Quadtree Generating Networks: Efficient Hierarchical Scene Parsing with Sparse Convolutions

Kashyap Chitta¹, Jose M. Alvarez², Matrial Hebert³

¹ MPI-IS and University of Tübingen | ² NVIDIA | ³ Carnegie Mellon University

Semantic Segmentation

Problem
• Segmentation is a memory-intensive task due to high resolution, dilation-based operations that maintain high resolution activations, and quadratic scaling
• Leads to (1) poor training due to low batch sizes, (2) latency at inference
• Goal: improve performance-memory trade-offs with sparse output representations

Quadtree Generating Networks

Architecture
• Encoder-decoder with skip connections and no dilated convolutions
• Decoder activations stored as hash tables (x, y, feature)
• Prediction layers at each block that control propagation

Propagation Scheme
• If prediction = composite class, activation propagated to next layer, else set to zero
• Can adjust propagation to trade-off performance and memory consumption

Results

Results on Cityscapes
• QGN-All: high-memory propagation scheme with all pixels propagated
• QGN-PC: low-memory with only composite class propagated
• Good trade-off between accuracy and memory

Results on ADE20k
• QGN can be combined with recent backbone architectures as a drop-in replacement for dilated convolutions

Datasets
• In Cityscapes, SUN-RGBD and ADE20k, more than 50% of the pixels belong to 32*32 blocks

Conclusions
• Segmentation predictions can be made hierarchically using quadtree representations
• Results competitive to state-of-the-art with 2x-4x less memory consumption
• Flexible approach that can be adapted at inference without retraining

Quadtrees
• Hierarchical representation of 2D grids
• Advantage: Memory scales sub-quadratically, based on the number of pixels at class boundaries

Segmentation Maps as Quadtrees
• Starting from full-resolution, recursively group 4 neighboring pixels (children) to a single node (parent)
• Composite class assigned if children belong to different classes (grey in illustration)

References