SEMANTIC IMAGE SEGMENTATION BY GRAPH BASED CLUSTERING OF DEEP NEURAL NETWORK PREDICTIONS

Author
Cristian Ion

Scientific Coordinator
Cornelius Croitoru, Prof., Dr.

Content

- Computer Vision Trends
- Image Segmentation with pixel groupings
- Capturing Image Semantics with CNN
- Dense CNN Predictions. Semantic Segmentation
- Refining Results. Final Architecture
- Framework
- Future Work
- Conclusions

Computer Vision Trends

- Scene understanding
- Augmented Reality
- Image Enhancement

Image Segmentation with pixel groupings

- Group pixels into disjoint regions - clustering
- Bottom-Up versus Top-Down

Image Segmentation with pixel groupings

- Graph representation
- Pixels are graph nodes
- Edges represent pairwise similarity between pixels
- Normalized graph cuts on the graph create pixel groups - clusters

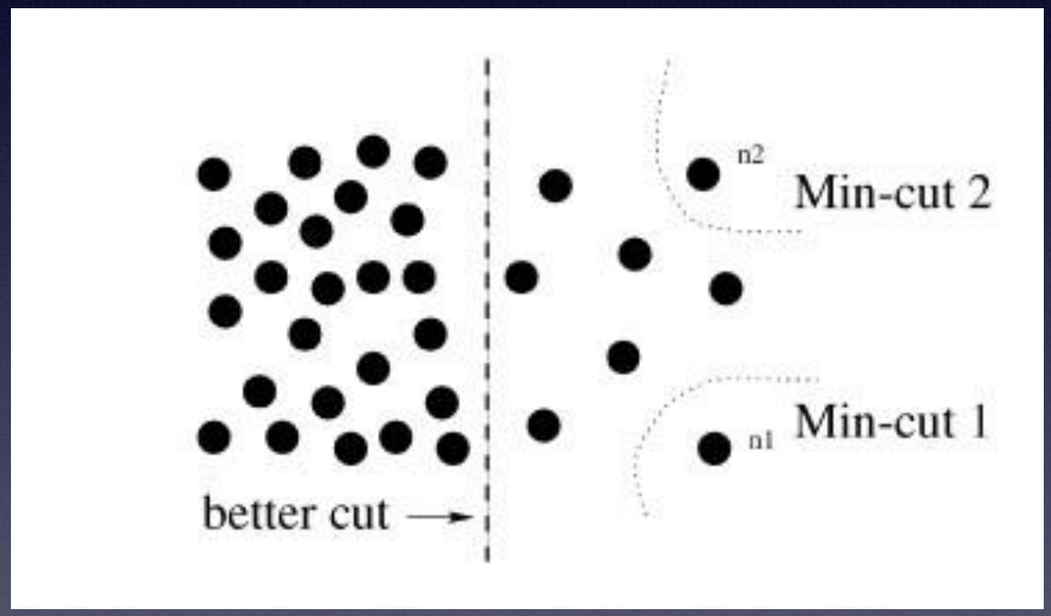


Image Segmentation with pixel groupings

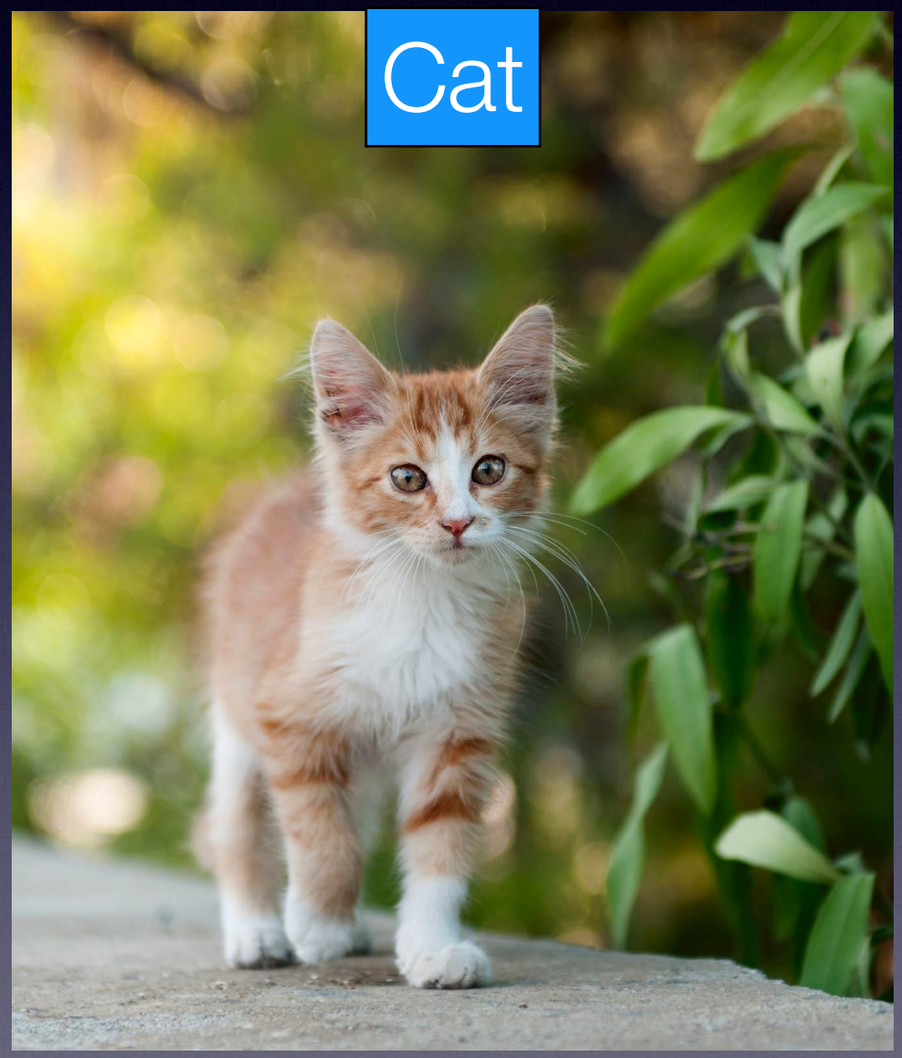
- The similarity matrix is defined by assigning an weight to each edge in the graph

- Weights are assigned based on texture, brightness, color and crossing object boundaries.

- Find groups (normalized cuts) using the algorithm proposed by *Jianbo Shi* and *Jitendra Malik*.

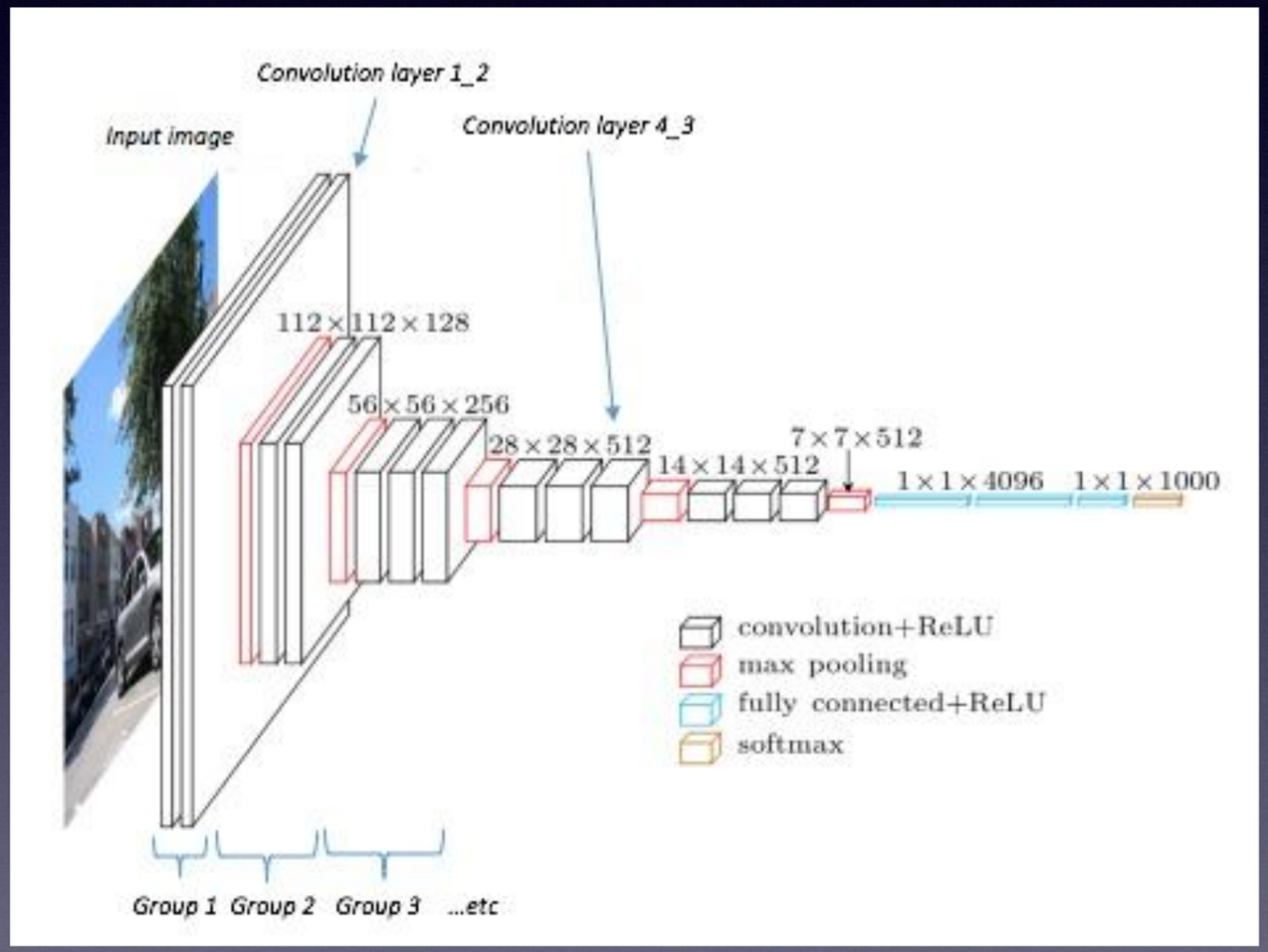
Capturing Image Semantics with CNN

- Image Classification - object classes found in an image, without localization
- ImageNet Database
- VGG, ResNet - CNN models used as the starting point for most image classification problem, with pre-trained values



Capturing Image Semantics with CNN

- Deep Neural Networks
- Supervised Learning
- Convolution
- Convolutional Neural Networks
- Image Classification

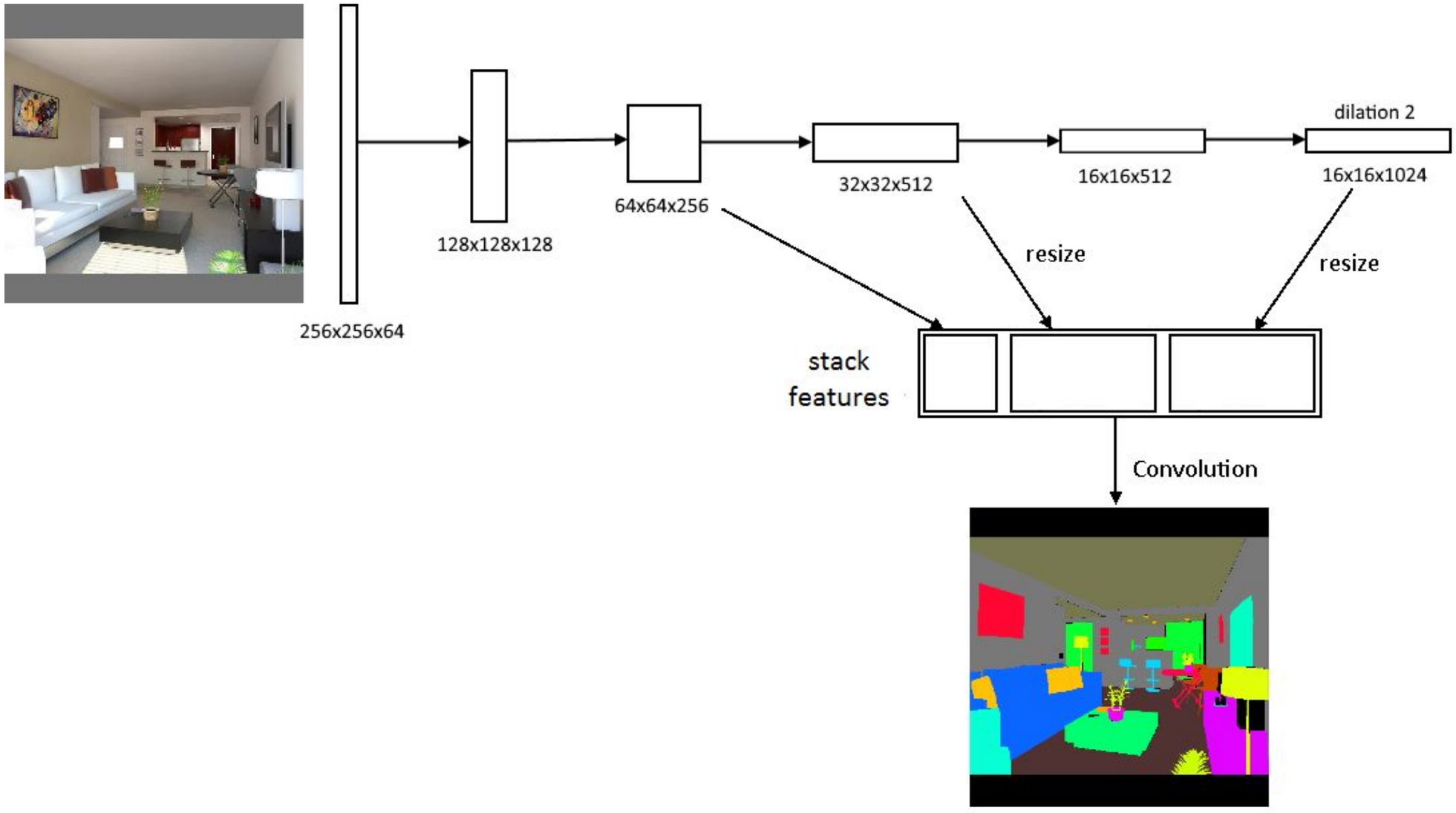


Dense CNN Predictions. Semantic Segmentation

- Training dataset with samples formed of images and their corresponding segmentation
- Fully Convolutional Architecture
- End-to-End Training

Example from the training dataset with image and expected output

Neural Network Model



Refining Results. Final Architecture

- The neural network is initialized with pre-trained parameters

- After training the neural network, results are easily enhanced through graph cuts

Framework

- Deep Learning framework specialized for images, built on top of Tensorflow
- Training was done using Stochastic Gradient Descent, accelerated with Adam optimizer
- Learning rate set at 0.001, batch size of 32 and trained for 20K steps
- Data augmentations include vertical and horizontal flipping and rotation
- Trained on Pascal segmentation dataset with accuracy on validation of 89%. Testing set is not provided

Future Work

- Explore results with a newer image classifier - ResNet
- Luminosity augmentations
- Train the neural network in the LAB color space
- Image segmentation from object classifiers without providing expected segmentations to the network.

Conclusions

- Deep neural networks provide the best image classification results
- Model parameters are transferable for a segmentation network
- Large dataset is required, which can be expensive to create
- Solutions exists, achieved through clustering methods built on top of image classifiers