1. L0 finds nonzero parameters. L1 measures the sparsity without considering the inter-element correlations. L2 can prevent overfitting a model
2. Since the images have sharp edges I would want to focus on capturing the edges so my formulation would be argmin $\int \| \nabla I(x) - f(x) \|^2_2 \, dx$
3. Since the images are wildlife, the edges aren't as important so my formulation would simply be p=Hf