LEARNABLE CONTEXTUAL REGULARIZATION FOR SEMANTIC SEGMENTATION OF INDOOR SCENE IMAGES

Jun Chu[1], *Xu Xiao*[2,3], *Gaofeng Meng*[3], *Lingfeng Wang*[3] *and Chunhong Pan*[3]

1.Institute of Computer Vision, Nanchang Hangkong University
2.School of Software, Nanchang Hangkong University
3.National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences
{chujun99602@163.com, xiaoxu64@qq.com, {gfmeng, lfwang, chpan}@nlpr.ia.ac.cn }

ABSTRACT

Semantic segmentation of indoor scene images has a wide range of applications. However, due to a large number of classes and uneven distribution in indoor scenes, mislabels are often made when facing small objects or boundary regions. Technically, contextual information may benefit for segmentation results, but has not yet been exploited sufficiently. In this paper, we propose a learnable contextual regularization model for enhancing the semantic segmentation results of color indoor scene images. This regularization model is combined with a deep convolutional segmentation network without significantly increasing the number of additional parameters. Our model, derived from the inherent contextual regularization on the indoor scene objects, benefits much from the learnable constraint layers bridging the lower layers and the higher layers in the deep convolutional network. The constraint layers are further integrated with a weighted L1-norm based contextual regularization between the neighboring pixels of RGB values to improve the segmentation results. Experimental results on NYUDv2 indoor scene dataset demonstrate the effectiveness and efficiency of the proposed method.

Index Terms— Deep convolutional neural networks, Semantic segmentation, Contextual constraints, End-to-end training

1. INTRODUCTION

As one of the important branches of image scene understanding, semantic image parsing has become a hot research topic in the fields of image processing and computer vision. Semantic segmentation is to assign a category level label to each pixel in an image. Recent several years have witnessed an increasing interest on pixel-wise image labeling [1–9]. Many recent approaches [5–8] have tried to directly adopt deep architectures designed for category prediction to pixel-wise labeling and achieved very encouraging results. However, image segmentation tasks in different scenes confront with different challenges. Semantic segmentation for indoor scenes still faces many difficulties, such as large variations of semantic categories, occlusions and overlaps between multiple indoor objects, lack of distinctive features and illumination changes and so on.

With the success of deep convolutional neural networks, CNNs are very popular in many visual recognition problems and also have been applied to the semantic segmentation problems. These models can be roughly divided into three categories. The first one is the currently most successful model based on fully convolution networks(FCN) [2]. This model can be trained in an end-to-end and pixels-to-pixels manner, which enables the model to adaptively combine sematic information. This information is from a deep and coarse layer with appearance information obtained from a shallow and fine layer to produce accurate and detailed segmentations. Unfortunately, due to its large receptive field and the lack of space constraints, each neuron corresponds to a large area of the original image (for example, FCN-32s has a 32-times magnification). Therefore, the abilities of FCNs to delineate boundaries, structures and shape are actually very poor. In order to integrate more contextual information, other approaches, e.g., [4, 8, 10, 11], propose to use techniques from graphical models, such as conditional random field (CRF), to introduce global context and structured information into FCNs.

The second class of models is to learn an encoding-decoding network for pixel-wise predictions, for example, SegNet [6], DeconvoNet [3] and semi-supervised Decoupled network [12]. DeconvNet applies the trained network to each proposal in an input image and construct the final semantic segmentation map by combining the results from all proposals in a simple manner. The encoder network of DeconvNet and decoupled network consist of the fully connected layers transplanted from the VGG-16 [13] network. A large number of parameters of the entire network often make their training very difficult and thus require additional steps such as the use of region proposals to enable training tractable. Moreover, time complexity will be increased significantly during inference the proposals. In comparison, SegNet, also trained end-to-end, does not use multi-stage training [2] or region proposals [3]. The key component of SegNet is the decoder network, which consists of a hierarchy of decoders that are one-to-one correspondence to each encoder.

The third class of methods for semantic segmentation are based on multi-scale deep architectures [7, 11]. Input images are resized into many scales and the corresponding multiple scales features are extracted or combined with feature maps from different layers of deep architecture [5, 14]. Multiple scales features can provide both local and global context [15], while early layers retain more details of class boundaries. However, parameter numbers in early layers and late layers are very different. Thereby, the training of multi-scale architectures is very difficult. Besides, the inference is also expensive with multiple convolutional pathways for feature extraction. Several of the recently proposed deep architectures for segmentation are not feed-forward in inference time [3, 12, 16]. They require either MAP inference over a CRF [10, 11] or region proposals [3] for inference.

Although region proposals can make the results more accurate, since they help to establish benchmarks that are more easily repeat-

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61663031, 61370039), Jiangxi Province Key Research Project(Grant No.20161BBE50085), the Beijing Nature Science Foundation (Grant No. 4162064)

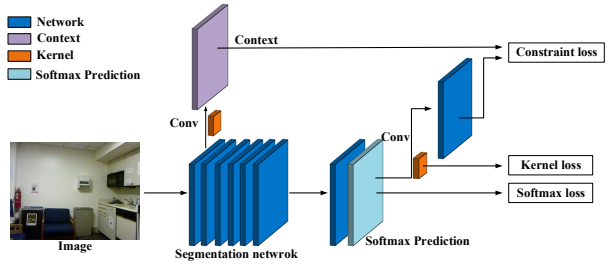

Fig. 1. The proposed architecture of deep neural networks with learnable contextual regularization for semantic segmentation of indoor scene images.

able, overlaps and occlusions that commonly exist in indoor scene images often introduce local confusions, making the accurate segmentation much difficult. To address this issue, we introduce a learnable contextual regularization in this paper to help clarify these local confusions. Our method benefits from three main contributions:

- A set of weight matrixes are constructed from the input images. These weight matrixes are computed in the form of convolution by some learnable kernels. To regularize the last classification layer with contextual information, a constraint loss layer is introduced, which plays the same role as softmax loss layer, for supervising the training of the network.
- These weight matrixes can impose additional contextual regularization of all the pixels on the segmentation network output, without relying on the labels. In addition, a kernel regularization is introduced to supervise the weight matrixes learning for local optimization. The learnable kernels help capturing more inherent and adaptive contextual weight matrixes to constrain the coarse segmentation results.
- Our method can also be applied to general models without significantly introducing additional parameters during network training. For this reason, the contextual regularization does not introduce much computational overhead versus training and evaluating a standard network, e.g., FCNs, while improving the performance significantly.

2. PROPOSED MODEL

Our proposed model for semantic segmentation of indoor scene images is shown in Fig. 1. They are trained end-to-end to optimize the output semantic segmentation quality. In this figure, the contextual regulation is modeled as a loss layer with some convolutional kernels to constrain the softmax predictions. In the architecture, the kernel of the contextual regulation layer is learnable, enabling the layer more robust and flexible to different types of indoor scenes. We will learn the context relationship both from the lowest layer and the highest segmentation output layer by means of weight sharing.

2.1. Contextual Regularization

An image can be considered as a combination of image patches in orders. Based on this assumption, we derive a contextual regularization from an input RGB image and use it to constrain the segmentation results of a deep segmentation network.

To this end, a weighting function $W(i,j)$ is introduced to model the contextual constraints of the segmentation results, i.e.,

$$W(i,j)(f(i)-f(j)), \tag{1}$$

where i and j are two neighboring pixels, and f is the last output that will be regularized. The weighting function permits candidate values from 0 to 1. For $W(i,j)=0$, the contextual constraint between $f(i)$ and $f(j)$ will be terminated. Intuitively, the weight function plays a switch role to control whether the constraint between i and j will be canceled. By learning a reasonable weighting function from lower layers of a segmentation network and applying it to the output layer of the network to regularize the segmentation results, we actually use this weighting function as an intermediary to transit low-level contextual information to upper-level semantic results.

Now the question is how to choose a reasonable $W(i,j)$. Obviously, the optimal $W(i,j)$ is closely related to the RGB value difference between pixel i and j. If two neighboring pixels in original image have similar RGB value, weight function will put significant constraints on the two pixels. Here we consider 8-neighbourhood of every pixel. In another word, the more difference between the neighboring pixels the smaller value of the $W(i,j)$ is. Consequently, the color difference of local pixels are necessary elements to construct the weight matrixes. Here below is the construction of such weighting function. It bases on the squared difference between the color vectors of two neighboring pixels [17], given as below:

$$W(i,j)=e^{-||\mathbf{I}(i)-\mathbf{I}(j)||^2/2\sigma^2}, \tag{2}$$

where σ is a prescribed parameter, $\mathbf{I}$ is the RGB image.

Integrating the weighted contextual constraints in the whole image domain leads to the following,

$$\sum_{i,j\in I} W(i,j) \mid f(i)-f(j) \mid, \tag{3}$$

where I is the index set of image. Each pixel in the results are weighted by contextual regularization.

2.2. Learnable Constraint Loss Layer

We construct neighboring pixel affinities to regularize the contextual information of all pixels in the coarse score maps. Contextual regularization means that building a learnable weight matrix from the input RGB image and then combine them in a global manner.

The learnable constraint loss function evident from Eq. (3), we employ L_1-norm which is more robust to outliers than L_2-norm and the boundary effect is better. To facilitate the computation, we introduce a set of convolutional operators and exchange summation order. We further give the discrete form of Eq. (3) as below:

$$\sum_{j\in\omega_i}\sum_{i\in I}\omega_{ij}|(D_j\otimes f)|, \tag{4}$$

or more compactly

$$\sum_{j\in\omega}||W_j\circ(D_j\otimes f)||_1, \tag{5}$$

where ω_{ij} is the discrete versions of $W(i,j)$, ω_{i} is the index set of neighbourhood of pixel i, D_j is the convolutional kernel, ω is an index set of the convolutional operators, $\circ$ represents the element-wise multiplication operator, $\otimes$ stands for the convolution operation.

The learnable constraint loss layer takes the context of the coarse score maps compared with the learnable contextual features and weighted them. The layer takes two inputs: (1) the current contextual maps of softmax output, and (2) the contextual weight feature maps learned from the input RGB image. Both of the two inputs are convolutional by j-dimensional kernels $D=(D_1\cdots D_{\mathrm{j}})$ with N-dimensional input $f=(f_1\cdots f_{\mathrm{N}})$.

The contextual weight maps weighted on softmax output. This is because softmax function is a generalization of the logistic function that maps a length-p vector of real values to a length-K vector of values, and normalizes classification probability value between 0 and 1. The expression of weight maps is as follows:

$$\mathrm{W}_j(i) = e^{-\left(\sum_c |(D_j \otimes I^c)_i|^2\right)/2\sigma^2}, \tag{6}$$

where σ is a prescribed parameter.

The last classification layer produces a bank of N feature maps for N category, and thus has a N-channels output for softmax. Using this notation, with N-dimensional input $f = (f_1 \cdots f_{\mathrm{N}})$ in our layer we can then formulate the constraint loss objective as:

$$l_c = \lambda \sum_{n=1}^{N} \sum_{j \in \omega} ||W_j \circ (D_j \otimes f_n)||_1, \tag{7}$$

where the constraint loss l_c is the accumulation of all pixels of the N-maps and λ is the super parameter.

Various convolutional kernels lead to different feature maps which represent relevant context information. Learnable kernels are more robust and appropriate to obtain particular context for the current task. In deep convolutional network, the lower layer contains more contextual boundary information than higher layer.

2.3. Optimization Of Our Model

A) Global Optimization The last step in our method is to combine the contextual weight maps with the coarse segmentation from the network softmax layer to produce an improved segmentation results. The traditional segmentation network only use a softmax loss layer to optimize. The softmax loss in semantic segmentation considers pixel-wise loss for the right classified while ignores the wrong classified neighboring pixels that share similar colour values. We introduce the contextual constraint loss, which can better utilize the local relationship of neighboring pixels and improve the initial network segmentation. We do this by introducing a global energy function that utilizes the contextual constraint loss and softmax loss as:

$$L = l_s + l_c, \tag{8}$$

where l_s is softmax loss, and l_c is our contextual regularization loss. We use back propagation and chain rule to compute derivatives with respect the input data.

B) Local Optimization For our purpose, we need control the kernels to learn the right context regularization features corresponding to the segmentation task. With more comprehensive consideration, we import a convolution kernel loss layer. Let l_k be the convolution kernel loss:

$$l_k = \sum_{i \in \omega} \{\alpha(\sum d_i^2 - 1)^2 + \beta \sum d_i^2\}, \tag{9}$$

where α, β are the regularization parameters for balancing and usually set very large. d_i is the i-th convolution kernel, the total number of kernels is ω.

3. EXPERIMENTS

3.1. Dataset and Metrics

Other challenges such as Pascal VOC12 [18] salient object segmentation have occupied researchers more, but indoor scene segmentation is more challenging and has more practical applications such as in robotics. To evaluate our proposed method, we implemented a series of experiments on the public NYUDv2 dataset. This dataset is an RGB-D collected using the Microsoft Kinect. It has 1,449 RGB-D images, both RGB image and depth image. HHA images [19] encoder the depth image of three channels (horizontal disparity, height above ground, and the angle the pixel's local surface normal). HHA images have enough common structure with RGB images that a CNN network can learn a suitable representation for them. We evaluate our method on semantic class sets with 4 and 40 labels, described in [20] and [19] respectively. The 4-class segmentation task uses high- level category labels floor, structure, furniture and props, while the 40-class tasks use different sets of more fine-grained categories. We report results on the standard split of 795 training images and 654 testing images.

Matrices Four metrics are used to evaluate our metod. For the 4-class task we use pixel accuracy and mean accuracy. For the 40-class task, we report metrics from common semantic segmentation and scene parsing evaluations that are variations on pixel accuracy and region intersection over union (IU) introduced in FCNs [2].

- pixel accuracy: $\sum_i n_{ii} / \sum_i t_i$
- mean accuracy: $(1/n_{cl}) \sum_i n_{ii}/t_i$
- mean IU: $(1/n_{cl}) \sum_i n_{ii}/(t_i + \sum_j n_{ji} - n_{ii})$
- frequency weighted IU: $(\sum_k t_k)^{-1} \sum_i t_i n_{ii}/(t_i + \sum_j n_{ji} - n_{ii})$

3.2. Compared Methods and Network Parameters

Two models are compared with our method: FCNs [2] and SegNet [**?**]. As described in Section2, our learnable constraint layer is added to the last coarse classification layer of the two models. We employ the VGG-16 network [13] which has been pre-trained on ImageNet and used SGD [21] with a fixed learning rate and momentum for training. We implement the proposed network based on Caffe [22] framework on NVIDIA K20 GPU with 12-GB memory. The constraints parameters are $\sigma = 0.25, \lambda = 100, \alpha = 1000, \beta = 1000$.

FCNs We use momentum of 0.99 and a batch size of one , a weight decay of 0.0005, the learning rate is fixed to e-9 with no normalize of softmax loss. Note that 100 zero-padding operation is used to FCNs, so the first layer output size is bigger than input image size and many nonessential information in outputs. We chose the second convolutional layer and resize the output by bilinear interpolation to have the same spatial resolution with the softmax output.

SegNet The learning rate was fixed to 0.001 and momentum to 0.9, the mini-batch size is four.

3.3. Experimental Results

Experiments on NYUDv2 dataset are conducted to prove the effective of our method. Satisfyingly, contextual regularization turns out can benefit for accuracy improvement in RGB and HHA images.

We compare the segmentation performances with FCNs and SegNet, and some representative results are shown in Fig. 2. Fewer mislabels are made for large-sized objects in FCNs results. Typically, small objects can also be correctly detected even embedded into other classes. The 32-times upsample in FCNs makes the boundary coarser, and even the structure of objects would be damaged seriously. Our method can distinguish small objects from large-sized ones. By adding contextual regularization, objects in the adjacent area can also be correctly classified, resulting in more accurate boundaries. The accuracies of 40-class and 4-class segmentation present them in Table 1 and Table 2. Our method improves fw-IU about 3.9% and mean-IU about 2.0% point with respect to FCN(32s-RGB).

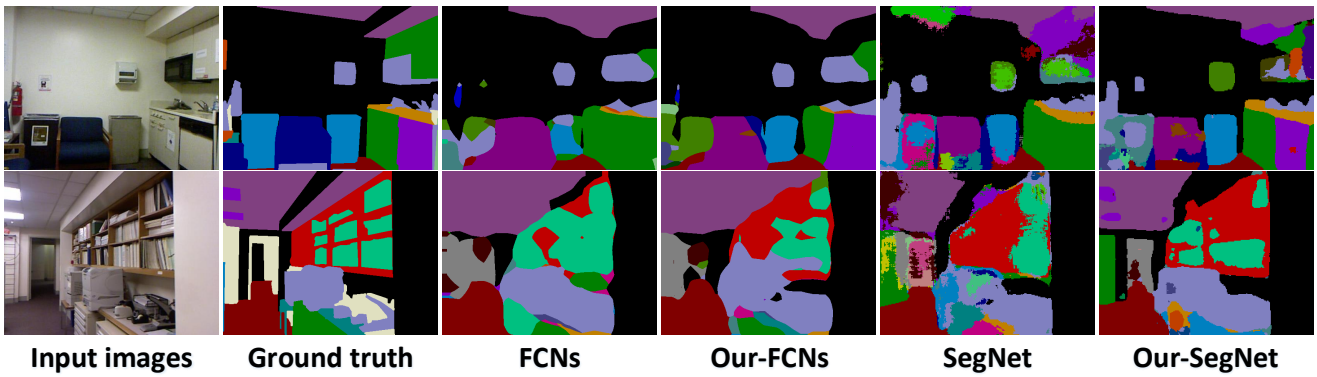

Fig. 2. Examples of semantic segmentation results on the NYUDV2 dataset RGB images. Respectively, Our-FCNs and Our-SegNet are the results by our method based on FCNs and SegNet. Compared with original FCNs, our method can distinguish small objects from large-sized ones. By adding contextual regularization, objects in the adjacent area can also be correctly classified, resulting in more accurate boundaries.

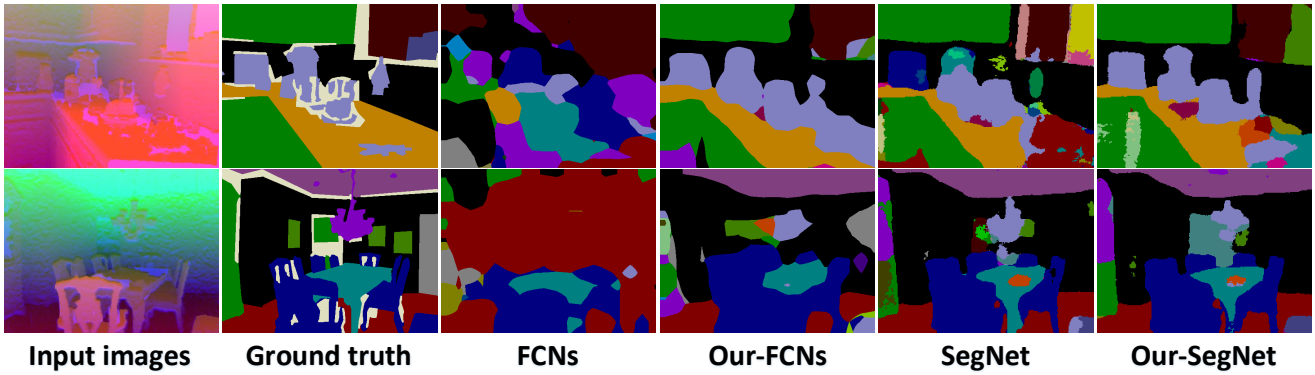

Fig. 3. Examples of semantic segmentation results on the NYUDV2 dataset HHA images. HHA images [19] encoder the depth image of three channels (horizontal disparity, height above ground, and the angle the pixel's local surface normal). Compared with original SegNet, local constraints are used in our method to achieve more consistent classification results in adjacent regions which proves that our method is also effective for HHA images.

Table 1. FCN-32s 40-class segmentation accuracy

	pixel-acc	mean-acc	mean-IU	fw-IU
FCN(32s-RGB) [2]	61.8	44.7	31.6	46.0
FCN(32s-HHA) [2]	58.3	35.7	25.2	41.7
Gupta et al [23]	60.3	-	28.6	47
our(32s-RGB)	**62.5**	**46.3**	**33.6**	**49.9**
our(32s-HHA)	60.7	45.3	32.0	47.9

Table 2. FCN-32s 4-class segmentation accuracy

	pixel-acc	mean-acc
Couprie et al [24]	64.5	63.5
Khan et al [25]	69.2	65.6
Stuckler et al [26]	70.9	67.0
Muller et al [27]	72.3	71.9
Gupta et al [23]	78	-
our(32s-RGB)	**81.1**	**80.3**

Due to max-unpooling operation in SegNet, scattered points would exist around boundary regions. Compared with SegNet, local constraints are used in our method to achieve more consistent classification results in adjacent regions. As for HHA images, the depth information can be learned in the training process. Especially, our constraints can follow spatial distributions of original HHA images in the regions where depth information changes severely. The class average accuracy and mean I/U metric are little poor, also at the same level as the hand engineered method which input image size is 425×520. Motivated by batch normalization layer [28], the more parameters and memory are required for running SegNet. Consider the restrictions of the GPU memory, we resized the image to 200×264 by original proportion for the minimum batch size. The quantitative results of SegNet with 40-class are presented in Table 3, where our method improves pixel accuracy about 8.3% (in RGB images) and 1.5% (in HHA images) point respectively.

Table 3. SegNet 40-class segmentation accuracy

	pixel-acc	mean-acc	mean-IU	fw-IU
SegNet-RGB [6]	46.8	22.3	14.2	33.4
our(SegNet-RGB)	55.1	**32.1**	**22.3**	39.1
SegNet-HHA	54.1	30.5	21.0	38.5
our(SegNet-HHA)	**55.6**	31.7	21.8	**39.9**

4. CONCLUSIONS

In this paper, we proposed a learnable contextual regularization for semantic segmentation of indoor scene images. This regularization term can be flexibly used in many segmentation networks, such as FCNs and SegNet. We consider not only the low-level information but also the upper-level information. We construct a set of learnable weight matrixes from the low-level that can impose additional contextual constraints of all the pixels on the segmentation network output, not limited to the label pixels. Our method helps models adapt to different segmentation tasks. In future work, we will try to use more multiple information, and standardize the middle tier effectively to guide the net learning more richer context and get more effective segmentation results.

5. REFERENCES

[1] Clement Farabet, C. Couprie, Laurent Najman, and Yann Lecun, "Learning hierarchical features for scene labeling," *IEEE Transactions on Software Engineering*, vol. 35, no. 8, pp. 1915–1929, 2013.

[2] Evan Shelhamer, Jonathon Long, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," vol. 79, no. 10, pp. 1337–1342, 2015.

[3] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han, "Learning deconvolution network for semantic segmentation," pp. 1520–1528, 2015.

[4] Shuai Zheng, Sadeep Jayasumana, Bernardino Romera-Paredes, Vibhav Vineet, Zhizhong Su, Dalong Du, Chang Huang, and Philip HS Torr, "Conditional random fields as recurrent neural networks," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1529–1537.

[5] Wei Liu, Andrew Rabinovich, and Alexander C Berg, "Parsenet: Looking wider to see better," *arXiv preprint arXiv:1506.04579*, 2015.

[6] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *arXiv preprint arXiv:1511.00561*, 2015.

[7] David Eigen and Rob Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2650–2658.

[8] Liang-Chieh Chen, Jonathan T Barron, George Papandreou, Kevin Murphy, and Alan L Yuille, "Semantic image segmentation with task-specific edge detection using cnns and a discriminatively trained domain transform," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4545–4554.

[9] Nico H?ft, Hannes Schulz, Sven Behnke, and Nico H?ft, "Fast semantic segmentation of rgb-d scenes with gpu-accelerated deep neural networks," in *German Conference on Artificial Intelligence*, 2014, pp. 80–85.

[10] Alexander G Schwing and Raquel Urtasun, "Fully connected deep structured networks," *arXiv preprint arXiv:1503.02351*, 2015.

[11] Guosheng Lin, Chunhua Shen, Anton van den Hengel, and Ian Reid, "Efficient piecewise training of deep structured models for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3194–3203.

[12] Seunghoon Hong, Hyeonwoo Noh, and Bohyung Han, "Decoupled deep neural network for semi-supervised semantic segmentation," in *Advances in Neural Information Processing Systems*, 2015, pp. 1495–1503.

[13] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[14] B. Hariharan, P. Arbelez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," pp. 447–456, 2015.

[15] Mohammadreza Mostajabi, Payman Yadollahpour, and Gregory Shakhnarovich, "Feedforward semantic segmentation with zoom-out features," pp. 3376–3385, 2014.

[16] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," *arXiv preprint arXiv:1412.7062*, 2014.

[17] Gaofeng Meng, Ying Wang, Jiangyong Duan, Shiming Xiang, and Chunhong Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *IEEE International Conference on Computer Vision*, 2013, pp. 617–624.

[18] Mark Everingham, S. M. Ali Eslami, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman, "The pascal visual object classes challenge: A retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98–136, 2015.

[19] S. Gupta, P. Arbelaez, and J. Malik, "Perceptual organization and recognition of indoor scenes from rgb-d images," 2013, pp. 564–571.

[20] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus, "Indoor segmentation and support inference from rgbd images," in *European Conference on Computer Vision*, 2012, pp. 746–760.

[21] Lon Bottou, *Large-Scale Machine Learning with Stochastic Gradient Descent*, Physica-Verlag HD, 2010.

[22] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[23] Saurabh Gupta, Ross Girshick, Pablo Arbelez, and Jitendra Malik, *Learning Rich Features from RGB-D Images for Object Detection and Segmentation*, Springer International Publishing, 2014.

[24] Camille Couprie, Clment Farabet, Laurent Najman, and Yann Lecun, "Indoor semantic segmentation using depth information," *Eprint Arxiv*, 2013.

[25] Salman Hameed Khan, Mohammed Bennamoun, Ferdous Sohel, and Roberto Togneri, *Geometry Driven Semantic Labeling of Indoor Scenes*, Springer International Publishing, 2014.

[26] J?rg Stckler, Benedikt Waldvogel, Hannes Schulz, and Sven Behnke, "Dense real-time mapping of object-class semantics from rgb-d video," *Journal of Real-Time Image Processing*, vol. 10, no. 4, pp. 599–609, 2015.

[27] A. C. Muller and S. Behnke, "Learning depth-sensitive conditional random fields for semantic segmentation of rgb-d images," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 6232 – 6237.

[28] Sergey Ioffe and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.