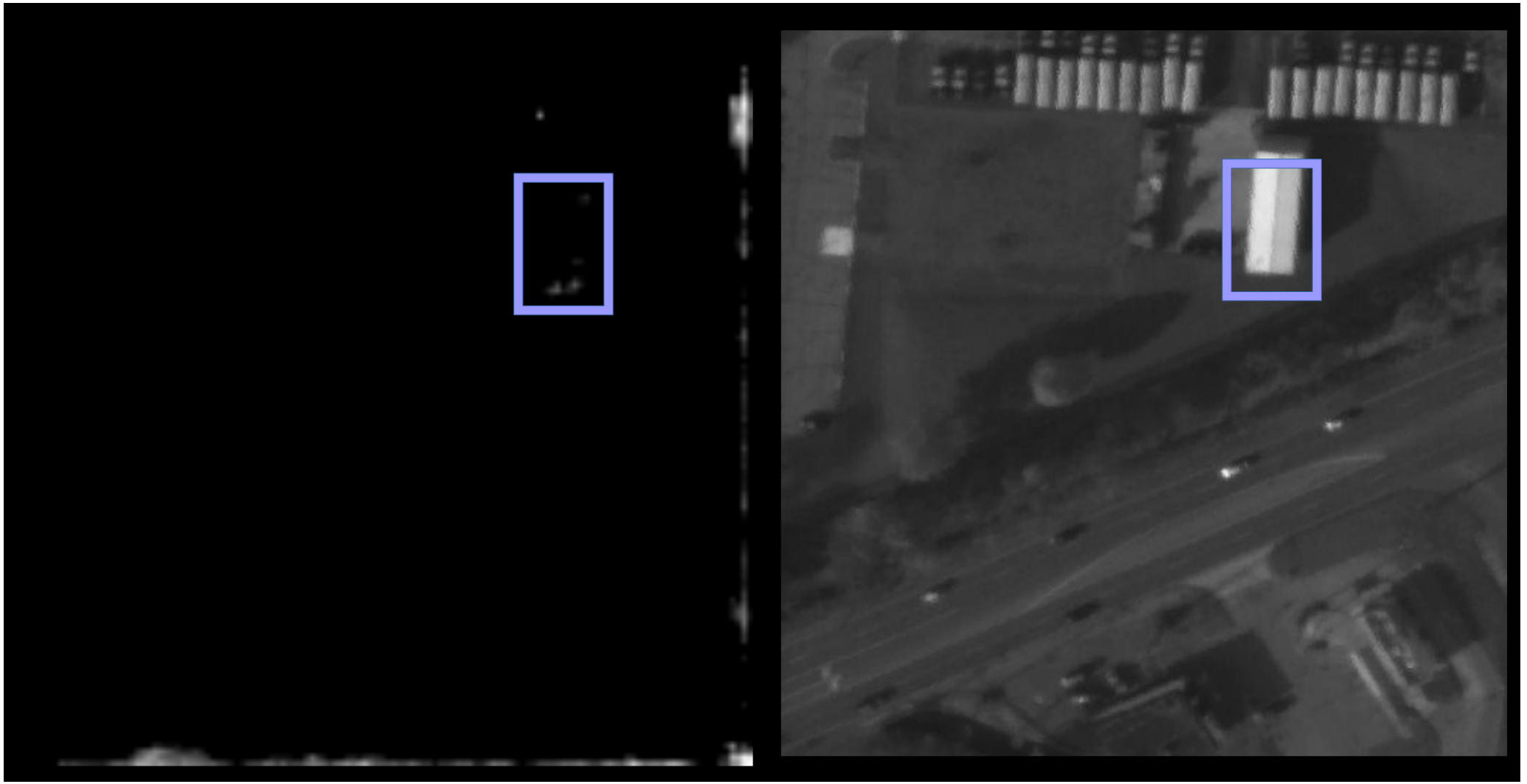
Experiment 2: 114 images, Better Annotations

Recognizes some edges and corners, but too much overall detection and borders problematic



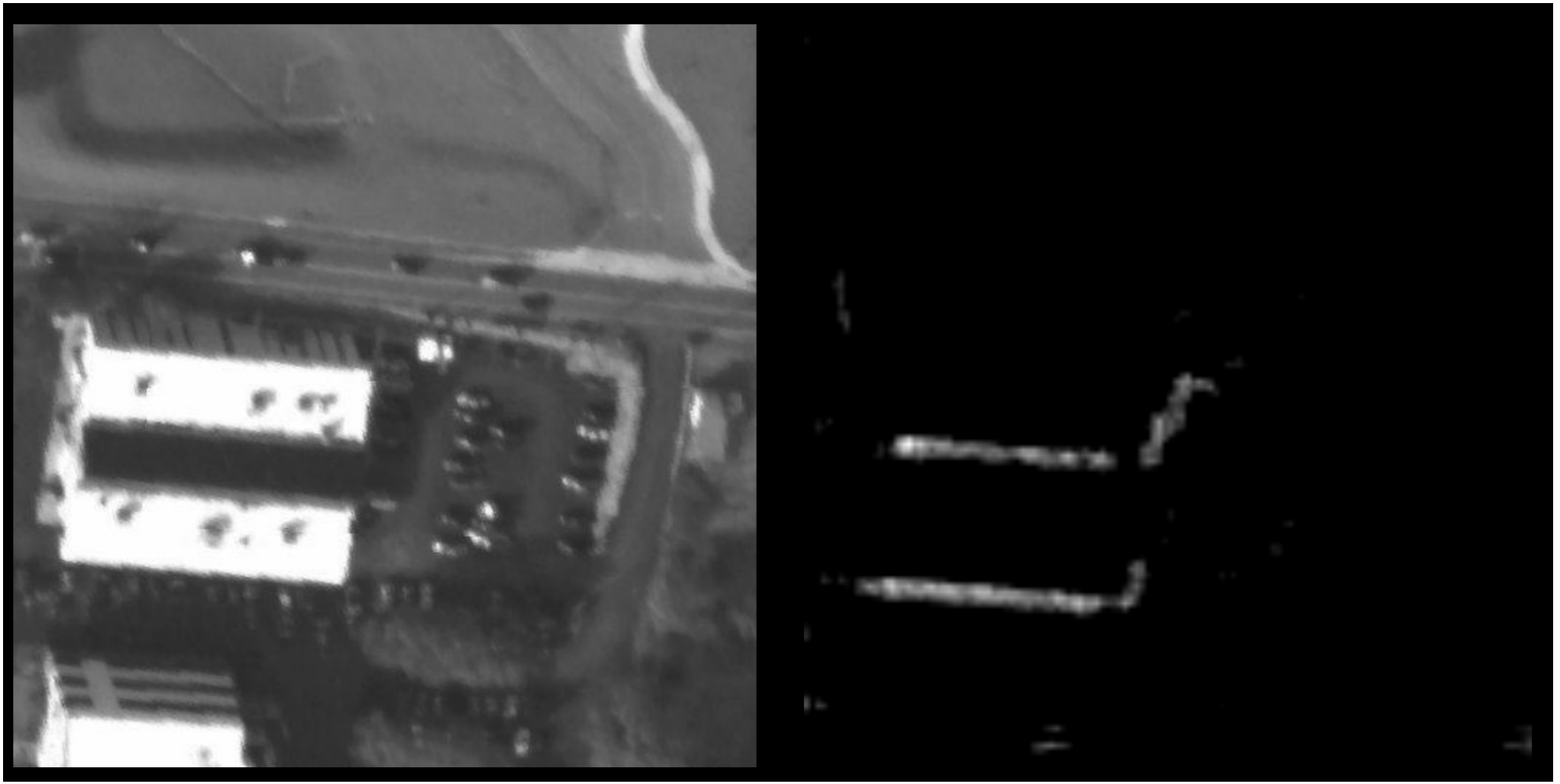
Experiment 3: 114 images, Increased Gaussian to 35 x 35 pixels on Heatmap

- Results were mostly blank
- Some detections of lines and/or corners



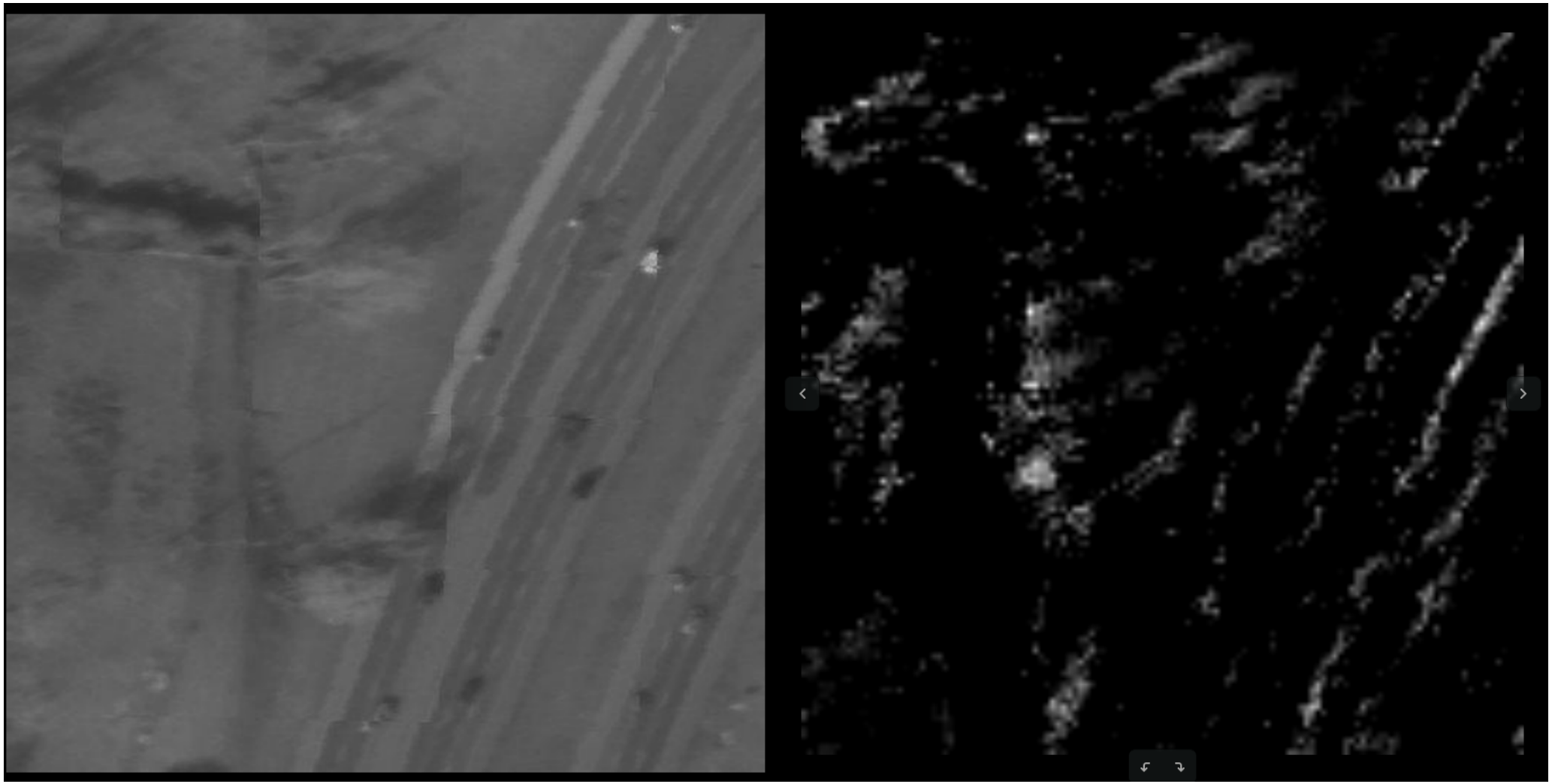
Experiment 4: 3000 images, 30 epochs

Reduced overall amount of detection, but recognition pattern appears unchanged



Experiment 5: 3000 Images, Image Border Handling, Data Preprocessing

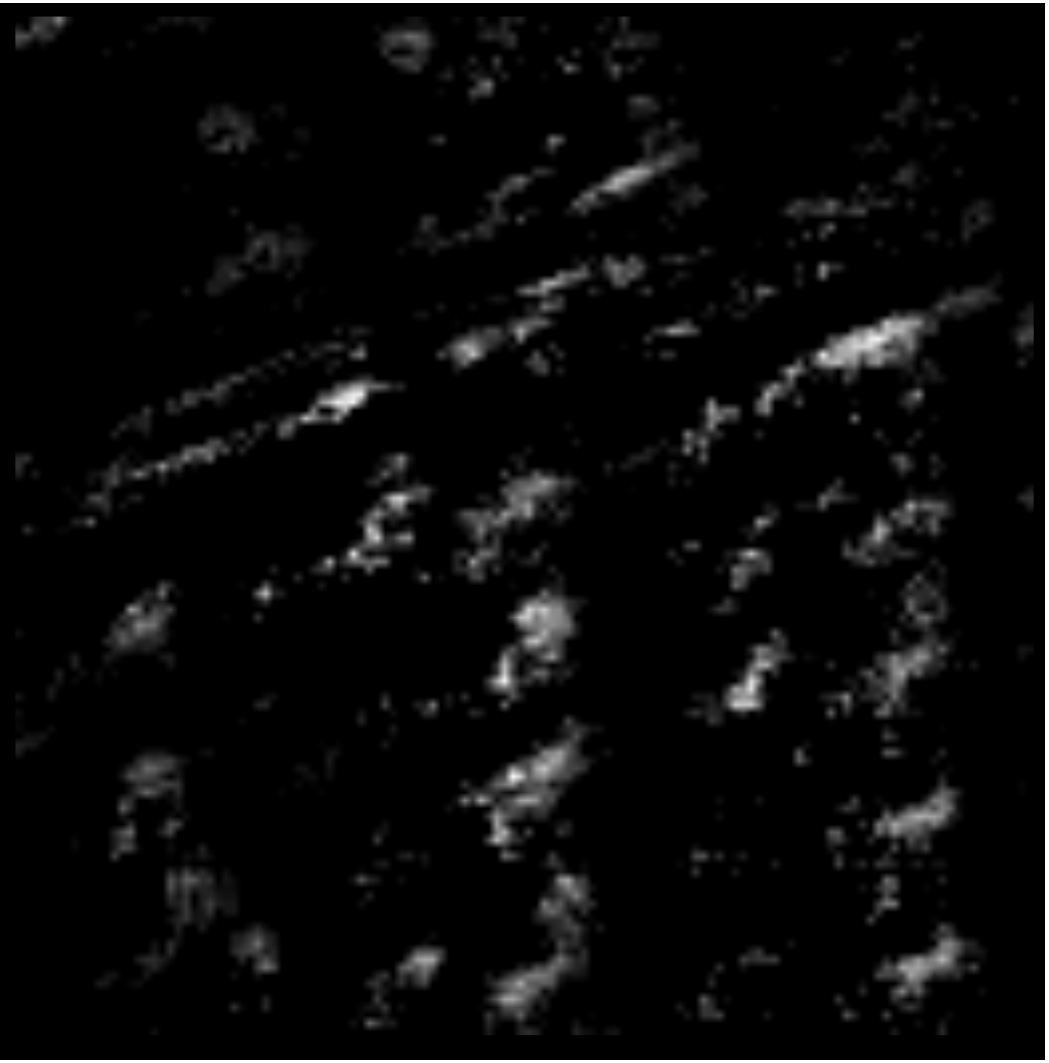
Overall results improved, but detection does not seem to be meaningful



Experiment 5: 3000 Images, Image Border Handling, Data Preprocessing

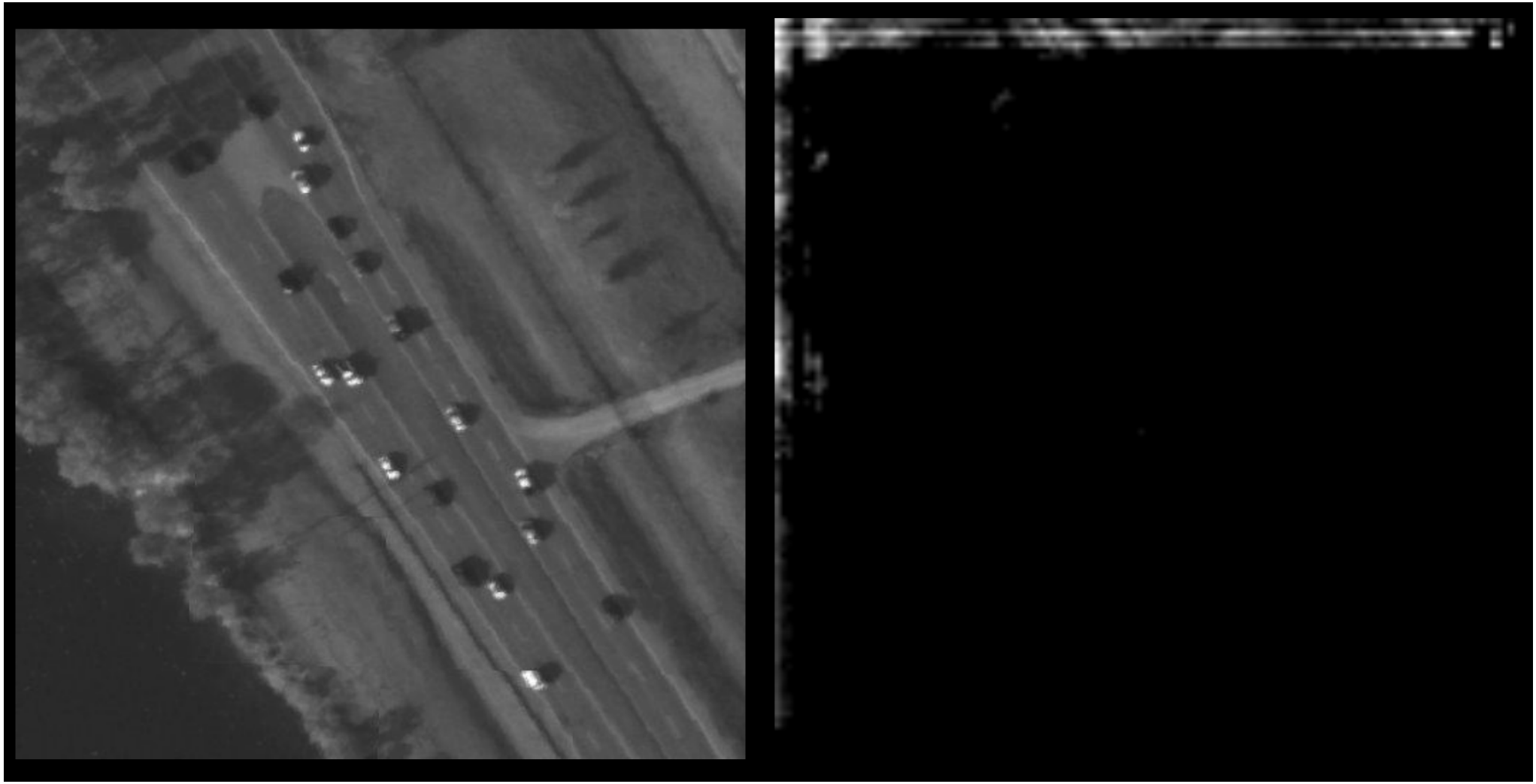
Overall results improved, but detection does not seem to be meaningful





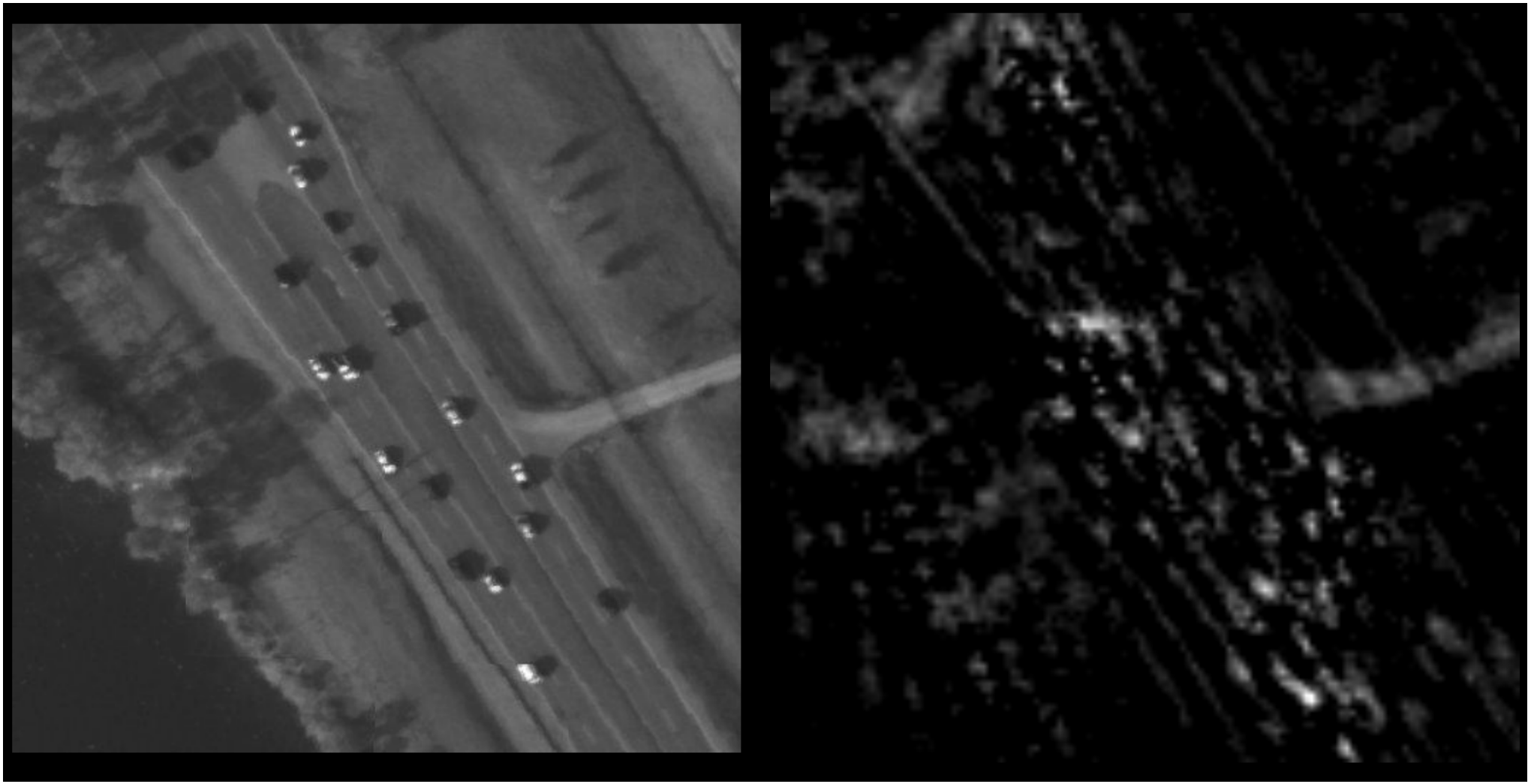
Experiment 6: 103 Images, Parked Vehicles Removed

Simply removing the parked vehicles lowered the overall detection



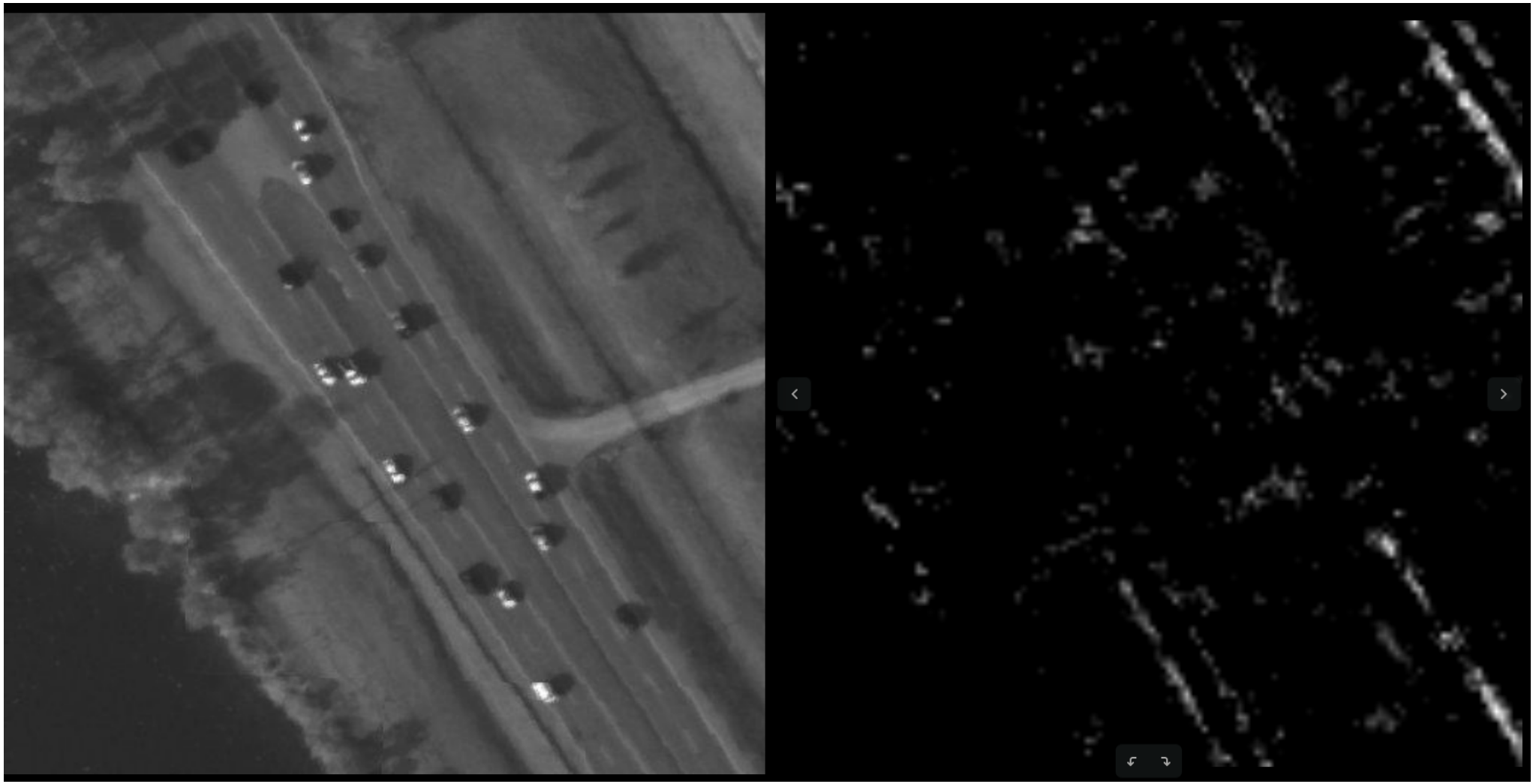
Experiment 7: 103 Images, Parked Vehicles Removed, Image Border Handling

Changing the heatmap borders significantly increased overall detection and seems to improve accuracy



Experiment 8: 103 Images, Parked Vehicles Removed, Image Border Handling, Data Preprocessing

Lower overall detection, unclear at this point if the added preprocessing improves accuracy



Conclusions

- WPAFB 2009 dataset annotations may not be sufficient

- Added relu activation layers, border handling, and removing images with parked vehicles appears to improve results, but data preprocessing may or may not

Future Work

- Use PVLabs dataset: better annotations, images are already registered

- Use 3D CNN and take advantage of frame differencing