Train on {**person**, **motorbike**, **car**}

Step *t-1*

Image   Pixel-Level Label   Expected Output

Train on {**cow**}

Step *t*

Image   Image-Level Label   Expected Output

"**COW**"