Pixel-Adaptive Convolutional Neural Networks
Hang Su1, Varun Jampani2, Deqing Sun2, Orazio Gallo2, Erik Learned-Miller1, Jan Kautz2
1UMass Amherst 2 NVIDIA

1 Motivation

Spatial convolution • Fundamental building block of modern neural networks • Provides an efficient and effective way to propagate and integrate features across image pixels

It has advantages ...
• Simple formulation → efficient parallelization
• Spatial sharing → capturing invariance
• Allows hierarchical features learning

But the filters are content-agnostic ...
• Loss gradient are pooled across pixel locations.
• Once learned, same filter banks are used across all images and pixels locations.

Goal: Make convolutions adaptive to image content

2 Pixel-Adaptive Convolution (PAC)

Making convolution content-adaptive ...
• Kernel function K modifies filters according to f.
• In contrast, kernel prediction networks [1,2,3] generate filters directly: \( \mathbf{v}' = \sum_{j} \mathbf{w}_{j} \mathbf{p}_{j} \). We use \( K(f_i, f_j) = \exp(-1/\sigma(\|f_i - f_j\|^2)) \) for our experiments.
• f and v can both be learned through backpropagation.

PAC generalizes many existing operations
• Spatial convolution
• Bilateral Filter
\( K(f_i, f_j) = 1 \)
\( f = (r, g, b) \)
\( W[p_i, p_j] = \exp(-1/\sigma(\|p_i - p_j\|^2)) \)

• Average pooling
• Detail-preserving pooling [4]

PAC is a content-adaptive operation that generalizes spatial convolutions \( \mathcal{O} \) https://suhangpro.github.io/pac (code available)

3 Deep Joint Upsampling Networks

Depth

Depth (NYU Depth v2)

Optical flow (Sintel)

Results on NYU Depth v2

4 Efficient CRF with PAC

Fully Connected CRF [10]

\( p(x,y) = \exp \left( \sum_{i} \mathbf{w}_{i,j} - \sum_{i} \mathbf{w}_{i,j} K(f_i, f_j) \right) \)

• Pairwise term:
\( \mathbf{w}_{i,j} = \mathbf{w}_{i,j} K(f_i, f_j) \)
• Learned features in K: \( K(f_i, f_j) = \exp(-1/\sigma(\|f_i - f_j\|^2)) \)
• Local connections with multiple dilation factors

\( \Omega_i \)

PAC-CRF

• A more general pairwise term:
\( \exp \left( \sum_{i} \mathbf{w}_{i,j} K(f_i, f_j) \right) \)
• Mean-field update rule:
\( \mathcal{Q}^{(t)}(\mathbf{v}) = \frac{1}{Z} \exp \left( \sum_{i} \mathbf{w}_{i,j} K(f_i, f_j) \right) \)

\( \Omega_i \)

5 Layer Hot-Swapping with PAC

A modification to pre-trained networks that:
• Replace certain Conv layers with PAC counterparts
• Retain the pre-trained weights

Experiments
• -2% improvement w/ minimal runtime overhead
• Complementary with PAC-CRF improvement

References