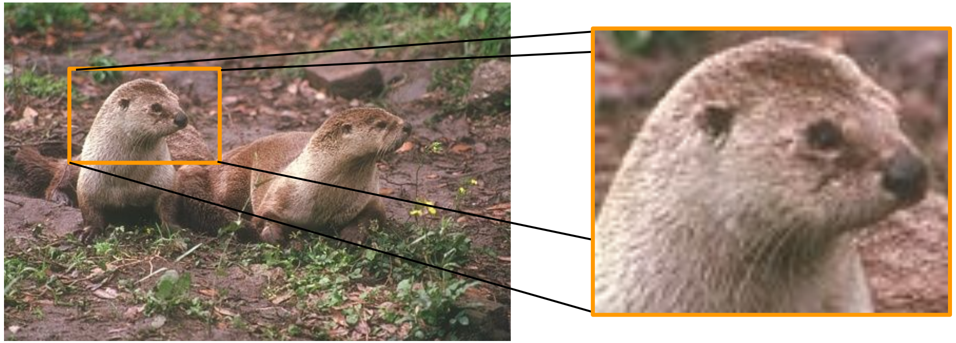
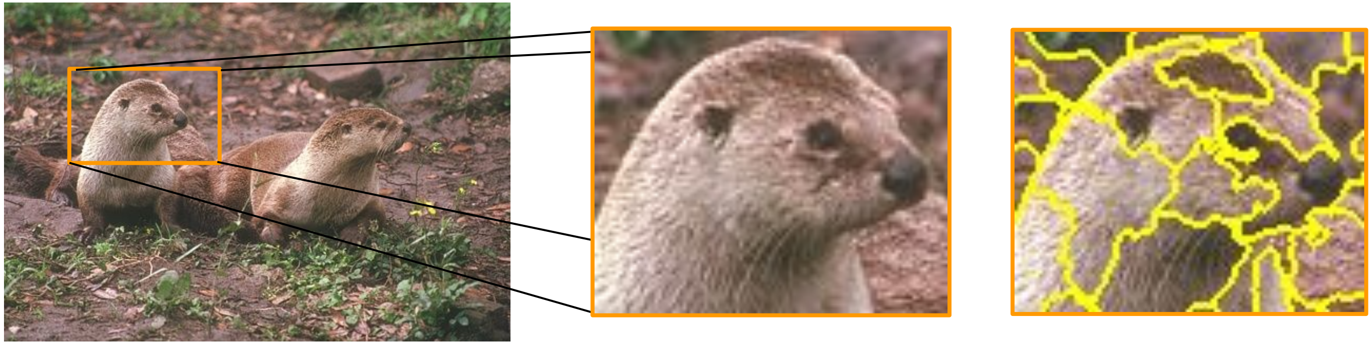
Soft Superpixel Neighborhood Attention

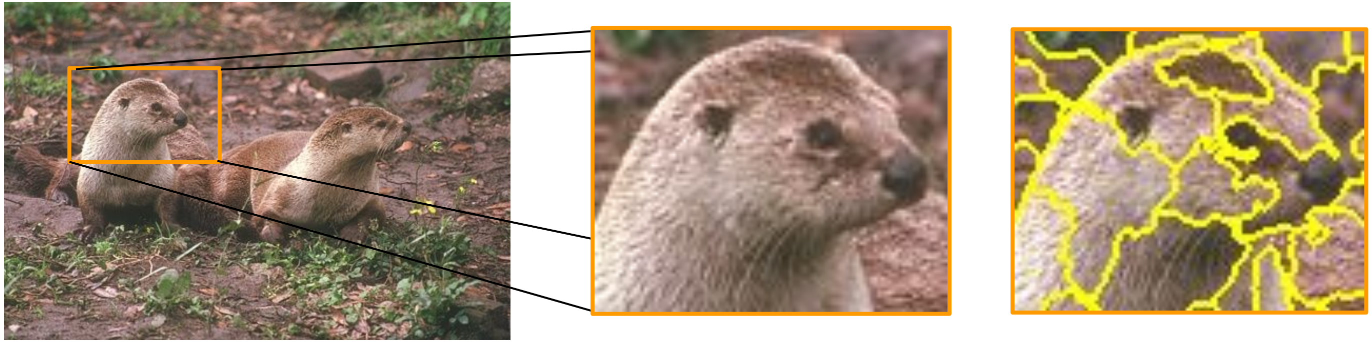
Images contain
deformable boundaries



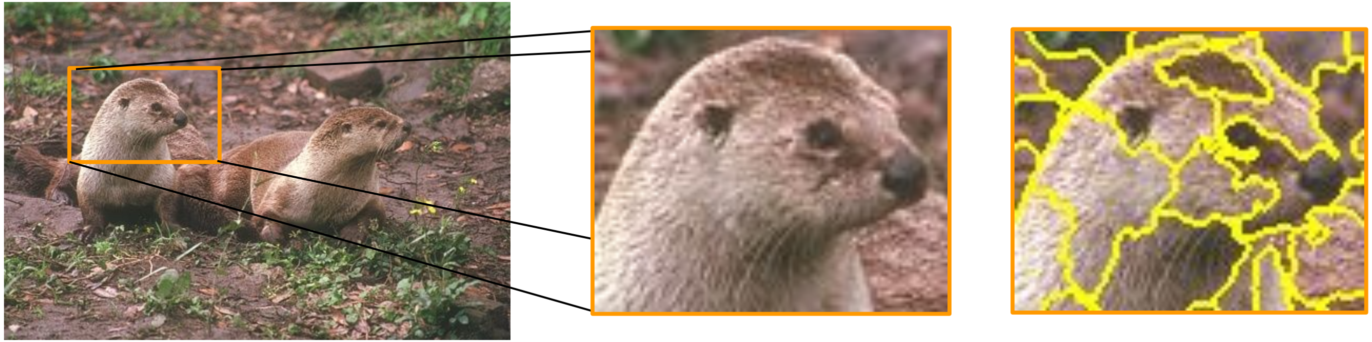
Images contain
deformable boundaries



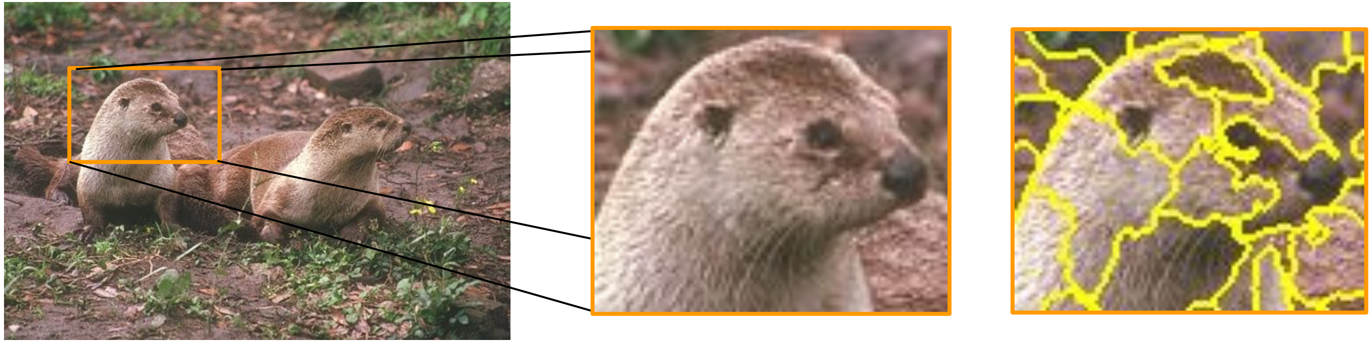
Images contain
deformable boundaries



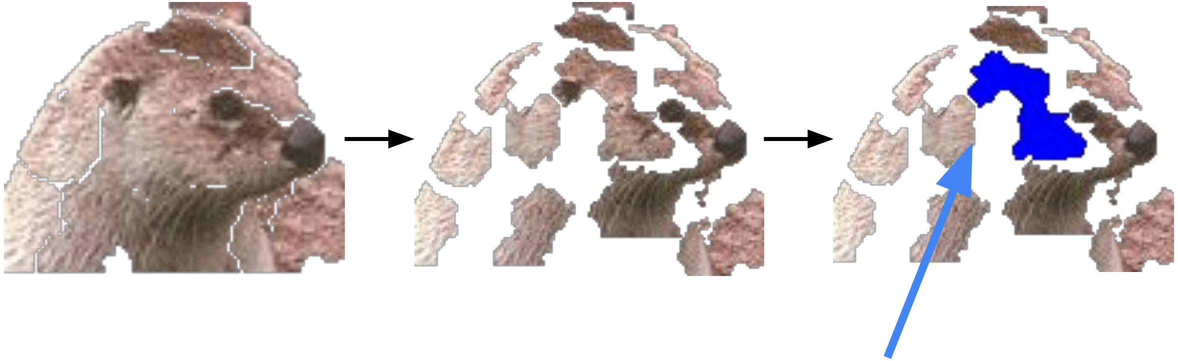
Images contain
deformable boundaries



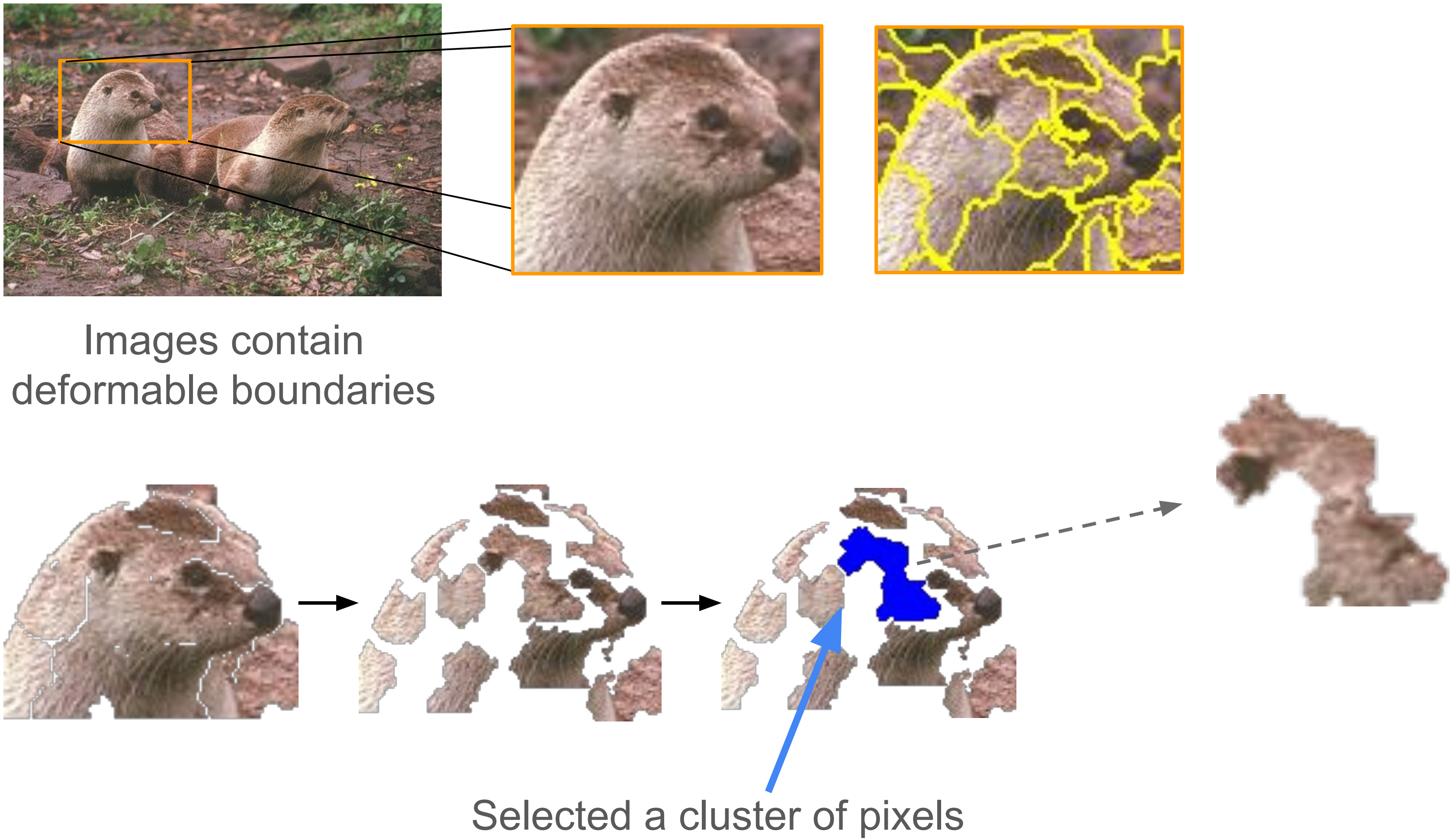
Images contain deformable boundaries

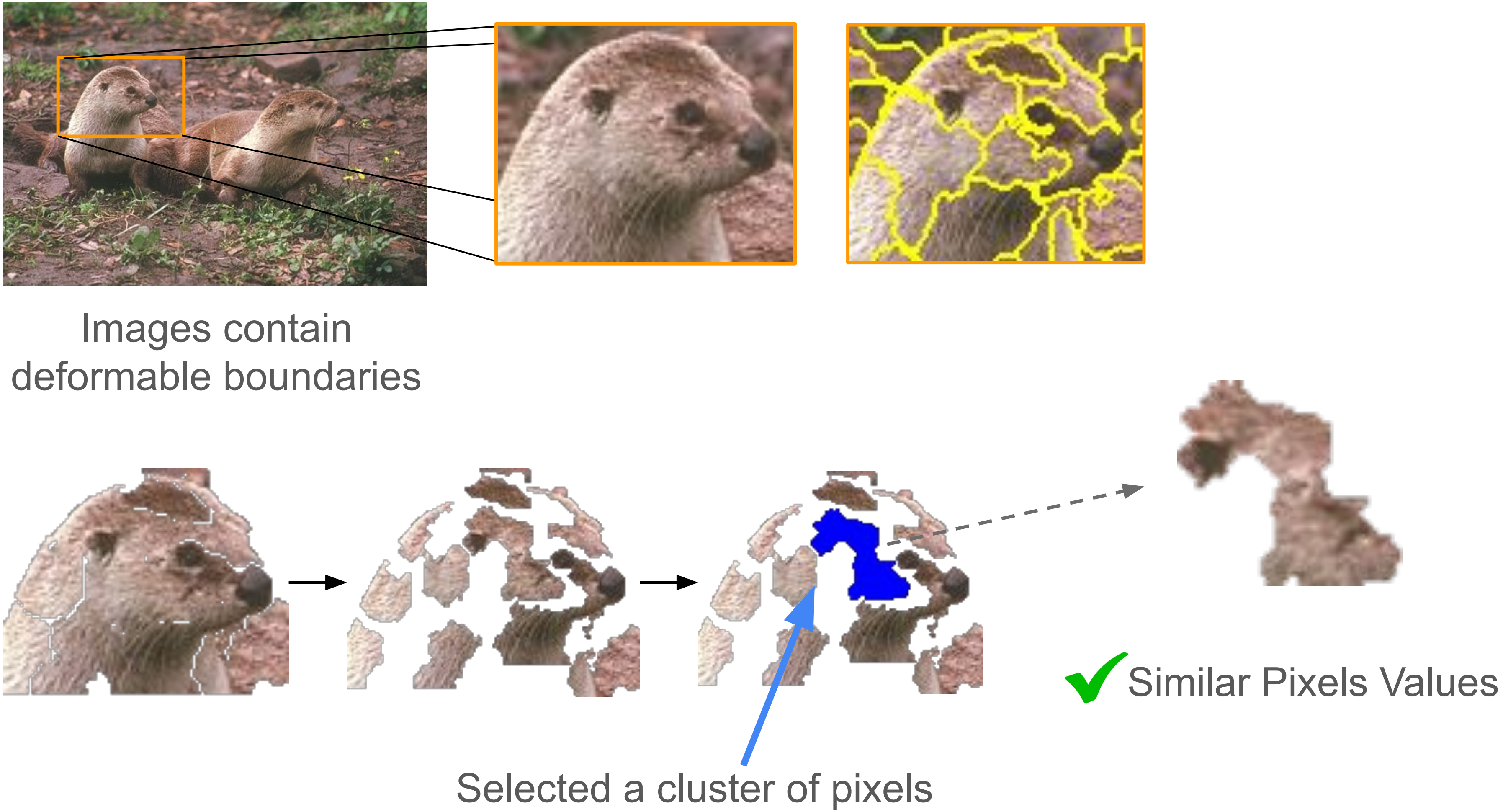


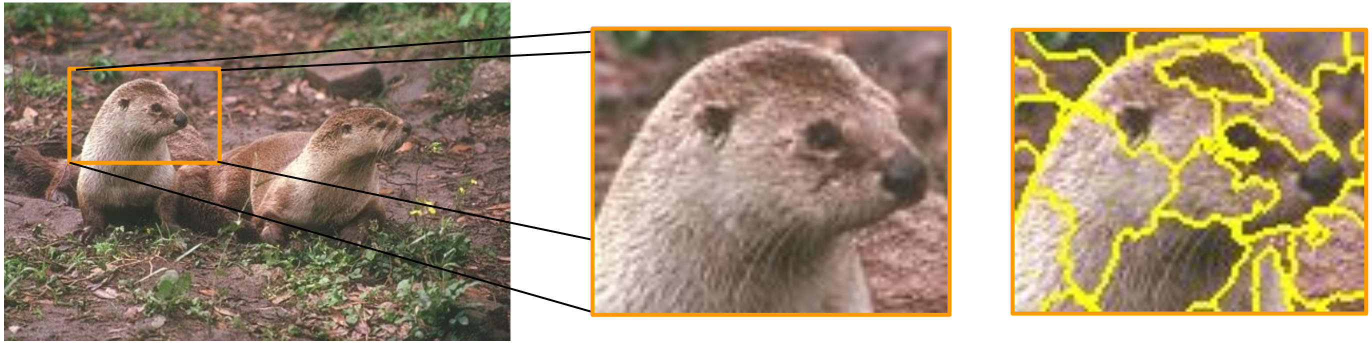
Images contain deformable boundaries



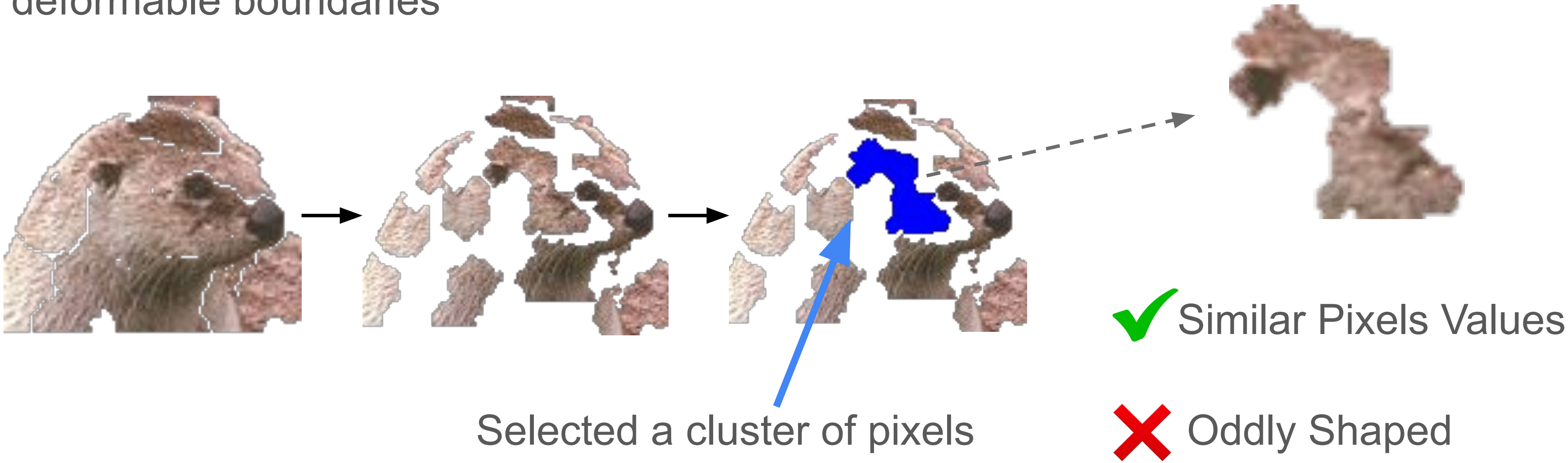
Selected a cluster of pixels

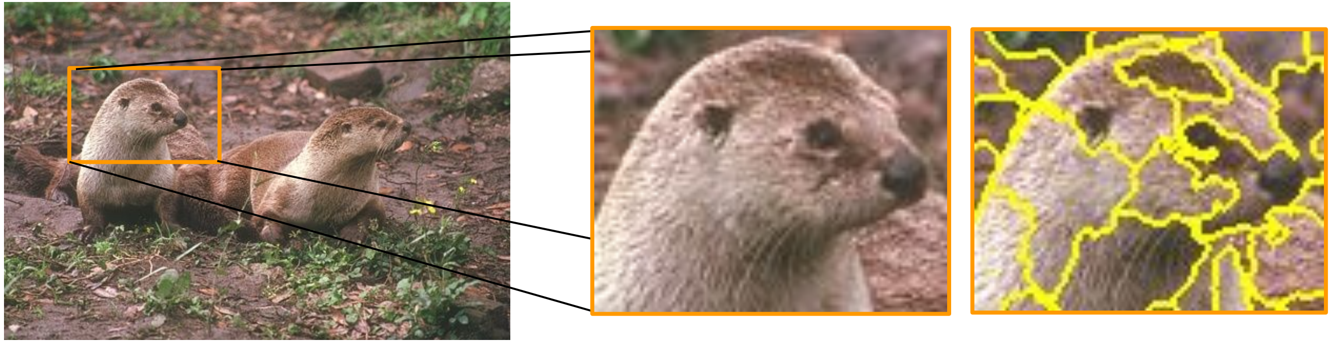

Images contain deformable boundaries
Selected a cluster of pixels


Images contain
deformable boundaries
Similar Pixels Values
Selected a cluster of pixels



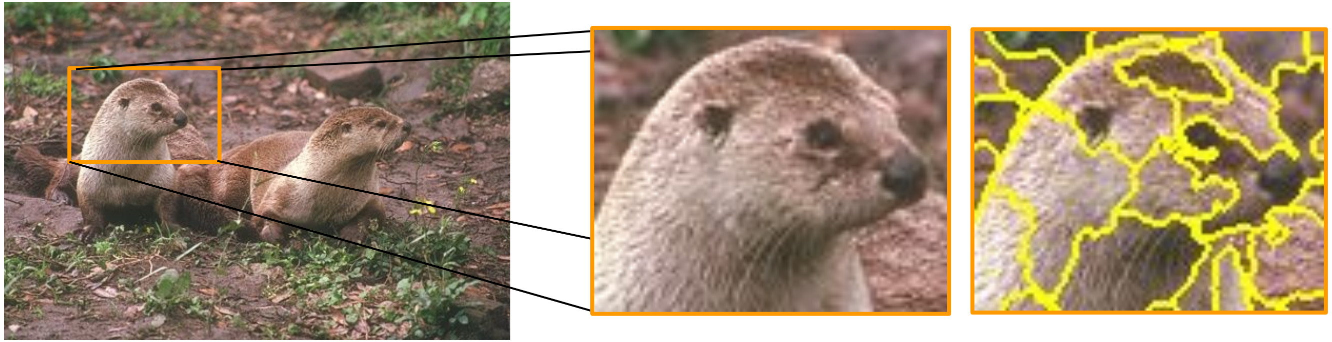
Images contain deformable boundaries





Similar Pixels Values

Oddly Shaped

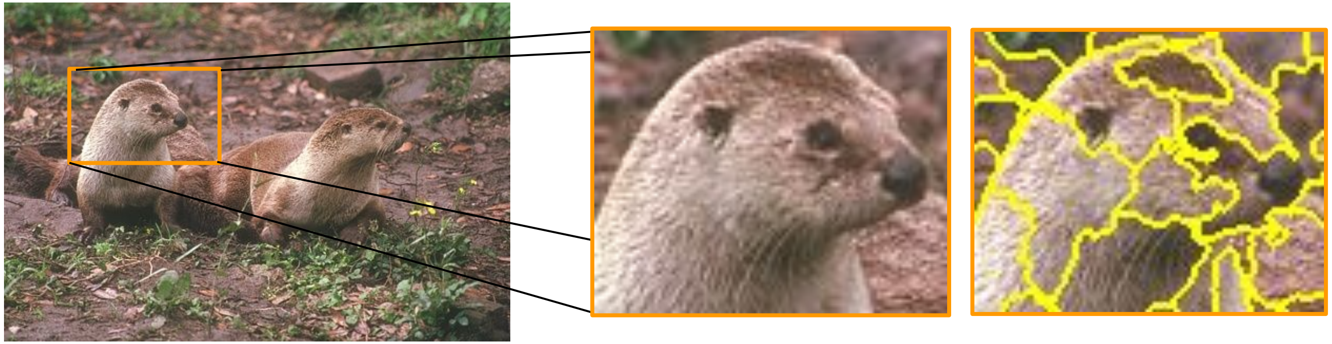


✓ Similar Pixels Values

✗ Oddly Shaped

Operators act on grids

Q1. How do we principally incorporate superpixels into DNN modules?



Similar Pixels Values

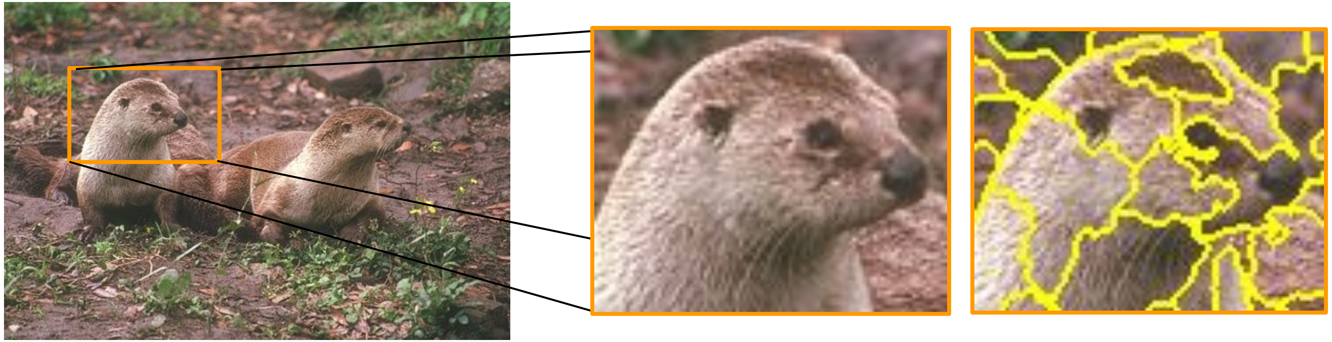
Oddly Shaped

Operators act on grids

Superpixels are inconsistent

Q1. How do we principally incorporate superpixels into DNN modules?

Q2. How do we address the inconsistency among superpixel assignments?



✔ Similar Pixels Values

✘ Oddly Shaped

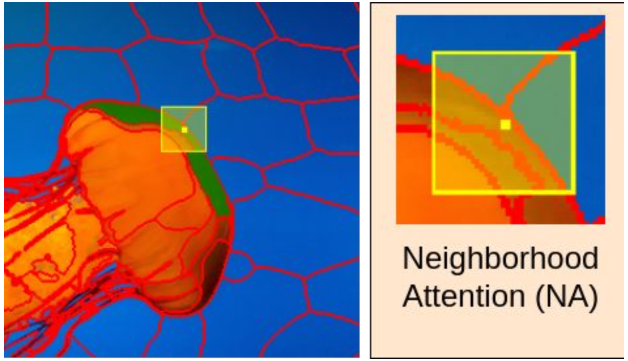
Operators act on grids

Superpixels are inconsistent

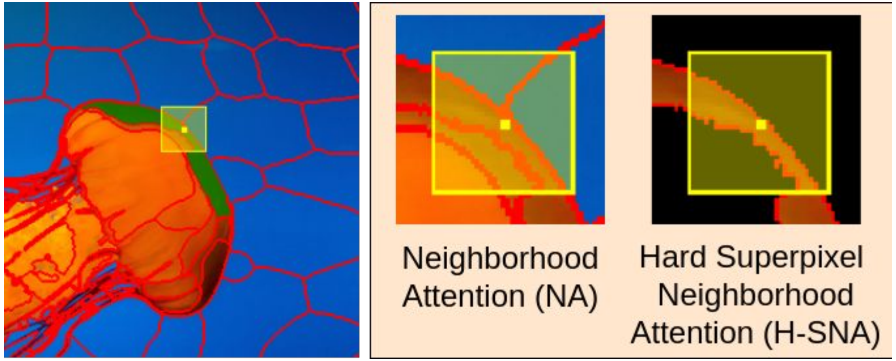
Q1. How do we principally incorporate superpixels into ~~DNN modules?~~

Neighborhood Attention

Q2. How do we address the inconsistency among superpixel assignments?



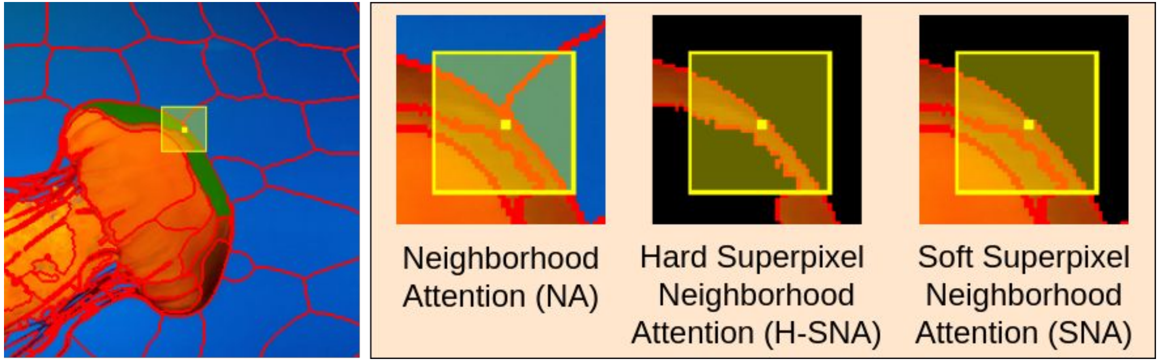
$$f_{\mathrm{NA}}^{(i)}(\boldsymbol{x}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\exp\left(\lambda_{\mathrm{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\mathrm{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$



$\hat{\boldsymbol{s}}$ Superpixel Assignment

$$f^{(i)}_{\text{NA}}(\boldsymbol{x}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$

$$f^{(i)}_{\text{H-SNA}}(\boldsymbol{x}; \hat{\boldsymbol{s}}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\mathbb{1}[\hat{s}_i = \hat{s}_j] \cdot \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \mathbb{1}[\hat{s}_i = \hat{s}_{j'}] \cdot \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$



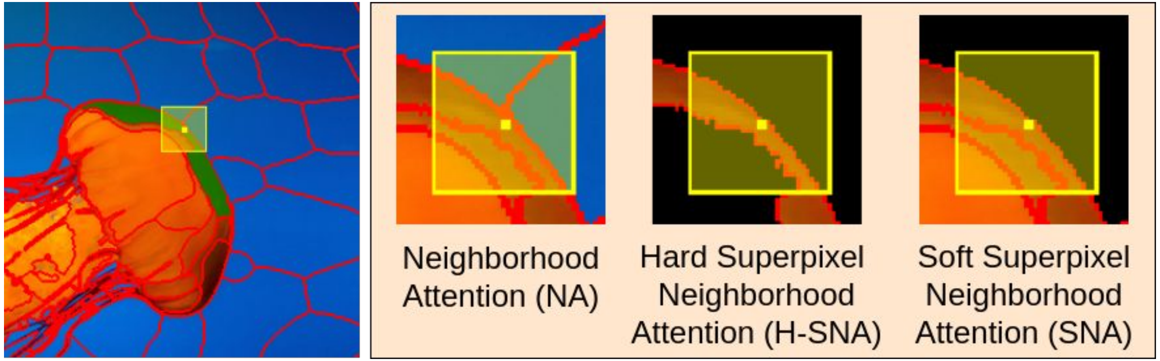
$\hat{\boldsymbol{s}}$ Superpixel Assignment

$\hat{\boldsymbol{\pi}}$ Superpixel Probabilities

$$f_{\text{NA}}^{(i)}(\boldsymbol{x}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$

$$f_{\text{H-SNA}}^{(i)}(\boldsymbol{x}; \hat{\boldsymbol{s}}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\mathbb{1}[\hat{s}_i = \hat{s}_j] \cdot \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \mathbb{1}[\hat{s}_i = \hat{s}_{j'}] \cdot \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$

$$f_{\text{SNA}}^{(i)}(\boldsymbol{x}; \hat{\boldsymbol{\pi}}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right) \sum_{s=1}^{N_{\text{sp}}} \hat{\boldsymbol{\pi}}^{(i,s)} \hat{\boldsymbol{\pi}}^{(j,s)}}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right) \sum_{s=1}^{N_{\text{sp}}} \hat{\boldsymbol{\pi}}^{(i,s)} \hat{\boldsymbol{\pi}}^{(j',s)}}$$



$\hat{\boldsymbol{s}}$ Superpixel Assignment

$\hat{\boldsymbol{\pi}}$ Superpixel Probabilities

$$f_{\text{NA}}^{(i)}(\boldsymbol{x}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right)}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right)}$$

$$f_{\text{H-SNA}}^{(i)}(\boldsymbol{x}; \hat{\boldsymbol{s}}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad w_{i,j} = \frac{\mathbb{1}[\hat{s}_i = \hat{s}_j] \cdot \exp\left(\lambda_{\text{at}} d \right.}{\sum_{j' \in \mathcal{N}(i)} \mathbb{1}[\hat{s}_i = \hat{s}_{j'}] \cdot \exp}$$

How did we get this?

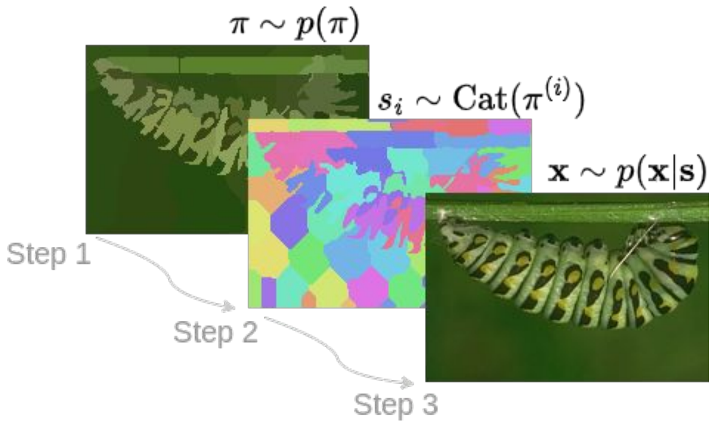
$$f_{\text{SNA}}^{(i)}(\boldsymbol{x}; \hat{\boldsymbol{\pi}}) = \sum_{j \in \mathcal{N}(i)} w_{i,j} \boldsymbol{v}_j, \qquad \boxed{w_{i,j} = \frac{\exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_j)\right) \sum_{s=1}^{N_{\text{sp}}} \hat{\boldsymbol{\pi}}^{(i,s)} \hat{\boldsymbol{\pi}}^{(j,s)}}{\sum_{j' \in \mathcal{N}(i)} \exp\left(\lambda_{\text{at}} d(\boldsymbol{q}_i, \boldsymbol{k}_{j'})\right) \sum_{s=1}^{N_{\text{sp}}} \hat{\boldsymbol{\pi}}^{(i,s)} \hat{\boldsymbol{\pi}}^{(j',s)}}}$$

Sample superpixel probabilities	$\boldsymbol{\pi} \sim p(\boldsymbol{\pi})$
Sample the superpixel assignment given the probabilities	$\boldsymbol{s}_i \| \boldsymbol{\pi}^{(i)} \sim p(\boldsymbol{s}_i \| \boldsymbol{\pi}^{(i)})$
Sample the image pixel given the superpixel assignment	$\boldsymbol{x} \| \boldsymbol{s} \sim p(\boldsymbol{x} \| \boldsymbol{s})$

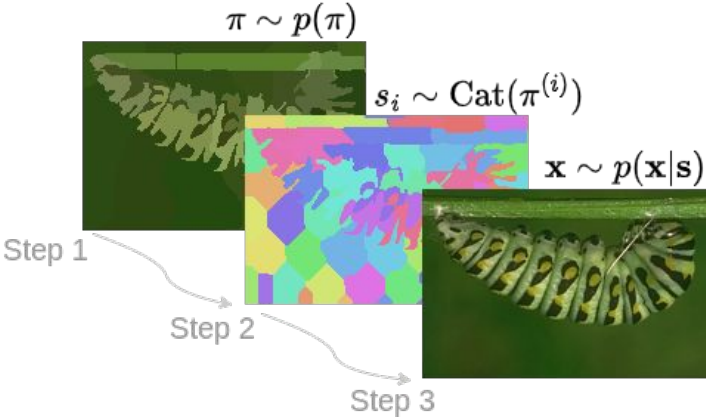
Sample superpixel probabilities $\boldsymbol{\pi} \sim p(\boldsymbol{\pi})$

Sample the superpixel assignment given the probabilities $\boldsymbol{s}_i | \boldsymbol{\pi}^{(i)} \sim p(\boldsymbol{s}_i | \boldsymbol{\pi}^{(i)})$

Sample the image pixel given the superpixel assignment $\boldsymbol{x} | \boldsymbol{s} \sim p(\boldsymbol{x} | \boldsymbol{s})$



Sample superpixel probabilities	$\boldsymbol{\pi} \sim p(\boldsymbol{\pi})$
Sample the superpixel assignment given the probabilities	$\boldsymbol{s}_i \vert \boldsymbol{\pi}^{(i)} \sim p(\boldsymbol{s}_i \vert \boldsymbol{\pi}^{(i)})$
Sample the image pixel given the superpixel assignment	$\boldsymbol{x} \vert \boldsymbol{s} \sim p(\boldsymbol{x} \vert \boldsymbol{s})$



$$f^{(i)}_{\mathrm{SNA}}(\tilde{\boldsymbol{x}}; \boldsymbol{\pi}) = D^*(\tilde{\boldsymbol{x}}_i; \sigma, \boldsymbol{\pi}) = \arg\min_D \left[\| D(\tilde{\boldsymbol{x}}_i; \boldsymbol{\pi}, \sigma) - \boldsymbol{x}_i \|^2 \right]$$

Gaussian Denoising

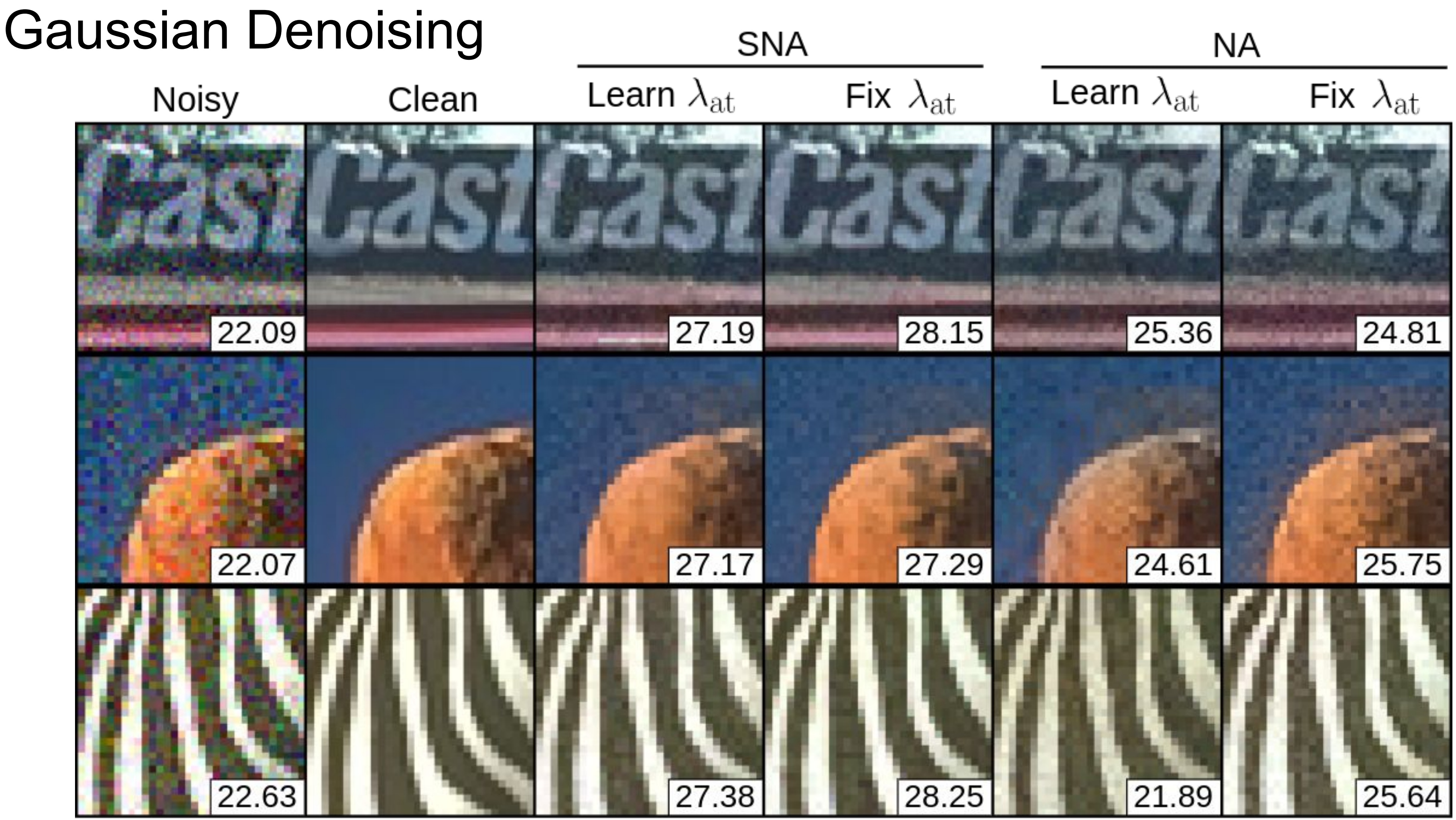
Attn.	SNA				H-SNA	NA [2]	
Learn λ_{at}	✓	✓				✓	
Sp. Model	$g_{\phi,\text{Deep}}$	$g_{\phi,\text{SLIC}}$	$g_{\phi,\text{Deep}}$	$g_{\phi,\text{SLIC}}$	$g_{\phi,\text{SLIC}}$		
σ	31.96	32.08	32.07	**32.19**	30.88	30.87	31.10
10	0.869	0.871	0.865	**0.871**	0.810	0.850	0.850
20	29.01	28.72	28.77	**29.08**	25.56	27.12	26.96
	0.838	0.815	0.819	0.804	0.630	0.774	0.743
30	**27.70**	26.94	27.25	27.51	22.37	25.69	24.91
	0.805	0.777	0.764	0.763	0.512	0.743	0.687
Deno Params (θ)	195	195	195	195	195	195	195
Aux Params (ϕ)	8.8k	8.8k	4.4k	4.4k	0	4.4k	0
Fwd Time (ms)	30.20	45.05	27.06	40.58	28.86	4.64	**2.08**
Bwd Time (ms)	38.72	80.93	40.00	51.35	32.54	6.06	**4.67**
Fwd Mem (GB)	1.90	2.30	1.87	2.28	1.96	0.23	**0.21**
Bwd Mem (GB)	3.27	3.68	3.25	3.66	3.13	0.27	**0.25**

$$\hat{\boldsymbol{y}}_{\text{Deno}} = \text{Simple Network}_{\theta,\phi}(\boldsymbol{x}) = f_{\text{Attn}}\left(\boldsymbol{x}\boldsymbol{W}_0, g_\phi(\boldsymbol{x}\boldsymbol{W}_0)\right)\boldsymbol{W}_1 + \boldsymbol{x}\boldsymbol{W}_0\boldsymbol{W}_1$$

Gaussian Denoising

Attn.	SNA				H-SNA	NA [2]	
Learn λ_{at}	✓	✓				✓	
Sp. Model	$g_{\phi,\text{Deep}}$	$g_{\phi,\text{SLIC}}$	$g_{\phi,\text{Deep}}$	$g_{\phi,\text{SLIC}}$	$g_{\phi,\text{SLIC}}$		
σ	31.96	32.08	32.07	**32.19**	30.88	30.87	31.10
10	0.869	0.871	0.865	**0.871**	0.810	0.850	0.850
20	29.01	28.72	28.77	**29.08**	25.56	27.12	26.96
	0.838	0.815	0.819	0.804	0.630	0.774	0.743
30	**27.70**	26.94	27.25	27.51	22.37	25.69	24.91
	0.805	0.777	0.764	0.763	0.512	0.743	0.687
Deno Params (θ)	195	195	195	195	195	195	195
Aux Params (ϕ)	8.8k	8.8k	4.4k	4.4k	0	4.4k	0
Fwd Time (ms)	30.20	45.05	27.06	40.58	28.86	4.64	**2.08**
Bwd Time (ms)	38.72	80.93	40.00	51.35	32.54	6.06	**4.67**
Fwd Mem (GB)	1.90	2.30	1.87	2.28	1.96	0.23	**0.21**
Bwd Mem (GB)	3.27	3.68	3.25	3.66	3.13	0.27	**0.25**

$$\hat{\boldsymbol{y}}_{\text{Deno}} = \text{Simple Network}_{\theta,\phi}(\boldsymbol{x}) = f_{\text{Attn}}\left(\boldsymbol{x}\boldsymbol{W}_0, g_\phi(\boldsymbol{x}\boldsymbol{W}_0)\right)\boldsymbol{W}_1 + \boldsymbol{x}\boldsymbol{W}_0\boldsymbol{W}_1$$


Gaussian Denoising
SNA
NA
Noisy
Clean
Learn λ_{at}
Fix λ_{at}
Learn λ_{at}
Fix λ_{at}
22.09
27.19
28.15
25.36
24.81
22.07
27.17
27.29
24.61
25.75
22.63
27.38
28.25
21.89
25.64