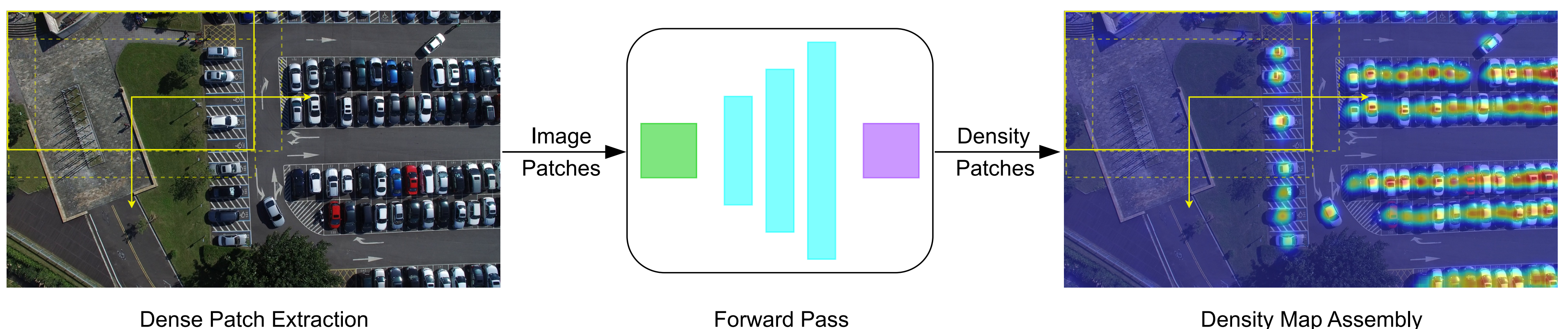


Counting Vehicles in Image and Video

Workflow



RGB patches are densely extracted from input image. The proposed Convolutional Neural Network processes them and produces density patches. Density patches are 8x smaller than input patches due to two max-pooling layers in the VGG-16 front-end. Therefore normalized (to preserve count) upscaling takes place. Final density map must be also normalized since extracted input patches are overlapping. Each position (pixel) in a final density map is an average of patches that cast a prediction in it.

Prediction Examples

