



1

2

3

4

5

6



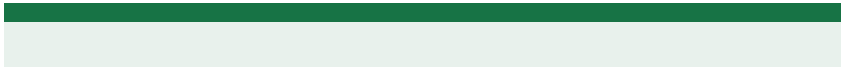
oooo

ooo



o

oo



(Semantic Image Segmentation)

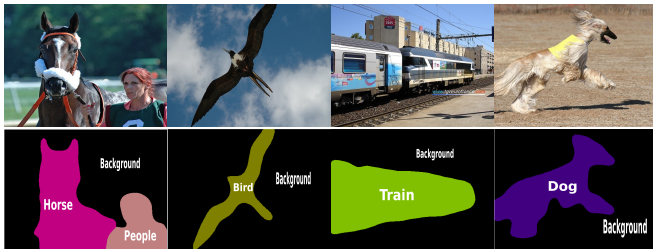
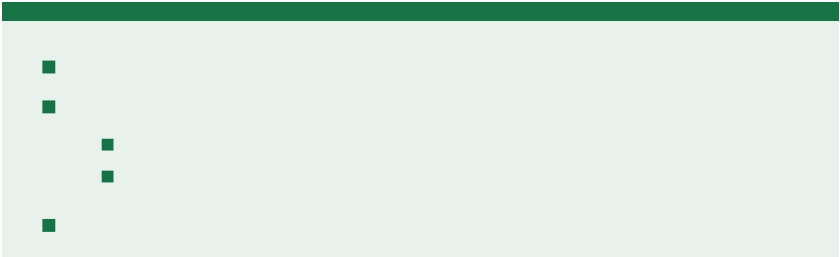
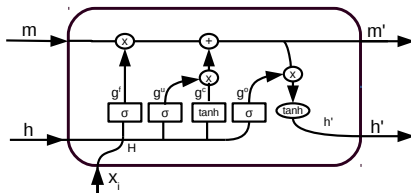
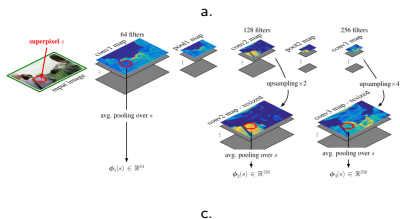
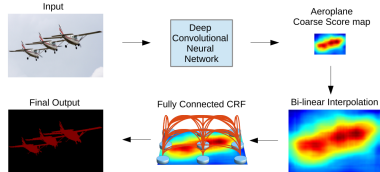
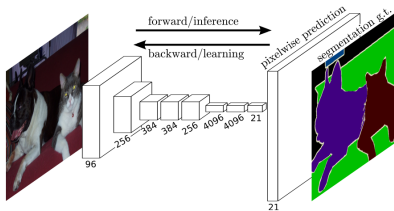


Figure 1:



}



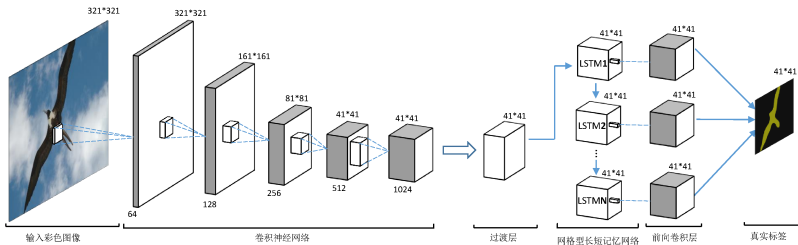


Figure 2:

-
-
-
-



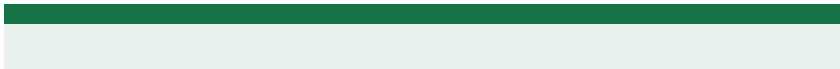
oooo

ooo



o

oo



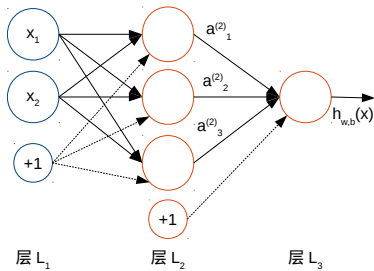


Figure 3:

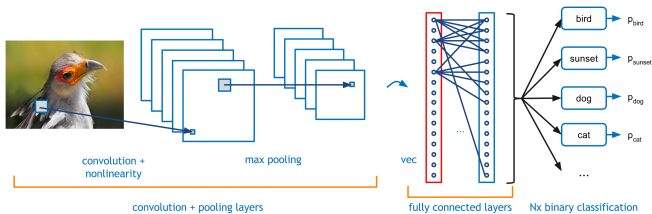
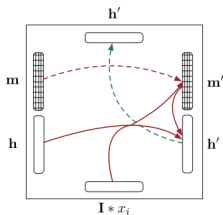


Figure 4:

-
-
-



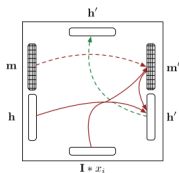
Standard LSTM block

Figure 5:

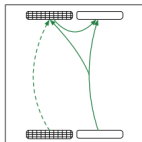
$$\begin{aligned}
 g^u &= \delta(W^u * H) \\
 g^f &= \delta(W^f * H) \\
 g^o &= \delta(W^o * H) \\
 g^c &= \tanh(W^c * H) \\
 m' &= g^f \odot m + g^u \odot g^c \\
 h' &= \tanh(g^o \odot m') \\
 H &= \begin{bmatrix} I * x_i \\ h \end{bmatrix}
 \end{aligned} \tag{1}$$

$$(h', m') = \text{LSTM}(H, m, W)$$

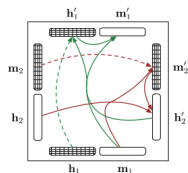
$$W, W^u, W^f, W^o, W^c$$



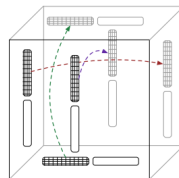
Standard LSTM block



1d Grid LSTM Block



2d Grid LSTM block



3d Grid LSTM Block

Figure 6:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_N \end{bmatrix} \quad (2)$$

$$\begin{aligned} (\mathbf{h}'_1, \mathbf{m}'_1) &= \text{LSTM}(\mathbf{H}, \mathbf{m}_1, \mathbf{W}_1) \\ &\vdots \\ (\mathbf{h}'_N, \mathbf{m}'_N) &= \text{LSTM}(\mathbf{H}, \mathbf{m}_N, \mathbf{W}_N) \end{aligned} \quad (3)$$

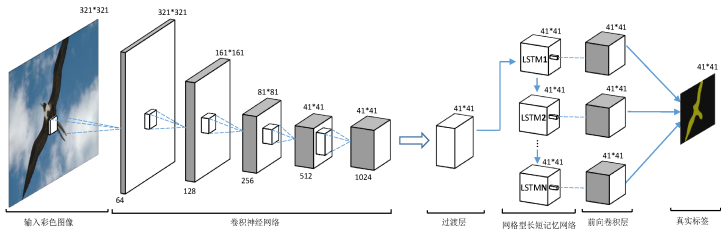


Figure 7:

-
-

- VGG₁₆¹,
- 6.5²

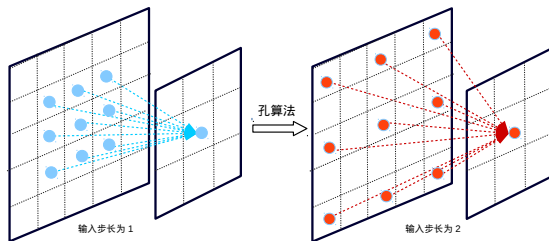


Figure 8: ”

¹Simonyan & Zissermanet, Very deep Convolutional Networks For Large-scale Image Recognition, ICLR 2015

²Chen et al, DeepLab-LargeFOV, ICLR 2015

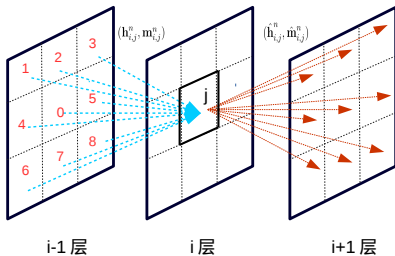


Figure 9:

$$\begin{aligned}
 (\hat{\mathbf{h}}_{i,j}^0, \hat{\mathbf{m}}_{i,j}^0) &= \text{LSTM}(\mathbf{H}_{i,j}, \mathbf{m}_{i,j}^0, \mathbf{W}_i) \\
 (\hat{\mathbf{h}}_{i,j}^1, \hat{\mathbf{m}}_{i,j}^1) &= \text{LSTM}(\mathbf{H}_{i,j}, \mathbf{m}_{i,j}^1, \mathbf{W}_i) \\
 &\vdots \\
 (\hat{\mathbf{h}}_{i,j}^N, \hat{\mathbf{m}}_{i,j}^N) &= \text{LSTM}(\mathbf{H}_{i,j}, \mathbf{m}_{i,j}^N, \mathbf{W}_i) \\
 \mathbf{H}_{i,j} &= [\mathbf{h}_{i,j}^0 \ \mathbf{h}_{i,j}^1 \ \dots \ \mathbf{h}_{i,j}^N]^T
 \end{aligned} \tag{4}$$

-
-
-



VOC 2012

SIFT FLOW

Pascal VOC 2012 & SIFT FLOW



Figure 10: VOC 2012



Figure 11: SIFT FLOW

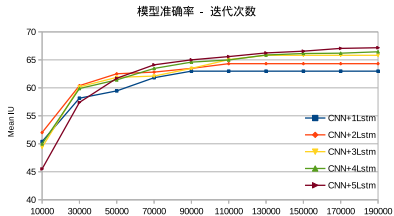
-
-

$$n_{ij} \quad i \quad j \quad n_{cl} \quad t_i = \sum_{j=1}^{n_{cl}} n_{ij} \quad i$$

$$= \sum_{i=1}^{n_{cl}} n_{ii} / \sum_{i=1}^{n_{cl}} t_i$$

$$= \frac{1}{n_{cl}} \sum_{i=1}^{n_{cl}} (n_{ii} / t_i) \quad (5)$$

$$\text{Mean IU} = \frac{1}{n_{cl}} \sum_{i=1}^{n_{cl}} \frac{n_{ii}}{t_i + \sum_j^{n_{cl}} n_{ji} - n_{ii}}$$



†
† 7.5%

Figure 12:



Figure 13:

Method	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	shep	sofa	train	tv	mIoU
SDS ³	63.3	25.7	63.0	39.8	59.2	70.9	61.4	54.9	16.8	45.0	48.2	50.5	51.0	57.7	63.3	31.8	58.7	31.2	55.7	48.5	51.6
FCN-8s ⁴	76.8	34.2	68.9	49.4	60.3	75.3	74.7	77.6	21.4	62.5	46.8	71.8	63.9	76.5	73.9	45.2	72.4	37.4	70.9	55.1	62.2
TTI-zoomout-16 ⁵	81.9	35.1	78.2	57.4	56.5	80.5	74.0	79.8	22.4	69.6	53.7	74.0	76.0	76.6	68.8	44.3	70.2	40.2	68.9	55.3	64.4
DeepLab-CRF ⁶	78.4	33.1	78.2	55.6	65.3	81.3	75.5	78.6	25.3	69.2	52.7	75.2	69.0	79.1	77.6	54.7	78.3	45.1	73.3	56.2	66.4
CNN+5LSTM	80.2	35.3	74.1	54.4	64.4	87.3	81.1	80.6	22.7	73.6	58.8	73.9	73.7	78.7	77.4	50.2	80.0	47.9	76.5	63.1	67.9

Table 1:



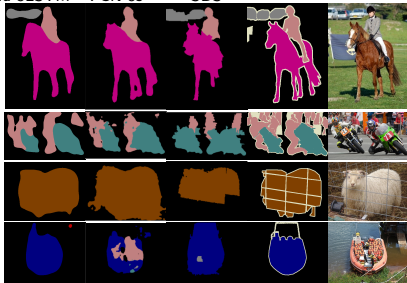
³Simultaneous Detection and Segmentation, ECCV 2014

⁴Fully convolutional networks for semantic segmentation, CVPR 2015

⁵Feedforward semantic segmentation with zoom-out features, CVPR 2015

⁶Semantic image segmentation with deep convolutional nets and fully connected crfs, ICLR 2015

Grid-5LSTM FCN-8s SDS



(a)



(b)

Figure 14: Grid-5LSTM

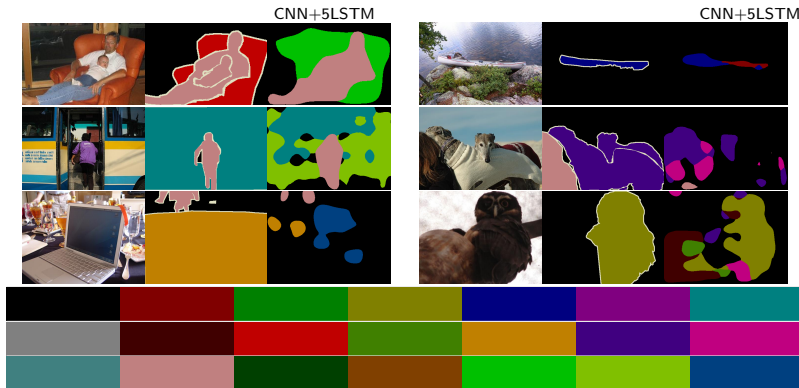


Figure 15:

SIFT FLOW

Method	Pixel Acc.	Mean Acc.	Mean IU.
Liu et al. ⁷	76.7	-	-
Tighe et al. ⁸	78.6	39.2	-
FCN-16s ⁹	85.2	51.7	39.5
Deeplab-LargeFOV ¹⁰	85.6	51.2	39.7
Grid-5LSTM	86.2	51.0	41.2

Table 2:

⁷ Sift flow: Dense correspondence across scenes and its applications, PAMI 2011

⁸ Finding things: Image parsing with regions and per-exemplar detectors, CVPR 2013

⁹ Fully convolutional networks for semantic segmentation, CVPR 2015

¹⁰ Semantic image segmentation with deep convolutional nets and fully connected crfs, ICLR 2015







Q & A

Questions?

Thank you!