Supplementary materials for "Spatial-aware Graph Relation Network for Large-scale Object Detection"

Paper ID: 4388

Additional diagram for understanding feature transformation

Our SGRN learns an interpretable sparse graph structure to encode relevant contextual information between regions. A novel spatial graph reasoning module with learnable spatial Gaussian Kernels is designed to perform graph inference with spatial awareness. We add an illustrative diagram for better understanding of feature transformation as Figure 1. Each node represents input x_i from different regions. Pseudo coordinates are computed according to the spatial information. Then K Gaussian weights are computed from the pseudo coordinates to guide graph convolution transformations of the features from the neighbors green nodes to the "blue" node.

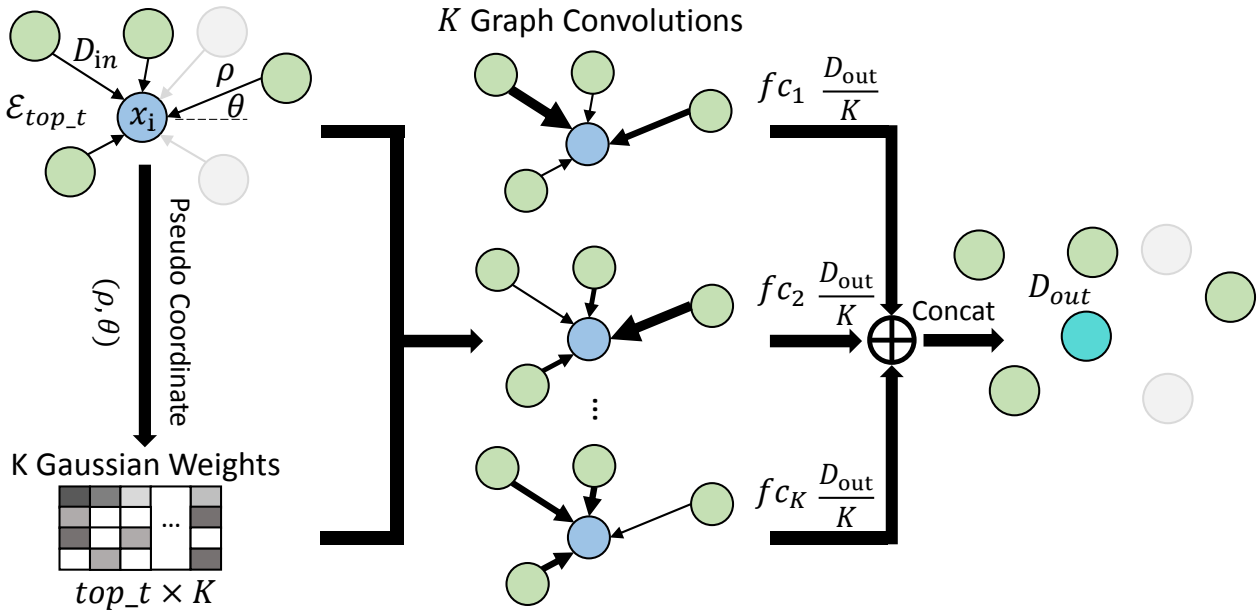

Figure 1: Additional pictorial diagram for understanding feature transformation. Each node represents input x_i from different regions. Pseudo coordinates are computed according to the spatial information. Then K Gaussian weights are computed from the pseudo coordinates to guide graph convolution transformations of the features from the neighbors green nodes to the "blue" node.

Additional Qualitative Results

More qualitative comparisons on VG_{1000} between FPN and our method SGRN can be found in Figure 2. From the comparisons, objects with occlusion, ambiguities and rare category can be detected and localized well by our method, while the FPN fails to detect them. Our method is more accurate than the baseline method due to the help of the encoded relation. Figure 3 shows some additional examples of the learned graph structures from our SGRN.The centers of regions are plotted and connected by the

learned graph edges. Edges thickness correspond to the strength of the graph edge weights. Our method learned interpretable edges between regions.

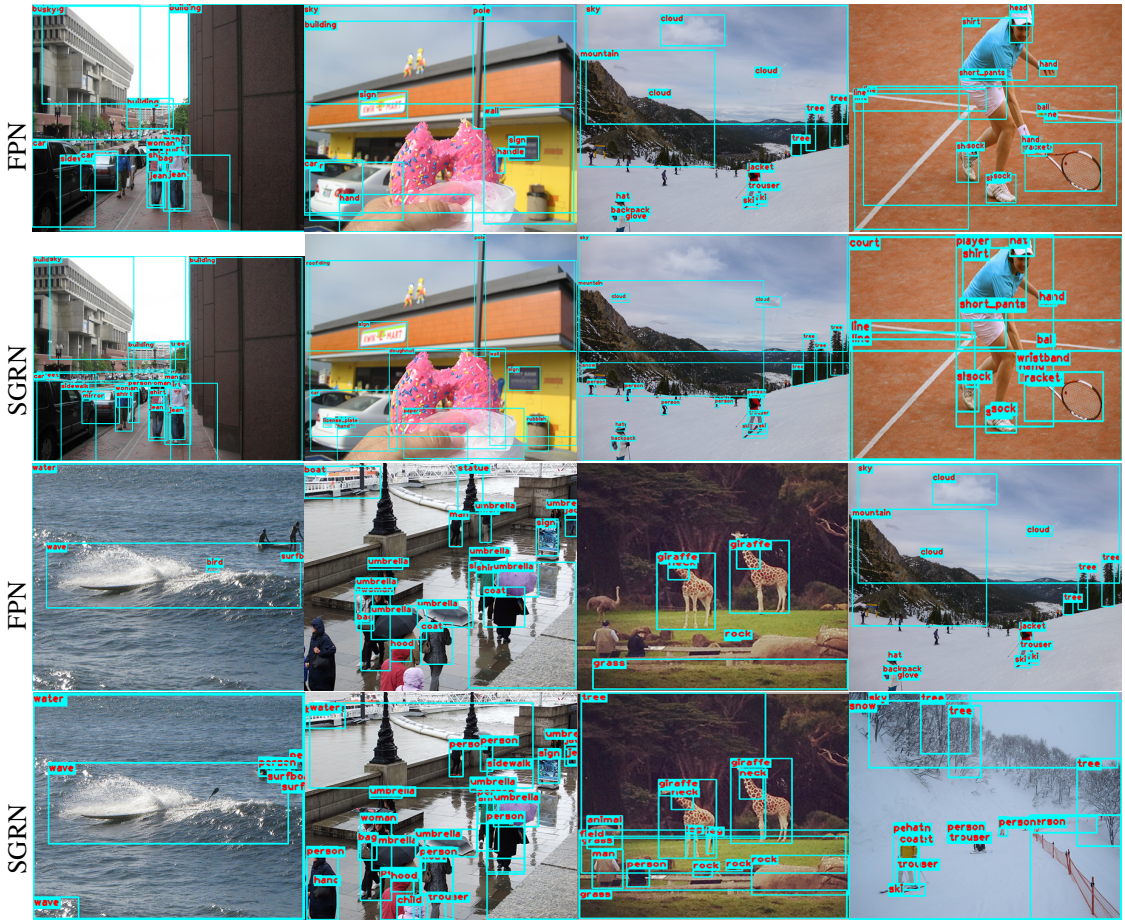

Figure 2: More qualitative results comparison on VG_{1000} between FPN and SGRN. Our method is more accurate than the baseline method due to the help of the encoded relation.

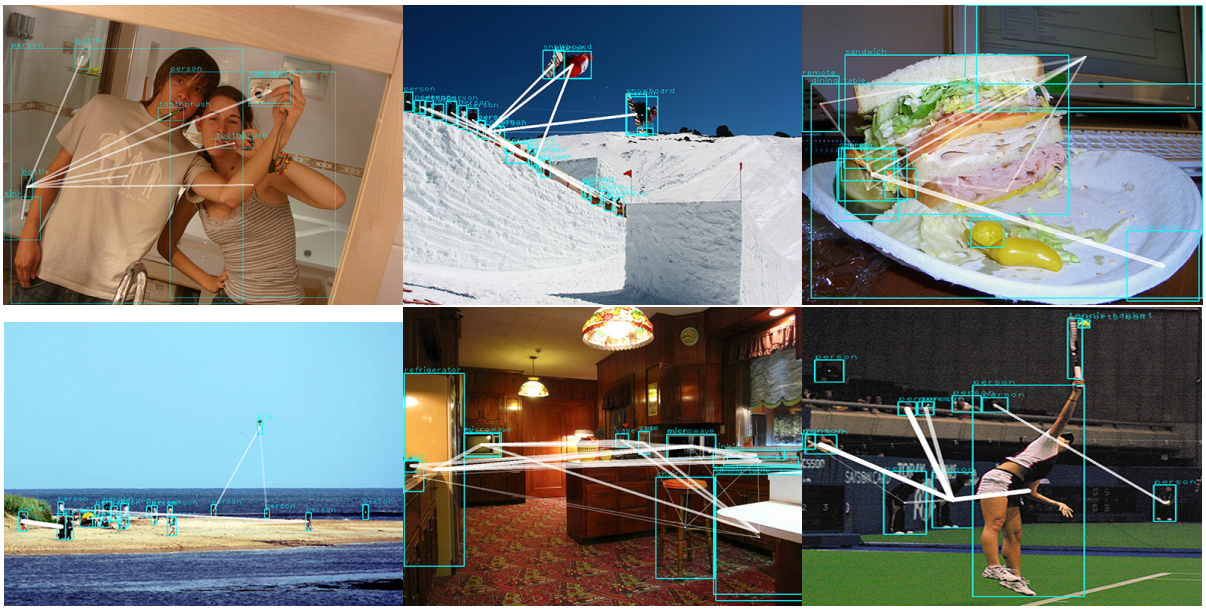

Figure 3: Examples of the learned graph structures from our SGRN. The centers of regions are plotted and connected by the learned graph edges. Edges thickness correspond to the strength of the graph edge weights.